National Logistics Hubs and the Geography of Supplier Networks: From a Market Integration Perspective

Tian Luo*

*School of Finance, Renmin University of China, Beijing 100872, China;

ARTICLE HISTORY
Compiled April 4, 2026

ABSTRACT
We exploit China's National Logistics Hub (NLH) construction program as a quasi-natural experiment to identify how logistics infrastructure shapes the geographic scope of firm-level supplier networks. Using a multi-period difference-in-differences framework applied to 3,001 A-share firm-year observations from CSMAR supply chain disclosures spanning 2012–2024, we find that NLH designation expands firms' weighted-average supplier distance by approximately 117.7 km. The effect survives a battery of robustness checks and endogeneity corrections. Mechanism analysis reveals three complementary channels: (1) supplier substitution, in which firms replace nearby high-cost suppliers with remote low-cost ones (fixed matching-cost channel); (2) share concentration, in which purchasing share concentrates toward remote efficient suppliers (variable transport-cost channel); and (3) vertical disintegration, in which firms increase outsourcing of previously in-house inputs (make-or-buy boundary channel). We additionally document significant policy accumulation effects and spatial spillover effects of the NLH program. Heterogeneity analysis shows private firms, non-high-tech firms, and firms with higher intermediate input dependence or stronger supply chain risk perception exhibit larger treatment effects. At the city level, NLH construction raises both the breadth and intensity of cross-provincial procurement links. Hub complementarity estimates reveal differentiated coordination effects: Core-type and Regional-type hub pairs exhibit the strongest bilateral trade linkages, yet Regional hubs exert weaker trade-driving effects than their Core counterparts.

KEYWORDS
National Logistics Hubs; supplier geography; supply chain; market integration

JEL No. H54, L23, D22

1. Introduction

China's logistics costs amount to roughly 14 percent of GDP—nearly double the share in the United States or Germany—a persistent gap that has drawn attention from both policymakers and economists [1]. A central source of this inefficiency is the geographic fragmentation of procurement: firms frequently source inputs from nearby but high-cost suppliers rather than from more productive suppliers located elsewhere in the country, because transport, search, and contracting frictions make remote sourcing prohibitively expensive. Reducing these frictions could fundamentally reshape how firms organize their supplier networks across space, with first-order consequences for production costs, firm competitiveness, and the allocation of resources within a large

and heterogeneous economy.

The spatial organization of supply chains is a first-order determinant of production costs, firm competitiveness, and regional development. A vast literature documents that trade costs (including transport, search, and contracting frictions) constrain the geographic range over which firms can profitably source inputs [2, 3, 4]. As governments worldwide invest in logistics infrastructure to reduce these frictions, a natural question arises: does such investment meaningfully reshape the geography of supplier networks, and through what channels?

This paper answers these questions using China's National Logistics Hub (NLH) construction program as a large-scale quasi-experiment. Launched under a 2019 Ministry of Transport policy directive and continued through successive annual designation rounds, the NLH program identifies and upgrades key intermodal logistics nodes (combining highway, rail, waterway, and air freight connectivity) in strategically selected cities. The program operates through two complementary dimensions. On the *hardware* side, it supports dedicated railway lines, multimodal freight facilities, and trunk-line transport infrastructure, substantially reducing firms' transport costs and lead times, with direct implications for input procurement, inventory management, and production organization. On the *software* side, it promotes logistics information platforms that facilitate upstream-downstream supply chain communication, lower search and matching frictions, and raise the efficiency of cross-regional supplier connections. Crucially, the program's eligibility and timing display quasi-random variation relative to pre-existing firm-level supply chains, enabling credible causal identification.

Our baseline analysis draws on a panel of 3,001 A-share firm-year observations constructed from CSMAR supply chain disclosure data spanning 2012–2024. The outcome variable is the *weighted-average supplier distance* (*Disw*), defined as the Haversine-distance-weighted average across a firm's top-five suppliers in a given year, where weights equal each supplier's procurement value share. This measure captures both the extensive margin (whether a firm sources from remote suppliers) and the intensive margin (how much it spends on them). For robustness checks and supplementary analyses, we additionally use 16.63 million procurement transaction records from China's public bidding transparency platform (1999–2025).

Our baseline multi-period difference-in-differences (DID) specification exploits staggered NLH designation across 127 cities and 6 cohorts (2019–2024). The main coefficient of interest, the effect of NLH designation on *Disw*, is estimated at approximately 1.177 (SE ≈ 0.453, $p < 0.01$), translating to a 117.7-km expansion in weighted-average supplier distance. Parallel-trends tests over five pre-treatment periods show no differential pre-trends, and a 500-iteration placebo test confirms that the effect is unlikely to be a statistical artifact. We further document significant policy accumulation effects (measured by cumulative hub counts within a city) and spatial spillover effects that extend to cities within 75 km and 75–150 km of designated hub cities.

We further document three economic mechanisms. First, NLH designation reduces fixed supplier matching costs, inducing firms to substitute nearby suppliers with more productive but remote ones, consistent with Antràs et al.'s [2017] fixed matching-cost framework. Second, reduced transport costs cause purchasing expenditure to concentrate toward efficient remote suppliers, consistent with the variable-cost gravity mechanism in Eaton and Kortum [5]. Third, lower contracting costs at logistics nodes ease the make-or-buy boundary, encouraging vertical disintegration and hence expansion of the upstream supplier pool, consistent with Williamson [2] and Fort [6].

Existing research on national logistics hubs has focused primarily on urban and regional dimensions, examining effects on city logistics resilience and hub-city network

competition, and on firm-level outcomes such as enterprise innovation performance. However, few studies approach the question from a supply chain perspective: how do logistics hubs, as core public infrastructure, reshape the geographic distribution of firms' supplier networks? The spillover effects and downstream economic consequences of such restructuring remain largely unexamined. A related body of literature on production networks [7, 8, 9, 4, 10] establishes that firms' output depends not only on internal factor endowments but critically on their supplier portfolio, and that changes in distribution conditions prompt firms to re-optimize supplier networks along spatial and trade-network margins [5]. These findings motivate a supply-chain perspective on NLH policy that the prior literature has not yet provided.

Our findings contribute to three strands of literature. First, we add to the infrastructure–trade nexus literature [11, 12, 13, 14, 15] by focusing on logistics hubs, a distinct form of infrastructure that operates through intermodal connectivity and third-party logistics services rather than pure route expansion. Second, we extend the firm-level supply chain literature [3, 4, 6] to a developing-country setting with a focus on domestic sourcing geography. Third, we contribute to the Chinese market integration literature [16, 17, 18] by showing that logistics infrastructure is a quantitatively important driver of cross-provincial supplier linkages.

The paper's contributions are threefold. *Theoretically*, we construct a procurement-sourcing model that explicitly distinguishes fixed matching costs from variable transport costs, and derive all three transmission channels (supplier substitution, share concentration, vertical disintegration) analytically within a unified framework. *Empirically*, this is the first study to bring logistics hub policy and firm-level supplier selection into a common empirical framework; identification relies on a multi-period difference-in-differences design validated against large-scale bidding data, propensity-score matching, a Bartik instrumental variable, and Heckman selection correction. *In terms of policy implications*, our results offer micro-level evidence that logistics infrastructure functions not merely as a cost-reduction channel but as an institutional mechanism for breaking local protectionism and expanding firms' access to the national market, with direct implications for supply chain resilience and domestic market integration.

The remainder of the paper proceeds as follows. Section 2 introduces the NLH policy background. Section 3 presents a theoretical framework with three impact channels. Section 4 describes the research design and data. Section 5 presents the empirical analysis, including baseline difference-in-differences estimates, parallel trends and placebo tests, robustness checks, and mechanism analysis. Section 6 presents further analysis covering heterogeneity effects, market integration consequences, and hub-type complementarity. Section 7 concludes with policy implications.

2. Policy Background

China's logistics sector has expanded substantially over the past two decades, yet structural challenges persist. Despite increased capacity, logistics infrastructure agglomeration effects and cross-regional spillovers remain limited. Fragmented infrastructure, inadequate long-distance facilities, and uneven interregional accessibility continue to generate frictions in the flow of production factors and in the geographic matching of supply and demand [19, 20].

To address these challenges, in December 2018 the National Development and Reform Commission (NDRC) and the Ministry of Transport jointly issued the *National*

Logistics Hub Layout and Construction Plan, which formally defined the concept and functional positioning of national logistics hubs (NLHs), designated 127 eligible cities, and classified hubs into six types: land-port, seaport, production-service, commerce-service, land-border-crossing, and airport hubs. The plan set a phased target of designating approximately 150 NLHs nationwide by 2025, marking the transition of China's logistics system construction from a "point-scattered" phase to a "corridor-and-network" phase.

The construction of NLHs exhibits a pattern of gradual expansion with successive waves of designation. Looking at successive designation batches, the 2019–2020 first two batches of 45 hubs were dominated by corridor-type (land-port and seaport) hubs, establishing the foundational skeleton of the national logistics network. The three batches of 2021–2023, comprising 80 hubs, progressively achieved full coverage of all 31 provincial-level regions nationwide, significantly improving the spatial equity of cross-regional logistics systems. Of the 56 newly added hubs in 2024–2025, production-service and commerce-service types together account for over 64% of the total. This gradual shift in designation trajectory marks the evolution of China's logistics hub operations from the "single-node construction" phase toward the "corridor-building" phase, and it is now crossing into the "network connectivity" stage [21].

This node-to-network hub layout not only physically connected China's national logistics corridors, but also produced far-reaching economic effects. On one hand, high-capacity hub nodes, through trunk-line transportation and multi-modal transfer, substantially reduced cross-regional commodity trade costs; on the other hand, supply chain management services at hubs effectively lowered the friction costs for firms to search, evaluate, and match suppliers. These combined effects—lower transport and matching costs, broader geographic coverage, and improved information coordination—constitute the economic channels through which the NLH program is expected to reshape firm-level supplier geography, the central question of this paper.

Data on NLH business revenues by region in the first half of 2025 show varied but collectively robust growth rates: central and northeastern regions grew the fastest, while eastern and western regions also maintained stable expansion. This demonstrates that NLH construction, by consolidating cross-regional infrastructure and bridging regional logistics capacity gaps, has provided solid physical infrastructure and service support for manufacturing firms to step beyond local markets and optimise supply chain networks nationally [22].

In summary, the staged, networked construction of national logistics hubs provides an ideal quasi-natural experimental setting for studying how hub construction reshapes the spatial distribution of supply chain networks.

3. Theoretical Model and Research Hypotheses

This section integrates the global sourcing model of Antràs et al. [3] with the production network model of Bernard et al. [4] to construct a theoretical framework incorporating heterogeneous firm–supplier matching. Within a unified framework of spatially varying fixed matching costs and variable transport costs, we introduce a supplier spatial density function to characterise the microeconomic mechanisms through which NLH construction expands firms' geographic supplier distribution.

3.1. Model Setup

Assume a closed economy with a representative consumer whose preferences are described by a standard CES utility function. Consumer demand for firm i's product is:

$$q_i = A p_i^{-\sigma}, \tag{1}$$

where p_i is the price of firm i's product, A denotes the market expenditure level (taken as exogenous), and $\sigma > 1$ is the elasticity of substitution between products.

Following Melitz [23], firm i has core productivity φ_i. The firm integrates a unit continuum of intermediate goods $\omega \in [0, 1]$ to produce final output Q_i, using firm-specific input quantity $x_i(\omega)$. The production function is:

$$Q_i = \varphi_i \exp\left(\int_0^1 \ln x_i(\omega)\, d\omega \right). \tag{2}$$

The firm's suppliers are distributed in geographic space, described by a spatial density function $g(d) \geq 0$ at distance d. Each supplier draws idiosyncratic productivity z independently from a Fréchet distribution:

$$F(z) = e^{-T_0\, z^{-\theta}}, \tag{3}$$

where $T_0 > 0$ parameterises the overall technology level of the supplier population, and $\theta > 1$ governs productivity dispersion.

The firm sources intermediate good ω from a supplier located at distance d with productivity z. Logistics infrastructure L enters through two cost components. First, the *fixed matching cost* of establishing a supplier relationship:

$$f(d, L) = \mu(L)\, d, \qquad \mu'(L) < 0, \tag{4}$$

where $\mu(L)$ is the per-unit-distance matching cost; logistics improvement (e.g., information platform adoption) reduces μ. Second, the *variable transport cost* follows the iceberg formulation of Samuelson [24]:

$$\tau(d, L) = 1 + \kappa(L)\, d, \qquad \kappa'(L) < 0, \tag{5}$$

where $\kappa(L)$ is the per-unit-distance transport cost coefficient; hub construction (e.g., highway improvement) reduces κ. The effective unit procurement price from a supplier at distance d with productivity z is therefore:

$$p(d, z) = \frac{w}{z}\, \tau(d, L), \tag{6}$$

where w is the supplier-side factor price, normalised to 1 for simplicity. The firm's total cost function is:

$$TC_i = \int_0^1 p_i(\omega)\, x_i(\omega)\, d\omega + \int_0^{d_i^*} \mu(L)\, s\, ds, \tag{7}$$

where the first term is variable procurement cost and the second is the total fixed matching cost incurred over the entire search range $[0, d_i^*]$.

The firm optimally selects the minimum-price supplier for each intermediate good ω. By the extreme-value properties of the Fréchet distribution,[1] the firm's *purchasing ability* is $\Phi_i = \int_0^{d_i^*} T_0\, g(s)\, [\tau(s, L)]^{-\theta}\, ds$, and the intermediate price index is:

$$P_i = \gamma\, \Phi_i^{-1/\theta}, \tag{8}$$

where γ is a constant determined by distributional parameters. P_i can be interpreted as the unit intermediate input cost for firm i's final production, which is tightly linked to marginal cost $c_i = P_i/\varphi_i$. Under monopolistic competition, $p_i = [\sigma/(\sigma-1)]\, c_i$.

Firm i's profit function is:

$$\pi_i(\varphi_i, D_i) = \frac{A}{\sigma}\left(\frac{\sigma\gamma}{\sigma-1}\right)^{1-\sigma} \varphi_i^{\sigma-1}\, \Phi_i^{(\sigma-1)/\theta} - \int_{s\in D_i} f(s, L)\, ds, \tag{9}$$

where the first term is variable profit and the second is total fixed matching costs over search set D_i. The firm's objective is to choose $D_i^* = [0, d_i^*]$ to maximise profit. Concavity of the profit function implies the optimal search set is a contiguous interval from zero to some optimal boundary d_i^*, determined by the first-order condition:

$$\frac{\partial \pi_i}{\partial d_i^*} = 0 \implies \frac{A(\sigma-1)}{\sigma\theta}\left(\frac{\sigma\gamma}{\sigma-1}\right)^{1-\sigma} \varphi_i^{\sigma-1}\, \Phi_i^{(\sigma-1)/\theta-1} T_0\, g(d_i^*)\, [\tau(d_i^*, L)]^{-\theta} = f(d_i^*, L). \tag{10}$$

This condition equates the marginal revenue of expanding search (left) to the marginal fixed matching cost (right). Rewriting in compact form as:

$$MB(d_i^*) = \mu\, d_i^*, \tag{11}$$

where $MB(d_i^*)$ collects the left-hand terms of (10). The optimal search boundary d_i^* is a function of firm productivity φ_i and logistics level L, endogenously determining the geographic scope of the firm's supplier network.

3.2. Impact Mechanisms of Logistics Hub Construction

We now conduct comparative statics to analyse how NLH construction (modelled as an improvement in L) affects firms' optimal sourcing decisions and, through those decisions, the geographic distribution of suppliers.

Channel 1: Fixed Matching Cost Reduction Expands Geographic Search Range. For simplicity, hold variable transport costs fixed at κ. Consider the effect of μ alone. Rewrite the first-order condition (11) as $H(d_i^*, \mu) = MB(d_i^*) - \mu\, d_i^* = 0$. By the implicit function theorem:

$$\frac{\partial d_i^*}{\partial \mu} = -\frac{\partial H/\partial \mu}{\partial H/\partial d_i^*} = \frac{-d_i^*}{\partial H/\partial d_i^*}. \tag{12}$$

The profit-maximising second-order condition requires $\partial^2 \pi_i/(\partial d_i^*)^2 < 0$, which implies

[1] Space constraints preclude the full derivation; see Online Appendix A.

$\partial H/\partial d_i^* < 0$. Hence $\partial d_i^*/\partial \mu < 0$. Since logistics improvement means $\mu'(L) < 0$:

$$\frac{\partial d_i^*}{\partial \mu} \cdot \mu'(L) > 0. \tag{13}$$

This result shows that logistics infrastructure improvement, by reducing fixed matching costs, lowers the cost barrier for firms to enter remote-geography markets, thereby incentivising firms to expand their geographic search range.

As d_i^* rises, firms begin searching regions they previously did not reach. In these new regions, firms may discover highly productive suppliers drawn from the Fréchet distribution's upper tail. Specifically, as the search boundary expands beyond d_i^*, a newly discovered supplier located remotely beyond d_i^* with high efficiency $z > d_i^{*\text{old}}$ can offer a unit price $p(d, z)$ lower than the firm's previous best. At this point the firm substitutes suppliers, replacing nearby incumbents with the newly found remote ones.

To characterise how this boundary expansion affects the weighted-average distance $Y_i = \int_0^{d_i^*} d \cdot \lambda_i(d)\, ds$ (with $\lambda_i(d)$ denoting the purchasing share at distance d and $\Phi_i = \int_0^{d_i^*} V(s, \kappa)\, ds$, where $V(d, \kappa) \equiv T_0 g(d)[1+\kappa d]^{-\theta}$), we differentiate Y_i with respect to d_i^* by Leibniz's rule:

$$\frac{dY_i}{dd_i^*} = \frac{d_i^* \cdot V(d_i^*, \kappa)}{\Phi_i} + \int_0^{d_i^*} d \cdot \frac{\partial(1/\Phi_i)}{\partial d_i^*} V(d, \kappa)\, ds. \tag{14}$$

Since $\partial \Phi_i/\partial d_i^* = V(d_i^*, \kappa)$, we have $\partial(1/\Phi_i)/\partial d_i^* = -V(d_i^*, \kappa)/\Phi_i^2$. Substituting and simplifying:

$$\frac{dY_i}{dd_i^*} = \frac{V(d_i^*, \kappa)}{\Phi_i} \left[d_i^* - \int_0^{d_i^*} \frac{d \cdot V(d, \kappa)}{\Phi_i}\, ds \right] = \frac{V(d_i^*, \kappa)}{\Phi_i} \big(d_i^* - Y_i \big) > 0, \tag{15}$$

where the last inequality holds because d_i^* is the upper bound of the search support, while Y_i is the weighted average over that support, so $d_i^* > Y_i$ strictly. Combining with $\partial d_i^*/\partial \mu < 0$ and $\mu'(L) < 0$:

$$\left. \frac{dY_i}{dL} \right|_{\kappa \text{ fixed}} = \frac{dY_i}{dd_i^*} \cdot \frac{\partial d_i^*}{\partial \mu} \cdot \mu'(L) > 0. \tag{16}$$

This proves that logistics infrastructure improvement, via fixed matching cost reduction, drives search boundary expansion and thereby raises firms' weighted-average supplier distance.

Channel 2: Variable Transport Cost Reduction Concentrates Purchasing Shares toward Remote Suppliers. Following Eaton and Kortum [5] and Antràs et al. [3], the firm's purchasing expenditure share allocated to suppliers at distance d within its search range D_i^* equals the probability of obtaining the lowest-cost supply from that distance. Using Fréchet properties, this share is:

$$\lambda_i(d) = \frac{T_0\, g(d)\, [\tau(d, L)]^{-\theta}}{\Phi_i}. \tag{17}$$

This expression clearly shows how the firm allocates purchases across geographic locations. Now, holding d_i^* fixed, consider how $\lambda_i(d)$ changes as κ decreases (transport

improves):

$$\frac{\partial \lambda_i(d)}{\partial \kappa} = \lambda_i(d) \left[\frac{\partial \ln V(d,\kappa)}{\partial \kappa} - E_s\left[\frac{\partial \ln V(s,\kappa)}{\partial \kappa} \right] \right], \tag{18}$$

where $E_s[\cdot]$ denotes the expectation weighted by $\lambda_i(s)$. The gradient of log-value with respect to κ is:

$$\frac{\partial \ln V(d,\kappa)}{\partial \kappa} = -\frac{\theta d}{1+\kappa d}. \tag{19}$$

Define $h(d) = -\theta d/(1+\kappa d)$. It is straightforward to verify that $h(0) = 0$ and $h'(d) < 0$, i.e., h is strictly decreasing in d. By the property of the strictly monotone function, there exists a unique $d_0 \in (0, d_i^*)$ such that $h(d_0) = E_s[h(s)]$. From equation (18), when $d > d_0$, $h(d) < E_s[h(s)]$, so $\partial \lambda_i(d)/\partial \kappa < 0$, meaning the share of remote suppliers *rises* when κ falls. Conversely when $d < d_0$, the nearby share falls. The re-allocation therefore systematically shifts purchasing weight from nearby to remote suppliers.

To establish the net effect on Y_i (holding d_i^* fixed):

$$\left.\frac{\partial Y_i}{\partial \kappa}\right|_{d_i^* \text{ fixed}} = \int_0^{d_i^*} d \cdot \frac{\partial \lambda_i(d)}{\partial \kappa}\, ds = E_s\big[d \cdot h(d)\big] - E_s[d] \cdot E_s[h(s)] = \mathrm{Cov}_s(d,\, h(d)). \tag{20}$$

Since $h(d)$ is strictly decreasing in d, the covariance $\mathrm{Cov}_s(d, h(d)) < 0$, and therefore $\partial Y_i/\partial \kappa < 0$. Since $\kappa'(L) < 0$:

$$\left.\frac{\partial Y_i}{\partial L}\right|_{d_i^* \text{ fixed}} = \frac{\partial Y_i}{\partial \kappa} \cdot \kappa'(L) > 0. \tag{21}$$

This proves that, even without changing the search boundary, variable transport cost reduction alone systematically reweights purchasing toward remote, efficient suppliers, thereby raising Y_i.

Channel 3: External Transaction Cost Reduction Raises Vertical Specialisation. Consider the firm's vertical organisation decision. Following Fort [6], assume the firm can either self-produce or outsource each intermediate good. Incorporating the self-production option, the total cost function becomes:

$$TC_i = \int_0^{i^*} \theta\, t\, x_i(t)\, dt + \int_{i^*}^1 p_i(\omega)\, x_i(\omega)\, d\omega + \int_0^d \mu(L)\, s\, ds + F(i^*, L), \tag{22}$$

where $\theta\, t$ is the unit self-production cost for input t (increasing in the input index), i^* is the threshold intermediate good below which the firm self-produces and above which it outsources, and $F(i^*, L)$ represents administrative control costs for managing in-house production. The firm sets i^* by the first-order condition:

$$\big(\theta i^* - p_i(i^*)\big)\, x_i(i^*) + \frac{\partial F(i^*, L)}{\partial i^*} = 0. \tag{23}$$

Since $\partial^2 F(i^*, L)/\partial i^* \partial L \approx 0$ (external transaction cost improvements affect sourcing costs but not administrative control costs directly), and substituting equations (5)

and (6), external procurement price $p_i(i^*)$ decreases as L improves. This implies i^* increases with L: more goods are outsourced. As vertical specialisation rises, the firm's external supplier pool expands, and by the gravity logic of equation (17), this disproportionately adds remote suppliers, further expanding Y_i.

A crucial feature of this channel concerns the role of supply chain stickiness. The sticky production network theory [25] posits that relational frictions between firms and their incumbent suppliers — such as contractual lock-in, trust-building costs, and asset-specific investments — cause existing supplier relationships to persist even in the face of external shocks, dampening the geographic reallocation of procurement. However, the new suppliers that arise from vertical disintegration are drawn fresh from the full distribution $G_i(p) = 1 - e^{-\Phi_i p^{\theta}}$ and are therefore *not* subject to these incumbent stickiness forces. As a result, when a firm expands its outsourcing boundary, the newly added external suppliers are selected purely on cost grounds, disproportionately attracting high-efficiency, remote suppliers. This stickiness-free selection mechanism ensures that vertical disintegration reliably expands Y_i, complementing the boundary-expansion effect derived above.

3.3. Research Hypotheses

The three core mechanisms above jointly show that NLH construction, through lowering fixed matching costs, variable transport costs, and external transaction costs, systematically raises firms' weighted-average supplier distance Y_i. The overall tendency is from "nearby" to "remote." Therefore, we get the mechanism diagram Figure 1.

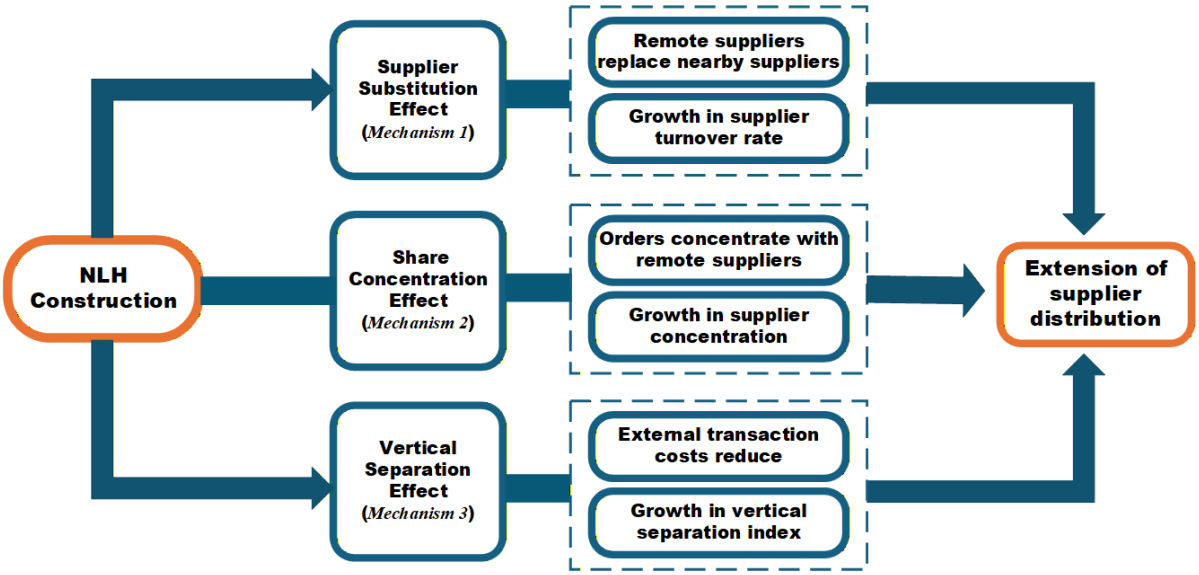


Figure 1. Mechanism Diagram

We therefore propose:

Hypothesis 1: *NLH construction significantly expands the geographic distribution of firm suppliers.*

From the three-channel mechanism, we derive the following subsidiary hypotheses:

Hypothesis 2: *NLH construction expands firms' geographic supplier distribution*

through improved supplier matching quality (supplier substitution effect).

Hypothesis 3: *NLH construction expands firms' geographic supplier distribution through increased purchasing share concentration toward remote efficient suppliers (share concentration effect).*

Hypothesis 4: *NLH construction expands firms' geographic supplier distribution through increased vertical specialisation and outsourcing (vertical disintegration effect).*

4. Research Design

4.1. Sample Selection and Data Sources

This paper uses A-share listed firms in China from 2012 to 2024 as the sample. Firm basic information and supply chain data come from the CSMAR (China Stock Market and Accounting Research) database. Data on NLH policy implementation are manually compiled from official policy documents. City-level macro controls come from the *China Statistical Yearbook* and *China City Statistical Yearbook*. Based on these raw data, we exclude financial-sector firms, firms with ST or *ST status during the sample period, and observations with missing key variables; we further winsorise all continuous variables at the 5% level. The final sample contains 3,001 firm-year observations.

4.2. Model Design

We use a multi-period DID design to estimate the effect of NLH policy on firms' supplier geographic distribution. The model is:

$$Disw_{i,t} = \alpha + \beta_1 \, DID_{c,t} + \gamma \, Controls_{c,t} + \mu_i + \delta_t + \varepsilon_{i,t}, \tag{24}$$

where $Disw_{i,t}$ is firm i's weighted-average supplier distance in year t. The key explanatory variable $DID_{c,t}$ is a dummy variable that equals 1 for firms whose city c has been included in the national NLH designation list in year t or earlier, and 0 otherwise. β_1 is the core coefficient of interest; under Hypothesis 1 it should be positive, reflecting that NLH construction expands the geographic distribution of firms' suppliers. μ_i and δ_t denote firm and year fixed effects, respectively. $\varepsilon_{i,t}$ is the error term.

4.3. Variable Definitions

(1) Dependent Variable. Following Lu et al. [26] and Tian et al. [27], we construct the supplier weighted-average distance index *Disw* as a proxy for the geographic distribution of firm suppliers, measuring the spatial distribution of a firm's top-five suppliers:

$$Disw_{i,t} = \sum_j Dis_{i,j,t} \times Ratio_{i,j,t}, \tag{25}$$

where $Dis_{i,j,t}$ denotes the distance between firm i and its supplier j in year t, and $Ratio_{i,j,t}$ denotes supplier j's share of firm i's total procurement value in year t. This index captures two dimensions simultaneously: a higher absolute distance to suppliers reflects the extensive-margin expansion of geographic sourcing; a higher purchasing

share to remote suppliers reflects the intensive-margin deepening. Both channels are consistent with the theoretical predictions of equations (16) and (21).

(2) Explanatory Variable. Consistent with the multi-period DID design, the explanatory variable $DID_{c,t}$ equals 1 if firm i's city c is included in the national NLH designation list in year t or any prior year, and 0 otherwise.

(3) Control Variables. Following Tian et al. [27], we select firm-level controls including: return on equity (*Roe*), return on assets (*Roa*), asset-liability ratio (*Lev*), number of employees (*Emp*), accounts receivable turnover (*Art*), log board size (*Board*), independent director ratio (*Indep*), digital transformation index (*Dig*), and fixed-asset ratio (*Fir*). We additionally include the three-year rolling industry-average ROA (*Indroe*) to absorb industry-level confounds. City-level controls include regional development level (*Dev*, GDP per capita), industrial structure (*Ind*, secondary-industry share), and log urban population (*Pop*). Descriptive statistics are presented in Table 1.

Table 1. Descriptive Statistics

VarName	*N*	Mean	SD	Min	Median	Max
Disw	3,001	7.966	6.119	0.000	7.183	20.458
Roe	3,001	0.040	0.042	−0.049	0.037	0.128
Roa	3,001	0.075	0.083	−0.127	0.073	0.234
Emp	3,001	9170.729	12677.344	345.000	3899.000	49890.000
Lev	3,001	0.485	0.186	0.157	0.490	0.788
Art	3,001	45.288	109.348	1.402	6.150	463.960
Board	3,001	2.162	0.171	1.792	2.197	2.485
Indep	3,001	0.368	0.040	0.333	0.364	0.444
Dig	3,001	1.472	1.296	0.000	1.386	4.043
Fir	3,001	0.226	0.155	0.020	0.197	0.570
Indroe	3,001	0.027	0.066	−0.124	0.039	0.131
Dev	3,001	1282.286	994.002	166.733	1021.447	3701.710
Ind	3,001	0.530	0.090	0.368	0.531	0.685
Pop	3,001	6.578	0.593	5.426	6.659	7.478

5. Empirical Analysis

5.1. Baseline Regression Analysis

5.1.1. Baseline DID Results

Table 2 reports the baseline regression results of the effect of NLH construction on firms' supplier geographic distribution. As shown in column (1), without adding any control variables, the estimated coefficient of NLH construction (*DID*) on supplier geographic distribution is 1.100 and is significant at the 5% level. To ensure robustness, we progressively add control variables. In columns (2)–(4), the core explanatory variable *DID* remains consistently positive and significant. The column (4) result shows that the *DID* coefficient is 1.177 and is significant at the 1% level, indicating that, holding other conditions constant, NLH designation raises a firm's weighted-average

supplier distance by approximately 117.7 km. NLH construction improved the logistics infrastructure of hub cities, reduced transport and matching costs, and enabled firms to search for better suppliers over a wider geographic range, thereby expanding the geographic distribution of firm supply chains, validating Hypothesis 1.

5.1.2. Parallel Trends Test

A key prerequisite for evaluating policy effects using DID is that treated and control groups must exhibit parallel pre-treatment trends, i.e., the parallel trends assumption. To validate this assumption and further examine the dynamic treatment effects of NLH construction, we construct a dynamic effects model [28], omitting the one-period-prior baseline and plotting parallel trends graphs (Figure 4). The estimated coefficients over the five pre-treatment periods (-5 to -2; period -1 omitted as baseline) all have 95% confidence intervals that include zero, with no systematic divergence between treated and control firms' supplier distance. This confirms that the DID model's key assumption is satisfied.

Furthermore, the dynamic coefficient estimates are monotonically increasing from the policy implementation year (period 0) onward over the subsequent four periods, with confidence intervals lying entirely above zero. This shows that NLH construction has a significant and lasting positive effect on expanding firms' supplier geographic scope over a certain time horizon.

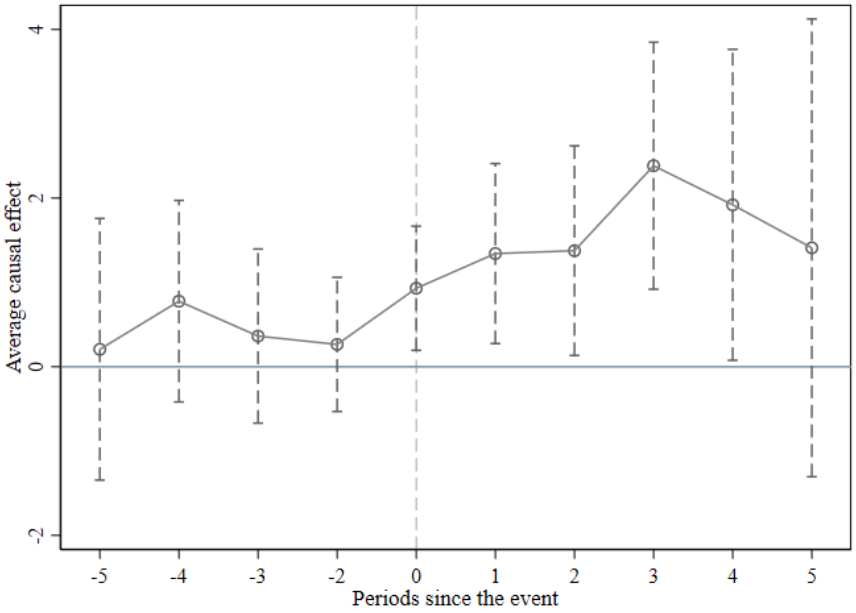


Figure 2. Parallel Trends Test

5.1.3. Placebo Test

To further rule out the possibility that unobservable omitted variables or random factors drive the estimated results, we construct a random policy-assignment placebo test. Figure 5 reports estimated coefficients from 500 random-assignment iterations. The distribution of these coefficient estimates is approximately centred at zero, only a very small fraction of placebo coefficients are significant at the 10% level, and all estimates lie to the left of the grey reference line corresponding to the baseline estimate.

This indicates that the other omitted variables do not affect the research conclusions of this paper.

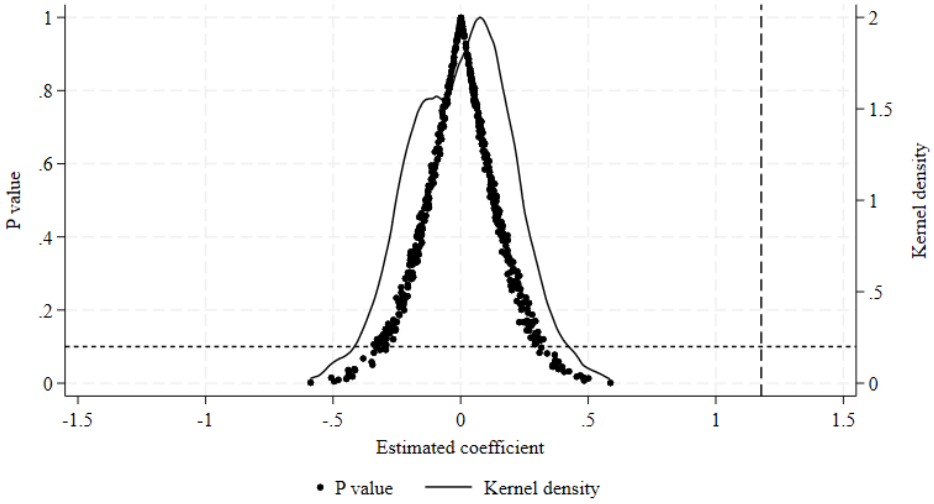


Figure 3. Placebo Test.

Notes: Distribution of estimated coefficients from 500 random policy-assignment iterations. The dashed vertical line marks the baseline DID estimate; the distribution is approximately centred at zero.

5.1.4. Policy Intensity and Spatial Spillover Test

Logistics infrastructure resource allocation is determined not only by whether a city has entered the policy treatment list, but also by the density of resources invested within that policy framework. To this end, we treat the hub designation count as a continuous variable representing policy treatment intensity (*Intensity*). Each additional NLH node designated to a city generates incremental effects on top of the baseline treatment, further expanding firms' supplier geographic distribution. Table 3 column (1) shows the regression results for policy intensity effects. The results show that the policy tool has a significant positive additive effect.

Furthermore, to test for spatial spillovers, following Coşar and Demir [29], we construct spatial ring variables based on the baseline model. *Ring_1* indicates whether the firm's city falls within the 0–75 km inner ring of a hub city, and *Ring_2* indicates whether it falls within the 75–150 km outer ring. The results in column (2) show that the first-ring variable (*Ring_1*) has an estimated coefficient of 2.315, which is significantly positive. The second-ring variable (*Ring_2*) has a negative but statistically insignificant coefficient. This indicates that the 0–75 km cities are notably affected by the policy spillover effect, suggesting that the hub city's logistics dispersion and trunk transport efficiency effectively reduces the overall logistics costs of adjacent regions, driving nearby related firms to similarly expand their supplier geographic scope. However, this network spillover dividend has an evident diminishing boundary.

Table 2. Baseline Regression Results

VARIABLES	(1) Base_Reg	(2) Base_Reg	(3) Base_Reg	(4) Base_Reg
DID	1.100**	1.169***	1.172***	1.177***
	(0.453)	(0.449)	(0.449)	(0.453)
Roe		7.358	7.233	6.659
		(10.277)	(10.246)	(10.247)
Roa		−2.737	−2.484	−2.182
		(4.490)	(4.444)	(4.470)
Lev		−1.249	−1.314	−1.474
		(1.479)	(1.471)	(1.455)
Indep		3.345	3.352	2.715
		(5.644)	(5.648)	(5.761)
Art		−0.000	−0.000	−0.001
		(0.003)	(0.003)	(0.003)
Board		0.389	0.423	0.367
		(1.448)	(1.450)	(1.480)
Emp		−0.000	−0.000	−0.000
		(0.000)	(0.000)	(0.000)
Dig		−0.086	−0.087	−0.072
		(0.188)	(0.188)	(0.192)
Fir		1.919	1.915	1.915
		(2.016)	(2.010)	(2.011)
Indroe			−1.397	−1.468
			(1.769)	(1.768)
Dev				0.000
				(0.001)
Ind				6.602
				(6.697)
Pop				−1.489
				(2.024)
Year FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
N	3,001	3,001	3,001	3,001
Constant	7.712***	5.913	5.899	12.115
	(0.110)	(4.697)	(4.696)	(12.950)
Adjusted R^2	0.748	0.749	0.749	0.749

Notes: Standard errors clustered at the firm level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Same applies to all subsequent tables.

Table 3. Policy Intensity and Spatial Spillover Effects

VARIABLES	(1) Policy Intensity	(2) Spatial Spillover
Intensity	0.585**	
	(2.17)	
DID		1.314**
		(2.34)
Ring_1		2.315**
		(2.49)
Ring_2		−0.846
		(−1.47)
Controls	YES	YES
Firm FE	YES	YES
Year FE	YES	YES
N	2,507	2,507
Adjusted R^2	0.749	0.752

Notes: Intensity is the cumulative number of NLH designations received by city c by year t. *Ring_1* (*Ring_2*) indicates whether the firm's city is within 0–75 km (75–150 km) of the nearest NLH-designated city.

5.2. Robustness Checks

Table 4 reports robustness results. To rule out potential confounds from firms' financial condition, we add the net cash flow ratio (*CASH*) as an additional control in column (1); the *DID* coefficient remains significantly positive. Column (2) uses PSM-DID: following Heckman et al. [30] and Wooldridge [31], we construct a quasi-natural framework by taking the DID treatment variable as the matching treatment variable and matching on the control variables in the baseline regression; after matching, the standard deviation of the matched covariates falls substantially and the parallel-trends assumption holds better. Re-estimating on the matched sample gives a *DID* coefficient that remains significantly positive, substantially alleviating concerns about sample self-selection bias and confirming that the effect of NLH construction on the expansion of firms' supplier geographic distribution is robust. To further mitigate endogeneity from reverse causality, column (3) introduces the one-period lag of the NLH policy variable; the lagged *DID* coefficient is still significantly positive, ruling out reverse causation and confirming robustness. Column (4) re-estimates using micro-level bidding data: following Lu et al. [26], we systematically collected approximately 16.63 million raw bidding transaction records from the public bidding transparency platform (1999–2025), matched them to a business registration database and city-level macro databases, and constructed a firm-level panel dataset of approximately 745,000 observations. The core conclusion remains robust.

Table 5 reports three further identification strategies. Column (1) applies Double/Debiased Machine Learning [32]: using the Lasso cross-fitting algorithm, we further introduce squared terms of all control variables on top of the baseline control set to capture potential nonlinear relationships and residual omitted variables through flexible functional-form estimation; DDML executes 3-fold cross-validation with 3 repetitions. The *DID* coefficient (1.255, significant at 1%) confirms baseline robustness and rules out high-order confounding interactions. Column (2) employs a Bartik shift-

Table 4. Robustness Check Results 1

VARIABLES	(1) Add. Controls	(2) PSM-DID	(3) Lag 1 Period	(4) Bidding Data
DID	1.1462**	1.1404**	1.1350*	7.6680*
	(2.5300)	(2.4017)	(1.8045)	(3.0430)
CASH	0.0000			
	(1.0054)			
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
N	2,507	2,199	1,160	257,291
Adjusted R^2	0.6572	0.6746	0.7452	0.2930

share instrument, following Goldsmith-Pinkham et al. [33]: $IV_{ct} = RLD_{c0} \times P_t$, where RLD_{c0} is the 1993 city railway density and P_t is the annual national NLH designation count. The exclusion restriction is supported on two grounds: first, 1993 railway density predates the sample period by nearly two decades and, conditional on contemporary infrastructure controls, has no plausible direct channel to 2012–2024 firm-level supplier distance; second, the national designation count P_t is determined by central government planning at the aggregate level and is not influenced by the procurement decisions of any individual firm or city, ensuring the macro shift component is exogenous to local economic conditions. The Kleibergen–Paap rk LM test p-value equals 0.000 and the Kleibergen–Paap Wald rk F-statistic far exceeds the Stock–Yogo critical value, confirming no weak instrument problem. Column (3) uses the Heckman two-stage approach to test for sample-selection bias arising from whether firms choose to disclose supplier information; the inverse Mills ratio (imr) in the second stage is not significant, providing no evidence of sample selection bias, confirming the robustness of estimation results.

Table 5. Robustness Check Results 2

VARIABLES	(1) DDML	(2) Bartik IV	(3) Heckman
DID	1.255***	2.753*	2.472**
	(0.33)	(1.92)	(2.05)
imr			−2.861
			(−1.03)
Controls (linear)	YES	YES	YES
Controls (quadratic)	YES	NO	NO
Year FE	YES	YES	YES
Firm FE	YES	YES	YES
N	3,001	2,498	2,179

5.3. Mechanism Analysis

The preceding analysis confirms that NLH construction can reduce cross-regional transport and logistics costs, and reshape firms' supplier geographic distribution. To

test the three transmission mechanisms, we follow a two-step causal chain approach. In the first step, each proposed mediating variable is regressed on $DID_{c,t}$ with the same controls and fixed effects as the baseline; a significant coefficient confirms that NLH construction activates that channel. In the second step, the main regression is re-estimated including the mediator; the DID coefficient remaining positive and significant validates that the channel is operative.

5.3.1. Supplier Substitution Effect

Research Hypothesis 2 indicates that NLH construction reduces firms' cross-regional matching costs, making it easier for firms to find better, more competitive suppliers from afar. Drawing on classical information economics and transaction cost theory [34, 2, 35], in an environment of underdeveloped logistics infrastructure, nearby suppliers enjoy a de facto "regional monopoly" advantage, and firms are compelled to maintain costly incumbent relationships to safeguard production continuity. NLH construction substantially reduces iceberg trade costs [24, 36] and matching friction, raising the fluidity of supplier relationships. Furthermore, following the trade-network mechanism of Chaney [37], as firms increase supplier turnover, the supply chain search radius expands in a network-like fashion: when an old supplier relationship is dissolved, the firm's reach extends outward to a broader pool of potential suppliers on the national market, where efficient remote suppliers far outnumber those in any local market. Newly matched suppliers with cost or technological advantages thus tend to be located at greater distances, and this dynamic "replace nearby with remote" substitution process fundamentally broadens the spatial reach of the procurement network. This manifests at the firm level as a rise in supplier turnover. Following Bernard et al. [4], we construct the supplier turnover rate (*Supplier Turnover Rate*) as the proportion of a firm's top-five suppliers in year t that changed from year $t-1$. Column (1) of Table 6 shows that the *DID* coefficient is 0.4929, significant at the 1% level, confirming that NLH construction significantly raises firms' supplier turnover rate, validating Hypothesis 2.

5.3.2. Share Concentration Effect

NLH construction reduces firms' variable transport costs, enabling firms to concentrate more purchasing budget toward remote, efficient suppliers. From a new economic geography perspective, in an environment of limited logistics infrastructure, firms are compelled to adopt a dispersed, nearby-sourcing strategy — maintaining a large number of small local backup suppliers — to guard against regional supply chain disruptions. This forces overall procurement concentration to remain low. NLH construction substantially enhances the reliability of long-distance delivery, eliminating the need for firms to maintain local redundant capacity. The resulting reduction in transport costs triggers a "Matthew effect" in resource allocation across the national unified market: firms progressively concentrate their orders on a small number of leading suppliers that enjoy the strongest scale economies and technological advantages. Constrained by economies of scale and the logic of industrial agglomeration [36], these lead suppliers are typically clustered in distant, specialised industrial hubs. As firms redirect orders previously placed with nearby small and medium suppliers toward these remote lead suppliers [4], the weighted-average geographic distance of the supply chain is extended. Column (2) of Table 6 shows that *DID* significantly raises firms' supplier concentration degree (*Supplier Concentration*), validating Hypothesis 3.

5.3.3. Vertical Disintegration Effect

Reducing transaction costs between firms is the fundamental goal of enterprise value-chain reconfiguration [38, 39]. NLH construction lowers the uncertainty and friction of external transactions. When the operating environment permits, firms are more inclined toward vertical specialisation and outsourcing non-core intermediate goods or primary processing links. Traditional approaches to measuring vertical disintegration typically rely on value-added-based financial indices (e.g., the VAS index). However, such financial measures have two well-known limitations: they are backward-looking and thus poorly capture management's forward-looking strategic intent, and they are susceptible to earnings management and accounting policy changes that can mask genuine organisational decisions. To precisely capture the management-layer strategic orientation toward vertical disintegration, following Tian et al. [27], we use a Large Language Model (LLM) to extract information from listed firms' annual report MD&A sections, constructing the "Vertical Separation Index" (*VSI*). Column (3) of Table 6 shows the *DID* coefficient is 0.0171, significant at the 5% level, confirming that NLH construction significantly promotes the management-layer tendency toward vertical disintegration, expanding firms' supplier network geographic boundary, validating Hypothesis 4. As a robustness check, we also verify the channel using a traditional VAS-based financial measure of vertical integration; the results remain consistent.

Table 6. Mechanism Analysis Results

VARIABLES	(1) Supplier Turnover Rate	(2) Supplier Concentration	(3) VSI
DID	0.4929***	0.0353*	0.0171**
	(3.6100)	(1.7425)	(2.4034)
Controls	YES	YES	YES
Firm FE	YES	YES	YES
Year FE	YES	YES	YES
N	1,850	2,507	2,266
Adjusted R^2	0.6119	0.3372	0.6628

6. Further Analysis

6.1. Heterogeneity Analysis

The baseline regressions have confirmed that NLH construction can significantly broaden firms' supplier geographic distribution. However, the heterogeneous effects on firms with different enterprise attributes, industry types, and risk perceptions remain unclear. From the perspective of economic consequences, whether hub construction has also broken down inter-regional trade barriers (especially inter-provincial) and promoted market integration at the city level requires further investigation. This section further analyses heterogeneity effects, economic consequence effects, and hub-type complementarity effects.

6.1.1. Firm Ownership Heterogeneity

State-owned enterprises typically enjoy administrative ties and financing advantages; their supply chains often carry constraints tied to administrative boundaries or specific relationships. By contrast, non-state-owned enterprises face stronger market competition and budget constraints, and are more sensitive to cost changes. The NLH construction programme primarily promotes firms' "de-localisation" of suppliers, channeling supplier procurement toward higher-efficiency regions. Following Lin et al. [40], we split the sample by enterprise ownership type into SOE and non-SOE subsamples and re-estimate. Table 7 columns (1)–(2) show that the core coefficient *DID* is significantly positive only in the non-SOE group, while it is insignificant in the SOE group. This indicates that NLH construction mainly promotes non-SOE firms' supplier "de-localisation". For non-SOE firms, the hub programme primarily uses logistics cost reduction to break through the market barriers that confine supply chains to local areas, creating room for these firms to search for lower-cost suppliers from afar and improving their competitive position. For SOE firms, supply chain decisions are subject to local administrative intervention and exhibit institutional path dependence: SOEs bear responsibilities for regional employment stability and local fiscal revenues that constrain their ability to replace local suppliers with remote ones regardless of logistics cost improvements. This institutional rigidity, rather than market indifference, explains why the NLH program's effects have not yet extended to SOEs.

6.1.2. High-Tech Industry Heterogeneity

High-tech enterprises typically lie at the high-value end of the industrial and innovation chain, with high dependence on specialised intermediate inputs. Such customised or complex intermediate goods often have limited local suppliers, and the reduction in transport costs generates smaller supply-cost savings for these specialized inputs; firms' supply choices are more governed by technical matching requirements than by pure transport cost reductions. Non-high-tech enterprises have higher demand for standardised intermediate goods, often inclining toward nearby procurement, and logistics cost improvements are more effective for promoting remote high-quality procurement. Following Aghion et al. [41], we classify the sample into high-tech and non-high-tech industry subsamples. Table 7 columns (3)–(4) show that in the non-high-tech industry group, the *DID* coefficient is significantly positive; while in the high-tech industry group, the coefficient is insignificant. This shows that the effect of NLH construction on expanding firms' supplier geographic distribution is significantly more pronounced in non-high-tech enterprises, confirming the above reasoning.

Table 7. Heterogeneity Analysis Results (I)

VARIABLES	(1) SOE	(2) Non-SOE	(3) Hi-Tech	(4) Non-Hi-Tech
DID	0.8350	1.5504***	0.7958	2.0434**
	(1.11)	(2.63)	(1.4117)	(2.4089)
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Observations	1,163	1,325	1,548	899
Adjusted R^2	0.655	0.670	0.666	0.649

6.1.3. Intermediate Input Dependence and Supply Chain Risk Heterogeneity

The degree of enterprise dependence on intermediate inputs directly determines the extent to which logistics cost reductions can benefit from geographic expansion. Enterprises with high input costs, due to the large proportion of external procurement in total operating costs, have greater space to benefit from geographic expansion and cost optimisation. Columns (1)–(2) of Table 8 show that the *DID* coefficient is 1.1777 and significantly positive in the high-dependence group, while insignificant in the low-dependence group. For enterprises with higher dependence on intermediate inputs, intermediate procurement costs account for a high proportion of total costs, so the transport cost reduction brought by NLH construction is even more prominent, effectively motivating such enterprises to expand their supply radius to more remote regions with more abundant intermediate materials and component supply, and procure inputs more efficiently and at higher quality.

Facing external shocks, adjusting supplier geographic distribution is an important means for enterprises to disperse risk and strengthen supply chain resilience [42, 43]. We use Large Language Models and Natural Language Processing technology to deeply mine listed firms' annual report "management discussion and analysis (MD&A)" texts, identifying the frequency of occurrence of supply chain risk-related terms, constructing an "enterprise supply chain risk perception index". Dividing the sample into high- and low-risk groups at the median, results are shown in columns (3)–(4) of Table 8: both groups yield significantly positive *DID* coefficients (low-risk: 1.3709, $p < 0.10$; high-risk: 1.5775, $p < 0.05$), but the high-risk group coefficient is larger and more precisely estimated. A Chow test confirms that the treatment effects differ significantly between the two groups, indicating a structural difference. This indicates that when firms perceive higher supply chain uncertainty (e.g., supply disruption or price volatility risk), their incentive to exploit logistics improvements from the hub program is stronger, reflecting a strategy of trading proximity for resilience: firms accept greater supplier distance in exchange for a more geographically diversified supplier portfolio, reducing concentration risk and strengthening supply chain resilience.

Table 8. Heterogeneity Analysis Results (II)

VARIABLES	(1) Low Input Dep.	(2) High Input Dep.	(3) Low SC Risk	(4) High SC Risk
DID	0.9477	1.1777**	1.3709*	1.5775**
	(1.24)	(2.20)	(1.72)	(2.40)
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Observations	1,077	1,088	1,103	1,055
Adjusted R^2	0.637	0.786	0.678	0.688

6.2. Market Integration Effects

Optimising logistics infrastructure layout is a key policy instrument for breaking regional market segmentation and reducing cross-regional trade costs [16, 17, 44]. A well-integrated domestic market requires that goods and factors flow freely across regional boundaries without administrative or infrastructural barriers [18]. Transportation infrastructure and logistics networks are the physical prerequisites for breaking regional

barriers, reducing cross-regional trade costs, and facilitating the national market's internal circulation. To further examine the macro-level role of NLH construction policy in building a national unified market, this section examines NLH construction's effects on both the breadth and intensity of city-level upstream procurement, drawing on the "corporate procurement" perspective of Lu et al. [26] and Fan et al. [45].

6.2.1. Market Integration Breadth

Market integration breadth is from the city's overall perspective, measuring whether the city's upstream procurement is moving from a provincial market toward the national market. We take the share of the city's upstream procurement transactions crossing provincial boundaries as the dependent variable. Results are shown in Table 9. From columns (1)–(2), NLH construction significantly raises the cross-provincial proportion of procurement transactions by 1.7%, and the cross-provincial proportion of procurement amounts by 4.9%. This indicates that logistics hub construction brings iceberg cost and matching cost reductions, which not only at the micro-level promotes enterprises to choose to source from more remote suppliers, but also at the macro-level breaks down provincial barriers in market segmentation, thereby expanding market integration breadth.

Table 9. Market Integration Breadth Effects

VARIABLES	(1) Cross-Prov. Trans. Share	(2) Cross-Prov. Amount Share
DID	0.017**	0.049*
	(0.007)	(0.026)
Observations	3,026	2,993
R-squared	0.727	0.331
Controls	YES	YES
City FE	YES	YES
Year×Province FE	YES	YES

6.2.2. Market Integration Intensity

A city sourcing a larger share of inputs from outside its borders indicates broader trade linkages with external markets; however, upstream procurement links may still be confined to neighbouring cities and may not represent a truly broad-based connection to external markets. To measure market integration intensity from the perspective of cities toward cross-city trade, we further examine whether market integration is strengthened from a one-city to cross-city perspective. We construct a triple-difference model as follows:

$$Intensity_{ijt} = \alpha + \beta_1 \, DID_{i,t} \times k_{ij} + \gamma \, Controls + \mu_i + \tau_j + \delta_t + \varepsilon_{ijt}, \qquad (26)$$

where $Intensity_{ijt}$ denotes the trade linkage intensity between city i and city j in year t, measured by the bilateral transaction count (or amount) as a share of total bilateral transactions (or amounts). $DID_{i,t}$ is city i's policy variable; k_{ij} is a dummy indicating whether the two cities are in different provinces ($k_{ij} = 1$) or not ($k_{ij} = 0$). Results are shown in Table 10.

The DDD coefficient (0.488–0.538) shows that NLH construction significantly raises the cross-provincial transaction count share by 49% and the cross-provincial transaction amount share by 54%, with both results showing the same significance. The negative DID coefficient (−0.280): 60% of the upward movement in cross-provincial transaction count share originates from the diversion of intra-provincial trade toward inter-provincial trade, confirming the "supplier substitution effect" and showing that NLH construction is capable of strengthening cross-city trade links, thereby promoting supplier substitution at the city level. The strengthening of these linkages further confirms the role of logistics hub construction as a physical prerequisite for breaking regional barriers and promoting a national unified market.

Table 10. Market Integration Intensity Effects

VARIABLES	(1) Cross-Prov. Trans. Share	(2) Cross-Prov. Amount Share
DDD ($DID \times k_{ij}$)	0.488***	0.538**
	(0.074)	(0.208)
DID	−0.280***	−0.287
	(0.076)	(0.200)
Observations	217,863	217,777
R-squared	0.896	0.365
Controls	YES	YES
City$_i$ FE	YES	YES
City$_j$ FE	YES	YES
Pair FE	YES	YES
Year FE	YES	YES

6.3. Hub-Type Complementarity Effects

We classify logistics hubs into three types based on the number and mix of hub functions. *Core hubs* are cities that simultaneously develop four or more hub types; they occupy strategically important positions in the national logistics network, integrating production-service and commerce-service functions, with Chongqing as a representative example. *Regional hubs* are cities that develop exactly three hub types, at least one of which must be either production-service or commerce-service; they serve as regional logistics centers. *Functional hubs* are cities that develop fewer hub types than the preceding two categories, playing a more specialised supporting role in the broader logistics network. Functional hubs constitute the majority of designated cities in the network, though their construction intensity is relatively weaker than Core and Regional hubs.

To study the different complementarity effects between the three types of hubs, we use the city-pair subsample for grouping tests. Specifically, for the Core hub group, Regional hub group, and Functional hub group, we conduct two-way DID on their pairwise combinations (i, j), each testing the two indicators of cross-provincial transaction count share and cross-provincial transaction amount share twice. Results are shown in Table 11.

Table 11. Hub-Type Complementarity Effects

Panel A: Core ↔ Regional				
	(1) Co→Re Cnt	(2) Co→Re Amt	(3) Re→Co Cnt	(4) Re→Co Amt
DID	0.450***	1.690***	0.362***	0.642*
	(0.087)	(0.434)	(0.119)	(0.363)
Observations	5,666	5,666	6,186	6,186
R-squared	0.697	0.155	0.786	0.199
Panel B: Core ↔ Functional				
	(5) Co→Fu Cnt	(6) Co→Fu Amt	(7) Fu→Co Cnt	(8) Fu→Co Amt
DID	0.244***	0.359**	0.062***	−0.170
	(0.052)	(0.128)	(0.023)	(0.167)
Observations	54,754	54,753	9,675	9,675
R-squared	0.721	0.182	0.787	0.180
Panel C: Regional ↔ Functional				
	(9) Re→Fu Cnt	(10) Re→Fu Amt	(11) Fu→Re Cnt	(12) Fu→Re Amt
DID	0.239***	−0.155	0.109***	0.197
	(0.060)	(0.329)	(0.027)	(0.306)
Observations	54,244	54,243	8,751	8,751
R-squared	0.724	0.180	0.694	0.142
Controls	YES	YES	YES	YES
City i FE	YES	YES	YES	YES
City j FE	YES	YES	YES	YES
Pair FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: Co = Core; Re = Regional; Fu = Functional; Cnt = transaction count share; Amt = transaction amount share. For example, Panel A column (1) of Table 11 shows the effect of Core hub NLH policy (DID variable) on transaction count between the two hub types. Cnt = count; Amt = amount.

From Panel A, the trade linkage between Core and Regional hub cities is notably strong. Core hub construction brings a 0.45% increase in transaction count share with Regional hub cities, and a 1.69% increase in transaction amount share; the bilateral relationship is likewise significant. The bilateral relationship's further strengthening reflects coordination effects between hub types. When both cities belong to key construction targets, the two cities' logistics facility standards are simultaneously elevated, the inter-city transport cost reduction magnitude is greater than with other cities, and from the perspective of strategic coordination, both localities are more inclined to adopt competitive and cooperative logistics channels, forming more competitive trade links.

Panel B and Panel C show a different pattern of change. In Panel B, Core hub construction significantly drives trade linkage strength and quantity increase with Functional hub cities in a meaningful positive direction; but Panel C shows that Regional hubs have no significant effect on transaction amount improvement with Functional hubs. We conjecture that this difference stems from a structural asymmetry in hub functions. Core hubs typically combine both production-service and commerce-service hub types: the former directly serves goods production while the latter faces consumer markets, enabling Core hubs to simultaneously link the supply side and the demand side, forming *dual-channel* trade ties with Functional hubs — one channel in raw materials (as buyer) and one in consumer goods (as seller). Regional hubs, by contrast, do not simultaneously develop both types; they serve only one end of the supply chain. Consequently, their trade ties with Functional hubs are confined to either raw materials *or* consumer goods, but not both, and the bilateral trade structure remains shallow and unsystematic. In summary, the key shortcoming of Regional hubs is their lack of the dual positive externalities — connecting both the production and consumption sides — that give Core hubs their stronger trade-driving capacity.

7. Conclusion and Policy Implications

7.1. Research Conclusions

This paper uses China's National Logistics Hub program as a quasi-natural experiment to study how logistics infrastructure investment reshapes the geographic scope of firm-level supplier networks. We build a domestic supplier selection model incorporating self-production and outsourcing choices, and using 2012–2024 A-share listed firms' data and enterprise bidding data, carefully examine the mechanisms through which NLH construction shapes firms' supplier geographic distribution. The main research conclusions are as follows.

First, NLH construction significantly broadens firms' supplier geographic distribution, with significant policy additive and spatial spillover effects. Empirical results show that NLH construction effectively reduces cross-regional transport costs, driving firms' procurement to shift systematically toward more geographically remote suppliers, as reflected in the significant increase in weighted-average supplier distance. The DID coefficient remains consistently positive and stable across a battery of robustness tests and endogeneity corrections, confirming Hypothesis 1.

Second, logistics hub construction promotes firms' supplier geographic distribution expansion through three channels — the "supplier substitution effect", the "share concentration effect", and the "vertical disintegration effect" — capturing the micro-level structural transformation of firms' procurement networks. Logistics infrastructure im-

provement breaks the regional monopoly advantage of nearby suppliers, raises supply chain mobility and competitive selection efficiency, drives purchasing shares toward high-productivity remote lead suppliers, and reduces external transaction costs that lead firms to outsource non-core production stages.

Third, the distributional effects of NLH construction are heterogeneous across firm types. Non-state-owned enterprises and non-high-tech enterprises — which face stronger market competition and tighter cost constraints — respond more strongly to logistics cost reductions, exhibiting larger treatment effects. Moreover, firms with higher intermediate-input dependence and stronger supply chain risk perception show an even greater motivation to leverage the hub network to build geographically diversified supplier portfolios, reflecting a strategy of trading proximity for resilience.

Fourth, logistics hub construction plays a positive role at the macro level in promoting a national unified market, expanding market integration breadth and increasing market integration intensity. The construction of logistics hubs not only promotes enterprises to source from more remote suppliers at the micro level, but also breaks down provincial barriers in market segmentation at the macro level. The triple-difference model shows that NLH construction effectively strengthens cross-city trade links, and the deepening of these connections further confirms the role of logistics hub construction as a physical prerequisite for breaking regional barriers.

Fifth, logistics hub construction generates differentiated complementarity effects based on hub type. The bilateral trade linkage between Core-type and Regional-type hubs is the strongest, with a clear bilateral synergy effect; the trade-driving role of Regional-type hubs is nonetheless weaker than that of Core hubs, owing to the absence of the dual production-service and commerce-service functions that give Core hubs their cross-regional integration capacity.

7.2. Policy Implications

Our findings carry several policy implications. First, the documented accumulation and spillover effects suggest that the social returns to NLH investment grow with network density: each additional hub generates benefits that extend beyond its host city, implying that further spatial expansion—particularly into central and western regions—may yield disproportionate aggregate returns. Complementary investments in multimodal connectivity between hub nodes would amplify these network-wide effects.

Second, the heterogeneity results indicate that the program operates primarily through market-cost channels, with stronger effects among private and non-high-tech firms. This pattern implies that administrative barriers to inter-regional supplier competition—rather than logistics capacity constraints alone—may limit the program's reach to state-owned and high-technology firms, pointing to the value of complementary reforms in procurement market access. Finally, the finding that Core hubs generate stronger bilateral trade linkages than Regional hubs, owing to their dual production- and commerce-service functions, suggests that hub functional design is a first-order determinant of network-wide integration effects and warrants attention in future designation decisions.

8. Appendices

Appendix A. Theoretical Model Supplement: Derivation of Parameter Φ_i

Firm i selects a supplier to minimise its unit input price $p(d, z)$. To enter a market at distance d, the firm must pay a fixed distribution cost $f(d, L)$. It then searches for the highest-efficiency supplier in that market—i.e. the one offering the lowest price.

At distance d, supplier efficiency z follows a Fréchet distribution

$$F(z) = \exp\big(-T_0\, g(d)\, z^{-\theta}\big),$$

where $g(d)$ is the local supplier density and T_0 is a technology-level parameter. Accounting for the delivery cost $\tau(s, L)$ between location s and the firm at L, the minimum price available in a market of density $g(d)$ follows the cumulative distribution

$$G_d(p) = 1 - e^{-T_0\, g(d)\, [\tau(s,L)]^{-\theta}\, p^{\theta}}.$$

Proof. The probability that the minimum price exceeds p is the probability that all suppliers offer prices greater than p:

$$\begin{aligned}
\Pr\big[p(d, L)/z > p \text{ for all } z\big] &= \Pr\big[z < w\tau(d, L)/p \text{ for all } z\big] \\
&= \prod F\big(w\tau(d, L)/p\big) \\
&= \big(e^{-T_0\, g(d)\, (w\tau(d,L))^{-\theta}}\big)^{g(d)} \\
&= e^{-T_0\, g(d)\, [\tau(s,L)]^{-\theta}\, p^{\theta}}.
\end{aligned}$$

Thus the CDF of the minimum price at distance d is

$$G_d(p) = 1 - e^{-T_0 g(d) [\tau(s,L)]^{-\theta} p^{\theta}}.$$

The firm optimally enters all markets D_i that offer positive expected profit. Because price distributions are independent across markets, the overall price drawn from the best supplier across all sourcing locations follows

$$G_i(p) = 1 - e^{-\Phi_i\, p^{\theta}},$$

where the composite parameter is defined as

$$\Phi_i = \int_{s \in D_i} T_0\, g(s)\, [\tau(s, L)]^{-\theta}\, ds \tag{A1}$$

The parameter Φ_i aggregates the effective procurement capability of firm i: it captures the firm's geographic sourcing radius D_i, the local supplier density $g(s)$, the technological frontier T_0, and the trade costs $\tau(s, L)$ connecting each sourcing region s to the firm. A reduction in trade costs τ directly raises Φ_i, expanding the pool of potential low-price suppliers and thereby lowering the firm's expected input cost.

Appendix B. PSM Covariate Balance

B.1. Common Support

Figure B1 displays the propensity-score distributions of the treated and control samples before matching. The substantial overlap between the two distributions confirms that the common-support assumption is satisfied; only a negligible fraction of observations fall off support and are dropped.

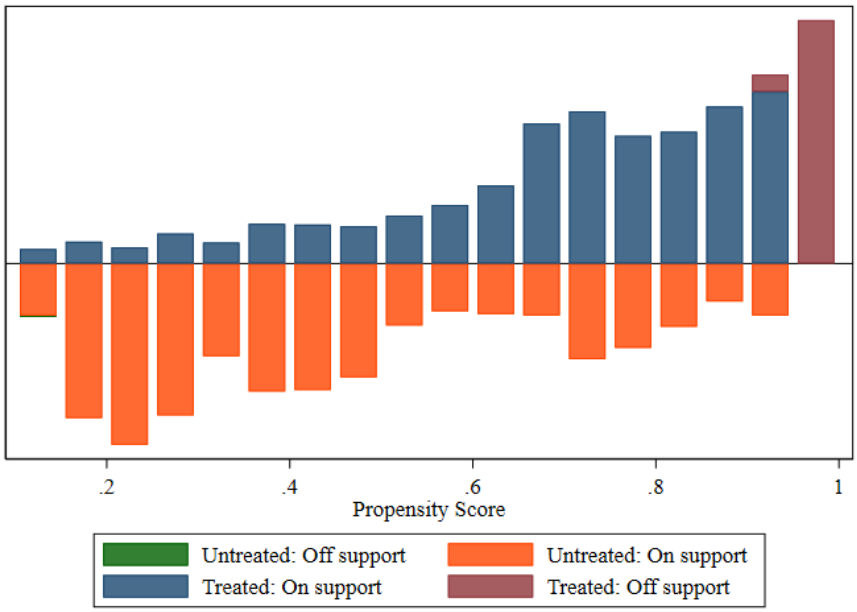


Figure B1. Common Support: Propensity-Score Distributions Before Matching

Notes: Histograms of estimated propensity scores for treated (Treated: On Support / Off Support) and control (Untreated: On Support / Off Support) firms. The large overlap confirms the common-support assumption. Off-support observations are excluded from the matched sample.

B.2. Standardized Bias Before and After Matching

Figure B2 plots the standardized percentage bias (SPB) for each covariate before and after PSM. Before matching, several covariates (particularly population scale *Pop* and economic development level *Dev*) exhibit SPBs well in excess of 50%. After matching, all covariates fall within the conventional $\pm 10\%$ threshold, confirming that the matched sample achieves adequate balance and supports causal attribution of the estimated treatment effects.

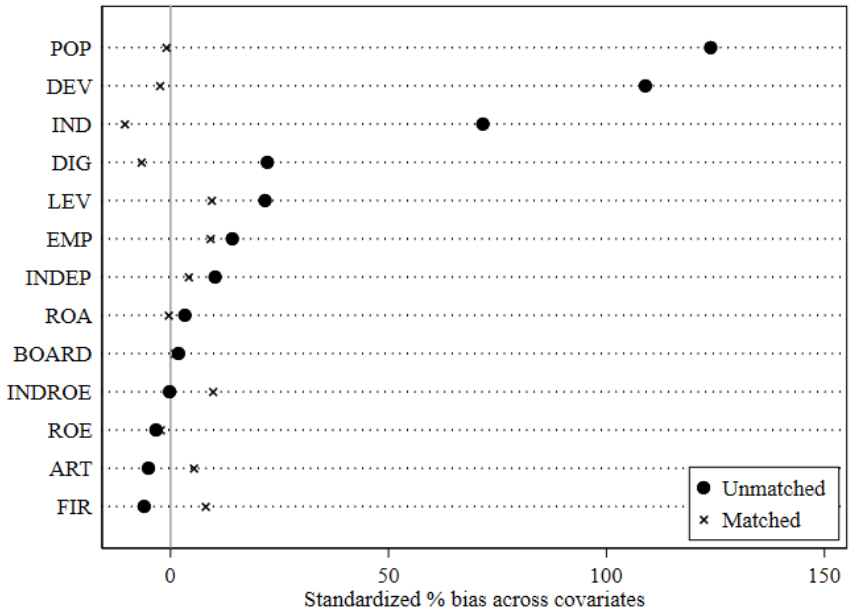


Figure B2. Standardized Bias Before and After Matching (Love Plot)

Notes: Standardized percentage bias (SPB) for each covariate before (circles) and after (triangles) PSM. After matching, all covariates fall within the conventional ±10% threshold, confirming adequate covariate balance.

Appendix C. VSI Construction: LLM-Based Vertical Separation Index

To overcome the limitations of traditional financial indicators in capturing the extent of firm vertical disintegration, this paper employs natural language processing techniques applied to the Management Discussion and Analysis (MD&A) sections of listed-firm annual reports for 2012–2024.

C.1. Keyword List Construction and Initial Screening

Firm-level statements on vertical strategy are highly heterogeneous in expression. We compile a structured keyword dictionary in three categories:

(1) **Integration keywords** (indicating vertical consolidation): terms such as "consolidation," "vertical integration," "wholly owned," "merger and acquisition," "holding," "full industry chain," "end-to-end control," etc.
(2) **Decomposition keywords** (indicating outsourcing/disintegration): terms such as "decomposition," "division of labour," "divestiture," "outsourcing," "spin-off," "asset disposal," "contract manufacturing," "OEM," "subcontracting," "co-operative production," "reduction," "elimination," "service migration," "third-party," "toll processing," "asset-light," "simplification," "decentralisation," etc.
(3) **Ambiguous keywords** (requiring contextual disambiguation): terms such as "layout optimisation," "structural adjustment," "partnership," "alliance," "platform," "ecosystem," "collaborative innovation," "supply chain sharing," "seeking," "transformation," etc.

Sentences containing at least one keyword from any category are extracted from MD&A texts, and punctuation markers (commas, periods, semicolons) are used for segmentation.

C.2. Sample Annotation and LLM Labelling

To provide high-quality annotations for the subsequent classification model, we randomly draw 2,000 sentences from the keyword corpus. The DeepSeek large language model API is used for annotation with the following system prompt: "You are a rigorous financial and economic text analysis expert. Please judge whether the following sentence reflects the firm's *vertical integration* tendency or *vertical disintegration* tendency. Assign the sentence to one category: 'vertical integration' or 'vertical disintegration.' If the model output conflicts or the judgment is ambiguous, that sample is excluded."

C.3. Text Classification Fine-Tuning

The labelled sentences are converted into a classification dataset. To avoid class imbalance, effective labelled samples are first subjected to class-balancing, then split 80:20 into training and validation sets. We fine-tune the Chinese pre-trained language model[46]. Key training parameters: learning rate $= 3 \times 10^{-5}$, batch size $= 8$, epochs $= 3$. After fine-tuning, the model predicts, for each sentence in the full MD&A corpus, the probability of belonging to the "vertical disintegration" class versus the "vertical integration" class.

C.4. VSI Generation

For each sentence j in firm i's MD&A in year t, the sentence-level VSI score is defined as:

$$Sentence_VSI_{i,j,t} = \Pr(\text{Disintegration})_{i,j,t} - \Pr(\text{Integration})_{i,j,t}.$$

The firm-year VSI is the arithmetic mean across all relevant sentences:

$$VSI_{i,t} = \frac{1}{N_{i,t}} \sum_{j=1}^{N_{i,t}} Sentence_VSI_{i,j,t}, \tag{C1}$$

where $N_{i,t}$ is the total number of MD&A sentences for firm i in year t that contain at least one keyword. This index is a continuous variable: a higher value indicates a stronger tendency toward vertical disintegration, outsourcing of non-core activities, and decomposition of the value chain.

Appendix D. Endogeneity Discussion and IV Test for Policy Stacking

The second-stage returns reported in the main text show that after removing endogeneity via the instrumental variable strategy, policy intensity (*Intensity*) significantly increases firm-level average supplier distance (*Disw*). Table D1 reports the IV estimation of the policy-stacking effect.

Table D1. IV Test of Policy Stacking Effect.

VARIABLES	(1) *Disw*
Intensity	1.584*
	(1.92)
N	2,498
R^2	−0.004
Controls	YES
Year FE	YES
Firm FE	YES
K–P F-stat	81.55

Notes: t-statistics in parentheses. Standard errors clustered at the city level. The Kleibergen–Paap rk Wald F-statistic of 81.55 exceeds the Stock–Yogo critical value of 16.38 for 10% maximal IV size. $^*p < 0.10$.

Appendix E. Construction of Additional City-Level Control Variables

Building on the existing literature, this paper adds four city-level macro control variables: economic density (*Eco_Den*), industrial structure (*Ind*), migration rate (*Mig_Rate*), and human capital (*Human_Cap*). City-level data span 2001–2024, sourced from the *China City Statistical Yearbook* and individual city statistical bulletins.

E.0.0.1. Economic Density (Eco_Den).. Regional GDP divided by the administrative area of the prefecture, log-transformed to reduce skewness:

$$Eco_Den = \ln\left(\frac{\text{Regional GDP}}{\text{Administrative Area (km}^2)}\right).$$

E.0.0.2. Migration Rate (Mig_Rate).. The difference between the permanent resident population and the registered (hukou) population, expressed as a share of the registered population:

$$Mig_Rate = \frac{\text{Permanent Residents} - \text{Registered Population}}{\text{Registered Population}}.$$

This indicator captures the city's attractiveness to in-migrants and the degree of population mobility.

E.0.0.3. Human Capital (Human_Cap).. The ratio of enrolled students in regular higher-education institutions to the year-end total population:

$$Human_Cap = \frac{\text{Enrolled Students in Regular Higher Education}}{\text{Year-End Total Population}}.$$

This measures the potential supply of high-quality labour and the degree of talent agglomeration.

References

[1] Qifan Huang. Keynote address on logistics cost reduction and supply chain development. Keynote address at the 13th China Logistics Investment & Financing Conference and 2nd Chengdu–Chongqing Region Supply Chain Development Conference, May 2023. Delivered 26 May 2023. Speaker: former Vice Chairman of the Finance and Economic Affairs Committee of the 12th National People's Congress and former Mayor of Chongqing.

[2] Oliver E. Williamson. *The Economic Institutions of Capitalism*. Free Press, New York, 1985.

[3] Pol Antràs, Teresa C. Fort, and Felix Tintelnot. The margins of global sourcing: Theory and evidence from US firms. *American Economic Review*, 107(9):2514–2564, 2017. .

[4] Andrew B. Bernard, Andreas Moxnes, and Yukiko U. Saito. Production networks, geography, and firm performance. *Journal of Political Economy*, 127(2):639–688, 2019. .

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

[6] Teresa C. Fort. Technology and production fragmentation: Domestic versus foreign sourcing. *Review of Economic Studies*, 84(2):650–687, 2017. .

[7] Federico Huneeus. Production network dynamics and the propagation of shocks. *Working Paper*, 2018. Princeton University.

[8] Frances E. Bowen, Paul D. Cousins, Richard C. Lamming, and Adam C. Faruk. The role of supply management capabilities in green supply. *Production and Operations Management*, 10(2):174–189, 2001. .

[9] Lorenzo Caliendo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte. The impact of regional and sectoral productivity changes on the US economy. *Review of Economic Studies*, 85(4):2042–2096, 2018. .

[10] Vasco M. Carvalho and Alireza Tahbaz-Salehi. Production networks: A primer. *Annual Review of Economics*, 11:635–663, 2019. .

[11] Dave Donaldson. Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934, 2018. .

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

[13] Benjamin Faber. Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies*, 81(3):1046–1070, 2014. .

[14] Ejaz Ghani, Arti Grover Goswami, and William R. Kerr. Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing. *Economic Journal*, 126(591):317–357, 2016. .

[15] Stephen J. Redding and Esteban Rossi-Hansberg. Quantitative spatial economics. *Annual Review of Economics*, 9:21–58, 2017. .

[16] Sandra Poncet. A fragmented China: Measure and determinants of Chinese domestic market disintegration. *Review of International Economics*, 13(3):409–430, 2005. .

[17] Alwyn Young. The razor's edge: Distortions and incremental reform in the People's Republic of China. *Quarterly Journal of Economics*, 115(4):1091–1135, 2000. .

[18] Laura Hering and Sandra Poncet. Market access and individual wages: Evidence from China. *Review of Economics and Statistics*, 92(1):145–159, 2010. .

[19] Costas Arkolakis. Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy*, 118(6):1151–1199, 2010. .

[20] Costas Arkolakis, Federico Huneeus, and Yuhei Miyauchi. Spatial production networks. Working Paper 30646, National Bureau of Economic Research, 2023.

[21] China Logistics & Purchasing Association, Logistics Hub and Park Branch. Decoding national logistics hub construction: Evolution of seven designation batches, development characteristics and future trends. Online Report [EB/OL], 2025. Published 2025-06-23, accessed 2026-02-19. Available at: `http://yqzwh.chinawuliu.com.cn/xwdt/202506/23/651818.shtml`.

[22] China Logistics & Purchasing Association, Logistics Hub and Park Branch. National logistics hub operations monitoring report for the first half of 2025. Online Report [EB/OL], 2025. Published 2025-09-15, accessed 2026-02-19. Available at: http://yqzwh.chinawuliu.com.cn/xwdt/202509/15/656132.shtml.

[23] Marc J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, 2003. .

[24] Paul A. Samuelson. The transfer problem and transport costs, II: Analysis of effects of trade impediments. *Economic Journal*, 64(254):264–289, 1954. .

[25] Kevin Lim. Endogenous production networks and the business cycle. *Working Paper*, 2016. University of Toronto.

[26] Yi Lu, Jin Wang, and Lianming Zhu. Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. *American Economic Journal: Economic Policy*, 11(3):325–360, 2019. .

[27] Can Tian, Xin Ye, and Xiaodong Zhang. Production networks and firm-level elasticities of substitution. *Journal of Political Economy*, 131(4):1011–1059, 2023. .

[28] Clément de Chaisemartin and Xavier D'Haultfœuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020. .

[29] A. Kerem Coşar and Banu Demir. Domestic road infrastructure and international trade: Evidence from Turkey. *Journal of Development Economics*, 118:232–244, 2016. .

[30] James J. Heckman, Hidehiko Ichimura, and Petra E. Todd. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4):605–654, 1997. .

[31] Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2nd edition, 2010.

[32] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, 21(1):C1–C68, 2018. .

[33] Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, 2020. .

[34] George J. Stigler. The economics of information. *Journal of Political Economy*, 69(3): 213–225, 1961. .

[35] Charles A. Weber, John R. Current, and W. C. Benton. Vendor selection criteria and methods. *European Journal of Operational Research*, 50(1):2–18, 1991. .

[36] Paul Krugman. Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–499, 1991. .

[37] Thomas Chaney. The network structure of international trade. *American Economic Review*, 104(11):3600–3634, 2014. .

[38] Laura Alfaro, Pol Antràs, Davin Chor, and Paola Conconi. Internalizing global value chains: A firm-level analysis. *Journal of Political Economy*, 127(2):508–559, 2019. .

[39] Pol Antràs and Alonso De Gortari. On the geography of global value chains. *Econometrica*, 88(4):1553–1598, 2020. .

[40] Justin Yifu Lin, Fang Cai, and Zhou Li. The China miracle: Development strategy and economic reform. *Revised Edition, Chinese University Press*, 2003. Hong Kong: Chinese University Press.

[41] Philippe Aghion, Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros. Industrial policy and competition. *American Economic Journal: Macroeconomics*, 7(4):1–32, 2015. .

[42] Vasco M. Carvalho, Makoto Nirei, Yukiko U. Saito, and Alireza Tahbaz-Salehi. Supply chain disruptions: Evidence from the Great East Japan Earthquake. *Quarterly Journal of Economics*, 136(2):1255–1321, 2021. .

[43] Barthélémy Bonadio, Zhen Huo, Andrei A. Levchenko, and Nitya Pandalai-Nayar. Global supply chains in the pandemic. *Journal of International Economics*, 133:103542, 2021. .

[44] Trevor Tombe and Xiaodong Zhu. Trade, migration, and productivity: A quantitative

analysis of China. *American Economic Review*, 109(5):1843–1872, 2019. .

[45] Haichao Fan, Yao Amber Li, and Mingzhi Xu. Trade and domestic production networks. *Review of Economic Studies*, 90(6):2942–2981, 2023. .

[46] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.