The Distributional Impacts of Transportation Networks in China

Lin Ma and Yang Tang

Singapore Management University and Nanyang Technological University

2022-05-23, ABFER

Outline

Motivation

- Infrastructure investment, especially transportation networks, is widely seen as a driver of economic prosperity.
- The case of China is particularly interesting as the country has rapidly built up its transportation networks since the 1990s.

Motivation

- Infrastructure investment, especially transportation networks, is widely seen as a driver of economic prosperity.
- The case of China is particularly interesting as the country has rapidly built up its transportation networks since the 1990s.
- However, the distributional impacts are not clear, for two reasons:
  1. From a **theory** point of view, better roads improve the mobility of both **goods** and **factor**:
     - Better goods mobility improves the relative market access of the small and remote cities;
     - However, better factor mobility allows people to move out from the small and remote cities at the same time.
     - Dynamic responses also imply that the impacts might not be realized immediately.

Motivation

- Infrastructure investment, especially transportation networks, is widely seen as a driver of economic prosperity.
- The case of China is particularly interesting as the country has rapidly built up its transportation networks since the 1990s.
- However, the distributional impacts are not clear, for two reasons:
  1. From a **theory** point of view, better roads improve the mobility of both **goods** and **factor**:
     - Better goods mobility improves the relative market access of the small and remote cities;
     - However, better factor mobility allows people to move out from the small and remote cities at the same time.
     - Dynamic responses also imply that the impacts might not be realized immediately.
  2. **Empirically**, to do a quantitative analysis, we also need **panel data** on transportation networks; this is lacking in the context of China.

This project: data

We construct a new panel dataset on the transportation networks in China between 1994 and 2017:

1. Measuring the **quality** of the roads and railroads using design speed.
    - Across **time**: a "highway" built in 1994 only allows for a fraction of travel speed as compared to a highway in 2017.
    - Across **space**: roads in the rugged terrains are built for half the design speed as those in the eastern flood plains.

This project: data

We construct a new panel dataset on the transportation networks in China between 1994 and 2017:

1. Measuring the **quality** of the roads and railroads using design speed.
    - Across **time**: a "highway" built in 1994 only allows for a fraction of travel speed as compared to a highway in 2017.
    - Across **space**: roads in the rugged terrains are built for half the design speed as those in the eastern flood plains.
2. Multiple modes: road, railroad, and waterway.
3. Distinguish the types of traffic on the railroads: passenger, freight, or mixed.

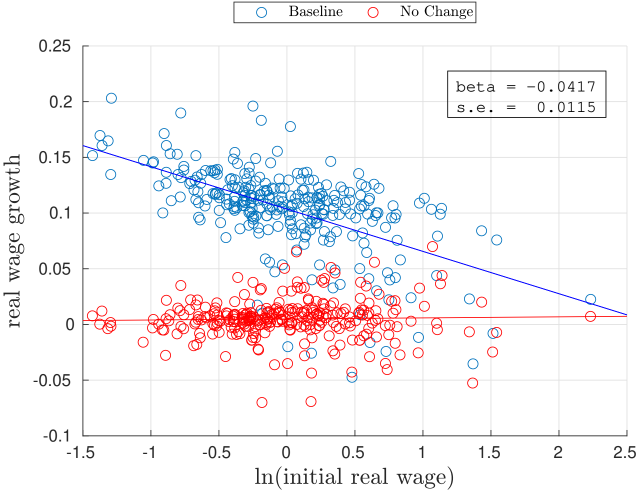
This project: model

Start from a dynamic discrete choice model (CDP 2019) with forward-looking migration. In our model:

- Both trade and **migration costs** respond to the changes in the observed transportation networks.
- Route choices of migration and trade, a la Allen and Arkolakis (2022), extended to multiple transportation modes.
  - Moment conditions to structurally estimate the mode-specific costs and elasticities on a transition path.
  - Dynamic returns to transportation networks.

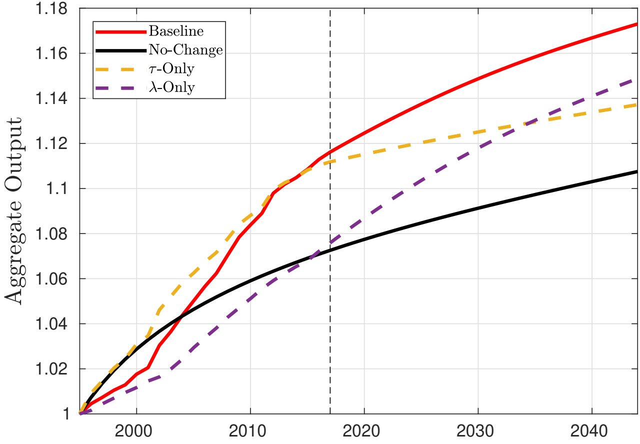
Results: Distributional Impacts

Without the expansion of the network, the spatial inequality would have stayed the same over time.



Results: Aggregate and Dynamic Impacts

- Return to trade liberalization is immediate but wanes in the long run.
- Return to migration liberalization could be negative in the short-run but grows over time.



Literature

- Quantitative evaluation of infrastructure projects:
  - Allen and Arkolakis (2014, 2022), Redding and Turner (2015), Donaldson and Hornbeck (2016), Alder (2016), Allen and Donaldson (2018), Donaldson (2018).
- In the Chinese context:
  - Bai and Qian (2010), Banerjee, Duflo, and Qian (2012), Faber (2014), Qin(2016), Lin (2017), Xu (2017), Baum-Snow, Henderson, Turner, Zhang and Brandt (2020), Fan, Lu and Luo (2021), Alder, Song and Zhu (2021).
- Migration:
  - Artuc, Chaudhuri and McLaren (2010), Caliendo, Dvorkin and Parro (2019), Tombe and Zhu (2019), Fan (2019), Ma and Tang (2020).

Outline

Measuring the Quality of Infrastructure

- Existing transportation data in China are binary and ignore quality variations:
  - USGS 1992 map (USGS, CDC at University of Michigan);
  - State Bureau of Surveying and Mapping (China) 2008 map;
  - ACASIAN Data Center at Griffith University from 1992 to 2010;
  - Baum-Snow et al. (2020).

Measuring the Quality of Infrastructure

- Existing transportation data in China are binary and ignore quality variations:
  - USGS 1992 map (USGS, CDC at University of Michigan);
  - State Bureau of Surveying and Mapping (China) 2008 map;
  - ACASIAN Data Center at Griffith University from 1992 to 2010;
  - Baum-Snow et al. (2020).
- The quality of roads and railroads differ by vintage, rate, and terrain, as stipulated by the official design codes:
  - Older, lower-rated, or built over rugged terrains have limited speed and capacity.
    - Nine revisions in road construction codes, recent ones in 1988, 1997, 2003, and 2014.
    - Recent revisions in railway design codes in 1985, 1999, 2006, 2012.
  - Conditional on revision and rate, quality vary by terrain.

Design Speed Depends on Time, Rate, and Terrain

Terrains | More Railway Design Codes

Revision	Plains	LRH	Hills	Mountains	Plains	LRH	Hills	Mountains
	Highways				First-Rate Roads			
1988	120	120	100	60	100	100	60	60
1997	120	120	120	60	100	100	60	60
2003	120	120	120	80	100	100	80	80
2014	120	120	120	80	100	100	80	80

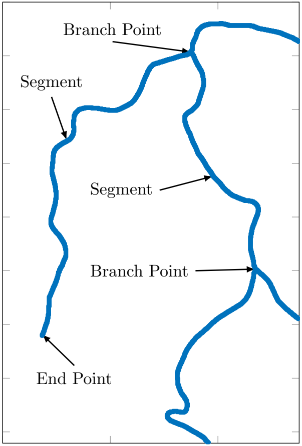
(a) Road Standards

Revision	Plains	LRH	Hills	Mountains	Plains	LRH	Hills	Mountains
	National I				National II			
1985	120	100	80	80	100	80	80	80
1999	140	120	80	80	100	80	80	80
2006	160	140	120	120	120	100	80	80

(b) Railroad Standards, Mixed-Use

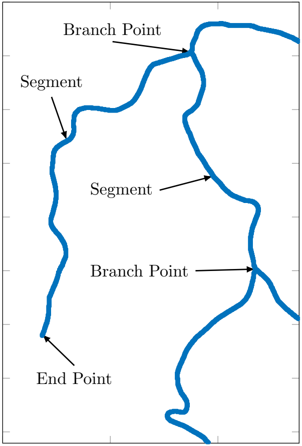
Table 1: Design Speed (km/h) of Roads and Railroads by Time, Rate, and Terrain

Compiling the Dataset



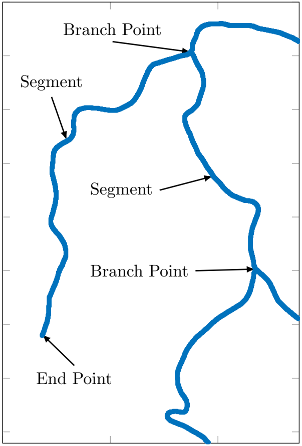
- We only geo-reference national maps that are larger than 1:6 million. Maps
  - To ensure consistent geo-referencing across the years.
  - Ten national maps met this threshold out of 24 years.
- Extract road, railroad, and waterway networks with color identification.
- Treat the network as a **graph**, and break it into "**segments**": the set of pixels between branch and end points.

Compiling the Dataset



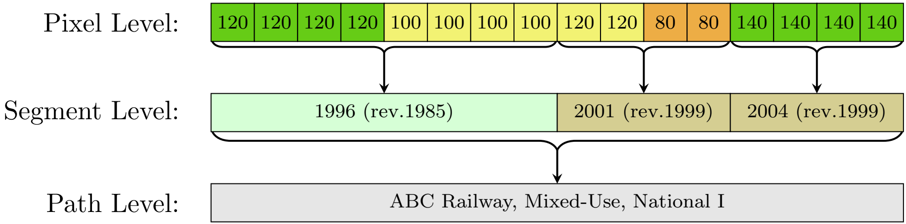
- For each segment, identify the **year of construction** (YOC).
  - Compare to the previous national map and determine if it is a new construction.
  - If new, then determine the YOC from the other sources.
  - For example, we only have national maps in 1994 and 1996. If a segment first showed up in 1996, it could be constructed in 1994, 1995, or 1996. Example
- References:
  - Provincial map collections
  - *Transportation Yearbooks*
  - *Railroad Yearbooks*
  - *Chronicles of Railroad Construction*

Compiling the Dataset



- Segments are parts of named roads or railroads that we call "**path**".
- For each parent "path", identify **rate** and **usage**, using the same references.
- Together with **YOC**, we can pin down the design code applicable to the segment.
- Within the segment, **pixels** differ by terrain. We can then determine the design speed of every pixel within the segment.

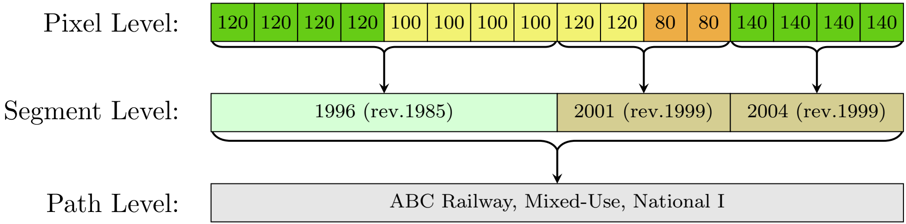
Structure of the Dataset



The full dataset contains three layers.

- The outer-most layer is a "path" that correspond to a named road or railroad, e.g. Beijing-Shanghai Highway, Longhai Railway, etc.
- We identify the **name**, **rate**, and **usage** of a path.
- Rate and usage are used to determine the applicable design codes.

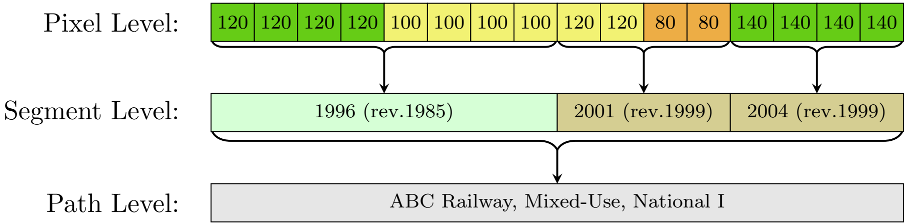
Structure of the Dataset



A “path” contains several “segments” built over multiple years.

- Within a path, segments differ in the **year of construction**.
- In this example, one segment was built in 1996 and was subject to the 1985 revision of the design codes.
- The two segments built in 2001 and 2004 were subject to the newer 1999 revision.

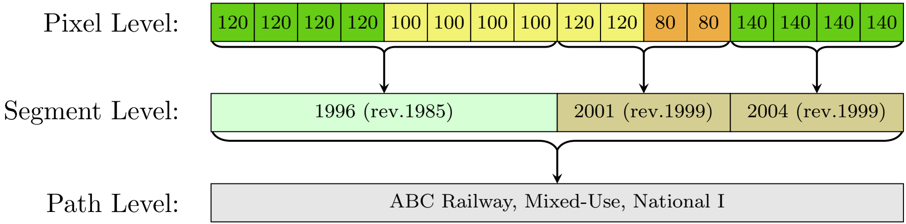
Structure of the Dataset



A “segment” contains many “pixels” that differ in **terrain**.

- = “plains”, = “low-rolling hills”, and = “hills”.
- Using the information on usage, rate, year of construction, and terrain, we determine the design speed of each pixel, shown in the box in the unit of “km/h”.

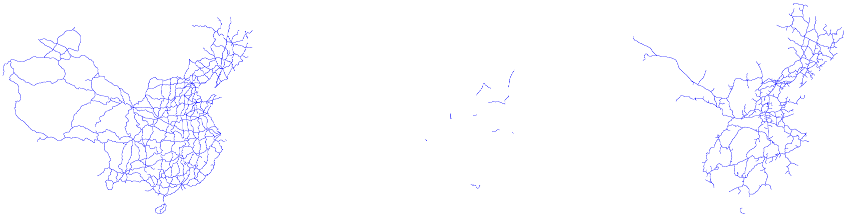
Structure of the Dataset



- Within a segment, design speed differs due to terrains.
- Within a terrain type, design speed differs due to the changes in applicable technical standards.
- From the design speed, we compute the distance between prefectures i, j in the unit of **travel time**.
- We do this for freight ($T_{ijt}^{m\mathbb{f}}$) and passengers ($T_{ijt}^{m\mathbb{p}}$) separately for mode m at year t.

Result: Changes in Connectivity

1994



First-Rate Roads | Highway | Railroad

Figure 1: Changes in Transportation Networks, 1994-2017

Result: Changes in Connectivity

1999

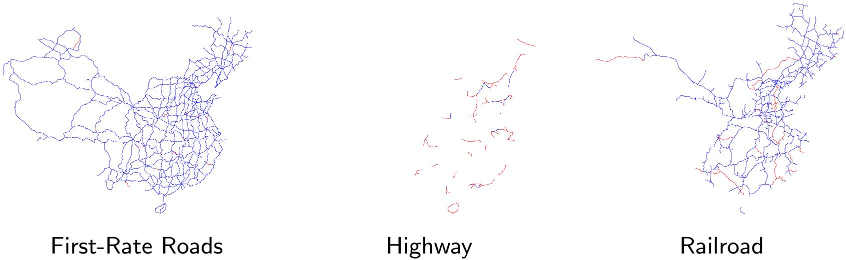
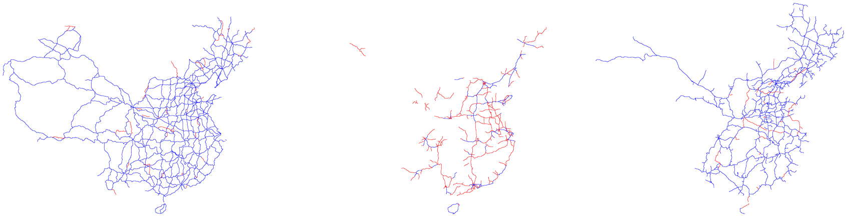


Figure 1: Changes in Transportation Networks, 1994-2017

Result: Changes in Connectivity

2004

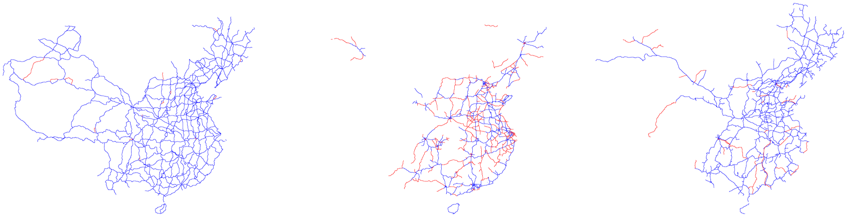


First-Rate Roads | Highway | Railroad

Figure 1: Changes in Transportation Networks, 1994-2017

Result: Changes in Connectivity

2009

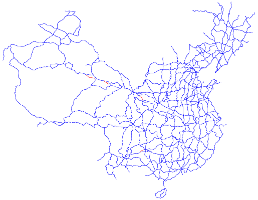


First-Rate Roads | Highway | Railroad

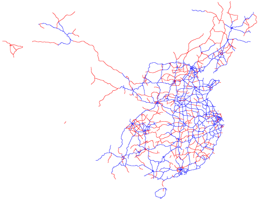
Figure 1: Changes in Transportation Networks, 1994-2017

Result: Changes in Connectivity

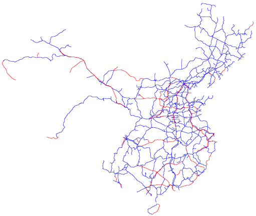
2014



First-Rate Roads



Highway

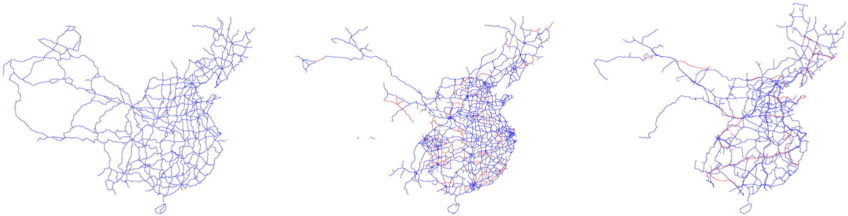


Railroad

Figure 1: Changes in Transportation Networks, 1994-2017

Result: Changes in Connectivity

2017

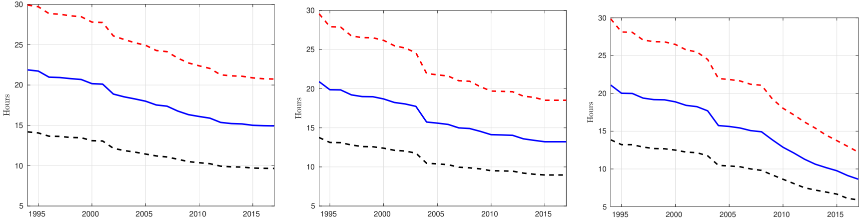


First-Rate Roads Highway Railroad

Figure 1: Changes in Transportation Networks, 1994-2017

Result: Changes in Travel Time

- The median travel time decline by 31 to 59 percent.
- Strongest improvements come from the passenger travel on railroads, due to HSR.



(a) Road (b) Railroad, Freight (c) Railroad, Passenger

Rich Variations Across Modes, Locations, and Time

Travel time of trips originating from Beijing and Guangzhou:

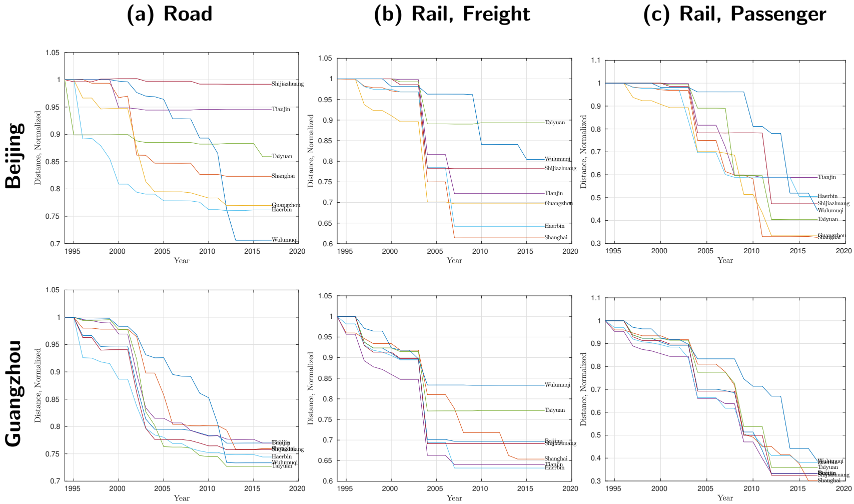
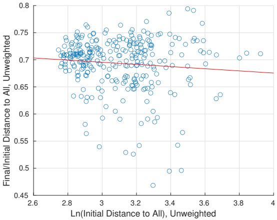


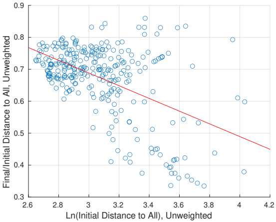
Figure 2: City Pairs

Infrastructure as an Equalizer

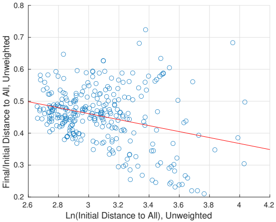
The initially remote cities received more reduction in transportation costs:



(a) Road



(b) Railroad, Freight



(c) Railroad, Passenger

Figure 3: Convergence of Travel Time

Outline

General Environment

Dynamic migration model from CDP, extended to allow for multiple modes of transportation, route choices, and time-varying migration costs.

- J geographically segmented domestic cities, indexed by $j = 1, 2...J$, and an additional ROW denoted as $j = 0$.
- Time is discrete and infinite, indexed as $t = 0, 1, 2, ...$

General Environment

Dynamic migration model from CDP, extended to allow for multiple modes of transportation, route choices, and time-varying migration costs.

- J geographically segmented domestic cities, indexed by $j = 1, 2 \ldots J$, and an additional ROW denoted as $j = 0$.
- Time is discrete and infinite, indexed as $t = 0, 1, 2, \ldots$.
- Migration is forward-looking, subject to frictions that depend on passenger transportation and policy barriers.

General Environment

Dynamic migration model from CDP, extended to allow for multiple modes of transportation, route choices, and time-varying migration costs.

- J geographically segmented domestic cities, indexed by $j = 1, 2...J$, and an additional ROW denoted as $j = 0$.
- Time is discrete and infinite, indexed as $t = 0, 1, 2, ...$
- Migration is forward-looking, subject to frictions that depend on passenger transportation and policy barriers.
- The **direct** cost of moving from j to i at time t using mode m, for passengers ($\mathbb{p}$):

$$d_{ijt}^{m\mathbb{p}} = a^{\mathbb{p}} \cdot a^{m\mathbb{p}} \cdot \left(T_{ijt}^{m\mathbb{p}}\right)^{h^{\mathbb{p}}}, \tag{1}$$

  - $T_{ijt}^{m\mathbb{p}}$ is the distance measured in travel time from our dataset.
  - New parameters to be estimated: average migration costs ($a^{\mathbb{p}}$), mode-specific costs ($a^{m\mathbb{p}}$), and the time elasticity ($h^{\mathbb{p}}$).

Route Choices in Dynamic Discrete Choice

Route choices following Allen and Arkolakis (2022):

- Migrants do not always use the direct route. They decide a mode of transportation and then a *route* r^m of K steps.
- The route incurs a **additive** cumulative cost of $\sum_{k=1}^{K} d^{m\mathbb{p}}_{r^m_{k-1}, r^m_k, t}$, where $r^m_k = 1, 2, \cdots, J$ is the location of the kth-step in the route r^m.
- $\mathcal{R}^m_{ij}$ denotes all the possible routes from j to i.

Route Choices in Dynamic Discrete Choice

Route choices following Allen and Arkolakis (2022):

- Migrants do not always use the direct route. They decide a mode of transportation and then a *route* r^m of K steps.
- The route incurs a **additive** cumulative cost of $\sum_{k=1}^{K} d^{m\mathbb{p}}_{r^m_{k-1}, r^m_k, t}$, where $r^m_k = 1, 2, \cdots, J$ is the location of the kth-step in the route r^m.
- $\mathcal{R}^m_{ij}$ denotes all the possible routes from j to i.
- Type-I GEV idiosyncratic preference shocks toward **each destination and route**, $\{\epsilon_{it,r^m}\}^{J}_{i=1, r^m \in \cup_{m=1}^{M} \mathcal{R}^m_{ij}}$, *i.i.d* across location, route, time, and individual.

The Recursive Migration Problem

$$v_{jt} = \log\left(\phi_{jt}\frac{w_{jt}}{P_{jt}}\right) + \max_{i,r^m \in \cup_{m=1}^{M} \mathcal{R}_{ij}^m} \left\{\delta \mathbb{E}[v_{i,t+1}] - \bar{d}_{it} - \sum_{k=1}^{K} d^{m\mathbb{P}}_{r^m_{k-1}, r^m_k, t} + \kappa \cdot \epsilon_{it,r^m}\right\}, \tag{2}$$

- $\phi_{jt} = \bar{\phi}_j L_{jt}^{\beta}$ is the endogenous amenity, w_{jt}/P_{jt} the real wage, and δ the discount rate.
- κ is the migration elasticity.
- $\bar{d}_{it}$ is the entry barrier, the hukou system.
- Two differences relative to CDP (2019):
  1. The dynamic choice set is destination-by-route: $i, r^m \in \cup_{m=1}^{M} \mathcal{R}_{ij}^m$.
  2. Migration costs are time-varying.

The Recursive Problem: Solution

- Let $V_{jt} \equiv \mathbb{E}[v_{jt}]$ denote the expectation of the continuation value. The optimization problem defined in equation (2) adopts the same solution as in CDP, conditional on λ_{ijt}:

$$V_{jt} = \log\left(\phi_{jt}\frac{w_{jt}}{P_{jt}}\right) + \kappa \log\left(\sum_{i=1}^{J} \exp\left(\delta V_{i,t+1} - \lambda_{ijt}\right)^{1/\kappa}\right), \tag{3}$$

- λ_{ijt} is the **expected** migration costs across all possible modes and routes from j, conditional on moving to i:

$$\lambda_{ijt} = \bar{d}_{it} - \kappa \log\left[\sum_{m=1}^{M} \sum_{r^m \in \mathcal{R}_{ij}^m} \exp\left(-\frac{\left(\sum_{k=1}^{K} d_{r_{k-1}^m, r_k^m, t}^{m\mathbb{p}}\right)}{\kappa}\right)\right]. \tag{4}$$

The Recursive Problem: Solution

- Let $V_{jt} \equiv \mathbb{E}[v_{jt}]$ denote the expectation of the continuation value. The optimization problem defined in equation (2) adopts the same solution as in CDP, conditional on λ_{ijt}:

$$V_{jt} = \log\left(\phi_{jt}\frac{w_{jt}}{P_{jt}}\right) + \kappa \log\left(\sum_{i=1}^{J} \exp\left(\delta V_{i,t+1} - \lambda_{ijt}\right)^{1/\kappa}\right), \tag{3}$$

- λ_{ijt} is the **expected** migration costs across all possible modes and routes from j, conditional on moving to i:

$$\lambda_{ijt} = \bar{d}_{it} - \kappa \log\left[\sum_{m=1}^{M} \sum_{r^m \in \mathcal{R}_{ij}^m} \exp\left(-\frac{\left(\sum_{k=1}^{K} d^{m\mathbb{P}}_{r^m_{k-1}, r^m_k, t}\right)}{\kappa}\right)\right]. \tag{4}$$

- The dynamic migration probabilities follow:

$$\mu_{ijt} = \frac{\exp\left(\delta V_{i,t+1} - \lambda_{ijt}\right)^{1/\kappa}}{\sum_{i'=1}^{J} \exp\left(\delta V_{i',t+1} - \lambda_{i'jt}\right)^{1/\kappa}}. \tag{5}$$

Solution to λ_{ijt}

- Following the methods in Allen and Arkolakis (2022), we can further express λ_{ijt} as a function of the transportation networks and policy barriers:

$$\lambda_{ijt} = \bar{d}_{it} - \kappa \log \left(\sum_{m=1}^{M} b_{ijt}^{m\mathbb{p}} \right), \tag{6}$$

- $b_{ijt}^{m\mathbb{p}}$ is the (i, j)th element of the matrix $\mathbf{B}_t^{m\mathbb{p}}$, which is a function of the observed transportation networks and the parameters of the model.

Solution to λ_{ijt}

- Following the methods in Allen and Arkolakis (2022), we can further express λ_{ijt} as a function of the transportation networks and policy barriers:

$$\lambda_{ijt} = \bar{d}_{it} - \kappa \log \left(\sum_{m=1}^{M} b_{ijt}^{m\mathbb{p}} \right), \tag{6}$$

- $b_{ijt}^{m\mathbb{p}}$ is the (i, j)th element of the matrix $\mathbf{B}_t^{m\mathbb{p}}$, which is a function of the observed transportation networks and the parameters of the model.

$$\mathbf{B}_t^{m\mathbb{p}} \equiv \left(\mathbf{I} - \mathbf{F}_t^{m\mathbb{p}}\right)^{-1} = \sum_{K=0}^{\infty} \left(\mathbf{F}_t^{m\mathbb{p}}\right)^{K},$$

$$\mathbf{F}_t^{m\mathbb{p}} \equiv \left[\exp\left(-\frac{d_{ijt}^{m\mathbb{p}}}{\kappa} \right) \right]$$

Production and Trade

- The production and trade side follows Allen and Arkolakis (2022): every city is able to produce every variety, ω. Market structure is perfect competition.
- Labor is the only input:

$$q_{jt} = A_{jt}\ell,$$

$A_{jt} = \bar{A}_{jt}\left(L_{jt}\right)^{\alpha}$ is the endogenous productivity.

- Consumers decide an origin and a **route** to source each variety. The direct **ad valorem** cost from moving from j to i for freight transportation ($\mathbb{f}$) is:

$$d_{ijt}^{m\mathbb{f}} = \exp\left(a^{\mathbb{f}} \cdot a^{m\mathbb{f}} \cdot \left(T_{ijt}^{m\mathbb{f}}\right)^{h^{\mathbb{f}}}\right), \tag{7}$$

- Moving from i to j through a *route* r^m of K steps incurs a **multiplicative** cost of $\prod_{k=1}^{K} d_{r_{k-1}^m, r_k^m, t}^{m\mathbb{f}} \geq 1$.

Production Cont'd

- Consumers in i face the following price if they source variety ω from location j through route r^m:

$$p_{ijt,r^m}(\omega) = \frac{w_{jt}}{A_{jt}} \frac{\prod_{k=1}^{K} d^{m\mathbb{f}}_{r^m_{k-1},r^m_k,t}}{z_{ijt,r^m}(\omega)},$$

$z_{ijt,r^m}(\omega)$ is i.i.d. Frechet shock across time, location, route, and variety and has a shape parameter θ.
- The probability that individuals in i source from j along route r^m is:

$$\pi_{ijt,r^m} = \frac{(w_{jt}/A_{jt})^{-\theta} \prod_{k=1}^{K} \left(d^{m\mathbb{f}}_{r^m_{k-1},r^m_k,t} \right)^{-\theta}}{\sum_{j'=0}^{J} \left(w_{j't}/A_{j't} \right)^{-\theta} \tau_{ij't}^{-\theta}}, \tag{8}$$

- τ_{ijt} is the expected transportation cost between i and j at time t:

$$\tau_{ijt} = \left[\sum_{m=1}^{M} \sum_{r^m \in \mathcal{R}^m_{ij}} \prod_{k=1}^{K} \left(d^{m\mathbb{f}}_{r^m_{k-1},r^m_k,t} \right)^{-\theta} \right]^{-\frac{1}{\theta}}. \tag{9}$$

Solution of τ_{ijt}

- Similar to λ_{ijt}, we can solve τ_{ijt} as:

$$\tau_{ijt} = \left(\sum_{m=1}^{M} b_{ijt}^{m\mathbb{f}} \right)^{-\frac{1}{\theta}}.$$

where $b_{ijt}^{m\mathbb{f}}$ the (i,j)th element of the matrix $\mathbf{B}_t^{m\mathbb{f}}$, which is a function of the observed freight transportation networks and parameters of the model:

$$\mathbf{B}_t^{m\mathbb{f}} \equiv \sum_{K=1}^{\infty} \left(\mathbf{F}_t^{m\mathbb{f}} \right)^K = \left(\mathbf{I} - \mathbf{F}_t^{m\mathbb{f}} \right)^{-1}$$

$$\mathbf{F}_t^{m\mathbb{f}} \equiv \left[\left(d_{ijt}^{m\mathbb{f}} \right)^{-\theta} \right].$$

Trade Balance and Equilibrium

- Standard trade balance condition holds every period:

$$w_{jt}L_{jt} = \sum_{i=0}^{J} \frac{(w_{jt}\tau_{ijt})^{-\theta}\left(A_j L_{jt}^{\alpha}\right)^{\theta}}{\sum_{j'=1}^{J}\left(w_{j't}\tau_{ij't}\right)^{-\theta}\left(A_{j'} L_{j't}^{\alpha}\right)^{\theta}} w_{it}L_{it} \qquad (10)$$

- The equilibrium:
  - the time-invariant fundamentals as $\Omega = \{A_j, \bar{\phi}_j\}$,
  - the time-variant fundamentals as $\Omega_t = \{\bar{d}_{jt}, \mathbf{D}_t^{m\mathbb{p}}, \mathbf{D}_t^{m\mathbb{f}}\}$,
  - and the sequence of endogenous variables as $\Upsilon_t = \{w_{jt}, L_{jt}, V_{jt}, \mu_{jt}\}$.
- Standard concept of steady state: $\Upsilon_t = \bar{\Upsilon}, \forall t$, and transition path to a steady state, on which:
  1. Individuals maximize their life-time utility by choosing a sequence of locations so that equations (5) hold.
  2. Firms maximize their profits in each period and trade balances, so that equation (10) holds.

Outline

Overall Setup

- We quantify the model to 291 prefecture cities in China plus 1 region representing the ROW.
- Calibration and estimation done on a 50-year transition path from 1995, $t = 1$. From 1995 to 2017, we apply the actual transportation networks and policy.
  - Assume that the transportation and policy stayed the same after 2017, and therefore the long-run steady state is the one implied by the 2017 variables.
- Do not need to assume that 1995 is in steady state. Instead, year 1995 is on a transition path towards its initial steady state.
- The transition path is solved in levels, which allows for flexibility in choosing the moments in quantification.

Hukou System

- We map the entry barrier in the migration frictions, $\bar{d}_{jt}$, to the hukou reform index constructed in Fan (2019).
- We extend the index to the 291 prefectures between 1994 and 2017 using the same archives.
- Denote the observed hukou index in prefecture j as $k_{jt} \geq 0$, we then model:

$$\bar{d}_{jt} = \exp\left(\psi k_{jt}\right).$$

- ψ is a parameter to be estimated.

Externally-Calibrated Parameters

Some parameters are taken from the literature:

name	value	source	notes
α	0.1	Redding and Turner (2015)	agglomeration elasticity
β	-0.3	Allen and Arkolakis (2022)	congestion elasticity
θ	6.2	Allen and Arkolakis (2022)	trade elasticity
κ	2.02	Caliendo et al. (2019)	migration elasticity
η	6.0	Anderson and van Wincoop (2004)	elasticity of substitution

Table 2: External Parameters

Two-Layered Nested Estimation

- All the other parameters are calibrated and estimated on the transition path. We adopt a nested estimation strategy.
- In the **outer layer**, we estimate the mode-specific costs and time elasticity of trade and migration costs using the GMM. These parameters are $\Theta^{\chi} = \left\{\{a^{m\mathbb{f}}, a^{m\mathbb{p}}\}_{m=1}^{M}, h^{\mathbb{f}}, h^{\mathbb{p}}\right\}$.
- All the other parameters are jointly-calibrated in the **inner layer**: $\Theta = \left\{\{\bar{A}_j, \bar{\phi}_j\}_{j=0}^{J}, a^{\mathbb{f}}, a^{\mathbb{p}}, \psi, \tau^*\right\}$.
- The inner-layer calibration is done for every guess of Θ^{χ} from the GMM estimation.

Inner-Layer Joint Calibration

name	value	target	notes
$\{\bar{A}_j\}$	-	output in 1995	fundamental productivity
$\{\bar{\phi}_j\}$	-	population in 1995	fundamental amenity
$a^{\mathbb{T}}$	0.815	internal-trade-to-GDP ratio, 2002	average internal trade costs
$a^{\mathbb{P}}$	5.823	average annual stay rate between 2000 and 2005	average migration costs
τ^*	1.614	average export-to-GDP ratio, 2000 to 2005	international trade barrier
ψ	0.039	average annual stay rate between 2010 and 2015	elasticity of $\bar{d}_{it}$ to the hukou reform index

- $\{\bar{A}_j, \bar{\phi}_j\}$ come from inverting the model in 1995, following Kleinman et al. (2021).
- Many target moments are not available in the initial year, but only on the transition path. Solving the model in levels allows us to utilize these moments.
- ψ captures the impact of hukou reform on internal migration, conditional on the changes in transportation.

The Outer Layer GMM: Moments, Freight

- The first set of moments is the average share of freight traffic that goes through the road, rail, and waterway networks each year across prefectures.
- Observed in the statistical yearbooks.
- These moments capture $\{a^{m\mathbb{f}}\}_{m=1}^{M}$. Normalize $a^{\text{river},\mathbb{f}} = 1$.
- The share of sales from j via mode m is the weighted average of pair-specific shares across destinations:

$$s_{jt}^{m\mathbb{f}} = \sum_{i=1}^{J} \left(\frac{b_{ijt}^{m\mathbb{f}}}{\sum_{m'=1}^{M} b_{ijt}^{m'\mathbb{f}}} \right) \frac{X_{ijt}}{X_{jt}}. \tag{11}$$

- The moment condition is the vector of average shares across prefecture each year between 1995 and 2017: $\left\{ \sum_{j=1}^{J} s_{jt}^{m\mathbb{f}} / J \right\}_{t=1}^{23}$.

The Outer Layer GMM: Moments, Passenger

- The second set of moments is for passenger flows. These moments capture $\{a^{m\mathbb{p}}\}_{m=1}^{M}$. Again, normalize $a^{\text{river},\mathbb{p}} = 1$.
- The share of population flow from j to all the other prefectures via mode m can be computed similarly as:

$$s_{jt}^{m\mathbb{p}} = \sum_{i=1}^{J} \left(\frac{b_{ijt}^{m\mathbb{p}}}{\sum_{m'=1}^{M} b_{ijt}^{m'\mathbb{p}}} \right) \frac{L_{ijt}}{L_{jt}}, \tag{12}$$

- The associated moment condition is $\left\{ \sum_{j=1}^{J} s_{jt}^{m\mathbb{p}} / J \right\}_{t=1}^{23}$.
- The asymmetry in the migration costs and entry barriers do not affect the choice of transportation mode because the route/mode choice is conditional on a destination location.

The Outer Layer GMM: Moments, CV

- The last set of moments is the **coefficient of variations (CV)** for traffic volumes. They capture the time elasticities $\{h^{\mathbb{f}}, h^{\mathbb{p}}\}$, the parameters that control the variations in τ_{ijt} and λ_{ijt} conditional on $\{\mathbf{T}_t^{m\chi}\}$.
- As h^{χ} declines, the trade or migration costs matrix becomes uniform and less dependent on the underlying geography summarized in $\{\mathbf{T}_t^{m\chi}\}$.
- The resulting variations in trade or migration traffic across prefecture will decline, leading to a smaller CV.

The Outer Layer GMM: Moments, CV

- The last set of moments is the **coefficient of variations (CV)** for traffic volumes. They capture the time elasticities $\{h^{\mathbb{f}}, h^{\mathbb{p}}\}$, the parameters that control the variations in τ_{ijt} and λ_{ijt} conditional on $\{\mathbf{T}_t^{m\chi}\}$.
- As h^{χ} declines, the trade or migration costs matrix becomes uniform and less dependent on the underlying geography summarized in $\{\mathbf{T}_t^{m\chi}\}$.
- The resulting variations in trade or migration traffic across prefecture will decline, leading to a smaller CV.
- The moment conditions are the vectors of the average CV across prefecture each year between 1995 and 2017: $\left\{\frac{\sum_{j=1}^{J} \mathsf{CV}_{jt}^{\mathbb{f}}}{J}\right\}_{t=1}^{23}$ and $\left\{\frac{\sum_{j=1}^{J} \mathsf{CV}_{jt}^{\mathbb{p}}}{J}\right\}_{t=1}^{23}$.

GMM

- Denote the **92 moment conditions** in the data as the vector $\bar{\mathcal{S}}$ to estimate **6 parameters**, and the counter-parts in the model as:

$$\mathcal{S}\left(\Theta^{\chi}\right)=\left\{\frac{\sum_{j=1}^{J} s_{jt}^{m\mathbb{f}}}{J}, \frac{\sum_{j=1}^{J} s_{jt}^{m\mathbb{p}}}{J}, \frac{\sum_{j=1}^{J} \mathsf{CV}_{jt}^{\mathbb{f}}}{J}, \frac{\sum_{j=1}^{J} \mathsf{CV}_{jt}^{\mathbb{p}}}{J}\right\}_{t=1}^{23}.$$

- The GMM estimator finds Θ^{χ} to minimize the distance between the data and the model moments:

$$\min_{\Theta^{\chi}}\left[\bar{\mathcal{S}}-\mathcal{S}\left(\Theta^{\chi}\right)\right]\mathbf{W}\left[\bar{\mathcal{S}}-\mathcal{S}\left(\Theta^{\chi}\right)\right]',$$

- $\mathbf{W}$ is the optimal weighting matrix, which is the inverse of the variance-covariance matrix of the data moments $\bar{\mathcal{S}}$, computed by bootstrapping the data.

Estimation Results

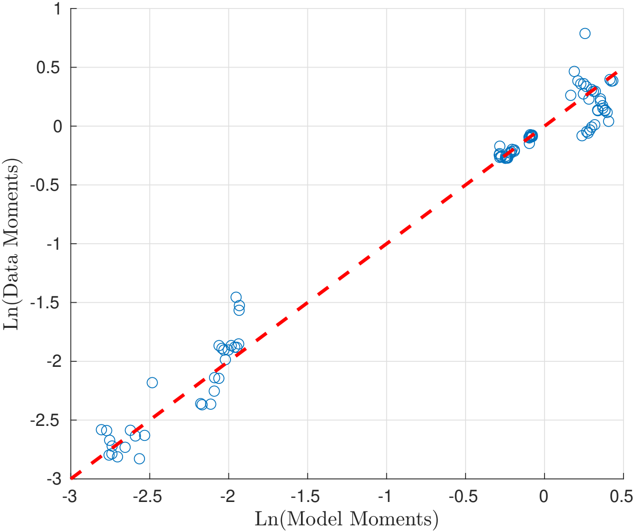


Figure 4: Goodness-of-Fit, GMM

Estimation Results

name	value	s.e.	notes
$a^{1,\mathbb{f}}$	0.984	0.078	road costs, freight
$a^{2,\mathbb{f}}$	1.151	0.085	rail costs, freight
$h^{\mathbb{f}}$	0.111	0.034	time elasticity, freight
$a^{1,\mathbb{p}}$	0.884	0.085	road costs, passenger
$a^{2,\mathbb{p}}$	1.275	0.093	rail costs, passenger
$h^{\mathbb{p}}$	0.320	0.051	time elasticity, passenger

Table 3: Parameters, Estimated

Outline

Baseline and No-Change Equilibrium

- Compare the steady states and transition paths between the "baseline" and the "no-change" counterfactual.
- In both cases, we start from 1995 and solve the model forward 50 years towards a long-run steady state.

Baseline and No-Change Equilibrium

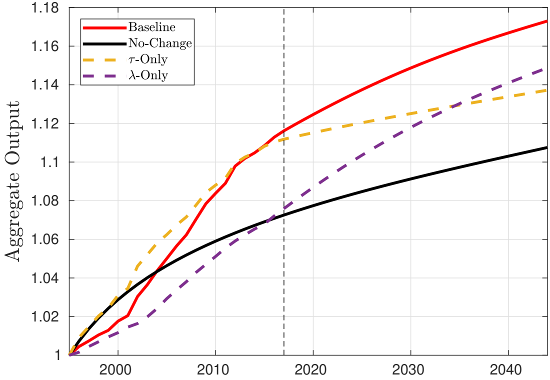
- Compare the steady states and transition paths between the "baseline" and the "no-change" counterfactual.
- In both cases, we start from 1995 and solve the model forward 50 years towards a long-run steady state.
- The **baseline** equilibrium is the same transition path that we used to estimate the model.
  - From 1995 to 2017, the actual data. No change in transportation and policy after 2017.
  - The steady state implied by the variables in 2017.

Baseline and No-Change Equilibrium

- Compare the steady states and transition paths between the "baseline" and the "no-change" counterfactual.
- In both cases, we start from 1995 and solve the model forward 50 years towards a long-run steady state.
- The **baseline** equilibrium is the same transition path that we used to estimate the model.
  - From 1995 to 2017, the actual data. No change in transportation and policy after 2017.
  - The steady state implied by the variables in 2017.
- The **no-change** counter-factual:
  - Fix the transportation networks and policy to those in 1995.
  - The steady state implied by the variables in 1995.

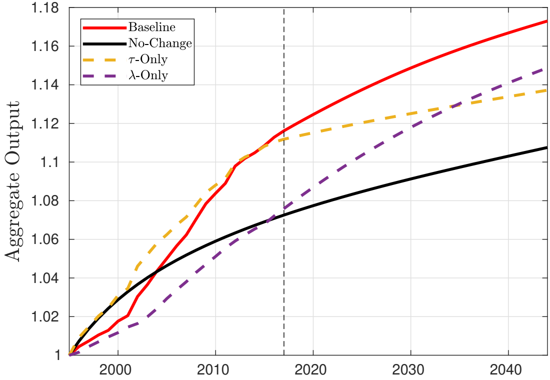
Results: Aggregate and Dynamic Impacts

- The expansion of transportation networks increases long-run output by about 60%.



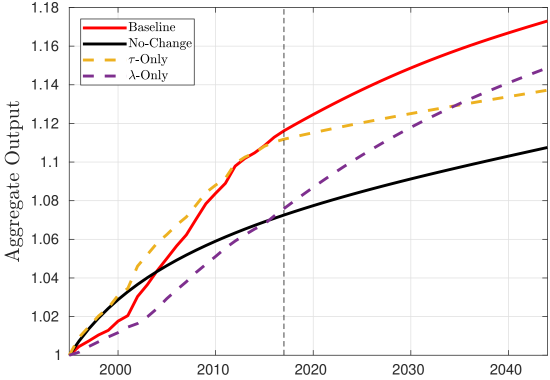
Results: Aggregate and Dynamic Impacts

- Return to trade liberalization is immediate but wanes in the long-run.
- By 2017, almost all the benefit comes from goods transportation.



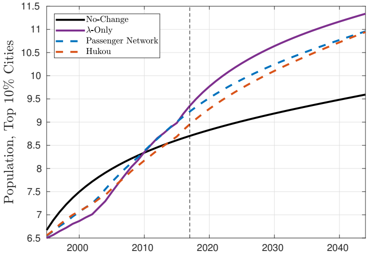
Results: Aggregate and Dynamic Impacts

- Return to migration liberalization could be negative in the short-run but grows over time.

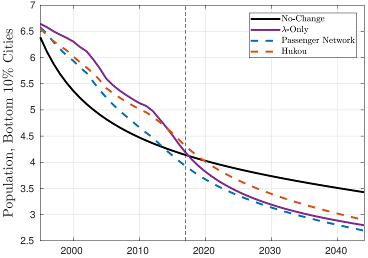


Delayed Return to Migration Liberalization

- The short-run negative return to migration liberalization comes from the delayed movements towards more productive locations.
- People expect better infrastructure and policy in the future, and thus they prolong their stay in the less productive prefectures.



(a) Top 10% Productive Locations

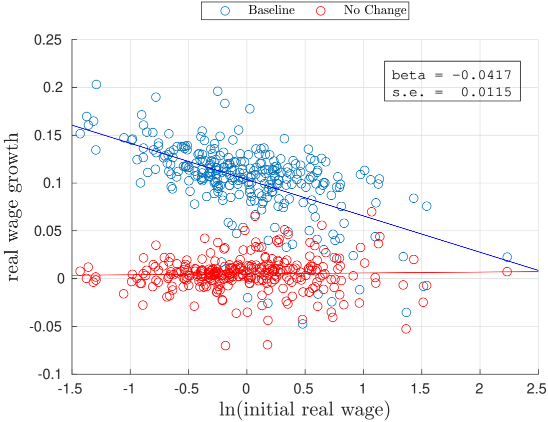


(b) Bottom 10% Productive Locations

Figure 5: The Dynamic Response to Migration Liberalization

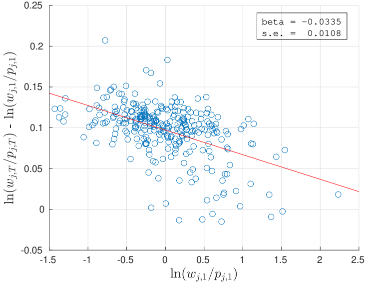
Results: Distributional Impacts

- β-convergence between 1995 and the long-run steady state in 2044.
- Without the expansion of the network, the spatial inequality would have stayed the same over time.

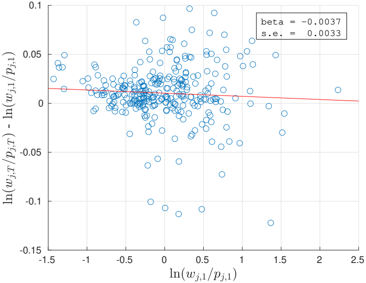


Internal Trade as an Equalizer

- Reduction in τ explains about 80% of the regional convergence.
- Changes λ is responsible for about 9%, and statistically insignificant.



(a) τ only



(b) λ only

Figure 6: The distributional impacts of trade and migration liberalization

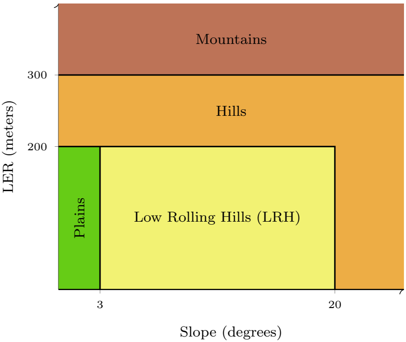
Conclusion

- We create the first panel dataset on the transportation networks that accounts for quality differences in China between 1994 and 2017.
- The expansion of transportation networks significantly lowered the average travel time, especially for initially remote locations.
- Evaluated through a dynamic GE framework, we show that:
  - Changes in transportation network increase the long run output by about 60%.
  - The dynamic return to trade and migration are different. Trade returns realize in the short-run, and migration returns in the long run.
  - The network expansion significantly reduces spatial inequality. Without the network expansion, spatial inequality would have stayed the same.
  - Internal trade liberalization explains around 80% of the reduction in spatial inequality.

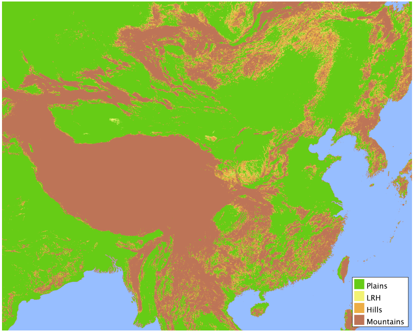
Outline

Terrain of China

Back



(a) Definition of Terrain



(b) Terrain of China

Figure 7: Definition of Terrains, and the Terrain of China

Railway Design Codes

Back

Codes:	*Code for Design of Railway Line*				CDRL (III,IV)[1]	CDSRLIF[2]
Revision:	1985	1999	2006	2017	2012	1987
Doc.Number:	GBJ90-85	GB50090-99	GB50090-2006	TB10098-2017	GB50012-2012	GBJ12-87
National I:	■	■	■	■	□	□
National II:	■	■	■	■	□	□
National III:	■	■	□	□	■	□
National IV:	□	□	□	□	■	□
Local I:	□	□	□	□	■	□
Local II:	□	□	□	□	■	□
Local III:	□	□	□	□	■	□
Industrial I:	□	□	□	□	□	■
Industrial II:	□	□	□	□	□	■
Industrial III:	□	□	□	□	□	■

Table 4: The Mapping Between Railroad Rates and the Codes of Railroad Design

[1] *Code for Design of III and IV Rated Railway Line.*

[2] *Code for Design of Standard Railway Line for Industrial Firms.*

List of Maps

Back

Year	National Maps Publisher	Scale	Projection	Provincial Map Publisher
1994	Sino Maps	1:6 million	Albers, 25N, 47E	Sino Maps
1995	N/A			Sino Maps
1996	Sino Maps	1:4.5 million	Albers, 25N, 47E	Global Maps
1997	N/A			Global Maps
1998	N/A			Xi'an Maps
1999	N/A			Xi'an Maps
2000	Sino Maps	1:6 million	Albers	Xi'an Maps
2001	N/A			Xi'an Maps
2002	Sino Maps	1:4.5 million	Albers, 25N, 47E	Xi'an Maps
2003	Sino Maps	1:6 million	Albers, 25N, 47E	Xi'an Maps
2004	N/A			Xi'an Maps
2005	N/A			Xi'an Maps
2006	N/A			Hunan Maps
2007	Guangdong Maps	1:6 million	Lambert, 24N, 46N, 110E	Dizhi
2008	N/A			Dizhi
2009	Sino Maps	1:4.5 million	Albers, 25N, 47E	Renmin Jiaotong
2010	N/A			Dizhi
2011	N/A			Dizhi
2012	Sino Maps	1:4.5 million	Albers, 25N, 47E	Dizhi
2013	Sino Maps	1:4.6 million	Albers, 25N, 47E	Renmin Jiaotong
2014	N/A			Sino Maps
2015	N/A			Sino Maps
2016	N/A			Sino Maps
2017	Sino Maps	1:6 million	Albers	Sino Maps

Table 5: List of References

Example of New Construction

Back

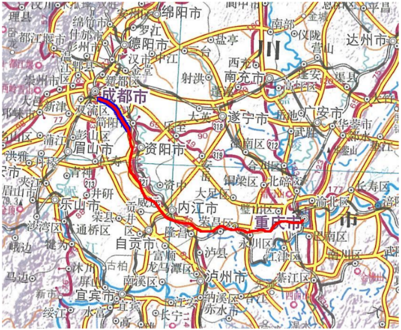


Figure B.2: Identification of New Construction

Notes: this is an example of a newly constructed highway. In the figure, the thin red line is a segment in $t_2 = 1996$, and the thick blue line is the nearest segment in the previous nodal year, $t_1 = 1994$. Based on this and the records in the yearbooks, we determine that the differences between the red and the blue lines form a new segment that was constructed between 1994 and 1996. The highway in question is the Chengyu Highway which connects Chengdu and Chongqing. The blue segment is the phase-one construction that was already finished by 1994, and the newly constructed red line is the phase-two construction that was finished in 1995, as recorded in the *Transportation Yearbook of China, 1996*, page 438.

Alder, Simon, Michael Zheng Song, and Zhitao Zhu, "Unequal Returns to China's Intercity Road Network," 2021. working paper.

Allen, Treb and Costas Arkolakis, "The Welfare Effects of Transportation Infrastructure Improvements," *The Review of Economic Studies*, 02 2022.

— **and Dave Donaldson**, "The geography of path dependence," Technical Report, working paper 2018.

Anderson, James E. and Eric van Wincoop, "Trade Costs," *Journal of Economic Literature*, September 2004, *42* (3), 691–751.

Artuc, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *The American Economic Review*, 2010, *100* (3), 1008–1045.

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.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, pp. 741–835.

Donaldson, Dave and Richard Hornbeck, " Railroads and American Economic Growth: A "Market Access" Approach *," *The Quarterly Journal of Economics*, 02 2016, *131* (2), 799–858.

Fan, Jingting, "Internal Geography, Labor Mobility, and the Distributional Impacts of Trade," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–88.

— **, Yi Lu, and Wenlan Luo**, "Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters," *The Review of Economics and Statistics*, 08 2021, pp. 1–46.

Kleinman, Benny, Ernest Liu, and Stephen J Redding, "Dynamic Spatial General Equilibrium," Working Paper 29101, National Bureau of Economic Research July 2021.

Ma, Lin and Yang Tang, "Geography, trade, and internal migration in China," *Journal of Urban Economics*, 2020, *115*, 103181.

Ma, Liqian, *The Chronicle of Railway Construction in China (in Chinese)*, China Railway Press, 1983.

Redding, Stephen J. and Matthew A. Turner, "Chapter 20 - Transportation Costs and the Spatial Organization of Economic Activity," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 of Handbook of Regional and Urban Economics, Elsevier, 2015, pp. 1339 – 1398.

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

Xu, Mingzhi, "Riding on the new silk road: Quantifying the welfare gains from high-speed railways," 2017.