Appendix For Online Publication

Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters

Jingting Fan
Pennsylvania State University

Yi Lu
Tsinghua University

Wenlan Luo
Tsinghua University

Contents

A Data and Empirics

A.1 Constructing Coordinates of Ports and Origin Cities

We define the location of a county by its center of mass using the geographic information in the 2010 census. We weight the coordinates of all counties making up a prefecture city by their population to calculate an average coordinate, which we then define as the location of a prefecture city. For the four provincial-level cities, Beijing, Shanghai, Tianjin, and Chongqing, we generate the coordinates by weighting the coordinates of their urban sub-divisions (districts). We exclude the rural sub-divisions in these provincial-level cities because their large rural areas have a disproportionate impact on the measured economic center.

In mapping the location of exporters to these coordinates, we use the origin city of export shipments from the customs data. An alternative definition is to use the registered address of the exporters. Using the former, instead of the latter, avoids potential measurement errors for the export of multi-plant firms.

A.2 Constructing Distance between Cities for Reduced-form Analyses

The raw road maps are in the form of line strings. For reduced-form analysis, we use the following procedure to find the shortest path between all pairs of cities for both 1999 and 2010.

In the first step, we split the entire main land China into $2km \times 2km$ squares. We define a square as 'on regular roads', if it intersects with any segments of regular roads. We define two *adjacent* squares as connected by regular roads, if both of them are 'on regular roads'. A path consisting of only regular roads is then a chain of connected squares, starting from the one containing the coordinates of the origin, ending with the one containing the coordinates of the destination.

We process the expressway networks in a similar way. We then overlay the two processed networks (regular roads and expressways). A path on this joint network is a set of connected squares, with each two adjacent squares connected by *either* regular roads *or* expressways. A square that is both 'on regular roads' and 'on expressways' is thus viewed as an intersection of regular roads and expressways, from which a trucker can switch from one type of road to the other.

Between any pair of cities, there could be many paths. In the second step, we search for the least-cost path. To this end, we assume each km along a regular road is twice as costly as a km on expressways and calculate the total regular road-equivalent length of all paths. We then use the Dijkstra's algorithm to find the path with the lowest length.

We do the above for both 1999 and 2010 road networks, generating the time-varying distances, $dist^t_{od}$.

A.3 Constructing City Network Graphs for the Routing Model

Our routing model treats individual cities as nodes in a network, connected by roads. Before the structural estimation, we prepare the data so that they are consistent with this model. To this end, we apply the following procedures separately to each of the three maps (expressways in 1999 and 2010, and regular roads in 2007 which are treated as time invariant).

- **Define connected cities.** In the first step, we identify the list of cities (prefectures) connected to the network. We define cities as 'connected' in a map, if the center of the city is within the 50 km radius of any roads on a map. Practically, it means measuring whether any of the coordinates characterizing roads from a map are within 50 km of the city center.

- **Define connections between cities.** We 're-base' the coordinates of 'connected' cities to the nearest coordinates of the road network. For each pair of connected cities, we search for the shortest path between them on the road network using the Dijkstra's algorithm. If the shortest path between two cities does not pass through another city, we define the pair to be 'directly connected'.
- **Construct the graph.** We construct the graph in which cities are the nodes and roads form the edges, through the following procedure. We draw an edge between two cities, if they are found to be 'directly connected' in the previous step. We define the length of the edge to be the length of the shortest path between the two cities.[1]

The left panel of Figure A.1 is the original digital maps. The right panel overlays their network representation, which is the output of the above process. Again, even though the edges are drawn as straight lines in the right panel, the length we assign to each edge is the length of the actual road.

We transform the right panel of Figure A.1 into adjacent matrices, $\mathbb{H}^{1999}$, $\mathbb{H}^{2010}$, and $\mathbb{L}$, respectively, for structural estimation. Element (k,l) in a matrix will be $\iota_{kl}^{-\theta}$, if cities k and l are adjacent and 'directly connected' in the road network represented by that matrix; otherwise (k,l) will be zero.

A.4 Backing Out Segment-Specific Road Construction Cost

We first cut expressways into 10-km segments. For each such segment, we check if it passes water and calculate the average slope of its terrains.[2] We calculate the *relative* construction cost of segment i following a simple function from the transport engineering literature:

$$cost_i = 1 + slope_i + 25 \times PassWater_i.$$

This specification is similar to the one used Faber (2014), except that we abstract from the measure of existing buildings due to the lack of data. According to this formula, the cost of constructing a segment passing water costs 26 times as much as on a dry plain. The *level* of the construction cost is determined such that the total cost of the segments constructed between 1999 and 2010 is 9.92% of the 2010 GDP.

The total cost (9.92% of the 2010 GDP) is 3983 billion 2010 CNY. The total dry-plain equivalent distance of all roads constructed during this period is 453,447 km, so each dry-plain equivalent km of expressway costs about 8.85 million 2010 CNY. The total length of expressway actually constructed during this period is 49,760 km, so the average cost for each kilometer is around 80 million 2010 CNY. This cost is much higher than the dry-plain equivalent cost, reflecting that most of the projects during this decade pass rugged terrain or water areas.

Figure A.2 shows the geographic features of China, which determine the cost estimates.

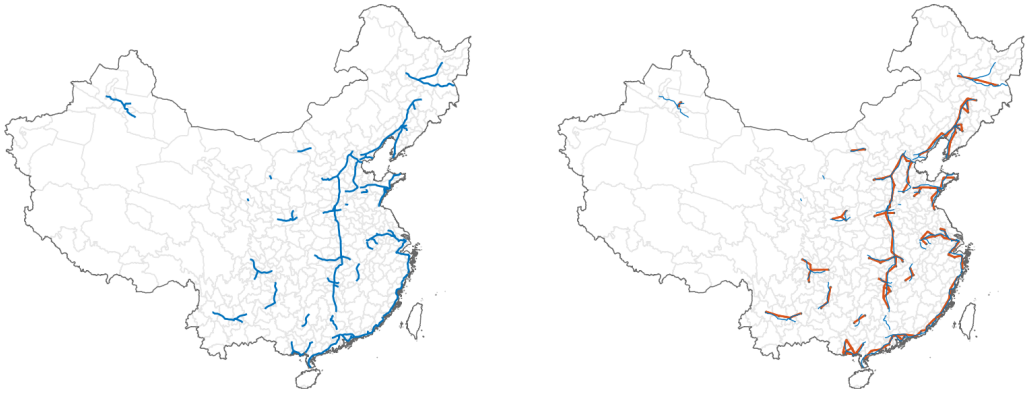
A.5 The Lists of Ports and Major Cities

List of seaports: Tianjin, Dalian, Shanghai, Ningbo, Fuzhou, Xiamen, Qingdao, Guangzhou, Shenzhen, Zhuhai, Shantou.

List of major cities: Beijing, Tianjin, Shijiazhuang, Tangshan, Handan, Xingtai, Baoding, Cangzhou, Shenyang, Dalian, Changchun, Haerbin, Shanghai, Xuzhou, Suzhou, Nantong, Yancheng, Hangzhou,

[1] The two expressway maps are digitized from the projection of published hard-copy maps, which introduce measurement errors that change the exact locations of roads. The same road might therefore has slightly different measured lengths from 1999 and 2010 expressway maps. We inspect all segments with less 5% change in length to rule out measurement errors.

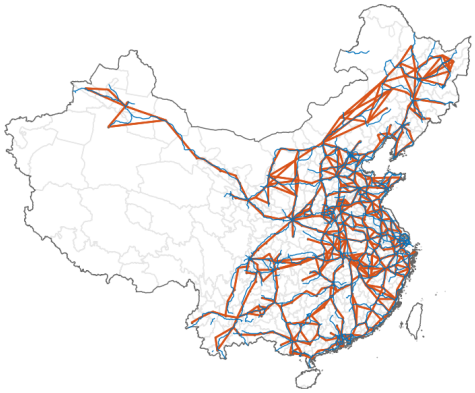
[2] 23.3% of the segments pass water areas.



(a) 1999 Expressway Map

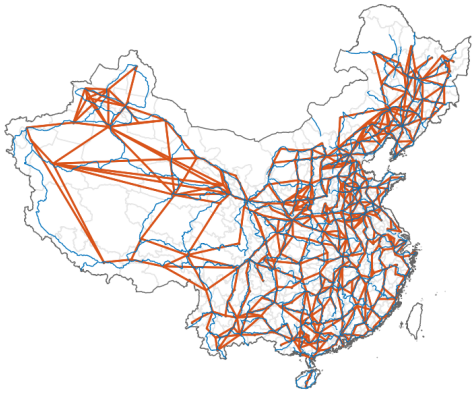
(b) 1999 Expressway Network

(c) 2010 Expressway Map



(d) 2010 Expressway Network

(e) Regular Road Map

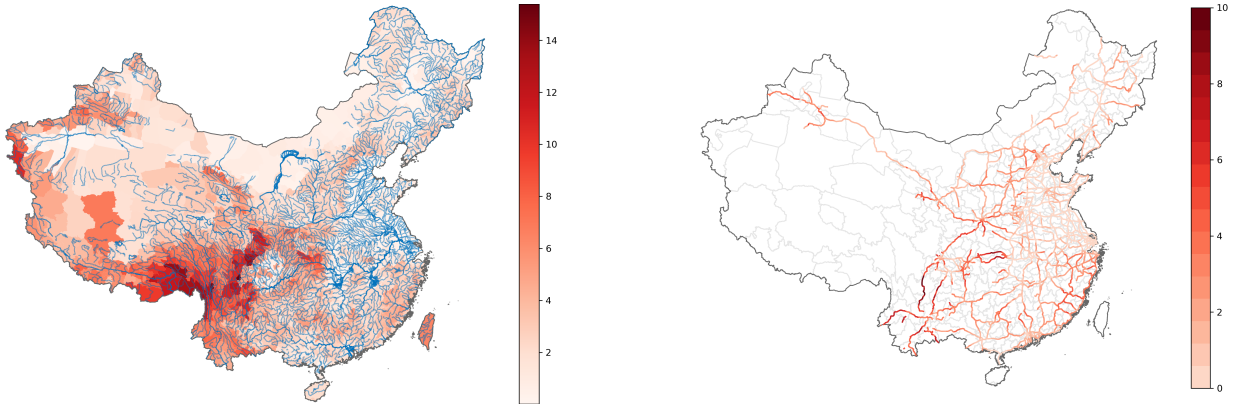


(f) Regular Road Network

Figure A.1: From Road Maps to Road Networks

Note: A city is defined as 'connected' on a road network, if the center of the city is within the 50 km radius of any roads of the network. Two cities are defined as connected on a road network if the shortest path connecting them on the road network does not pass a third city. The distance between two connected cities is then defined as the *road length* of the shortest path between them. The left three panels plot the maps of the three road networks. The right three figures overlay the connected city pairs; each solid line segment corresponds to a pair of connected cities.

Wenzhou, Fuyang, Suzhou, Liuan, Quanzhou, Ganzhou, Jinan, Qingdao, Yantai, Weifang, Jining, Linyi, Liaocheng, Heze, Zhengzhou, Luoyang, Xinxiang, Nanyang, Shangqiu, Xinyang, Zhoukou, Zhumadian,



(a) Slopes and Water Areas

(b) Slopes of Expressway Segments

Figure A.2: Geography and Expressway Construction Costs

Note: The left figure plots the slope of land and the geographic distribution of water areas. The right panel plots the expressways in 2010, indicating using color the average slope for each 10-km segment.

Wuhan, Huanggang, Changsha, Hengyang, Shaoyang, Changde, Guangzhou, Zhanjiang, Munidiqu, Chongqing, Chengdu, Nanchong, Zunyi, Bijiediqu, Xi'an.

A.6 Additional Analyses: Channels and Sectoral Level Results

This subsection compares our reduced-form estimates to the literature. It inspect the channels and presents additional robustness results.

Comparing the estimate to the literature. The closest empirical setting to ours is Coşar and Demir (2016), who estimate the impacts of regional road capacity on trade. In a semi-elasticity specification (p. 240), they find that upgrading from carriageway to expressways lead to a "reduction of travel costs around 27% on an average stretch of 820 km." Take our estimate from Column 4 of Table 2 (with a coefficient of 0.174), under the assumption that expressways are on average twice as fast as regular roads, our finding implies that the upgrade reduces the coefficient by $0.174/2 = 0.087$. This coefficient is the product of an elasticity and the percentage difference in trade cost between regular roads and expressways. Using the elasticity of 4 used in Coşar and Demir (2016), our baseline estimate implies that upgrading each hundred km reduces trade cost by $0.087/4 = 2.2\%$. An upgrade of 820 km would therefore reduce trade cost by around 18%. This estimate is lower than Coşar and Demir (2016), likely because many regular roads in China have two or more lanes, so the marginal gains from upgrading to expressway are not as important as the upgrade from single-lane carriageway in Turkey. Nevertheless, the two estimates are in the same order of magnitude and their confidence intervals overlay.

Export growth versus rerouting. By controlling for city-time fixed effects, our baseline estimate uses two sources of variation: organic growth in shipments over an existing route, and rerouting of city export through competing ports. Notice both forces reflect the change in route choice due to the change in domestic shipment cost, and are precisely the forces used to infer the trade cost elasticity that is of interest. To gauge the relative importance of the two forces, we use two complementary approaches that rely on different assumptions.

The first approach is to aggregate the export of a city into a few groups of ports based on the geographic location of ports. The idea is that, if an improvement in connection between city o and port

Table A.1: Understanding Channels and Results from Sectoral Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Aggregate Data			Sectoral Data					
	Growth v.s. Rerouting			Baseline			Growth v.s. Rerouting		
$dist^t_{od}$	-0.226***	-0.166***	-0.157***	-0.373***	-0.138***		-0.183***	-0.137***	-0.137***
	(0.050)	(0.041)	(0.041)	(0.017)	(0.044)		(0.044)	(0.039)	(0.039)
- on express 5						-0.075*			
						(0.039)			
- on regular						-0.137***			
						(0.044)			
log(export of o through $d' \neq d$)		0.171**						0.053**	
		(0.073)						(0.026)	
Specification	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Fixed Effects	og_d, ot, g_dt	od, p_ot, dt	od, p_ot, dt	oti, dti	odi, oti, dti	odi, oti, dti	og_di, oti, gti	odi, p_oti, dti	odi, p_oti, dti
Exclude Major Cities yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1048	2082	2082	20946	11758	11758	7060	12920	12920
R^2	0.932	0.870	0.869	0.593	0.896	0.896	0.926	0.850	0.850

Notes: The dependent variable is the log of total value of goods exported in city o through port d to the RoW. Columns 1 through 3 use aggregate data to explore whether response in export is due to organic growth or rerouting between ports (see text in Appendix A.6 for explanation). Columns 4 through 9 show the results are similar if we use data at the (2 digit) sectoral level. Columns 4 through 6 show that with sectoral data, controlling for city-port fixed effect also halves the cross-sectional estimate, and that expressways are less costly than regular roads. Columns 7 through 9 replicate Columns 1 through 3 using sectoral data. In 'Fixed Effects,' o stands for exporting city, d stands for port city, t stands for time, p_o stands for province of city o, g_d stands for geographic group of port d, i stands for sector.
Standard errors are clustered at city-port level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

d increases the export through d mostly by drawing in export of o through *other ports near* d, then by estimating the regressions at city-port group level, part of the rerouting effect would cancel out and the point estimate would be mainly about the export growth effect. To implement this approach, we group all ports into one of the three based on their geographic locations: North, Central, and South. We then aggregate the data to city-port group-time level, and estimate a similar panel-fixed effect specification as the baseline, controlling for city-port group, city-time, and port group-time fixed effects. Column 1 of Table A.1 finds the point estimate to be -0.226. If any, this is larger (although not statistically significantly) than the baseline estimate of -0.174 (Column 4 of Table 2), suggesting that it is growth of export in a city-port pair, rather than substitution between ports, that drives the baseline estimate.

The second approach seeks to directly control for export through other ports. Specifically, if the rerouting force is strong, then an improvement in the connection between city o and port d likely reduces the export of city o through all other ports. This implies that the export of city o through other ports ($d' \neq d$) is negatively correlated with the improving access between city o and port d, which in turn implies that if we control for city o's export via other ports, the estimated coefficient for distance will shrink. Columns 2 and 3 of Table A.1 implement this test. Column 2 includes export of a city through other ports as a control. Notice that $\sum_d v_{od}$ (total export of city o) is co-linear with $\sum_{d' \neq d} v_{od'}$ (the sum of export through other ports) if city-time fixed effects are included, and the model is not identified. Therefore, we control for province-time fixed effects instead, aiming to capture the overall export growth of a region that might be correlated with road connection.[3] The identifying assumption is that, to the extent that expressway expansion could be endogenous to the overall prospect of export growth in a city,

[3]More precisely, because we use the log of export, rather than the level of export, we can still include city-time fixed effects; however, the identification comes only from the difference between log and linear function forms.

once the major cities are excluded, such correlation is similar across smaller cities within a province and captured by the province-time fixed effects. We find that the coefficient for export through other ports is positive—inconsistent with a strong rerouting force. More importantly, the coefficient for regular-equivalent distance is -0.166, similar to the baseline estimate of -0.174. To rule out that such similarity is a coincidence under a different set of fixed effects from the baseline, Column 3 includes the province-time fixed effects as in Column 2 but excludes from independent variables the export through other ports. The estimated coefficient for regular-equivalent distance under this specification is similar.

While each of these two exercises requires stronger assumption than baseline specification—the first on the working of rerouting by geographic regions, the second on the identifying assumption—the consistent conclusion in both suggest that our finding is likely primarily driven by export growth, rather than rerouting.[4]

Robustness using sectoral level data. A remaining concern of the baseline specification is that it might be driven by changes in the sectoral composition of city export. For example, if as cities gain access to ports, they also become more specialized in export-intensive industries, such as textile, and if for some reason, export in the textile industry is concentrated among the ports that experienced disproportionate increases in expressway connectivity to the hinterland, then the correlation between the shipment share and the bilateral connectivity will be picked up by our regressions. We note that if the expressway expansion is truly exogenous to non-major cities, then this concern does not pose a threat to the IV estimate. Nevertheless, in Columns 4 through 6, we use shipment value at the sectoral level for a robustness check. Column 4 includes city-time-sector, port-time-sector (letter 'i' in the row 'Fixed Effects' denote sectors), and Column 5 further add city-time-sector fixed effects. They show that, as in the baseline regressions, using over-time variation estimates a much smaller coefficient compared to using cross-sectional variation. Column 6 further confirms that expressways are less costly than regular roads.

Finally, we examine the importance of growth versus rerouting using sectoral data. This is useful because if rerouting takes place within a sector (i.e., exporters of cloth used to go through Shanghai, now switch to Guangzhou), then the previous exercises using more aggregate data might be too blunt to detect such patterns. Columns 7 through 9 revisit the exercises in Columns 1 through 3 of Table A.1. Column 7 uses data at city-port group-sector level, and finds slightly larger estimate than the baseline estimator of 0.138. Columns 8 and 9 further show that include export of a city through other ports do not have a big impact on the distance coefficient. Together, these results corroborate the earlier finding that differential route-specific export growth accounts for most of the estimated effects.

A.7 PPML Specifications and Comparison to OLS

An alternative specification used in estimating the trade or shipment elasticity is Poisson Pseudo-Maximum Likelihood, which can address biases arising from heteroskedasticity in the error term. In this appendix, we show that our main points are robust if PPML is used.

Table A.2 reports the results. Columns 1 through 4 vary the set of fixed effects included as in the first four columns of Table 2. Importantly, the difference between Columns 2 and 3 confirms that once city-port fixed effects are controlled for, the estimate for distance shrinks by half. Column 5 further splits the total regular-equivalent distance into the length of regular road and expressway segments along

[4] A caveat in this conclusion is that during the sample period, China's export increased by five folds, and this could be part of the reason we find most effects were on the export growth margin.

Table A.2: Robustness using PPML

	(1)	(2)	(3)	(4)	(5)
$dist^t_{od}$	-0.670***	-0.706***	-0.350***	-0.488***	
	(0.039)	(0.042)	(0.054)	(0.077)	
-on express					-0.279***
					(0.096)
-on regular					-0.481***
					(0.071)
Fixed Effects	o, d, t	ot, dt	od, ot, dt	od, ot, dt	od, ot, dt
Exclude Major Cities				Yes	Yes
Observations	3668	3660	2838	2068	2038

Notes: This table reports the regressions of export shipment through a port on the distance between the city and the port. All specifications are estimated using Poisson Pseudo-Maximum Likelihood. The dependent variable is the total value of goods exported in city o through port d to the RoW. The independent variables are the regular-equivalent road distance between city o and port d along the shortest path (Columns 1-4); and the separate length of expressways and regular roads along the shortest path (Column 5).
Standard errors are clustered at city-port level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the shortest path between o and d. The estimated coefficient for regular roads is higher than that for expressways, consistent with findings from Table 2.

PPML and OLS comparison. The PPML specifications reported above generally produce larger estimates than linear regressions. To see what leads to this result, consider the PPML specification, in which the data generating process is assumed to be $E(v^t_{od}|dist^t_{od}) = \exp(\gamma_1 dist^t_{od} + \beta^t_o + \tilde{\beta}^t_d + \beta_{od})$. In this specification, β^t_o and $\tilde{\beta}^t_d$ are city-time and port-time fixed effects, respectively. Under the assumption that v^t_{od} follows a Poisson distribution, the first order condition with respect to the estimator, $\hat{\gamma}_1$, to maximize the pseudo-likelihood function is:

$$\sum_{o,d,t} dist^t_{od} \cdot (v^t_{od} - \hat{v}^t_{od}) = 0,$$

i.e., the choice of $\hat{\gamma}_1$ is as if minimizing the distance-weighted sum of the *level* difference between v^t_{od} and its predicted values, $\hat{v}^t_{od}$. In an OLS specification, in contrast, the coefficient is chosen to satisfy the following first order condition, in order to minimize the sum of the errors in log values:

$$\sum_{o,d,t} dist^t_{od} \cdot (\log(v^t_{od}) - \log(\hat{v}^t_{od})) = 0.$$

Comparison between the two first order conditions shows that, because PPML minimizes the level difference whereas the OLS minimizes the percentage difference, the PPML effectively places more weights on observations with larger export values. As shown in Figure 2a of the text, the distance gradient for export is larger among the city-port pairs that are particularly close to each other. Since these city-port pairs are the ones with the highest export volume, PPML results in a larger estimated distance effect. Columns 3 and 4 of Table A.3 exclude city-port pairs that are less than 100 km apart. They show that, first, excluding these observations indeed brings the PPML estimate much closer to the OLS estimates. Second, that the overt-time estimate is significantly smaller than cross-sectional estimate continues to hold.

Table A.3: Alternative PPML Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	Replicate baseline		Exclude dist<100 km		Include zeros	
$dist^t_{od}$	-0.706***	-0.350***	-0.423***	-0.239***	-0.714***	-0.321***
	(0.042)	(0.054)	(0.032)	(0.039)	(0.043)	(0.050)
Observations	3660	2838	3542	2694	5852	4328
Fixed Effects	ot, dt	od, ot, dt	ot, dt	od, ot, dt	ot, dt	od, ot, dt

Notes: This table reports additional results using PPML. The dependent variable is total value of export from city o to the RoW through port d. Columns 1 and 2 reproduce Columns 2 and 3 of Table A.2 for ease of comparison; Columns 3 and 4 exclude city-port pairs that are less than 100 km apart; Columns 5 and 6 include observations with zero export values.
Standard errors are clustered at city-port level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The role of zeros. To keep the specification consistent with OLS, we have excluded zeros in the PPML specifications. Columns 5 and 6 of Table A.3 show that this choice does not affect our estimates materially: including zeros increases the number of observations by around 50%, but the estimated coefficients are similar to those reported in Columns 1 and 2 of the table.

A.8 Measuring Shipment using Weights

Our estimation has used the value of goods to measure shipment, which is a commonly used measure in the trade literature. Below we show that all results are similar if we measure shipment by their weight, which is a theory-consistent measure when heterogeneity in transport costs is allowed.

Table A.4 replicates Table 2 with the log of shipment weight as the dependent variable. Although the coefficients change slightly compared to the baseline, the main points are robust: 1) the cross-sectional specifications overestimate the distance effect by as much as 100%; 2) the IV estimates are quantitatively similar to the OLS estimates; and 3) when both are included, the coefficient for regular roads is bigger than that for expressways.

A.9 Illustration of the IV

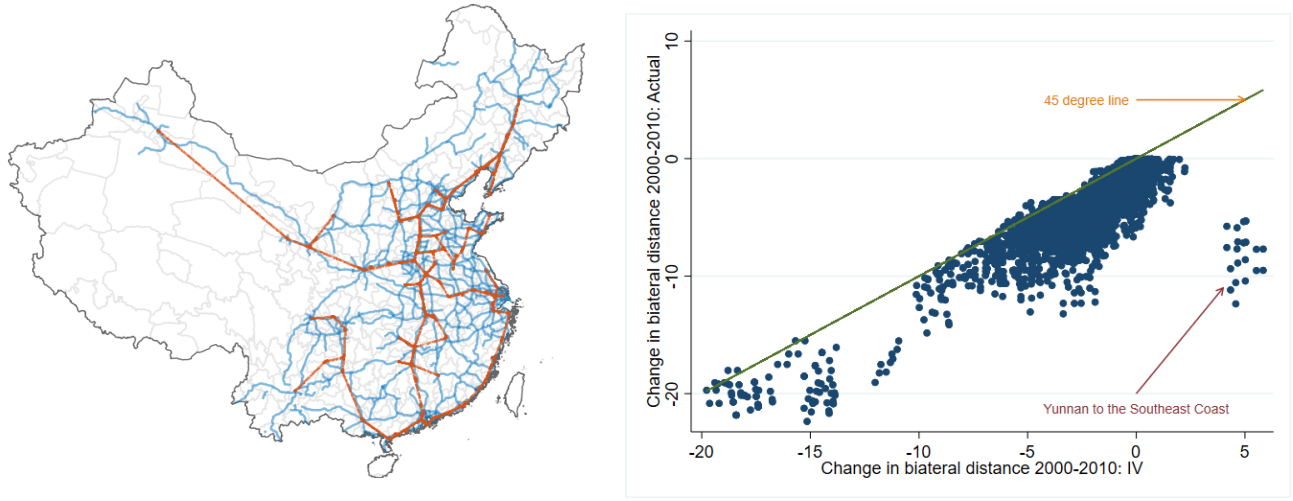
Figure A.3 shows the hypothetical expressway network used to construct the IV and illustrates the variation underlying the first stage regression. In the left panel, the blue lines indicate the actual expressways in 2010; the red lines indicate the minimum-length hypothetical expressways that connect all major cities. Like the actual network in 2010, the hypothetical network covers the entire country, but it consists of mostly straight lines, and is far less dense.

We use the hypothetical network to construct the IV for $dist^{2010}_{od}$ and use $dist^{2000}_{od}$ as an IV for itself. Given that the panel has exactly two periods, with city-port fixed effects controlled for, the first stage of the two-stage least square is essentially regressing the change in actual bilateral distance on the change predicted by the IV. Panel b of Figure A.3 illustrates the correlation between the two changes. The distance changes predicted by the IV are strongly correlated with the actual changes. The dots are mostly below the 45-degree line, reflecting that the minimum-spanning network is more sparse than the actual network. Finally, as annotated in the figure, the IV predicts that a small groups of cities—in the southwestern Yunnan Province—to have increased in distance to ports. This happens because the minimum-spanning network removes one road along the southern border of China that connects Yunnan to Guangzhou–according to the minimum-spanning algorithm, this link should not have been

Table A.4: Robustness with Weight of Shipment as the Dependent Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Effective Route Length and Export						By Type of Road		
	OLS				IV Reduced Form	2SLS	OLS	IV Reduced Form	2SLS
$dist^t_{od}$	-0.363***	-0.412***	-0.191***	-0.217***		-0.231***			
	(0.011)	(0.012)	(0.042)	(0.052)		(0.067)			
-on express							-0.088**		-0.163**
							(0.044)		(0.080)
-on regular							-0.215***		-0.248***
							(0.052)		(0.074)
IV $dist^t_{od}$					-0.268***				
					(0.079)				
-IV express								-0.161**	
								(0.063)	
-IV regular								-0.285***	
								(0.086)	
Fixed Effects	o,d,t	ot,dt	od,ot,dt	od,ot,dt	od,ot,dt	od,ot,dt	od,ot,dt	od,ot,dt	od,ot,dt
Exclude Major Cities				yes	yes	yes	yes	yes	yes
Observations	3612	3603	2786	2024	1996	1996	2024	1996	1996
R^2	0.606	0.680	0.893	0.884	0.884	0.023	0.884	0.884	0.018
First Stage K-P F Stat						1356.045			163.977

Notes: This table replicates Table 2 using the log of weight of export as the dependent variable. See notes under Table 2 for more information. Standard errors are clustered at city-port level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) The Minimum Spanning 2010 Expressway Network

(b) First Stage in Differences

Figure A.3: Illustration of the IV

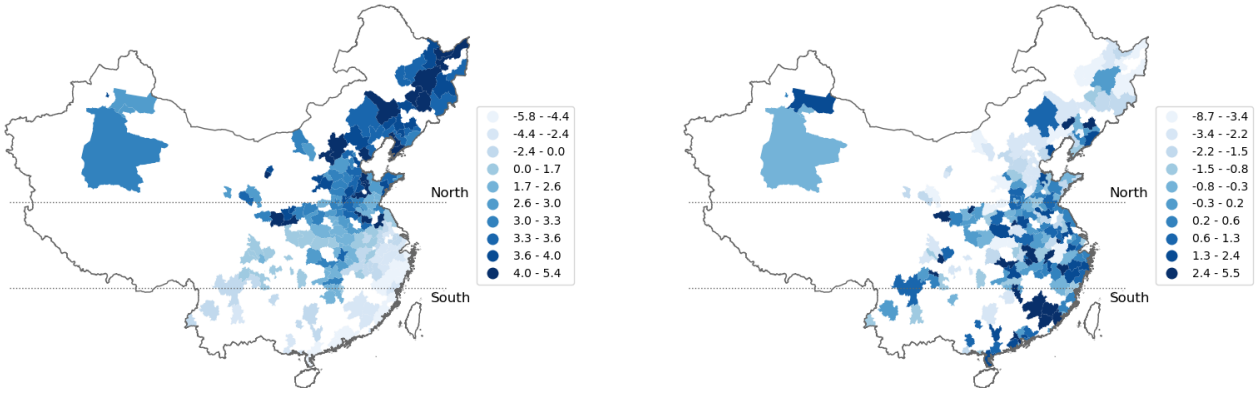
Note: The left panel overlays the minimum spanning tree (red) on the 2010 expressway network (blue); the right panel plots the first-stage regression.

built.

A.10 Visualizing Variation by Geographic Regions

Figure 2 presented in the text does not show the geographic dimension of the variation. In Figure A.4 we further show how each city' differential improvements in access to ports affect their choice of ports.

The left panel plots the *relative* reduction in cities' distances to two *groups* of ports due to the expressway construction during the decade. Focusing on the difference between the two port groups reduces the dimensionality of data from route-level to city-level, so the variation can be shown on a map. The



(a) Relative Change in Port Access: North Minus South

(b) Relative Growth in Export: North Minus South

Figure A.4: Relative Growth in Export: North Minus South

Note: The figure covers only cities exporting through both groups of ports in both periods.

first group comprises ports in the north (above the upper dotted line), while the second group comprises ports in the southeastern China (below the lower dotted line). For each city o, we first calculate its average regular-equivalent distance to ports in each of the two groups, denoted by $dist^t_{o,\text{North}}$ and $dist^t_{o,\text{South}}$, respectively. The left panel of Figure A.4 plots $(dist^{2010}_{o,\text{North}} - dist^{2000}_{o,\text{North}}) - (dist^{2010}_{o,\text{South}} - dist^{2000}_{o,\text{South}})$ by city. Dark colors indicate that an exporting city experienced a larger decrease in its distance to southern ports than to northern ports; light colors indicate the opposite. Northern cities tend to have much improved access to the ports in the south. Southern cities, which were already well connected to the southeastern coast before the massive expressway construction, experienced more substantial decreases in their distance to northern ports.

The right panel shows the relative growth in the export of city o through northern and southern ports, defined as $(\log(v^{2010}_{o,\text{North}}) - \log(v^{2000}_{o,\text{North}})) - (\log(v^{2010}_{o,\text{South}}) - \log(v^{2000}_{o,\text{South}}))$. Positive values indicate more rapid growth in export through northern ports than through southern ports; negative values indicate the opposite. The figure shows that southern cities saw more rapid export growth through northern ports than through southern ports. The two panels show that cities with dark colors on the right panel tend to have light colors in the left panel—-the cities experiencing larger improvements in access to a port group also export more through that group, suggestive evidence that export routing responds to domestic transport infrastructure improvements.

The contrast between the two panels in Figure A.4 highlights that the variation among individual cities in their access to different groups of ports is one source of identification. If this is the only source of variation exploited, however, a potential concern is that the increasing connectedness between southern cities and northern ports, and between northern cities and southern ports, could be driven by other macroeconomic trends that increased overall connectedness between broad geographic regions within China. The IV might not fully eliminate this concern as the minimum spanning network is also designed to connect broad geographic regions.

Table A.5: Result from Within-Region Variation Only

	(1)	(2)	(3)	(4)	(5)	(6)
	Aggregate Data			Sectoral Data		
$dist^t_{od}$	-0.471***	-0.173***		-0.501***	-0.187***	
	(0.024)	(0.060)		(0.028)	(0.055)	
-on express			-0.058			-0.080*
			(0.047)			(0.043)
-on regular			-0.208***			-0.193***
			(0.066)			(0.055)
Fixed Effects	$ot, dt, r_o g_d t$	$ot, dt, r_o g_d t, od$	$ot, dt, r_o g_d t, od$	$oit, dit, r_o g_d t$	$oit, dit, r_o g_d t, odi$	$oit, dit, r_o g_d t, odi$
Observations	2752	2068	2068	20903	11682	11682
R^2	0.706	0.898	0.879	0.630	0.901	0.901

Notes: This table shows that using only variation within geographic regions gives similar point estimates and also leads to the conclusion that the cross-sectional estimate is larger than the over-time estimate. All specifications are estimated using OLS. In 'Fixed Effects', o, d, and t stand for origin city, port, and time, respectively; r_o stands for the big region that city o belong to and g_d stands for the geographic group of port d; i stands for sector. Standard errors are clustered at city-port level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To address this potential concern, we show in Appendix Table A.5 that there is also variation among cities and ports *within* broad regions, which our baseline estimate also exploits. To show this, all specifications in Table A.5 also control for $r(o) - g(d) - t$ fixed effects, in which $r(o)$ stands for the geographic region that city o belongs to, $g(d)$ stands for the group a port d belongs to, and t stands for time. The grouping of cities into regions is based on a geographic classification that splits China into seven regions: north, northeast, east, central, south, southwest, and northwest, each containing on average 5 provinces. The grouping of ports into geographic groups is the same as in Figure A.4, in which there are north, central, and south three port groups. These region-port group-time fixed effects absorb all variation from changes in overall connectedness between broad geographic regions in China. Once they are controlled for, identification comes from the variation in distance to port *within* pairs of broad geographic regions, arising from changes in local expressway connections due to the expansion.

Columns 1 through 3 use the aggregate data and show that focusing entirely on local variation tells a consistent story: that moving from cross-sectional to over-time estimate significantly reduces the point estimate, and that regular roads are more costly than expressways. The coefficients are also around the same magnitudes as the baseline estimate. Columns 4 through 6 use sectoral level data with fixed effects related to sectors controlled for. The conclusions are similar.

Figure A.4 and Table A.5 together demonstrate the two sources of variation exploited in the empirical exercise: one from changes in the connectedness between broad geographic regions; one from local variation in expressway access within individual geographic regions.

B Model

B.1 Deriving Equation (5)

We prove that when truck drivers choose from two networks (regular roads represented by $\mathbb{L}$ and expressways represented by $\mathbb{H}$), the expected trade cost can be derived based on the combined adjacency matrix $\mathbb{A} = \mathbb{L} + \mathbb{H}$. Moreover, the expected trade cost is given by Equation (5).

We prove this by mathematical induction. First, consider the average cost of going from o to d among

all routes with only one edge.

$$\tau_{od,1} = \Gamma(\frac{\theta-1}{\theta})\left([\mathbb{L}_{(o,d)}] + [\mathbb{H}_{(o,d)}]\right)^{-\frac{1}{\theta}} = \Gamma(\frac{\theta-1}{\theta})\left([\mathbb{A}_{(o,d)}]\right)^{-\frac{1}{\theta}}.$$

Note also that if the od-th element of both $\mathbb{L}$ and $\mathbb{H}$ are zero, then $\tau_{od,1} = \infty$, meaning there is no feasible one-edge path from o to d.

Assuming that the sum of the $-\theta$th power of cost from o to d across all paths with exactly N steps is $[\mathbb{A}^N_{(o,d)}]$, then the sum across all paths with exactly $N+1$ steps is:

$$[(\mathbb{A}^N \cdot \mathbb{H} + \mathbb{A}^N \cdot \mathbb{L})_{(o,d)}].$$

The first part sums across all the paths that first gets to an adjacent city of d in exactly N steps and then goes on to d through an expressway; the second part sums across all the paths that gets to an adjacent city of d in N steps and then goes on to d through a regular road.

The above expression equals exactly $[\mathbb{A}^{N+1}_{(o,d)}]$. In other words, $[\mathbb{A}^{N+1}_{(o,d)}]$ is the sum across all paths that go from o to d in exactly $N+1$ steps. The average cost across all paths is thus:

$$\tau_{od} = \lim_{N\to\infty} \tau_{od,N} = \Gamma(\frac{\theta-1}{\theta})\left(\sum_{N=1}^{\infty} [\mathbb{A}^N]_{(o,d)}\right)^{-\frac{1}{\theta}} = \Gamma(\frac{\theta-1}{\theta})\mathbb{B}^{-\frac{1}{\theta}}_{(o,d)},$$

where $\mathbb{B} \equiv \left(\mathbb{I} - \mathbb{A}\right)^{-1}$, and $\mathbb{A} \equiv \mathbb{L} + \mathbb{H}$.

B.2 The Probability of Repeated Segments in a Route

The routing model in principle allows routes to have repeated segments. For example, a trucker going from LA to SD would choose an LA-SF-LA-SF-SD trip with a positive probability. This choice is strictly dominated absent trucker heterogeneity, because it involves repeated use of the LA-SF segment.

Proposition B.1 derives the probability that a segment is being used more than once in a trip, conditional on it is being used. It shows that at the estimated range of our parameters, the probability of such event happening is very small.

Proposition B.1. *Denote $\pi^{kl,(n)}_{od}$ the fraction of ground-transported shipment between o and d that passes edge kl for n time(s). Then,*

$$\pi^{kl,(n)}_{od} = \frac{\tilde{b}_{ok}(a_{kl}\tilde{b}_{lk})^{n-1}a_{kl}\tilde{b}_{ld}}{b_{od}}, \forall n \geq 1,$$

where $\tilde{b}_{od}$ is the od-th element of $\tilde{\mathbb{B}}$, with $\tilde{\mathbb{B}} = (\mathbb{I} - \tilde{\mathbb{A}})^{-1}$ and $\tilde{\mathbb{A}}$ being equal to $\mathbb{A}$ except that the kl-th element is set to zero, and a_{kl} is the kl-th element of $\mathbb{A}$.

Proof. Denote P_{od} the set of all paths from o to d. Denote $P^{kl,(n)}_{od,K}$ the set of paths from o to d of K steps that passes edge kl for n time(s). Then, under the assumption that the path-specific dis-utility follows the Fréchet distribution, $\pi^{kl,(n)}_{od}$ satisfies

$$\pi^{kl,(n)}_{od} = \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{p\in P^{kl,(n)}_{od,K}} r_p,$$

where $p = (p_0 = o, p_1, ..., p_{K-1}, p_K = d)$ denotes a path of K steps, with its k-th element being the k-th node of the path, and

$$r_p \equiv \prod_{k=1}^{K} a_{p_{k-1},p_k},$$

with a_{kl} being the kl-th element of $\mathbb{A}$ and b_{od} being the od-th element of $\mathbb{B}$. Now consider $\pi_{od}^{kl,(1)}$, it can be written as

$$\pi_{od}^{kl,(1)} = \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{B_1=0}^{K-1} \sum_{B_2=0}^{K-1} \mathbb{1}_{B_1+B_2=K-1} \times \Big(\sum_{p \in P_{ok,B_1}^{kl,not}} r_p \Big) \times a_{kl} \times \Big(\sum_{q \in P_{ld,B_2}^{kl,not}} r_q \Big)$$

where $\mathbb{1}_{B_1+B_2=K-1}$ is an indicator function that takes one if and only if $B_1 + B_2 = K - 1$, and $P_{od,B}^{kl,not}$ denotes the set of paths from o to d of step B that *does not* pass edge kl. Presenting the summation in a compact form, we can write $\pi_{od}^{kl,(1)}$ as

$$\pi_{od}^{kl,(1)} = \frac{\tilde{b}_{ok} a_{kl} \tilde{b}_{ld}}{b_{od}}, \tag{B.1}$$

where $\tilde{b}_{od}$ is the od-th matrix of $\tilde{\mathbb{B}}$, with $\tilde{\mathbb{B}} = (\mathbb{I} - \tilde{\mathbb{A}})^{-1}$ and $\tilde{\mathbb{A}}$ being equal to $\mathbb{A}$ except that the kl-th element is set to zero. Intuitively, the numerator of (B.1) enumerates r_p for all paths p that first take an arbitrary number of steps to go from o to k without passing kl, then pass kl, and next take an arbitrary number of steps to go from l to d without passing kl.

Similarly, we can write $\pi_{od}^{kl,(2)}$ as

$$\begin{aligned}\pi_{od}^{kl,(2)} &= \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{B_1=0}^{K-2} \sum_{B_2=0}^{K-2} \sum_{B_3=0}^{K-2} \mathbb{1}_{B_1+B_2+B_3=K-2} \times \Big(\sum_{p \in P_{ok,B_1}^{kl,not}} r_p \Big) \times a_{kl} \times \Big(\sum_{q \in P_{lk,B_2}^{kl,not}} r_q \Big) \times a_{kl} \times \Big(\sum_{s \in P_{ld,B_3}^{kl,not}} r_s \Big) \\ &= \frac{\tilde{b}_{ok} a_{kl} \tilde{b}_{lk} a_{kl} \tilde{b}_{ld}}{b_{od}}.\end{aligned}$$

Similarly, we can show that

$$\pi_{od}^{kl,(n)} = \frac{\tilde{b}_{ok} (a_{kl} \tilde{b}_{lk})^{n-1} a_{kl} \tilde{b}_{ld}}{b_{od}}, \forall n \geq 1.$$

□

We now apply Proposition B.1 to calculate the probability of passing an edge kl more than once, conditional on passing kl, denoted by Q_{od}^{kl},

$$\begin{aligned}Q_{od}^{kl} &= \frac{\sum_{m=2}^{\infty} \pi_{od}^{kl,(m)}}{\sum_{m=1}^{\infty} \pi_{od}^{kl,(m)}} \\ &= \frac{\sum_{m=2}^{\infty} (a_{kl} \tilde{b}_{lk})^{m-1}}{\sum_{m=1}^{\infty} (a_{kl} \tilde{b}_{lk})^{m-1}} = a_{kl} \tilde{b}_{lk} \equiv Q^{kl},\end{aligned}$$

which is irrelevant of od.

We calculate $Q^{kl}, \forall kl$ that forms an edge (i.e., $a_{kl} > 0$). Table B.1 reports the distribution of Q^{kl} across all edges. As shown, at the calibrated $\theta = 111.5$, the mean of Q^{kl} is 0.3%, the median is less than 0.1%, and the 95th percentile is only around 1.2%. This shows that the likelihood for passing an edge more than once is rather low. Other things equal, increasing θ further lowers the likelihood for repetitive passing and lowering θ increases the likelihood. But overall, the likelihood remains quite low for the range of θ estimated.

Table B.1: Conditional Probability of Passing an Edge More than Once

θ	Min	Max	Mean	Median	p95	p99	p99.9
80	0.000	0.768	0.017	0.000	0.080	0.352	0.767
111.5	0.000	0.332	0.003	0.000	0.012	0.054	0.324
200	0.000	0.045	0.000	0.000	0.000	0.002	0.043

B.3 Definition of Equilibrium

We define the competitive equilibrium as a set of prices and quantities that satisfy a set of conditions described below.

Definition 1. *Given fundamentals* $\{\tilde{\tau}^i_{od}, T^i_d, B_d, \overline{H}_d, L_{CHN}, L_{RoW}\}$,[5] *a competitive equilibrium is: consumer utility* U_d, *consumption of land* H_d *and sectoral final goods* C^i_d, *labor allocations* L_d *and* l^i_d, *quantities of sectoral final goods used as intermediate input* m^{ij}_d, *quantities of sectoral final goods produced* Q^i_d, *quantity of intermediate goods traded* $\tilde{q}^i_{od}$, *quantity of intermediate goods produced* q^i_d, *lump-sum transfers for domestic regions and RoW* (Tr, Tr_{RoW}), *rental prices of land* R_d, *prices of final goods* P^i_d, *import prices of intermediate goods* p^i_{od}, *unit production costs of intermediate goods* c^i_o, *and wages* w_d, *s.t.*

- *Consumers' optimality conditions hold:*

$$U_d = B_d[H_d]^{\alpha^0}\prod_{i=1}^{S}[C^i_d]^{\alpha^i},$$
$$\alpha^0 I_d = R_d H_d,$$
$$\alpha^i I_d = P^i_d C^i_d,$$

where $I_d = w_d + Tr$, $\forall d \in CHN$ *and* $I_{RoW} = w_{RoW} + Tr_{RoW}$.

[5] The solution to the transport mode choice and the drivers' routing problem have been incorporated through the trade cost matrix $\tilde{\tau}^i_{od}$.

- *Intermediate good producers' optimality conditions hold:*

$$
\begin{aligned}
q_d^i &= T_d^i [l_d^i]^{\beta^i} \prod_{j=1}^{S} [m_d^{ij}]^{\gamma^{ij}}, \\
c_d^i &= \kappa^i w_d^{\beta^i} \prod_{j=1}^{S} [P_d^j]^{\gamma^{ij}} / T_d^i, \\
P_d^j m_d^{ij} &= \gamma^{ij} c_d^i q_d^i, \\
w_d l_d^i &= \beta^i c_d^i q_d^i, \\
p_{od}^i &= c_o^i \tilde{\tau}_{od}^i,
\end{aligned}
$$

  where $\kappa^i = (\beta^i)^{-\beta^i} \prod_{j=1}^{S} (\gamma^{ij})^{-\gamma^{ij}}$.

- *Final good producers' optimality conditions hold:*

$$
\begin{aligned}
Q_d^i &= \Big(\sum_o [\tilde{q}_{od}^i]^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}}, \\
P_d^i &= \Big(\sum_o [p_{od}^i]^{1-\sigma} \Big)^{\frac{1}{1-\sigma}}, \\
\tilde{q}_{od}^i &= Q_d^i [\frac{p_{od}^i}{P_d^i}]^{-\sigma}.
\end{aligned}
$$

- *Markets clear for labor, land, final goods, and intermediate goods:*

$$
\begin{aligned}
\sum_i l_d^i = L_d, &\qquad \textit{(Labor markets clear)} \\
H_d L_d = \overline{H}_d, &\qquad \textit{(Land markets clear)} \\
\sum_d \tilde{\tau}_{od}^i \tilde{q}_{od}^i = q_o^i, &\qquad \textit{(Intermediate good markets clear)} \\
\sum_i m_d^{ij} + C_d^j L_d = Q_d^j. &\qquad \textit{(Final good markets clear).}
\end{aligned}
$$

- *Rents from land are rebated via lump-sum transfers:*

$$
\begin{aligned}
\sum_{d \in CHN} R_d H_d L_d &= Tr \cdot L_{CHN}, \\
R_{RoW} H_{RoW} L_{RoW} &= Tr_{RoW} \cdot L_{RoW}.
\end{aligned}
$$

- *Domestic workers are mobile:*

$$
\begin{aligned}
U_d = U_{d'}, \forall d, d' \in CHN. \\
\sum_{d \in CHN} L_d = L_{CHN}.
\end{aligned}
$$

We also state the definitions of other equilibrium objects used in the main text and the appendix, that can be written

as functions of the equilibrium objects defined above.

- *The total expenditure on intermediate goods in sector i of region d*

$$E_d^i \equiv P_d^i Q_d^i.$$

- *The value of trade flows from o to d in sector i*

$$X_{od}^i \equiv p_{od}^i \tilde{q}_{od}^i = E_d^i \pi_{od}^i,$$

where $\pi_{od}^i \equiv [\frac{p_{od}^i}{P_d^i}]^{1-\sigma}$.

B.4 Proof of Lemma 1 and Proposition 1

We first state a lemma characterizing the first order effect of the segment shipment cost on the trade cost.

Lemma B.1. *The entries of $\mathbb{A}$ and of its Leontief inverse $\mathbb{B} \equiv (\mathbb{I} - \mathbb{A})^{-1}$ satisfy*

$$\frac{\partial \log b_{od}}{\partial \log a_{kl}} = \frac{b_{ok} \cdot a_{kl} \cdot b_{ld}}{b_{od}},$$

where a_{kl} and b_{od} are the kl-th and od-th elements of $\mathbb{A}$ and $\mathbb{B}$, respectively.

Proof. Apply the the formula for the derivative of the inverse of a matrix we have

$$\begin{aligned}\frac{\partial \mathbb{B}}{\partial \log a_{kl}} &= -(\mathbb{I} - \mathbb{A})^{-1} \frac{\partial (\mathbb{I} - \mathbb{A})}{\partial \log a_{kl}} (\mathbb{I} - \mathbb{A})^{-1} \\ &= \mathbb{B}(\mathbb{E}_{kl} \circ \mathbb{A})\mathbb{B},\end{aligned}$$

where $\mathbb{E}_{kl}$ is a matrix of the same size as $\mathbb{A}$, with the kl-th element being one and other elements being zero. Therefore,

$$\frac{\partial \log b_{od}}{\partial \log a_{kl}} = \frac{b_{ok} \cdot a_{kl} \cdot b_{ld}}{b_{od}}.$$

□

Denote $\tilde{\pi}_{od}^{kl} \equiv \frac{b_{ok} \cdot a_{kl} \cdot b_{ld}}{b_{od}}$, we prove the following lemma.

Lemma B.2. *With $\tilde{\pi}_{od}^{kl}$ defined above, we have the following*

(1) $\dfrac{\partial \log \tilde{\tau}_{od}^i}{\partial \log \iota_{kl}} = \pi_{od}^{road} \tilde{\pi}_{od}^{kl},$

(2) $\dfrac{\partial^2 \log \tilde{\tau}_{od}^i}{\partial \log \iota_{kl} \partial \log \iota_{k'l'}} = \pi_{od}^{road} \tilde{\pi}_{od}^{kl} \Big(-\theta[\mathbb{1}(kl = k'l') + \tilde{\pi}_{ok}^{k'l'} + \tilde{\pi}_{ld}^{k'l'} - \tilde{\pi}_{od}^{k'l'}] - \theta_M (1 - \pi_{od}^{road}) \tilde{\pi}_{od}^{k'l'} \Big).$

Proof. First, from

$$\tilde{\tau}_{od}^i = \Gamma(\frac{\theta_M - 1}{\theta_M})[(\bar{\tau}_{od}^i)^{-\theta_M} + (\tau_{od}^i)^{-\theta_M}]^{-1/\theta_M},$$

we have

$$\frac{\partial \log \tilde{\tau}_{od}^i}{\partial \log \tau_{od}^i} = \pi_{od}^{road} = \frac{(\tau_{od}^i)^{-\theta_M}}{(\bar{\tau}_{od}^i)^{-\theta_M} + (\tau_{od}^i)^{-\theta_M}}. \tag{B.2}$$

Next, recall that $\tau_{od}^i = (\frac{h_i}{h_0})^{\mu} b_{od}^{-1/\theta}$ and $\iota_{kl} = a_{kl}^{-1/\theta}$. Applying Lemma B.1, we have

$$\frac{\partial \log \tau_{od}^i}{\partial \log \iota_{kl}} = \frac{\partial \log b_{od}}{\partial \log a_{kl}} = \frac{b_{ok} \cdot a_{kl} \cdot b_{ld}}{b_{od}} = \tilde{\pi}_{od}^{kl}.$$

This proves part (1).

Now from Equation (B.2), we have

$$\frac{\partial \log(\pi_{od}^{road})}{\partial \log \tau_{od}^i} = -\theta_M(1 - \pi_{od}^{road}).$$

Combining with

$$\frac{\partial \log \tau_{od}^i}{\partial \log \iota_{k'l'}} = \tilde{\pi}_{od}^{k'l'},$$

we have

$$\frac{\partial \log(\pi_{od}^{road})}{\partial \log \iota_{k'l'}} = -\theta_M(1 - \pi_{od}^{road})\tilde{\pi}_{od}^{k'l'}. \tag{B.3}$$

Next start with the definition of $\tilde{\pi}_{od}^{kl}$ we have,

$$\log \tilde{\pi}_{od}^{kl} = \log b_{ok} + \log a_{kl} + \log b_{ld} - \log b_{od}.$$

Take derivative with respect to $\log \iota_{k'l'}$ and apply Lemma B.1, we have

$$\frac{\partial \log \tilde{\pi}_{od}^{kl}}{\partial \log \iota_{k'l'}} = -\theta(\tilde{\pi}_{ok}^{k'l'} + \tilde{\pi}_{ld}^{k'l'} - \tilde{\pi}_{od}^{k'l'} + \mathbb{1}(kl = k'l')). \tag{B.4}$$

Combining (B.3) and (B.4) we arrive at

$$\frac{\partial(\pi_{od}^{road}\tilde{\pi}_{od}^{kl})}{\partial \log \iota_{k'l'}} = \pi_{od}^{road}\tilde{\pi}_{od}^{kl}[-\theta_M(1 - \pi_{od}^{Road})\tilde{\pi}_{od}^{k'l'} - \theta(\tilde{\pi}_{ok}^{k'l'} + \tilde{\pi}_{ld}^{k'l'} - \tilde{\pi}_{od}^{k'l'} + \mathbb{1}(kl = k'l'))].$$

This proves part (2). □

We are now ready to prove Lemma 1.

Proof for Lemma 1.

Proof. With Lemma B.1, it suffices to prove that $\tilde{\pi}_{od}^{kl}$ converges to the fraction of ground-transported shipments between o and d that passes edge kl, denoted by π_{od}^{kl}, as θ goes to infinity. In fact, we prove a stronger result below:

$$\lim_{\theta\to\infty} \frac{\pi_{od}^{kl} - \tilde{\pi}_{od}^{kl}}{\pi_{od}^{kl}} = 0, \text{ if } \pi_{od}^{kl} > 0.$$

That is, not only the level error but also the relative error converges to zero.[6]

First, under the assumption that the path-specific dis-utility follows the Fréchet distribution, π_{od}^{kl} satisfies

$$\pi_{od}^{kl} = \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{p\in P_{od,K}^{kl}} \prod_{k=1}^{K} a_{p_{k-1},p_k}, \tag{B.5}$$

where $p = (p_0 = o, p_1, ..., p_{K-1}, p_K = d)$ denotes a path of K steps, with its k-th element being the k-th node of the path, and $P_{od,K}^{kl}$ denotes the set of paths of K steps that pass edge kl. We partition $P_{od,K}^{kl}$ into two disjoint sets: $\overline{P}_{od,K}^{kl}$ and $\tilde{\mathcal{P}}_{od,K}^{kl}$, where $\overline{P}_{od,K}^{kl}$ denotes the subsets of paths that pass edge kl *only once*, and $\tilde{\mathcal{P}}_{od,K}^{kl}$ denotes the subsets of paths that pass edge kl *more than once*. Given any $\bar{p} \in \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}$, there exists $K' < \bar{K}$ and $p' \in \overline{P}_{od,K'}^{kl}$ for which p' is the path removing any loops in $\bar{p}$ that involve multiple passes of kl. Denote $\bar{a} = \max_{kl} a_{kl}$, then $\lim_{\theta\to\infty} \bar{a} = 0$ since $a_{kl} = \exp(-\kappa^L \theta \text{dist}_{kl}^L) + \exp(-\kappa^H \theta \text{dist}_{kl}^H)$. Therefore,

$$\frac{\prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}}{\sum_{K=0}^{\infty} \sum_{p\in P_{od,K}^{kl}} \prod_{k=1}^{K} a_{p_{k-1},p_k}} \leq \frac{\prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}}{\prod_{k=1}^{K'} a_{p'_{k-1},p'_k}} \leq \bar{a} \to 0, \tag{B.6}$$

as $\theta \to \infty$, where the first inequality shrinks the positive denominator, and the second inequality applies that p' is a path removing loops in $\bar{p}$ (i.e., $\bar{p}$ contains all segments in p' and additional detoured segments). Since $\cup_{\bar{K}=0}^{\infty} \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}$ is a countable set, multiply (B.6) by $\bar{K}$ and sum over all $\bar{p} \in \cup_{\bar{K}=0}^{\infty} \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}$, we have

$$\lim_{\theta\to\infty} \sum_{\bar{K}=0}^{\infty} \sum_{\bar{p}\in\tilde{\mathcal{P}}_{od,\bar{K}}^{kl}} \frac{\bar{K} \cdot \prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}}{\sum_{K=0}^{\infty} \sum_{p\in P_{od,K}^{kl}} \prod_{k=1}^{K} a_{p_{k-1},p_k}} = 0. \tag{B.7}$$

Now consider the summation

$$\varsigma_{od}^{kl} \equiv \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} \left(\sum_{p\in P_{ok,B}} \prod_{k=1}^{B} a_{p_{k-1},p_k} \right) \times a_{kl} \times \left(\sum_{q\in P_{kd,K-B-1}} \prod_{k=1}^{K-B-1} a_{q_{k-1},q_k} \right).$$

Then the cost of a path $\bar{p}$ of $\bar{K}$ steps that passes edge kl for $n \leq \bar{K}$ time(s), $\prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}$, will appear in the

[6] The stronger result allows $\pi_{od}^{kl} \to 0$ as $\theta \to \infty$.

above summation for exactly n times. Therefore,

$$\frac{\varsigma_{od}^{kl}}{b_{od}} \geq \pi_{od}^{kl} = \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{p \in P_{od,K}^{kl}} \prod_{k=1}^{K} a_{p_{k-1},p_k}$$
$$\geq \frac{1}{b_{od}} \left\{ \varsigma_{od}^{kl} - \sum_{\bar{K}=0}^{\infty} \sum_{\bar{p} \in \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}} \bar{K} \cdot \prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k} \right\},$$

where the first inequality applies that paths in $\tilde{\mathcal{P}}_{od,\bar{K}}^{kl}$ appear in the summation for more than one time, and the second inequality applies that paths in $\tilde{\mathcal{P}}_{od,\bar{K}}^{kl}$ appear in the summation no more than $\bar{K}$ times.

Notice that

$$\frac{\varsigma_{od}^{kl}}{b_{od}} = \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} \left(\sum_{p_1=1}^{N} \cdots \sum_{p_{B-1}=1}^{N} a_{o,p_1} \times \ldots \times a_{p_{B-1},k} \right) \times a_{kl} \times \left(\sum_{q_1=1}^{N} \cdots \sum_{q_{K-B-1}=1}^{N} a_{l,q_1} \times \ldots \times a_{q_{K-B-1},d} \right)$$
$$= \frac{1}{b_{od}} \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} \mathbb{A}_{ok}^{B} \times a_{kl} \times \mathbb{A}_{ld}^{K-B-1} = \frac{b_{ok} a_{kl} b_{ld}}{b_{od}} = \tilde{\pi}_{od}^{kl}.$$

Therefore,

$$\tilde{\pi}_{od}^{kl} \geq \pi_{od}^{kl} \geq \tilde{\pi}_{od}^{kl} - \frac{\sum_{\bar{K}=0}^{\infty} \sum_{\bar{p} \in \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}} \bar{K} \cdot \prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}}{b_{od}}$$
$$\Rightarrow 0 \geq \frac{\pi_{od}^{kl} - \tilde{\pi}_{od}^{kl}}{\pi_{od}^{kl}} \geq - \frac{\sum_{\bar{K}=0}^{\infty} \sum_{\bar{p} \in \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}} \bar{K} \cdot \prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}}{\sum_{K=0}^{\infty} \sum_{p \in P_{od,K}^{kl}} \prod_{k=1}^{K} a_{p_{k-1},p_k}}$$
$$\Rightarrow 0 \geq \lim_{\theta \to \infty} \frac{\pi_{od}^{kl} - \tilde{\pi}_{od}^{kl}}{\pi_{od}^{kl}} \geq - \lim_{\theta \to \infty} \frac{\sum_{\bar{K}=0}^{\infty} \sum_{\bar{p} \in \tilde{\mathcal{P}}_{od,\bar{K}}^{kl}} \bar{K} \cdot \prod_{k=1}^{\bar{K}} a_{\bar{p}_{k-1},\bar{p}_k}}{\sum_{K=0}^{\infty} \sum_{p \in P_{od,K}^{kl}} \prod_{k=1}^{K} a_{p_{k-1},p_k}} = 0,$$

where the second row applies (B.5) and that $\pi_{od}^{kl} > 0$, and the last equality applies (B.7). This completes the proof for $\lim_{\theta \to \infty} \frac{\pi_{od}^{kl} - \tilde{\pi}_{od}^{kl}}{\pi_{od}^{kl}} = 0$. □

To prove Proposition 1, we first prove the following series of lemmas. Define world welfare $\log \overline{W} \equiv \log W + \omega_{RoW} \log U_{RoW}$ with the Pareto weight $\omega_{RoW} = Y_{RoW}/Y$, where Y_{RoW} and Y are the RoW GDP and domestic GDP under the competitive equilibrium, respectively. Define Ω the expanded input-output matrix evaluated at the competitive equilibrium, encompassing all domestic and foreign final good producers, intermediate good producers, and traders.[7] Specifically, the kj-th entry of Ω is defined as

$$\Omega_{kj} \equiv \frac{\hat{p}_j \hat{q}_{kj}}{S_k},$$

where $\hat{p}_j$ is the price of good j, $\hat{q}_{kj}$ is the quantity of good j used by sector k, and S_k is the total revenue in sector k. Stack the productivities of all final-good producers, intermediate good producers, and traders

[7] Traders from place o to place d in sector i competitively convert intermediate goods of (o,i) to intermediate goods of (d,i). With such interpretations the iceberg trade costs are the inverse of the productivities of traders.

into a vector denoted by A, and the corresponding price into a vector denoted by $\hat{p}$. Define χ_j the total sales of sector j as the share of domestic GDP:

$$\chi_j \equiv \frac{\sum_k \hat{p}_j \hat{q}_{kj}}{Y}.$$

Lemma B.3 below associates the first-order effect of sectoral productivity on welfare with sectoral sales share, a result extending Hulten (1978) with international trade and domestic mobile labor.

Lemma B.3. *With $\overline{W}$, A_j and χ_j defined above,*

$$\frac{\mathrm{d}\log \overline{W}}{\mathrm{d}\log A_j} = \chi_j.$$

Proof. By Shephard's lemma,

$$\mathrm{d}\log(\hat{p}) = \Omega \mathrm{d}\log(\hat{p}) + \beta \mathrm{d}\log(w) - \mathrm{d}\log(A),$$

where β is the vector stacking the labor shares of all expanded sectors. So we have

$$\mathrm{d}\log(\hat{p}) = (\mathbb{I} - \Omega)^{-1}\Big(\beta \mathrm{d}\log(w) - \mathrm{d}\log(A)\Big). \tag{B.8}$$

Starting from the consumer utility in region d, $U_d = B_d \frac{I_d}{P_d}$, we have

$$\mathrm{d}\log U_d = \mathrm{d}\log(I_d) - \sum_{i=1}^{S} \alpha^i \mathrm{d}\log(P_d^i) - \alpha^0 \mathrm{d}\log(R_d), \forall d.$$

Combining the optimal housing expenditure and the housing market clearing conditions:

$$\alpha^0 I_d L_d = R_d \overline{H}_d, \forall d,$$

we have

$$\mathrm{d}\log U_d = (1-\alpha^0)\mathrm{d}\log I_d - \sum_{i=1}^{S} \alpha^i \mathrm{d}\log(P_d^i) - \alpha^0 \mathrm{d}\log(L_d). \tag{B.9}$$

Define Ξ_d a column vector, of which the k-th entry equals α^i if the k-th producer among all extended producers is the final good producer of sector i of region d. That is, Ξ_d maps final good producers of region d to the index of extended producers. With such definitions, (B.9) can be written as

$$\mathrm{d}\log U_d = (1-\alpha^0)\mathrm{d}\log I_d - \Xi_d' \mathrm{d}\log(\hat{p}) - \alpha^0 \mathrm{d}\log(L_d).$$

Plugging in (B.8) we have

$$\mathrm{d}\log U_d = (1-\alpha^0)\mathrm{d}\log I_d - \Xi_d'(\mathbb{I} - \Omega)^{-1}\Big(\beta \mathrm{d}\log(w) - \mathrm{d}\log(A)\Big) - \alpha^0 \mathrm{d}\log(L_d). \tag{B.10}$$

With domestic GDP $Y = \sum_{d \neq RoW} I_d L_d$ as the numeraire, multiply (B.10) by $I_d \cdot L_d$ both sides and sum

over d, and notice $U_d = W, \forall d \neq RoW$ by the domestic welfare equalization condition, we have

$$\begin{aligned} \mathrm{d}\log W + \frac{Y_{RoW}}{Y}\mathrm{d}\log U_{RoW} = \sum_d \Big[(1-\alpha^0) I_d L_d \mathrm{d}\log I_d - I_d L_d \Xi_d' (\mathbb{I}-\Omega)^{-1}\Big(\beta \mathrm{d}\log(w) - \mathrm{d}\log(A)\Big)\Big] \\ - \alpha^0 \sum_d \mathrm{d}\log(L_d) I_d L_d. \end{aligned} \tag{B.11}$$

The labor market clearing conditions imply:

$$\sum_d w_d L_d \mathrm{d}\log w_d - \sum_d I_d L_d \Xi_d' (\mathbb{I}-\Omega)^{-1}\beta \mathrm{d}\log(w) = 0. \tag{B.12a}$$

The normalization $Y = \sum_{d\neq RoW} I_d L_d = 1$ implies

$$\sum_{d\neq RoW} [I_d L_d \mathrm{d}\log I_d + \mathrm{d}\log(L_d) I_d L_d] = 0. \tag{B.12b}$$

The determination of domestic transfer and normalization implies

$$\begin{aligned} Tr = \frac{\alpha^0 \sum_{d\neq RoW} I_d L_d}{\sum_{d\neq RoW} L_d} \Rightarrow \mathrm{d}Tr = 0 \\ \Rightarrow \mathrm{d}I_d = \mathrm{d}(w_d + Tr) = \mathrm{d}w_d, \forall d \neq ROW \\ \Rightarrow I_d L_d \mathrm{d}\log I_d = w_d L_d \mathrm{d}\log w_d, \forall d \neq ROW. \end{aligned} \tag{B.12c}$$

The determination of RoW transfer implies

$$\begin{aligned} Tr_{RoW} = \alpha^0 I_{RoW} L_{RoW}, \quad I_{RoW} = Tr_{RoW} + w_{RoW} \\ \Rightarrow I_{RoW} = \frac{w_{RoW}}{1-\alpha^0} \Rightarrow \mathrm{d}I_{RoW} = \frac{1}{1-\alpha^0}\mathrm{d}w_{RoW}. \end{aligned} \tag{B.12d}$$

Plugging (B.12a)-(B.12d) and $\mathrm{d}L_{RoW} = 0$ to (B.11) we have

$$\begin{aligned} \mathrm{d}\log W + \frac{Y_{RoW}}{Y}\mathrm{d}\log U_{RoW} = \sum_d I_d L_d \Xi_d' (\mathbb{I}-\Omega)^{-1}\mathrm{d}\log(A) \\ = \sum_j \chi_j \mathrm{d}\log(A_j). \end{aligned}$$

□

Notice that the productivity of a trader from o to d in sector i is the inverse of the trade cost from o to d in sector i. And the sales of the trader is the sales of intermediate goods from o to d in sector i. Therefore, we have

Corollary 1. *With $\overline{W}$ defined above,*

$$\frac{\mathrm{d}\log \overline{W}}{\mathrm{d}\log \tilde{\tau}^i_{od}} = -\frac{X^i_{od}}{Y}.$$

We next characterize the exposure of RoW consumption to RoW import prices.

Lemma B.4. *Assume* $\mathrm{d}\log T^i_{RoW} = 0$ *(i.e., there is no change in RoW productivity). Then*

$$\mathrm{d}\log U_{RoW} = -\sum_{o,i} \frac{\Lambda^i_o}{Y_{RoW}} \mathrm{d}\log(p^i_{o,RoW}/Y_{RoW}),$$

where $\Lambda^i_o = Y_{RoW}[\alpha'\left(\mathbb{I} - \hat{\Omega}\right)^{-1}]_i \pi^i_{o,RoW}$*, in which* $\alpha' = (\alpha^1, \alpha^2, ..., \alpha^S)$*,* $\hat{\Omega}_{ij} = \gamma^{ij}\pi^i_{RoW,RoW}$*, and* $[x]_i$ *is the i-th element of row vector* x*.*

Proof. Denote $\hat{x} \equiv x/Y_{RoW}$, for $x = (P^i_{RoW}, c^i_{Row}, p^i_{o,RoW}, w_{RoW}, R_{RoW}, I_{RoW})$. Since $Y_{RoW} = I_{RoW}L_{RoW} = \frac{w_{RoW}L_{RoW}}{1-\alpha^0}$, and L_{RoW} is fixed, we have

$$\mathrm{d}\log \hat{w}_{RoW} = \mathrm{d}\log \hat{I}_{RoW} = 0.$$

By Shephard's lemma,

$$\mathrm{d}\log \hat{P}^i_{RoW} = \pi^i_{RoW,RoW}\mathrm{d}\log \hat{c}^i_{RoW} + \sum_{o'} \pi^i_{o',RoW}\mathrm{d}\log \hat{p}^i_{o',RoW}, \tag{B.13}$$

and since $\mathrm{d}\log T^i_{RoW} = 0$, we have

$$\mathrm{d}\log \hat{c}^i_{RoW} = \sum_j \gamma^{ij}\mathrm{d}\log \hat{P}^j_{RoW} + \beta^i \mathrm{d}\log \hat{w}_{RoW}. \tag{B.14}$$

Plugging (B.14) into (B.13) and applying $\mathrm{d}\log \hat{w}_{RoW} = 0$, in matrix form we have

$$\mathrm{d}\log \hat{P}_{RoW} = (\mathbb{I} - \hat{\Omega})^{-1}\Pi, \tag{B.15}$$

where $\log \hat{P}_{RoW} = (\log \hat{P}^i_{RoW}, ..., \log \hat{P}^S_{RoW})'$, and Π is a column vector with $\Pi_i = \sum_{o'} \pi^i_{o',RoW}\mathrm{d}\log \hat{p}^i_{o',RoW}$. Plug (B.15) to

$$\mathrm{d}\log U_{RoW} = \mathrm{d}\log \hat{I}_{RoW} - \alpha^0 \mathrm{d}\log \hat{R}_{RoW} - \alpha' \mathrm{d}\log \hat{P}_{RoW},$$

and note that $\mathrm{d}\log(\hat{R}_{RoW}) = 0$ since $R_{RoW}\overline{H}_{RoW} = \alpha^0 Y_{RoW}$ and $\overline{H}_{RoW}$ is fixed, we have the desired result. □

Proof of Proposition 1

Proof. Combine Corollary 1 and Lemma B.4 we arrive at

$$\frac{\mathrm{d}\log W}{\mathrm{d}\log \tilde{\tau}^i_{od}} = -\frac{X^i_{od}}{Y} - \frac{Y_{RoW}}{Y}\frac{\mathrm{d}\log U_{RoW}}{\mathrm{d}\log \tilde{\tau}^i_{od}},$$

where

$$\frac{\mathrm{d}\log U_{RoW}}{\mathrm{d}\log \tilde{\tau}^i_{od}} = -\sum_{o',i'} \frac{\Lambda^{i'}_{o'}}{Y_{RoW}} \frac{\mathrm{d}\log[p^{i'}_{o',RoW}/Y_{RoW}]}{\mathrm{d}\log \tilde{\tau}^i_{od}}.$$

Notice that $p^{i'}_{o',RoW} = \tilde{\tau}^{i'}_{o',RoW} \cdot c^{i'}_{o'}$, so

$$\frac{\mathrm{d}\log[p^{i'}_{o',RoW}/Y_{RoW}]}{\mathrm{d}\log\tilde{\tau}^{i}_{od}} = \mathbb{1}(i'=i, o'=o, d=RoW) + \frac{\mathrm{d}\log[c^{i'}_{o'}/Y_{RoW}]}{\mathrm{d}\log\tilde{\tau}^{i}_{od}},$$

and

$$\frac{\mathrm{d}\log U_{RoW}}{\mathrm{d}\log\tilde{\tau}^{i}_{od}} = -\sum_{o',i'}\frac{\Lambda^{i'}_{o'}}{Y_{RoW}}\frac{\mathrm{d}\log[p^{i'}_{o',RoW}/Y_{RoW}]}{\mathrm{d}\log\tilde{\tau}^{i}_{od}} = -\frac{\Lambda^{i}_{o}}{Y_{RoW}}\mathbb{1}_{d=RoW} - \sum_{o',i'}\frac{\Lambda^{i'}_{o'}}{Y_{RoW}}\frac{\mathrm{d}\log[c^{i'}_{o'}/Y_{RoW}]}{\mathrm{d}\log\tilde{\tau}^{i}_{od}}.$$

Therefore,

$$\frac{\mathrm{d}\log W}{\mathrm{d}\log\tilde{\tau}^{i}_{od}} = -\Big(\frac{X^{i}_{od}}{Y} - \frac{\Lambda^{i}_{o}}{Y}\mathbb{1}_{d=RoW}\Big) + \sum_{o',i'}\frac{\Lambda^{i'}_{o'}}{Y}\frac{\mathrm{d}\log[c^{i'}_{o'}/Y_{RoW}]}{\mathrm{d}\log\tilde{\tau}^{i}_{od}}.$$

Apply the first order Taylor expansion of $\log W$ with respect to $(\log\tilde{\tau}^{i}_{od})_{o,d,i}$, we have

$$\Delta\log W = -\sum_{o,d,i}\Big(\frac{X^{i}_{od}}{Y} - \frac{\Lambda^{i}_{o}}{Y}\mathbb{1}_{d=RoW}\Big)\Delta\log\tilde{\tau}^{i}_{od} + TOT + HO_T, \tag{B.16}$$

where $TOT = \sum_{o,d,i}\Big(\sum_{o',i'}\frac{\Lambda^{i'}_{o'}}{Y}\frac{\mathrm{d}\log[c^{i'}_{o'}/Y_{RoW}]}{\mathrm{d}\log\tilde{\tau}^{i}_{od}}\Big)\Delta\log\tilde{\tau}^{i}_{od}$, is the first order terms-of-trade effects, and HO_T is the higher order effect of trade costs on welfare. Further apply Taylor expansion of $\Delta\log\tilde{\tau}^{i}_{od}$ with respect to $(\Delta\log\iota_{kl})_{kl\in\mathbb{C}}$, we have

$$\Delta\log\tilde{\tau}^{i}_{od} = \sum_{kl\in\mathbb{C}}\frac{\partial\log\tilde{\tau}^{i}_{od}}{\partial\log\iota_{kl}}\Delta\log(\iota_{kl}) + \frac{1}{2}\sum_{kl\in\mathbb{C}}\sum_{k'l'\in\mathbb{C}}\frac{\partial^2\log\tilde{\tau}^{i}_{od}}{\partial\log\iota_{kl}\,\partial\log\iota_{k'l'}}\Delta\log(\iota_{kl})\Delta\log(\iota_{k'l'}) + \widetilde{HO}_R, \tag{B.17}$$

where $\widetilde{HO}_R$ is the effect of route costs on trade costs beyond the second order effect. Plugging (B.17) to (B.16), we have the desired result. □

B.5 Interaction Among Routes

We characterize the interaction among different routes and illustrate it through an example in the calibrated model.

We can view a large project as a *collection* of expressway segments. Proposition 1 shows that the second order effect from rerouting, denoted by SO_R, is

$$SO_R = -\frac{1}{2}\sum_{o,d,i}\Big(\frac{X^{i}_{od}}{Y} - \frac{\Lambda^{i}_{o}}{Y}\mathbb{1}_{d=RoW}\Big)\sum_{kl\in\mathbb{C}}\sum_{k'l'\in\mathbb{C}}\frac{\partial^2\log\tilde{\tau}^{i}_{od}}{\partial\log\iota_{kl}\,\partial\log\iota_{k'l'}}\Delta\log(\iota_{kl})\Delta\log(\iota_{k'l'}),$$

in which $\sum_{kl\in\mathbb{C}}\sum_{k'l'\in\mathbb{C}}\frac{\partial^2\log\tilde{\tau}^{i}_{od}}{\partial\log\iota_{kl}\,\partial\log\iota_{k'l'}}$ captures the own second-order effect (when $kl = k'l'$) as well as potential complementary and substitution among different routes (when $kl \neq k'l'$).

To see what the interaction effect entials, we use Lemma 1 to write the cross-derivative term in the

summand as:

$$\frac{\partial^2 \log \tilde{\tau}^i_{od}}{\partial \log \iota_{kl} \partial \log \iota_{k'l'}} \approx \pi^{road}_{od} \pi^{kl}_{od} \Big\{ \underbrace{-\theta[\mathbb{1}(kl = k'l') + \pi^{k'l'}_{ok} + \pi^{k'l'}_{ld} - \pi^{k'l'}_{od}]}_{\text{Rerouting of ground traffic}} - \underbrace{\theta_M (1 - \pi^{road}_{od}) \pi^{k'l'}_{od}}_{\text{mode switch}} \Big\}, \tag{B.18}$$

where $\mathbb{1}(kl = k'l')$ is the indicator function that takes one if $kl = k'l'$ and zero otherwise

Consider first the own effect (when $kl = k'l'$). The first term in the curly bracket captures the impact on shipment over $k \to l$ through rerouting *within* the road network. With $1 + \pi^{kl}_{ok} + \pi^{kl}_{ld} - \pi^{kl}_{od} > 0$, this force contributes negatively: a *decrease* in the cost on edge $k \to l$ *increases* the share of shipments taking this edge. The second term in the bracket captures the response in the mode choice—more shipment will be made via road in response to a decrease in the edge cost. Both forces work in the same direction and imply that as an expressway is added to $k \to l$, more trade flows will go through this edge.

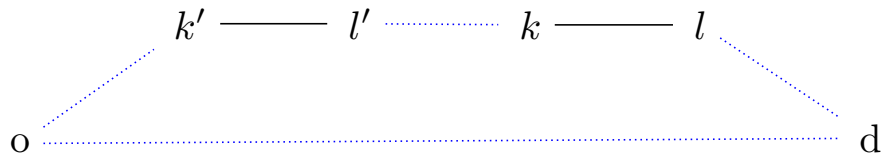


Figure B.1: Interactions Between Segments

Note: The diagram illustrates a case in which expressway in $k' \to l'$ and $k \to l$ complement each other.

Now consider the cross-derivative (when $kl \neq k'l'$). The response in mode choice has the same sign as before, but the first term in the bracket capturing the re-optimization of the ground traffic could be positive or negative, depending on the positions of $k' \to l'$ and $k \to l$ in the network. When $k' \to l'$ and $k \to l$ are on competing routes between o and d, shipments between o and k and between l and d are unlikely to pass through $k' \to l'$, so $\pi^{k'l'}_{ok}$ and $\pi^{k'l'}_{ld}$ are both small and $-\theta(\pi^{k'l'}_{ok} + \pi^{k'l'}_{ld} - \pi^{k'l'}_{od})$ is more likely to be positive. In such cases a reduction in $\iota_{k'l'}$ draws ground traffic *away* from $k \to l$. On the other hand, if $k' \to l'$ is en route of $o \to k \to l \to d$, as in the example given in Figure B.1, then the opposite can happen—reducing $\iota_{k'l'}$ increases the traffic passing through $k \to l$.[8]

To see the importance of own- and cross-routing effects, we can decompose SO_R into two terms.

$$SO_R = -\frac{1}{2} \sum_i \sum_{o \neq d} \frac{X^i_{od}}{Y} \sum_{kl \in \mathbf{C}} [\underbrace{\frac{\partial^2 \log \tilde{\tau}^i_{od}}{(\partial \log \iota_{kl})^2} (\Delta \log(\iota_{kl}))^2}_{\text{Own } SO_R} + \underbrace{\sum_{k'l' \in \mathbf{C}, k'l' \neq kl} \frac{\partial^2 \log \tilde{\tau}^i_{od}}{\partial \log \iota_{kl} \, \partial \log \iota_{k'l'}} \Delta \log(\iota_{kl}) \Delta \log(\iota_{k'l'})}_{\text{Cross } SO_R}].$$

While the sign of 'Own SO_R' is always negative, the sign of 'Cross SO_R' is ambiguous as discussed above. We illustrate two cases in Figure B.2 using the parameterized model. Consider one of the busiest expressway segments, the one between Laiwu and Linyi, colored solid black in the map. The colors of other edges indicate their cross derivative term with the one between Laiwu and Linyi. Cold colors indicate that the cross derivative is negative, in which case a new expressway between Laiwu and Linyi

[8] In the case illustrated, $\pi^{k'l'}_{ld}$ is close to zero and $\pi^{k'l'}_{ok}$ is close to one. The sum of the three terms is thus strictly positive as long as not all shipments between o and d go through the upper branch (i.e., $\pi^{k'l'}_{od} < 1$).

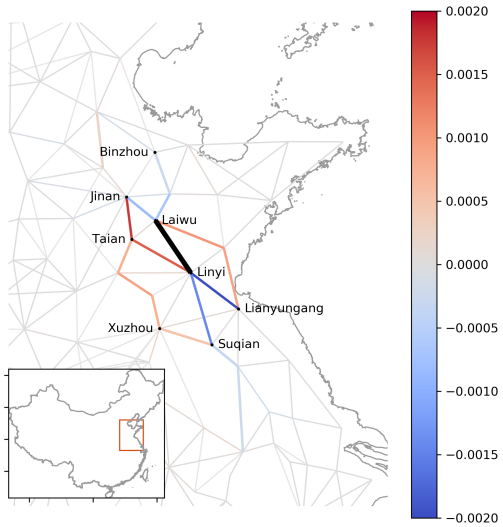


Figure B.2: Complementarity and Substitution between Segments: An Example

Note: The selected road segment is from Laiwu to Linyi, colored black. The map shows the cross derivative between each segment and the selected one (Laiwu to Linyi). Warm colors indicate that the cross derivative is positive, suggesting that an expressway between Laiwu and Linyi would draw traffic away from that segment. Cold colors indicate the opposite. Numbers are in percentage points of domestic GDP.

will increase the traffic on the segment. For example, the segment between Jinan and Laiwu.[9] On the other hand, warm colors indicate that a segment is a substitute to the expressway between Laiwu and Linyi. For example, the route from Jinan to Xuzhou. Importantly, most of these rerouting are traffic from cities to the north of Laiwu, to cities to the south of Linyi. This suggests when evaluating expressways, it is necessary to consider not only the direct trade between two cities connected by a segment, but also traffic that is merely passing by.

The importance of the cross-derivative force crucially depends on the segments being jointly evaluated. Equation (17) in the text provides a decomposition focusing on the 100 busiest expressway segments. It shows that the own second order effect accounts for -58% of the effect, whereas the cross-second order effect accounts for 8%. This implies that the expressway segments tend to be complementary.

[9] A negative cross derivative means that the second order effect has the opposite sign of the first order effect. So for an ex-post evaluation of the welfare looses from the removal of an expressway, it adjusts down the inferred first order effect. The opposite is true when the cross-derivative is positive.

C Quantification

C.1 Identification of Structural Parameters

Composite parameters that enter the route model. To understand how the composite parameters governing route choices—$\kappa^H\theta$, $\kappa^L\theta$ and $\frac{\theta_F}{\theta}$—are identified, it is useful to consider a limit case with $\theta = \infty$, under which the effective trade cost between o and d, τ_{od}, is simply the cost of the least-cost path. Slightly abusing notation, we use $dist_{od}^{H,t}$ and $dist_{od}^{L,t}$ to denote the *total* length of expressways and regular roads along the least-cost path at time t, respectively. Then the trade cost between o and d is $\kappa^H dist_{od}^{H,t} + \kappa^L dist_{od}^{L,t}$. With this we can write the structural routing equation as:

$$\log(\pi^t_{(o,RoW),d}) = -\theta_F(\kappa^H dist_{od}^{H,t} + \kappa^L dist_{od}^{L,t}) + \text{fixed effects.}$$

The above equation provides a micro-foundation for the reduced-form specification in Section 2. It also conveys two points on identification. First, route choices can identify the relative costs between the two types of roads, $\frac{\kappa^H}{\kappa^L}$. The identification comes from changes in compositions of regular roads and expressways in a route. Second, route choices can only identify *products of parameters*, $\theta_F\kappa^H$, and $\theta_F\kappa^L$. Intuitively, port choices reflect the combined effect of two forces: the marginal cost of additional distance; the marginal impact of cost on port choices. This is reminiscent of an result in gravity estimation that trade cost and trade elasticity cannot be separately identified using trade flows alone. Research in international trade has used price data (Eaton and Kortum, 2002) to overcome this problem. In the same spirit, we use the unit value information contained in the customs data, a step that we will return to below.

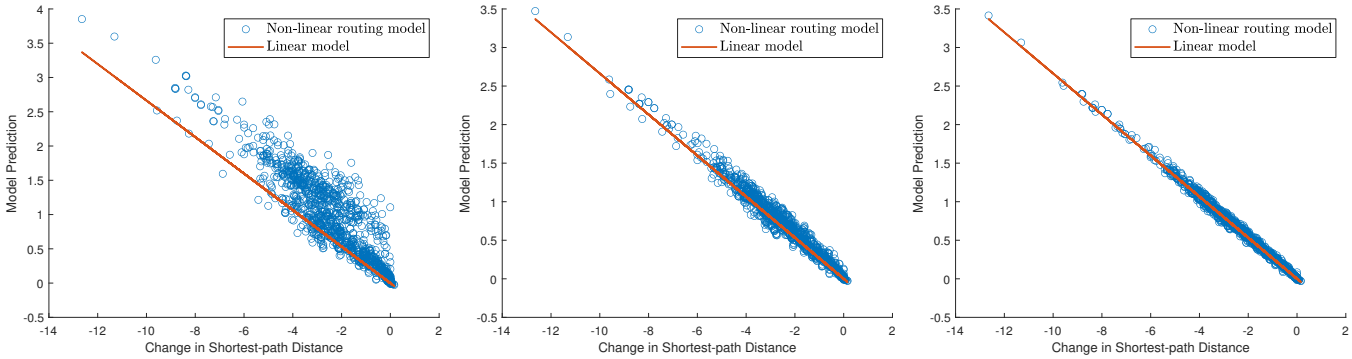
Moving from the above limit case back to our specification, *suppose we have already estimated θ_F using the* unit value information, what identifies θ? We can decompose the routing pattern from o to d implied by the model as below:

$$\begin{aligned}
&\log(\pi^t_{(o,RoW),d}) = \frac{\theta_F}{\theta}\log(\tilde{\mathbb{B}}^t_{(o,d)}) + \text{fixed effects} \\
&= -\theta_F\kappa^L \underbrace{(\frac{\kappa^H}{\kappa^L}dist_{od}^{H,t} + dist_{od}^{L,t})}_{\text{regular-equivalent distance}} - \underbrace{\left[\frac{\theta_F}{\theta}\log(\tilde{\mathbb{B}}^t_{(o,d)}) - \theta_F\kappa^L(\frac{\kappa^H}{\kappa^L}dist_{od}^{H,t} + dist_{od}^{L,t})\right]}_{\text{deviation}} + \text{fixed effects,}
\end{aligned}$$

i.e., $\log(\pi^t_{(o,RoW),d})$ encompasses the effect of the regular-equivalent length of the least-cost path, and a deviation term. Given estimated (κ^L, κ^H), the deviation term is a function of θ, and summarizes collective impacts of all routes that are inferior to the least-cost one. For a given network structure, θ affects the relative importance of the inferior routes. When θ is infinite, only the least-cost path matters, so the deviation term has no impact on routing; when θ is small, routes that are slightly inferior will be used more often and the deviation term will have more explanatory power for port choices.

Figure C.1 illustrates this intuition. In each panel, the horizontal axis is the change in the effective length of the least-cost path after the expressway construction.[10] The vertical axis denotes the change in $\log(\pi^t_{(o,RoW),d})$ predicted by the model. If all circles fall on the prediction by the linear component, it means that all truckers choose the least-cost paths and the deviation term defined above matters very

[10]The regular-road equivalent distance is based on the estimated value of (κ^L, κ^H).



(a) $\theta = 0.75 \times \theta^*$ (b) $\theta = \theta^*$, the calibrated value (c) $\theta = 1.5 \times \theta^*$

Figure C.1: Model Prediction with Varying θ

Note: The horizontal axis is the change in regular-equivalent distance between city pairs due to the expressway construction; the vertical axis is the model-predicted change in log shipments. As θ increases, changes in predicted shipments become closer to linearly correlated with changes in shortest distance.

little in the model; if dots are more spread out, it means that beyond the least-cost path, the structure of the entire network, captured by the deviation term, also matters. As expected, the figures show that as we increase θ from the estimated value (θ^*), the circles fall more tightly around the prediction by the linear component; on the other hand, the deviation from the linear prediction increases as θ decreases.

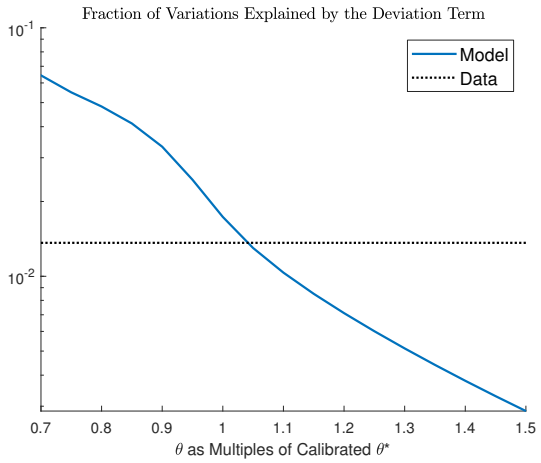


Figure C.2: Predictions of the Nonlinear Model Beyond the Shortest-path Distance

Note: The figure reports the fraction of variation in log export shipment explained by the deviation term, among total variation explained by the model. The deviation term is calculated as the difference between the prediction of the linear model and that of the nonlinear model under the calibrated θ. See the text for details.

This discussion makes clear that, *conditional on* θ_F, θ can be identified by the relevance of the structure of the road network beyond shortest-path distance in explaining changes in port choices. To see what the data tell us about this importance, we calculate the fraction of variation in log export shipment explained by the deviation term, among total variation explained by the road network structure.[11] The dashed line in Figure C.2 plots this fraction. As shown, the deviation term explains a non-zero fraction of data variation, so θ is not infinite. The solid line plots the implication of the model for this object as θ

[11] To do this, we estimate Equation (9) with only the length of the least-cost path and the fully nonlinear structure separately, and then calculate the percentage change of the residual sum of squared errors.

Table C.1: Transport Cost and Weight-to-Value Ratio

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	log price ratio				log price ratio		
Heaviness- HS2 Category	0.184***	0.184***	0.289***	0.187**			
	(0.062)	(0.062)	(0.089)	(0.087)			
Heaviness- HS4 Category					0.292***	0.352***	0.229***
					(0.044)	(0.047)	(0.030)
Fixed Effects	o, d, c	odc	fdc	fdc	fdc, i	$fdci$	$fdci$
Exclude major cities	yes	yes	yes	yes	yes	yes	yes
Exclude differentiated goods				yes			yes
Observations	3362494	3361110	3127626	236027	3127625	2017990	142835
R^2	0.058	0.069	0.330	0.427	0.374	0.570	0.582

Notes: This table reports the regressions of the log price ratio on log sector-level weight-to-value ratio. The dependent variable is the log of price ratio and is always computed within a city-destination country-HS8 category; the explanatory variable is the log of the weight-to-value ratio at HS2 category level (Columns 1-4) and HS4 category level (Columns 5-7). Letters o, d, c, f, i stand for origin city, port, destination country, firm, and HS2 category fixed effects, respectively.
Standard errors are clustered at HS2 category level (Columns 1-4) or HS4 category level (Columns 5-7).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

varies. At around our estimated θ^*, the model generates about the same prediction as in the data.[12] This argument identifies θ only conditional on θ^F, i.e., it identifies $\frac{\theta_F}{\theta}$. Below we use price variation contained in the customs data to identify the *level* of θ.

Price-heaviness elasticity μ. We estimate Equation (10) to identify μ and θ. Since these two parameters are identified from different variations, we estimate them separately, so more flexible controls can be included. Table C.1 reports our estimate of μ, the elasticity of the trade cost with respect to the weight-to-value ratio of a sector. The first four columns focus on the comparison of the log price differences across HS2 categories, with progressively more demanding fixed effects. The first and second columns control for city, port, and destination country fixed effects and city-port-country fixed effects, respectively. The estimated coefficient is around 0.184. Even within a city-port-country cell, some firms might systematically set prices differently. To account for this possibility, Column 3 control for firm-port-country fixed effects. The point estimate increases somewhat to 0.29 and is precisely estimated.

The set of fixed effects and the narrow definition of a product allows us to rule out many plausible alternative explanations. To the extent that the price ratio might still capture variations in qualities and markups, as long as they are not systematically correlated with the weight-to-value ratio, they will not affect our estimates. Nevertheless, Column 4 focuses only on the HS2 categories that are classified as non-differentiated goods (Rauch, 1999), which likely have a smaller scope for either quality differentiation or price discrimination. Reassuringly, despite that the sample is only a tenth of the baseline sample, the point estimate remains broadly in line.

One further concern is that our measure of 'heaviness', the weight-to-value ratio, might capture other characteristics of a sector that correlate systematically with prices. In Columns 5 through 7, we estimate the specification using the weight-to-value ratio at the HS4 category level. This allows us to control for

[12] The discussion here aims at visualizing the data patterns that identify θ. The intersection in the figure is not exactly at θ^* because θ^* is not chosen to target this fraction, but rather estimated via the nonlinear least square specified in Section 4.

Table C.2: Price Distance Regression

	(1)	(2)	(3)	(4)	(5)
	OLS		2SLS		Structural
$dist_{od}$	0.050***	0.057***	0.049***	0.058***	
	(0.011)	(0.012)	(0.011)	(0.012)	
$\log(\tilde{B}(\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta})_{(o,d)})$					-0.0090***
					(0.0022)
Fixed Effects	dci, oci	dci, oci	dci, oci	dci, oci	dci, oci
Exclude major cities	yes	yes	yes	yes	yes
Exclude differentiated goods		yes		yes	yes
Observations	3156133	279165	3156133	279165	279158
R^2	0.335	0.351	-	-	-
First Stage KP-F statistic			1191.648	935.979	1090.070

Notes: This table reports the regressions of the log price ratio on the distance between the origin city and the port (Columns 1-4) or the output of the routing model (Column 5). Letters o, d, c, i stand for origin city, port, destination country, and HS-8 product fixed effects, respectively.
Standard errors in Columns 1 through 4 are clustered at city-port level. Standard error in Column 5 is generated through bootstrapping. See Appendix C.2 for details.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the HS2 fixed effects. Column 5 of Table C.1 is our preferred specification, which is identified from within a city-port-country and HS2 cell, whether heavier goods are relatively more expensive when exported through a different seaport than own city. The point estimate suggests that a one-percent increase in the weight-to-value ratio of a good increases the ad-valorem shipping cost by around 0.3%.

Route elasticity θ. Table C.2 reports our estimates for θ. Since we do not aim to identify μ in this regression, we can absorb the category characteristics in fixed effects. We present results for two sets of regressors. The first is for the distance between city o and port d, which illustrates the variation in the data that identifies θ more transparently . The second is for $\log(\tilde{B}(\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta})_{(o,d)})$, which allows us to estimate θ directly.

The first two columns use OLS and control for port-HS8-destination country and city-HS8-destination country fixed effects, respectively. The former set of fixed effects captures, within a HS8 category, the overall tendency of some ports or destination countries to be involved in the export of more pricey goods; the latter controls for the overall tendency of a city in producing pricey good for exporting to specific countries. The point estimate of the first column, which uses all categories, suggests that the price ratio increases by around 5% as an additional hundred km regular-road equivalent distance is added. The second column restricts to non-differentiated varieties for robustness. This restriction significantly reduces the sample size but the point estimate remains similar. To alleviate the concern about the endogeneity of the road network, Columns 3 and 4 estimate a 2SLS specification using the IV generated from the minimum-spanning tree. The point estimates are in the range of 0.05 to 0.06, statistically indistinguishable from the OLS estimates.

These reduced-form results show that: the price of goods is more expensive when a port is further apart from the origin city. Importantly, the point estimate is robust across specifications with different fixed effects, sample restrictions, and the use of IV.

The variation exploited in these estimates can identify θ. Specifically, recall that from Equation (10), $-\frac{1}{\theta}$ is exactly the elasticity of log price ratio with respect to $\log(\tilde{B}(\kappa^H\theta, \kappa^L\theta)_{(o,d)})$. Having estimated $\kappa^H\theta$

and $\kappa^L\theta$ in the previous step, we can plug in their values and construct $\log(\tilde{B}(\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta})_{(o,d)})$. Column 5 in Table C.2 uses the same specification as in Column 4, but has $\log(\tilde{B}(\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta})_{(o,d)})$ as the dependent variable. The point estimate of $-\frac{1}{\theta} \approx -0.0090$ translates into $\theta \approx 111$, implying that different routes connecting the same pair of cities are highly substitutable. This estimate is close to the estimate of Allen and Arkolakis (2019) using routing information of domestic shipments.

C.2 Inference of Structural Parameters

Table 3 in the text reports point estimates and distributional statistics of the key structural parameters. Panel C of Table 4 reports statistics for additional parameters determined jointly in calibration. This section explains how we draw statistical inference for each of these parameters.

Step 1. The model does not incorporate structural errors at city-port level, so we assume that the source of uncertainty is due to measurement errors and use bootstrap to infer the size of uncertainty. For parameters in Panel A of Table 3, we resample with replacement by city 200 times. Each time we draw a sample, we estimate the nonlinear routing problem described in Equation (9) to obtain a new set of estimated $(\kappa^H\theta, \kappa^L\theta, \frac{\theta_F}{\theta})$. We then calculate from these 200 repetitions the distributional statistics of the composite parameters. With bootstrapping at the city-level, the standard errors calculated capture the potential correlation between different city-port pairs, so they tend to be more conservative than city-port level clustering in reduced-form analyses.

Step 2. For the inference of θ, on top of measurement error of the price data, the errors due to generated regressors $(\log(\tilde{B}(\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta})_{(o,d)}))$ also need to be taken into account. We therefore use a joint bootstrap procedure. Specifically, from each bootstrap sample in Step 1, we have obtained one estimate for $\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta}$. We use the corresponding $(\log(\tilde{B}(\widehat{\kappa^H\theta}, \widehat{\kappa^L\theta})_{(o,d)}))$ on a bootstrapped price sample for the regression reported in Column 5 of Table C.2. We obtain the distributional statistics of θ by repeating this procedure 200 times. As Column 5 shows, the standard error generated this way is similar in magnitude to those for the linear regressions (Columns 1 through 4) calculated using asymptotic theories.

To draw inference for μ, which is estimated using linear regressions, we use directly the standard error in Column 5 of Table C.1.

Step 3. The uncertainty in these structural parameters affect the estimation of the full model. Panel A of Table 4 summarizes the parameters estimated using micro data and their standard errors. Panel B is the parameters from external sources, which we take as given. Panel C is the parameters estimated jointly. To take into account the uncertainty of parameters in Panel A, we draw 200 realizations of Panel A parameters from their joint distribution; for each of these draws, we calibrate the remaining parameters to match the same targets. Reported in Panel C of Table A are the standard errors of calibrated h_0 and $\overline{\kappa}$. The standard errors tend to be small, reflecting that they are mostly determined by their own targets, rather than the parameters in Panel A.[13]

Three comments are in order on this procedure. First, when generating the joint distribution of the parameters in Panel A, we take into account that some of these parameters are estimated jointly and thus correlated. In bootstrap, we consider three sets of parameters. (1) The composite parameters estimated from the port choices, $\kappa^H\theta$, $\kappa^L\theta$, and $\frac{\theta^F}{\theta}$. (2) For each draw of these parameters, we estimate the linear regression in Column 5 of Table C.2 for θ, which, together with the three composite parameters,

[13]Note that in this procedure, each time we will also have different values for other parameters in Panel C such as T_d^i. We omit these parameters and their standard errors from the Table.

gives us all four parameters for the routing model. (3) We randomly draw a μ from its own asymptotic distribution. Each calibration of the model then uses one realization of these three sets of parameters.

Second, in evaluating the impacts of the expressway construction, we repeat the counterfactual experiment for each of the 200 model calibrations. The distributional statistics calculated in Table 5 are from these 200 counterfactual experiments.

Finally, in this entire procedure we take the parameters from external sources, such as the trade elasticity and the IO table parameters, as given. These parameters are either from the aggregate data, or estimated by a other studies with no consistent ways for statistical inference. We conduct sensitivity analyses to show how results vary with these parameters in Section C.6 of this appendix.

C.3 Model Validation

We validate the model by comparing its 'out-of-sample' predictions to the data.

Expressway and Export Growth. Given our use of export data in estimation, we first assess how the model fits city-level export in the data. This comparison is out-of-sample, because in calibration we absorb the level of export through city-port fixed effects and use only the within-variation from the patterns of routing. Figure (C.3) plots the model-implied city export against the data. To ensure that the fit is not due to city size, the plots are for residuals from a regression that controls for city-level employment. The figure shows that the model closely matches the export observed in the data well.

In the second validation test, we compare the model-predicted export growth led by the expressway network expansion to the actual export growth in the data. This is a joint test of two hypotheses: 1) whether the expressway expansion as large as the one seen in China over the decade led to differential growth of exports across cities; 2) when fed into the expressway expansion, whether the model can generate the changes in trade patterns in the data.

To implement this exercise, we feed in the 1999 expressway network to the model and solve a counterfactual equilibrium holding all other parameters at the calibrated values. We treat the export generated from this counterfactual equilibrium as the model export around 2000. We then compare the export at the city-sector level between the model and the data for 2000 and 2010. Table C.3 reports the results. The dependent variable is the log export in the data and the independent variable is its model counterpart. The first column presents the result based on a cross-sectional specification. The second and third columns control for sector-time and city-sector fixed effects, so the comparison is on export growth within a city-sector cell. The point estimates are highly statistically significant in both cases.

Importantly, all these regression models have a F statistic above the rule-of-thumb for bounding biases in IV estimates. Under the assumption that road networks affect city export only through improving the access of a city to ports, the model predictions can serve as an IV for export at the city-industry level. A growing literature has examined the impacts of Chinese export on its domestic economy. One IV commonly used in this literature is variation in tariffs due to the WTO accession, which is valid by assuming that the pre-WTO tariffs are exogenous (Facchini et al., 2019; Tian, 2019). The IV based on our model predictions vary across regions and over time and is valid under a different set of assumptions from existing studies. It could be of use for future research in this area.

Comparison with truck flow data. Our second set of validation exercises aims to show that the customs data capture the variation in domestic shipment well. Specifically, we obtain the number of bilateral truck movements between pairs of Chinese cities in 2019, collected by a digital logistic platform,

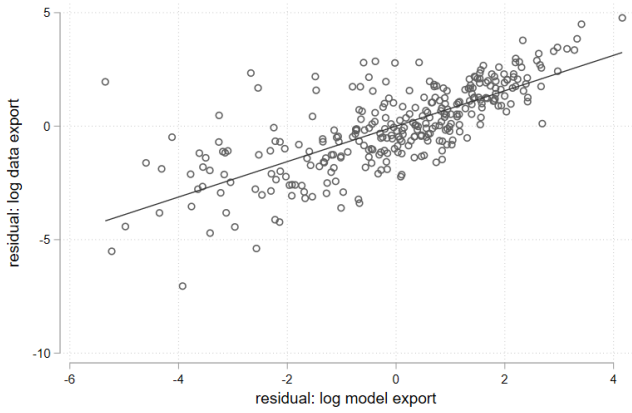


Figure C.3: City-level export: Model versus Data

Note: The figure plots model-implied city export against the data, netting out employment at the city level.

Table C.3: Predicting Export Growth

	(1)	(2)	(3)
Log(export), model	0.465***	0.953***	0.871***
	(0.048)	(0.189)	(0.199)
Fixed Effects	t	oi, it	oi, it
Exclude major cities	no	no	yes
Observations	8472	8472	6576
R^2	0.333	0.878	0.860
F-statistic	92.706	25.332	19.223

Notes: The dependent variable is the log city-sector export in the data; the independent variable is the log city-sector export in the model. Letters t, o, i, in 'Fixed Effects' stand for time, city, and sector (two-digit) fixed effects, respectively.
Standard errors (clustered by city) in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G7, which helps logistic companies and transportation firms to manage 1.3 million trucks in China.[14] Since the dataset is available only for long after expressways were built, we can not use it to estimate the specification exploiting over-time variation, but we can still estimate a cross-sectional specification. We find an estimate -0.432 (Column 1 of Table C.4) off the cross-sectional specification, which is very close to the baseline estimate of -0.384 (Column 2, Table 2) using customs data. This exercise shows that at least when the cross-sectional variation is used, customs data and domestic shipment data give similar estimates on the shipment-distance semi-elasticity.

We also compare the model-implied bilateral shipment to the data. Column 2 of Table C.4 shows that the model-implied shipment flows are highly correlated with truck flows, with a linear regression coefficient of 1.3. Of course, the raw correlation between the two variables could be driven by the size of origin and destination cities. Column 3 controls for the origin and destination fixed effects and shows that doing so does not diminish the importance of model-implied shipments in explaining the data.

Figure C.4 visualizes the close relationship between the two variables. After netting out origin and

[14]The company provides services to logistic companies. Among their main services is the management of an in-truck camera that monitors risky driving behaviors (such as drowsy driving and DUI). This in-truck device records the trip made by truckers. We do not observe whether a truck is loaded or not, so the measure of shipment is symmetric.

Table C.4: Validation: Data Truck Flow v.s Model Shipment Flow

Dependent variable	Log data truck flow		
	(1)	(2)	(3)
Effective distance	-0.432***		
	(0.003)		
Log model shipment flow		1.281***	1.668***
		(0.007)	(0.011)
Observations	54057	54057	54057
o and d fixed effects	yes	no	yes
R^2	0.627	0.435	0.597

Notes: The dependent variable is log of number truck flows between city pairs in the data (2019); the independent variable is the regular-road equivalent distance, and the log shipment flow between city pairs predicted by the model, calibrated to match the 2010 Chinese economy. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

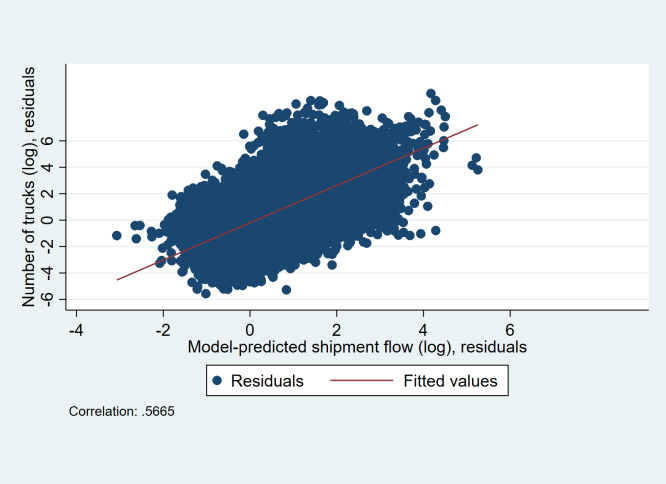


Figure C.4: Data Truck Flows v.s Model Shipment Flows

Notes: The figure plots the residual correlation between truck flows and model implied shipment flows after netting out origin and destination fixed effects.

destination fixed effects, the residual correlation between them is around 0.57. This test is 'out-of-sample' in two senses. First, our model is estimated using the export data, whereas truck flows capture mainly domestic shipment. Second, our estimation uses over-time variation between 1999 and 2010, whereas truck flows are from 2019, after a decade of rapid growth and transformation of economic landscape in China. Despite this, the fit is comparable to the 'in-sample' fit of models that are estimated to match domestic trade flows. For example, Allen and Arkolakis (2019) (Figure 2) shows a residual correlation of 0.60. This suggests that the customs data, combined with our routing model, can capture the bilateral shipment in the data well.

Beyond looking at the bilateral correlation, we also examine whether the model generates the pattern of shipment over different distances as in the data. The model implies that the value of shipment decreases in bilateral distance significantly when the distance is below 200 miles, and then the decrease becomes more gradual. This is consistent with what we find using the truck flow data. Interestingly, both the model-implied shipment flows and truck flows capture the salient features of domestic shipment documented in Hillberry and Hummels (2008) based on the U.S. data. These results are available

Table C.5: Correlation with Shipment

	(1)	(2)	(3)
Log(shipment), model	0.314***	0.177***	0.176***
	(0.040)	(0.035)	(0.041)
Log(employment)		0.594***	0.587***
		(0.059)	(0.062)
Observations	240	240	234
Fixed Effects	no	no	prov
R^2	0.234	0.488	0.636

Notes: The dependent variable is the log of shipment that passes a city in the data (2010); the independent variable is the log of shipment that passes a city in the model. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

upon request.

Transport hubs. As a final validation, we examine the model's prediction on shipment by city. Because of their central locations in the transport network, some cities become 'hubs' that shipments to other places go through. To validate the model, we can compare the model-inferred shipment that passes a city to its empirical counterpart, sourced from the 2010 yearbook for transportation.[15] Table C.5 reports the regression of log shipment in the data on the model prediction. The first column shows the raw correlation. The second column controls for city employment. The coefficient is still significant and meaningful. This suggests that the model prediction correlates with the data not only due to usual gravity forces, which predicts more trade for bigger cities, but also because it captures the traffic passing by. The third column further shows that including province fixed effects does not change the estimate. This implies that the prediction power comes from the network connections of a city shaped by the routing model, rather than the broad location of the city.

C.4 The Role of International Trade, Sector Heterogeneity, and Input-output Linkages

Our benchmark model differs from those used in the growing literature quantifying the impacts of transportation infrastructure (see, e.g., Asturias et al., 2018; Fajgelbaum and Schaal, 2020; Allen and Arkolakis, 2019) in three aspects. First, our structural estimation exploits changes in the route choice of exporters resulting from the domestic expressway network expansion, which naturally implies that the network expansion reduces trade costs not only for trade between domestic partners but also for trade between the hinterland and foreign countries; second, with sector-level information on production and export prices, we allow for regions to differ in sectoral specializations and sectors to differ in trade costs; third, we incorporate intermediate inputs.

This section shows that because these ingredients allow us to infer the distribution of shipment among different routes and the shipment values more accurately, they are important for the quantitative results. We parameterize a series of restricted models and compare the inferred welfare gains in these models to the baseline results. For transparency, throughout this subsection we recalibrate only the trade cost level parameter h_0, the amenity $\{B_d\}$, and the city-sector productivity $\{T_d^i\}$, to match the av-

[15] The data are aggregated by city; the National Bureau of Statistics surveys firms in the logistics industry to produce this statistic. The data series appear inconsistently defined over time, with frequent abrupt changes from one year to another, so we do not use the time dimension of the data.

Table C.6: Welfare Gains in Alternative Models, Matching Average Shipment Distance

	Baseline	Model (2)	Model (3)	Model (4)	Model (5)
International trade	✓				
Trade cost heterogeneity	✓	✓			
Regional specialization	✓	✓	✓		
Intermediate Input	✓	✓	✓	✓	
Welfare gains	5.10%	4.47%	4.29%	3.36%	0.89%

Note: For each alternative model, the trade cost level parameter h_0, amenity $\{B_d\}$, and city-sector productivity $\{T_d^i\}$ are re-calibrated to match the same average domestic ground shipment distance, population distribution, and city-sector sales (or city-level sales, depending on whether regional specialization is allowed).

erage domestic shipment distance, the population distribution, and the sales by either city or city-sector, depending on the restriction on the model. We keep other structural parameters in the routing model as in the benchmark. Below reports the changes in results as we sequentially eliminate the elements in the model.

Domestic transport costs in international trade. Column 2 of Table C.6 is the result from a model without international trade, i.e., with $\tau_{RoW}^i = \infty$, $\forall i$. The inferred gains from expressway construction in this model is about 12% (or 0.63 p.p.) smaller than in the baseline model (reproduced in Column 1).

We can understand the difference by inspecting the first-order effect on the aggregate welfare of expressway segments. As we show in Proposition 1, the welfare gains of trade cost reductions can be approximated by total cost savings on trade flows on segments directly affected, adjusted for the savings that are passed on the RoW. By matching the average shipment distance for goods within China, both the full model and the restricted model without international trade generate similar predictions for domestic trade flows, so they predict similar cost savings from domestic trade. Through the lens of the full model, however, these are only part of the benefits—the improvements in domestic infrastructure reduce the trade costs for the importers and exporters from the hinterland. Because part of these additional cost savings will accrue to the Chinese economy, overlooking international trade leads to smaller inferred gains.

Transportation intensity. In the second experiment, we set $\mu = 0$ and then recalibrate the model to match other moments. We then conduct the same exercise as before. Under the assumption of homogeneous transport cost across sectors, the inferred gains are down from Model (2) to 4.29%.

At first glance, this might seem surprising, as with a large enough number of regions and road segments, the law of large numbers should have kicked in and the heterogeneity in transport intensity across sectors could be washed out. The reason why sector heterogeneity is not simply washed out is, when calibrated to match the same average shipment distance, Model (2) infers systematically higher values of shipment compared to Model (3). More specifically, with heterogeneity in trade costs, for the same level of inter-city shipments, Model (2) will predict a higher fraction of them in lighter sectors (with lower weight-to-value ratios) because they incur lower shipping costs in Model (2) but not in Model (3). Because the welfare gains are, to the first order, proportional to the value of goods but not their weights, the model with sector transport intensities predicts larger welfare gains.

Regional comparative advantage. Chinese regions specialize in different broad sectors, e.g., manufacturing and service in the Southeast versus energy in the Northwest. To understand the importance of

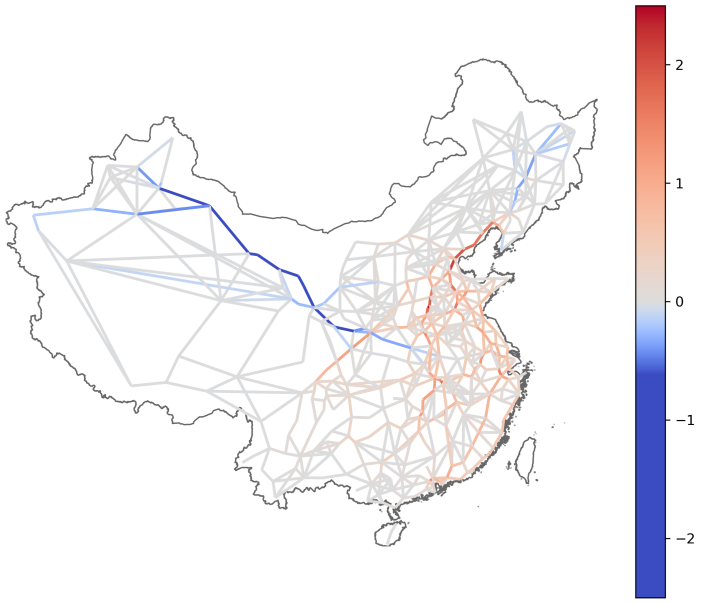


Figure C.5: Differences in Shipment Value Shares, 'No Specialization' Minus 'Baseline'

Note: The numbers are the differences in shipment value over GDP between a model with no specialization and the baseline. The values plotted include both expressway and regular road shipments. Cold colors indicate that there is less shipment in the model with no specialization than the baseline.

accounting for regional productivity differences, Column 4 reports the result from a recalibrated model without specialization. Specifically, we assume all sectors within a region have the same productivity, i.e., $T_o^i = T_o^j = T_o, \forall o, i, j$, and pin down $\{T_o\}$ by matching the total sales of each city in the data. The input-output structure is kept the same as in the baseline model. The inferred gains in this model are 22% smaller than an otherwise similar model with regional specialization (Column 3).

Patterns of regional specialization matter because they contain information for the distribution of trade flows across pairs of domestic partners. Because of the strong spatial clustering of production, the calibrated productivity in the full model has a spatial correlation, too. As a result, regions tend to trade with partners that are far away. When comparative advantages are eliminated, the spatial clustering also disappears, so inter-city trade in the restricted model shifts towards partners that are closer to each other. Although both models are calibrated to generate the same average shipment distance, this simple statistic does not capture all these patterns. Indeed, Figure C.5 plots the change in shipment intensities between city pairs from Model (3) to Model (4). The segments that see the biggest decrease in inferred shipments are the ones connecting the northwest and northeast—the energy producing area—to the center of the country with a heavy manufacturing presence; the segments that see an increase in inferred shipments are the ones connecting regions within the center and the east of China. As a result, Model (4) infers higher gains for expressway segments in the center of the country and lower gains for projects connecting the center to the northeast and northwest—regions with very different comparative advantages. Whether it underestimates or overestimates the return to a specific project thus depends crucially on where a project is. For the actual projects built during the decade, the balance comes down to an underestimation of the welfare gains by 22%.

Intermediate inputs. In the final comparison, we further shut down intermediate inputs in production by assuming the labor shares (β^i) are one and sectoral shares (γ^{ij}) are zero in all industries. The

welfare gains inferred by this model decline by three-quarters to around 0.9%. This difference can be understood by inspecting Equation (C.1).

$$\frac{X_{od}^i}{Y} = \frac{X_{od}^i}{\sum_i \sum_{o,d} X_{od}^i} \cdot \frac{\sum_i \sum_{o,d} X_{od}^i}{Y}. \tag{C.1}$$

For a simple example, assume that all regions o and d are symmetric, with positive but symmetric inter-regional transport costs. When calibrated to match the average shipment distance, Models (4) and (5) generate the same trade intensity, i.e., $\frac{X_{od}^i}{\sum_i \sum_{o,d} X_{od}^i}$. However, in the model without intermediate inputs, the overall absorption $\sum_i \sum_{o,d} X_{od}^i$ is equal to the GDP, whereas in the model with intermediate inputs, the overall absorption is several (around three in our calibration) times of the GDP. As a result, the inferred value of $\frac{\sum_i \sum_{o,d} X_{od}^i}{Y}$ is too small in the model without intermediate inputs. As we show in Proposition 1, by and large, the overall gains from the expressway construction are determined by $\frac{X_{od}^i}{Y}$. By assuming away intermediate inputs, the restricted model overlooks that goods are traded multiple times on the road, which amplifies the gains from the reduction in transport costs.[16]

To summarize, when restricted to a bare-bone one-sector model used in most of the literature, the welfare gains is only a small fraction of the baseline result.

C.5 Comparison to Existing Evaluations Using Other Approaches

We compare our assessment of the welfare impacts to existing evaluations by academia and policy institutions using three alternative approaches.

The first approach, which is also what we adopt in this paper, is to rely on quantitative models for simulations. The obstacle faced by this approach is the lack of reliable domestic trade data for disciplining the importance of transport infrastructure for trade and welfare. Roberts et al. (2012) uses a one-sector new economic geography model. The model expresses regional wages as a function of market access, which in turn depends on trade elasticity and transport costs. Cross-sectional wage variation can then be used to discipline trade and transport cost elasticity. Roberts et al. (2012) finds that the static welfare gains from the expressway network to be around 6%, slightly higher than our estimate.[17]

The second approach directly measures the return to expressway investment in the transport, logistic, and postal service sectors. The idea is that if transport infrastructure affects the aggregate economy

[16] Although it is well known that the inferred gains from international trade are larger when intermediate goods are introduced (Caliendo and Parro, 2015; Costinot and Rodríguez-Clare, 2014), we show that for the evaluation of domestic infrastructure projects, this insight matters at least as much, if not more. In recent work, Baqaee and Farhi (2019) shows that if the true underlying model is one with intermediate goods, and the researcher specifies a model without intermediate goods, then calibrating the specified model to match the trade over GDP ratio (as opposed to the theory-consistent target under this model, trade over absorption/production) gives a better approximation to the true gains from trade. In our setting, this approach (one that changes the target, but not the model) runs into two practical difficulties. First, reliable inter-regional trade data is lacking, so we cannot directly measure trade/value added at the regional level. Second, even when the data are available, at the micro level, this measure could be easily above one, which a model without input-output linkages cannot accommodate. In our baseline economy, for example, this ratio is around 1.45 for the tradable sector as a whole.

[17] In light of our finding that input-output linkages and sector heterogeneity amplify the welfare gains, the larger gains in Roberts et al. (2012) might be surprising. The reason for this finding is that, instead of targeting trade flows, Roberts et al. (2012) targets wage dispersions, under the assumption that the observed wage dispersions are entirely due to differences in market access arising from trade costs. Large wage disparities in the data are thus interpreted as large trade frictions, which in turn imply large gains from infrastructure investment. Compared with our evaluation, which allows regional productivity differences, Roberts et al. (2012) imply too low a volume of trade, but too large trade cost reductions led by transport infrastructure. These two forces turn out to bring their calculated gains similar to ours.

only through these sectors, then its return would be captured in the capital value added of these sectors. Return measured this way might be lower than the true social return of infrastructure for two reasons. First, China's transport infrastructure is likely under priced,[18] so the capital return to companies in charge of expressway operations might be lower than the social value of investment. Second, the transport, logistic, and postal service sectors are not the only sectors using the expressways. For example, transportation of goods and services by residents or manufacturing firms benefit from the expressway expansion, but their benefits do not necessarily show up in the value added of the transportation sector. Using sectoral value added data, Bai and Qian (2010) finds that the gross per-period return to expressway investment to be around 25-30%. This is smaller than the 51% static gross return to capital implied by our approach (5.1% welfare gains divided by 10% GDP capital investment), but in the same order of magnitude.

The third approach is to estimate a provincial-level production function, with provincial GDP being the output and expressway investment being one of many inputs. Given challenges to identification and that the studies generally use different measures of infrastructure, the literature has not reached a consensus.[19] For example, Shi and Huang (2014) finds that after 2001, investment in a broad category of transport infrastructures offers a lower gross return than private capital, but the estimate include different kinds of infrastructure so it is hard to compare this number to ours. On the other hand, focusing on roads, Fan and Chan-Kang (2005) estimates that each additional km of 'high-quality' roads generates 32% static gross return. Their focus, 'high-quality' roads, includes multi-lane paved roads that are not expressways. In addition, such an approach identifies only the different effects across regions while overlooks the general equilibrium effects, which likely improve welfare in all regions. These differences might explain why their estimate is smaller, but the conventional confidence intervals of their estimate covers our baseline estimate of a 51% return.

Overall, comparing with findings from existing studies using different approaches, our estimate appears reasonable.

C.6 Sensitivity Analyses

Table C.7: Sensitivity Analyses

Change in	(1) High Heterogeneity in Sectoral Trans. Cost	(2) High Substitution across Trans. Mode	(3) External Economy of Scale	(4) Immobile Labor	(5) Mobile Labor + Migration Costs
Aggregate welfare (%)	0.053	0.047	0.048	0.052	0.052
Log(Domestic trade)	0.117	0.119	0.144	0.140	0.138
Log(Exports)	0.097	0.102	0.097	0.113	0.108

Note: The table reports (the minus of) changes in model statistics as the economy moves from the calibrated equilibrium with the 2010 expressway network to the one with the 1999 expressway network. Alternative models in (1)-(5) are recalibrated to match the same targets as in Table 4.

[18] Many highway management companies are on the government support since they cannot self sustain operations.

[19] A challenge to this approach is that at the provincial level, stock of expressways might be endogenous to GDP growth. To circumvent the identification challenge, more recent empirical works focus on counties or prefecture cities, at which level it is possible to use 'exogenous' placement of expressways (see, e.g., Banerjee et al., 2020, Faber, 2014, and Baum-Snow et al., 2020). Because this approach overlooks the general equilibrium effects, and often estimates a coefficient associated with a dummy indicating whether being connected to expressways, it is difficult to convert such estimates to overall benefits of the macro economy.

We conduct a number of exercises to assess the sensitivity of the baseline results to alternative assumptions. We focus on four scenarios. The first is on the sector heterogeneity of transport costs. Instead of 0.3 in the baseline calibration, we now set μ to 1, which corresponds to a linear relationship of the iceberg trade cost in the weight-to-value ratio. The second robustness check increases the elasticity of substitution between road transportation and the outside mode, θ_M, from the benchmark value 2.5 to 14.2, an estimate by Allen and Arkolakis (2014). Our third robustness allows for industry-level agglomeration. Specifically, we set $T_d^i = \overline{T}_d^i [l_d^i]^\chi$ with $\chi > 0$. This assumption implies an external increasing return to scale to specialization. The estimates for χ in the literature, as surveyed in Combes and Gobillon (2015), range from 0.02 to 0.13. We set $\chi = 0.075$, which lies in the mid-range of the estimates. Finally, given the hukou reform in China that reduces migration frictions was gradual during this period, we conduct two exercises—one with immobile labor, the other with partially mobile labor. The additional model ingredients with immobile labor or with migration frictions are presented at the end of this subsection.

The first column of Table C.7 shows that when sectoral heterogeneity in transport costs is more important, the inferred welfare gains are slightly larger. The second column shows that when the elasticity of substitution between transport modes increases, the inferred welfare gains are slightly smaller. This is because after expressways are removed, traders can switch to the alternative mode more easily and incur smaller losses. Adding external economies of scale at the industry level leads to a modest decrease in the inferred gains. Finally, the welfare gains increase slightly if domestic workers are completely immobile or subject to migration frictions, once the models are recalibrated to match the same targets.

Model with immobile labor. With the assumption of immobile labor, the numbers of workers in domestic cities, $\{L_d\}_{d \in CHN}$, are fixed. The competitive equilibrium with immobile labor can be defined by including $\{L_d\}_{d \in CHN}$ as additional fundamentals, and removing the free labor mobility condition from Definition 1. The aggregate welfare is defined as the income weighted average consumer utility of domestic regions:

$$\log W \equiv \sum_{d \in CHN} \omega_d \log U_d,$$

where $\omega_d = \frac{I_d L_d}{Y}$, with (I_d, Y) evaluated at the calibrated equilibrium. This definition ensures that Proposition 1 still applies, so it is comparable to the welfare defined in the benchmark model with free labor mobility.

Model with mobile labor subject to migration costs. Assume domestic cities are endowed with initial numbers of workers $\overline{L}_o$. The utility a worker ς from city o would obtain by migrating to region d is:

$$\frac{U_d}{d_{od}} \epsilon_d(\varsigma),$$

where U_d is the utility living in city d that is specified in Subsection 5.1, d_{od} is the iceberg migration cost for migrating from o to d, and $\epsilon_d(\varsigma)$ is an idiosyncratic preference shock that is drawn from the Fréchet distribution and assumed to be i.i.d. across o, d, ς. Under the optimal migration decision, the fraction of

workers from city o that migrate to city d is thus

$$\pi^e_{od} = \frac{(\frac{U_d}{d_{od}})^{\theta_e}}{\sum_{d'}(\frac{U_{d'}}{d_{od'}})^{\theta_e}},$$

which implies that the total number of workers in region d satisfies

$$L_d = \sum_o \overline{L}_o \pi^e_{od}. \tag{C.2}$$

The competitive equilibrium with migration costs can be defined by including $(\overline{L}_o, d_{od})$ as additional fundamentals, and replacing the free labor mobility condition from Definition 1 with (C.2). For the calibration of the parameters that enter the migration model block, we set the elasticity of migration $\theta_e = 1.5$, taken from the estimate in Tombe and Zhu (2019). We specify the migration cost d_{od} as a function of geographic and cultural distances, and use the estimates of d_{od} from Fan (2019) for China. We recalibrate the amenities of domestic cities $\{B_d\}_{d\in CHN}$, along with other parameters, such that the equilibrium domestic labor allocations $\{L_d\}_{d\in CHN}$ match the population distribution from the 2010 Census—the same target used in the benchmark calibration. The aggregate welfare is defined as the income weighted average consumer utility of domestic regions:

$$\log W \equiv \sum_{d\in CHN} \omega_d \log U_d,$$

where $\omega_d = \frac{I_d L_d}{Y}$, with (I_d, L_d, Y) evaluated at the calibrated equilibrium. This definition ensures that the aggregate welfare agrees with the model with mobile labor if migration costs are set to zero ($d_{od} = 1, \forall o, d$), or agrees with the model with immobile labor if migration costs are set to infinity.

C.7 Numerical Implementation

Solve the competitive equilibria. We describe the design of the algorithm that makes it possible to load the most intensive part of the computation to GPUs. This enables us to solve equilibria robustly and efficiently, despite the size of the problem (our benchmark model has 323 regions and 25 sectors).[20] The large size of the problem also renders a well-known approach to solve/calibrate this type of model—Mathematical Programming with Equilibrium Constraint (Su and Judd, 2012)—less effective as the Jacobian matrix is a dense matrix with $(323 \times 25)^2$ entries. Our algorithm falls back to a fixed point algorithm described below.

With E^i_d being the total expenditure on intermediate goods in sector i of region d, the minimal system

[20] For example, to estimate the model and to conduct statistical inference, we need to solve the equilibria numerous times. And because of the sequential nature of many global optimization routines, paralleling this step is not straightforward, so speed is important.

of equations that can be used to solve the equilibrium is[21]

$$
\begin{aligned}
E_o^j &= \alpha^j (w_o + Tr_o) L_o + \sum_i \gamma^{ij} \sum_d \pi_{od}^i E_d^i, \\
w_o L_o &= \sum_i \beta^i [\sum_d \pi_{od}^i E_d^i], \\
P_d^i &= \Big(\sum_o [p_{od}^i]^{1-\sigma} \Big)^{\frac{1}{1-\sigma}},
\end{aligned} \tag{C.3}
$$

for unknowns (E_d^i, w_o, P_d^i), where $(p_{od}^i,\ \pi_{od}^i, L_d, Tr_o)$ are auxiliary variables and are evaluated according to[22]

$$
\begin{aligned}
p_{od}^i &= [\kappa^i w_o^{\beta^i} \prod_{j=1}^{S} [P_o^j]^{\gamma^{ij}} \tilde{\tau}_{od}^i] / T_{o'}^i \\
\pi_{od}^i &= \frac{[p_{od}^i]^{1-\sigma}}{[P_d^i]^{1-\sigma}}, \\
L_d &= \frac{B_d^{\frac{1}{\alpha^0}} \overline{H}_d \left[\frac{(w_d + Tr_d)^{1-\alpha^0}}{\prod_{i=1}^{S} (P_d^i)^{\alpha^i}} \right]^{\frac{1}{\alpha^0}}}{\sum_{d'} B_{d'}^{\frac{1}{\alpha^0}} \overline{H}_{d'} \left[\frac{(w_{d'} + Tr_{d'})^{1-\alpha^0}}{\prod_{i=1}^{S} (P_{d'}^i)^{\alpha^i}} \right]^{\frac{1}{\alpha^0}}} L_{CHN}, \forall d \in CHN, \\
Tr_o &= \frac{\alpha^0}{1-\alpha^0} \frac{1}{L_{CHN}} \sum_{d \in CHN} w_d L_d, \forall o \in CHN, \\
Tr_{ROW} &= \frac{\alpha^0}{1-\alpha^0} w_{RoW}.
\end{aligned} \tag{C.4}
$$

We design a nested fixed point algorithm according to the strength of the hardware. A key observation is that given (π_{od}^i, L_d, Tr_o), the first two equations of (C.3) give a (dense) system of linear equations for E_d^i and $w_d L_d$, for which GPUs are designed to solve efficiently. Based on this observation we design the nested fixed-point algorithm below:

Algorithm 1 Nested fixed-point algorithm for solving the competitive equilibrium using GPUs

1 Guess $(w_{d,Old},\ P_{d,Old}^i, Tr_{d,Old})$
2 Set flag_converged to **false**
while flag_converged is false **do**
  3 Construct (π_{od}^i, L_d) according to (C.4) based on $(w_{d,Old}, P_{d,Old}^i, Tr_{d,Old})$
  4 Solve the system of linear equations for E_d^i and $w_d L_d$ (with GPUs)
  5 Construct p_{od}^i, P_d^i, Tr_d according to (C.4) and (C.3)
  6 Set flag_converged to **true** if distance between (w_d, P_d^i, Tr_d) and $(w_{d,Old},\ P_{d,Old}^i, Tr_{d,Old})$ is small enough
  7 Update $(w_{d,Old},\ P_{d,Old}^i, Tr_{d,Old})$ according to (w_d, P_d^i, Tr_d)
end while

The step of solving the system of linear equations (line 4 in the algorithm) takes more than 90% of

[21] We describe the algorithm setting the exogenous deficits to zero. The model with exogenous deficits can be solved similarly.

[22] To see the determination of L_d, combine the consumer utility at the optimal choice $U_d \propto B_d \frac{w_d + Tr}{R_d^{\alpha^0} \prod_{i=1}^{S} (P_d^i)^{\alpha^i}}$, the land market clearing condition $R_d \overline{H}_d = \alpha^0 L_d (w_d + Tr)$, and the free mobility condition $U_d = U_{d'}, \forall d, d' \in CHN$.

the computation time in our benchmark model. Starting from an initial guess with uniform entries in (w_d, P_d^i, Tr_d), the benchmark equilibrium can be solved (under the convergence criterion of $1e-6$ in log difference) within a minute with a GTX1080Ti video card, compared to around 10 minutes with 2*Intel Xeon CPU E5-2650 v4.

Calibrate city-sector productivities T_d^i. The indirect inference estimation proceeds in a nested manner. In the inner loop, given other model parameters, we calibrate T_d^i for tradable sectors i such that the sectoral sales ratios between each city and the RoW in the model agree with those in the data. To do this, we treat sales ratios as observables, and solve T_d^i to generate the observable sales ratios while respecting the equilibrium conditions. Specifically, the minimal system of equations for calibrating T_d^i to match the sales ratios M_d^i[23] while respecting the equilibrium conditions are

$$\begin{aligned}
M_o^j &= \sum_d \frac{[c_o^j \cdot \tilde{\tau}_{od}^j]^{1-\sigma}}{\sum_{\tilde{o}}[c_{\tilde{o}}^j \cdot \tilde{\tau}_{\tilde{o}d}^j]^{1-\sigma}} \left(\alpha^j I_d + \sum_i \gamma^{ij} M_d^i\right) \text{ for all tradable sector } j, \\
w_d L_d &= \sum_i \beta^i M_d^i, \\
P_d^i &= \left(\sum_o [p_{od}^i]^{1-\sigma}\right)^{\frac{1}{1-\sigma}},
\end{aligned} \tag{C.5}$$

for unknowns $\left((T_d^i)_{i \text{ tradable}}, P_d^i, w_d\right)$[24], where $(I_d, c_d^i, p_{od}^i, Tr_o)$ are auxiliary variables and evaluated according to

$$\begin{aligned}
I_d &= (w_d + Tr_d)L_d + D_d, \\
c_d^i &= \kappa^i w_d^{\beta^i} \prod_{j=1}^{S} [P_d^j]^{\gamma^{ij}} / T_d^i, \\
p_{od}^i &= c_o^i \tilde{\tau}_{od}^i, \\
Tr_o &= \frac{\alpha^0}{1-\alpha^0} \frac{1}{L_{CHN}} \sum_{d \in CHN} w_d L_d, \forall o \in CHN, \\
Tr_{ROW} &= \frac{\alpha^0}{1-\alpha^0} w_{RoW},
\end{aligned}$$

with D_d being the exogenous trade deficits which are necessary to match the aggregate import and export shares. The above procedure is done by taking the targeted regional labor L_d as given. After the calibration, the relative amenities B_d are backed out combining the following: (1) the consumer utility at the optimal choice $U_d \propto B_d \frac{w_d + Tr_d}{R_d^{\alpha^0} \prod_{i=1}^S (P_d^i)^{\alpha^i}}$, (2) the land market clearing condition $R_d \overline{H}_d = \alpha^0 L_d (w_d + Tr)$, and (3) the free mobility condition $U_d = U_{d'}, \forall d, d' \in CHN$.

Calibrate remaining model parameters. With the inner loop inverting T_d^i to match M_d^i exactly, in the outer loop we search over other parameters to target the rest of the moments. These parameters include the sectoral international trade costs τ_{RoW}^i, the trade cost level parameter h_0, and the alternative mode

[23] M_d^i in the model is the total sales of intermediate goods from sector i of region d, and is linked to E_d^i defined before via $M_o^i = \sum_d E_d^i \pi_{od}^i$.

[24] Notice the system of Equation (C.5) is homogeneous of degree one in T_d^i for any i. That is, fixing i, scaling up T_d^i by the same factor scales nominal price and wage proportionally but does not affect real allocations. Therefore, we normalize $T_d^i = 1$ for a chosen region d for all i.

cost $\bar{\kappa}$. Since the number of parameters is equal to the number of moments, calibrating these parameters is to solve the system of equations such that the model moments are equal to their data counterparts listed in Table 4. We solve the system of equations using an iterative procedure based on a line search method. The equations are solved such that the maximum distance between the data moments and the model moments is less than 1%, and the maximum difference in the inner loop is smaller than $1e-5$.

References

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

_ and _ , "The Welfare Effects of Transportation Infrastructure Improvements," *NBER Working Paper No. 25487*, 2019.

Asturias, Jose, Manuel García-Santana, and Roberto Ramos, "Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India," *Journal of the European Economic Association*, 2018.

Bai, Chong-En and Yingyi Qian, "Infrastructure Development in China: The Cases of Electricity, Highways, and Railways," *Journal of Comparative Economics*, 2010, *38* (1), 34–51.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the Road: Access to Transportation Infrastructure and Economic Growth in China," *Journal of Development Economics*, 2020, p. 102442.

Baqaee, David and Emmanuel Farhi, "Networks, barriers, and trade," *NBER Working Paper No. 26108*, 2019.

Baum-Snow, Nathaniel, J Vernon Henderson, Matthew A Turner, Qinghua Zhang, and Loren Brandt, "Does Investment in National Highways Help or Hurt Hinterland City Growth?," *Journal of Urban Economics*, 2020, *115*, 103124.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2015, *82* (1), 1–44.

Combes, Pierre-Philippe and Laurent Gobillon, "The Empirics of Agglomeration Economies," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 247–348.

Coşar, A Kerem and Banu Demir, "Domestic Road Infrastructure and International Trade: Evidence from Turkey," *Journal of Development Economics*, 2016, *118*, 232–244.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade theory with numbers: Quantifying the consequences of globalization," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 197–261.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *70* (5), 1741–1779.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Facchini, Giovanni, Maggie Y Liu, Anna Maria Mayda, and Minghai Zhou, "China's "Great Migration": The Impact of the Reduction in Trade Policy Uncertainty," *Journal of International Economics*, 2019.

Fajgelbaum, Pablo D and Edouard Schaal, "Optimal Transport Networks in Spatial Equilibrium," *Econometrica*, 2020, *88* (4), 1411–1452.

Fan, Jingting, "Internal geography, labor mobility, and the distributional impacts of trade," *American*

Economic Journal: Macroeconomics, 2019, *11* (3), 252–88.

Fan, Shenggen and Connie Chan-Kang, *Road Development, Economic Growth, and Poverty Reduction in China*, Vol. 12, Intl Food Policy Res Inst, 2005.

Hillberry, Russell and David Hummels, "Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data," *European Economic Review*, 2008, *52* (3), 527–550.

Hulten, Charles R, "Growth accounting with intermediate inputs," *The Review of Economic Studies*, 1978, *45* (3), 511–518.

Rauch, James E, "Networks Versus Markets in International Trade," *Journal of International Economics*, 1999, *48* (1), 7–35.

Roberts, Mark, Uwe Deichmann, Bernard Fingleton, and Tuo Shi, "Evaluating China's Road to Prosperity: A New Economic Geography Aapproach," *Regional Science and Urban Economics*, 2012, *42* (4), 580–594.

Shi, Hao and Shaoqing Huang, "How Much Infrastructure is Too Much? A New Approach and Evidence from China," *World Development*, 2014, *56*, 272–286.

Su, Che-Lin and Kenneth L. Judd, "Constrained optimization approaches to estimation of structural models," *Econometrica*, 2012, *80* (5), 2213–2230.

Tian, Yuan, "International Trade Liberalization and Domestic Institutional Reform: Effects of WTO Accession on Chinese Internal Migration Policy," *Working Paper*, 2019.

Tombe, Trevor and Xiaodong Zhu, "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 2019, *109* (5), 1843–72.