

Vital Links, Divided Gains: How China's Transportation Boom Impacted Healthcare

Jing Li¹, Lin Ma¹, Xiao Lin Ong², Wei Yan³, Junjian Yi⁴

DSE @ HKU, December 2025

¹Singapore Management University

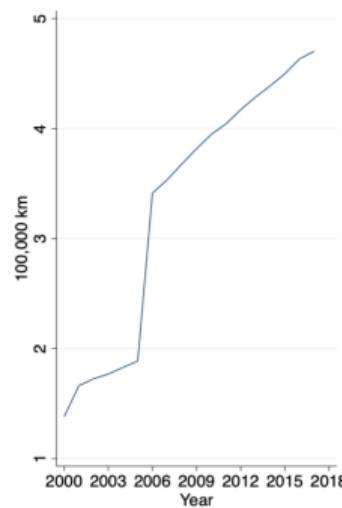
²University of Rochester

³Renmin University of China

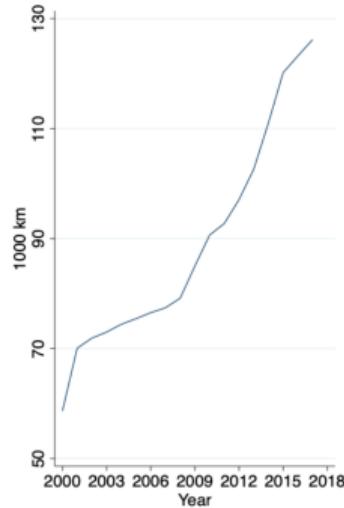
⁴Peking University

Motivation

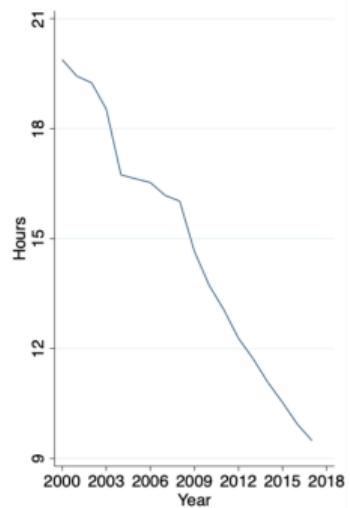
Better transportation networks improve the mobility of goods and labor and reshape the distribution of economic activities.



(a) Road Network



(b) Railway Network



(c) City-Pair Travel Time

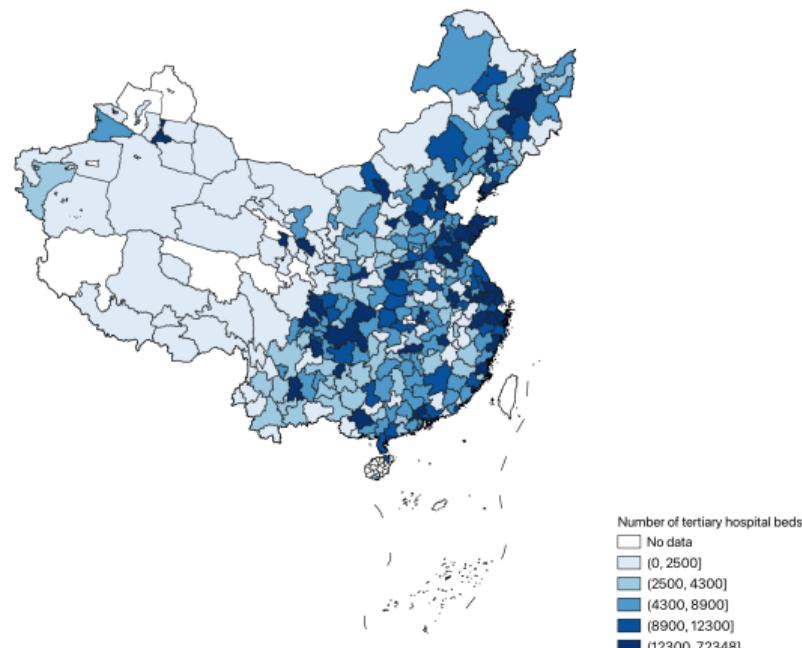
Figure 1: Transport Infrastructure and Travel Time in China (2000-2017)

Motivation

Better roads also facilitate cross-regional utilization of medical resources.

[Background](#)

Question: How did China's transportation boom impact healthcare?



This study

- We develop and structurally estimate a dynamic spatial model in which individuals choose treatment locations by weighing the expected effectiveness of care against travel and financial costs.
- Counterfactual analysis:
 1. Improvements in the national transport network between 2010 and 2018 reduce mortality due to cerebral-cardiovascular diseases (CCVD) by about 10,000 lives per year.
 2. Transportation improvements narrow geographic disparities in health outcomes.
 3. Health gains accrue disproportionately to high-income patients.

Contributions

- Urban economics literature: focus on the economic impacts of infrastructure through trade and labor mobility (Faber, 2014; Donaldson and Hornbeck, 2016; Donaldson, 2018; Banerjee et al., 2020; Asher and Novosad, 2020)
 - We examine how transportation networks shape healthcare-seeking behaviors and health outcomes (Li, 2013; Dingel et al., 2023).
- Health economics literature: importance of travel time in shaping healthcare demand (Ho, 2006; Ho and Pakes, 2014; Hackmann, 2019; Prager, 2020)
 - Based on a dynamic spatial model of hospital choice, we investigate the distributional consequences of reduced travel time for both healthcare utilization and health outcomes.

Outline

Data

Stylized Facts

Model

Estimation

Quantitative Results

Data

1. **Admission-Level Information:** hospital admission and discharge summary dataset (Sichuan Health Commission, 2017-2018)
 - Administrative hospitalization records for all patients with CCVDs: demographic characteristics (city of residence and demographics), medical details (hospital, severity, procedure, and mode of discharge in the form of recovery or death)
2. **Hospital-Level Information:** hospital annual report (Sichuan Health Commission, 2018)
 - Hospital name, tier, locations, and the amount of medical resources (number of doctors, nurses, and inpatient beds)
3. **City-Level Information:**
 - Number of tertiary hospital beds, number of admissions with CCVD, population
4. **City-Pair-Level Information:**
 - Travel time (Ma and Tang, 2024)

	Mean	SD
Panel A. Admission-level variables		
Female	0.457	0.498
Monthly income	8,353.027	6,450.316
Severe	0.387	0.487
Recovery	0.926	0.261
Mortality	0.065	0.247
Seeking out-of-town treatment	0.044	0.205
Number of admissions	611,575	
Panel B. Hospital-level variables		
<i>Tertiary hospitals</i>		
Number of beds	853.665	676.303
Number of observations	227	
<i>Secondary hospitals</i>		
Number of beds	217.880	186.449
Number of observations	652	
<i>Primary or ungraded institutions</i>		
Number of beds	71.740	90.731
Number of observations	1,656	
Panel C. City-level variables		
Number of tertiary hospital beds (10,000)	0.923	1.060
Number of admissions with CCVD (10,000)	1.586	1.562
Population (10,000)	466.056	291.599
Number of observations	42	
Panel D. City-pair-level variables		
Travel time by road (hours)	4.389	2.932
Travel time by railway (hours)	7.855	6.951
Minimum travel time (hours)	4.272	2.924
Probability of medical travel across cities	0.048	0.197
Number of observations	882	

Outline

Data

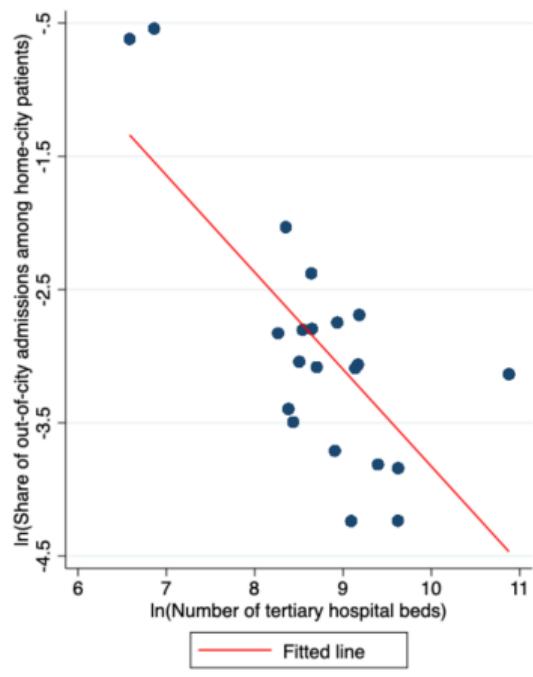
Stylized Facts

Model

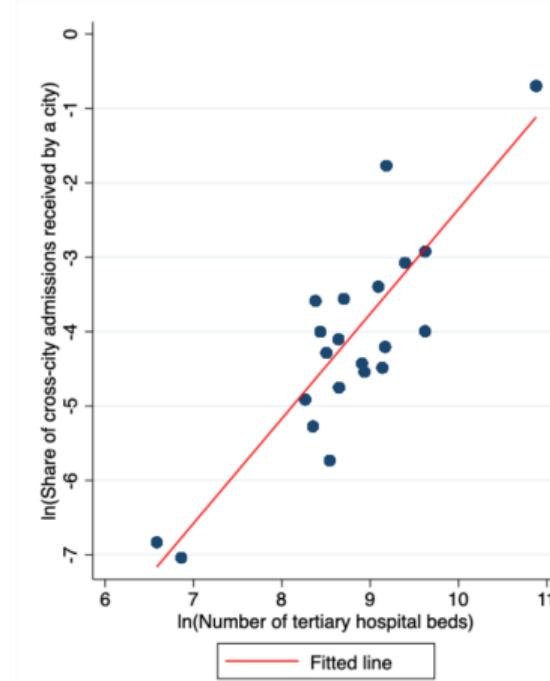
Estimation

Quantitative Results

Medical Travel and Medical Resources



(a) Home city: share of out-of-city admissions and number of tertiary hospital beds



(b) Destination city: share of cross-city admissions and number of tertiary hospital beds

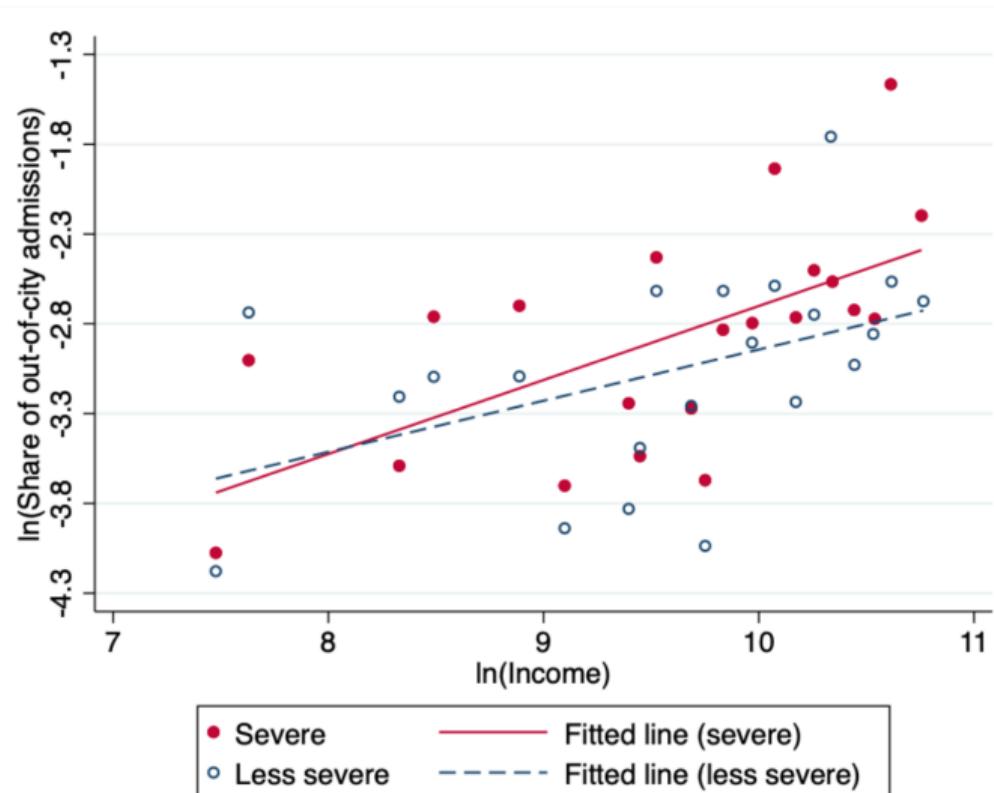
Medical Travel and Travel Time

Table 2: Estimation results for the gravity equation

	ln(Share of admissions outside the home city)					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Minimum travel time)	-2.031*** (0.278)			-2.362*** (0.334)		
ln(Travel time by rail)		-1.919*** (0.290)			-3.282*** (0.727)	
ln(Travel time by road)			-2.050*** (0.273)			-2.529*** (0.303)
Origin-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	653	653	653	653	653	653
R-squared	0.752	0.715	0.743	0.366	0.140	0.334

Notes: Standard errors clustered at the origin-by-destination-city level are reported in parentheses. * p < 0.1,
** p < 0.05, *** p < 0.01.

Medical Travel, Disease Severity, and Income



Outline

Data

Stylized Facts

Model

Estimation

Quantitative Results

Setup

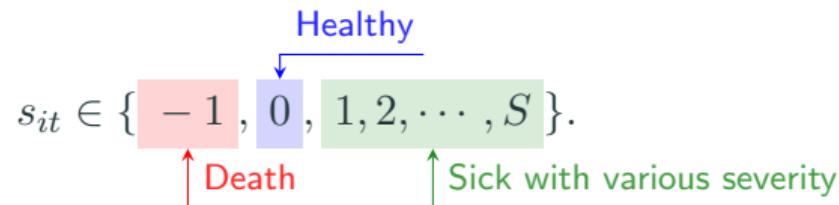
Discrete time $t = 0, 1, \dots, \infty$ and city $k = 1, 2, \dots, K$.

- Constant population size P_k
- All individuals reside permanently in their home city. Upon falling ill, they may seek hospital care either in their home city or in another city.

Setup: Two state variables for individuals

For individual i at time t ,

1. Health status:



- Health status evolves stochastically over time, with probability $\pi_{kt}^{s,\bar{s}}$ of transitioning from s in time t to \bar{s} in time $t + 1$.
- Probability of being sick in city k with severity s : $\pi_{kt}^{0,s}$ ($s > 0$).
- When receiving care in city k , mortality rate is $\pi_{kt}^{s,-1}$ and recovery rate is $\pi_{kt}^{s,0}$ ($\pi_{kt}^{s',-1} \geq \pi_{kt}^{s,-1}$ and $\pi_{kt}^{s',0} \leq \pi_{kt}^{s,0}$, $\forall s' > s > 0$).

Setup: Two state variables for individuals

2. **Productivity** z_{it} follows an AR(1) process:

$$\log z_{it} = \rho_z \log z_{i,t-1} + v_t, \quad v_t \sim \mathcal{N}(0, \sigma_v),$$

- Sickness reduces productivity by a factor of $\delta^s \in (0, 1]$; $\delta^{s'} \leq \delta^s$, $\forall s' > s > 0$.
- Income: $\omega_{ikt} = w_{kt}z_{it}$ (healthy); $\omega_{ikt}^s = \delta^s w_{kt}z_{it}$ (sick).
(w_{kt} : wage rate for one efficient unit of labor productivity in city k and time t)

Healthy Individual

Healthy individuals earn income, consume it fully in t , and then transition to either another healthy state ($s = 0$) or illness ($s > 0$) in $t + 1$.

Value function of individual i in city k at time t :

$$v_{kt}^0(z_{it}) = u(w_{kt}z_{it}) + \beta \mathbf{E} \left[\left(1 - \sum_{s=1}^S \pi_{kt}^{0,s} \right) v_{k,t+1}^0(z_{t+1}) + \sum_{s=1}^S \pi_{kt}^{0,s} v_{k,t+1}^s(z_{t+1}) \right].$$

- The expectation is taken over productivity shocks;
- $v_{k,t+1}^s(z_{i,t+1})$: value function of individual with health status s in time $t + 1$;
- The value of a deceased individual is normalized to zero: $v_{k,t+1}^{-1} = 0$.

Sick Individual

When an individual residing in city k becomes sick (i.e., $s > 0$), she chooses a destination (treatment) city l , by solving the recursive problem:

$$v_{kt}^s(z_{it}) = \max_{l \in \mathbb{F}_{kt}^s(z_{it})} \left\{ u \left(\delta^s w_{kt} z_{it} - \mathbb{1}(l \neq k) \cdot \lambda \right) - \tau_{klt} \right.$$

fixed out-of-city costs

disutility from travel costs

choice set

$$\left. + E \left[\beta \left[\pi_{lt}^{s,0} v_{k,t+1}^0(z_{i,t+1}) + \left(1 - \pi_{lt}^{s,0} - \pi_{lt}^{s,-1} \right) v_{k,t+1}^s(z_{i,t+1}) \right] + \kappa \cdot \varepsilon_l - \bar{\varepsilon}_{kt}^s(z_{it}) \right] \right\}.$$

cardinality correction

preference shock
(GEV-I)

treatment outcome at l

- $\pi_{lt}^{s,0}$ and $\pi_{lt}^{s,-1}$ are functions of medical resources (m_{lt}) and patient volume (p_{lt}).
- κ captures individuals' sensitivity to observable travel costs:
(higher κ , less sensitive to observable costs, more affected by ε_l)

Choice Set

- Individual-specific choice set
$$\mathbb{F}_{kt}^s(z_{it}) = \{l \in K | m_{lt} \geq m_{kt} \cap (\delta^s w_{kt} z_{it} - \mathbb{1}(l \neq k) \cdot \lambda) > 0\}$$
 1. Individuals only go to cities with better medical resources;
 2. Medical travel has to be feasible.
- Upside: model income-induced selection bias and match zeros in the data.
- Downside: value function differs mechanically due to the size of the choice set.
 - In discrete choice, individuals value options based on objective value and taste shock. With a large enough K , people might prefer being sick.
 - The last term in the value function, $\bar{\varepsilon}_{kt}^s(z_{it}) = \kappa \cdot \log \bar{\mathcal{F}}_{kt}^s(z_{it})$, corrects this, where $\bar{\mathcal{F}}_{kt}^s(z_{it})$ is the cardinality of the set $\mathbb{F}_{kt}^s(z_{it})$.

Timeline of Sick Individual's Decision

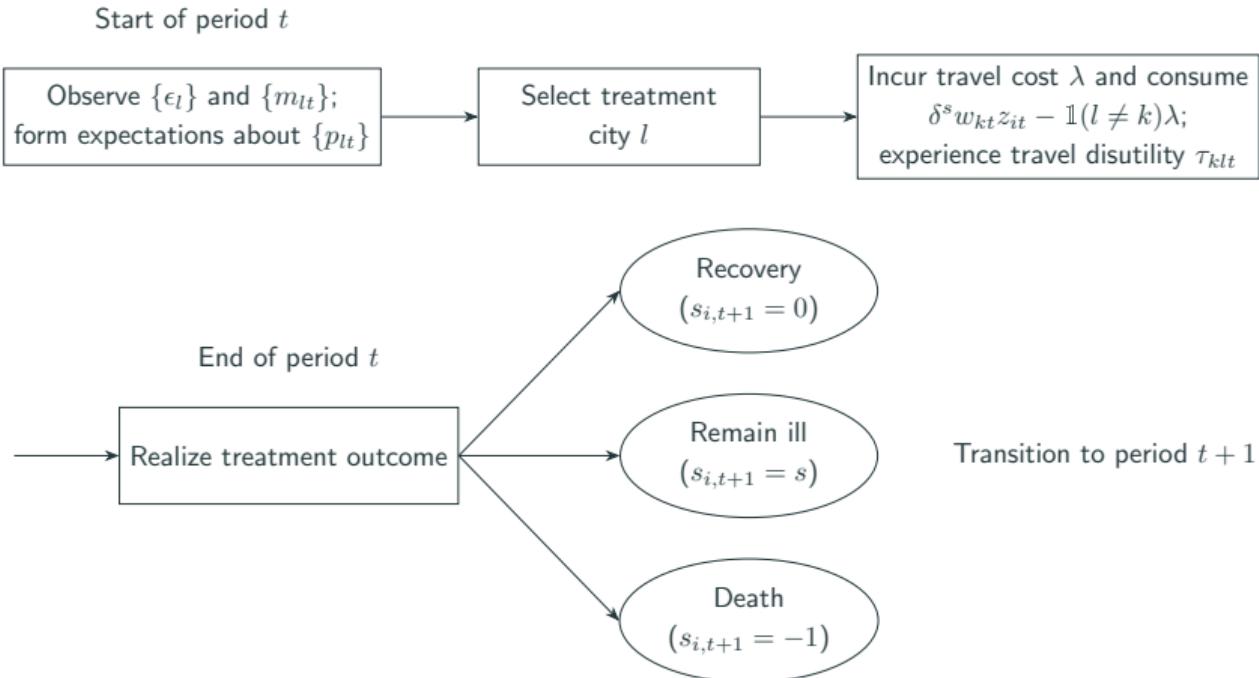


Figure 1: Timeline of Sick Individual's Decision and Health Transition at Time t

Choice Probabilities

Define the expected value of choosing treatment city l :

$$\begin{aligned} W_{kl,t+1}^s(z_{t+1}) &= \mathbf{E} \left[\pi_{lt}^{s,0} v_{k,t+1}^0(z_{t+1}) + \left(1 - \pi_{lt}^{s,0} - \pi_{lt}^{s,-1}\right) v_{k,t+1}^s(z_{t+1}) \right] \\ &= \int_0^\infty \mathbf{E}_{\epsilon'} \left[\pi_{lt}^{s,0} v_{k,t+1}^0(z_{t+1}) + \left(1 - \pi_{lt}^{s,0} - \pi_{lt}^{s,-1}\right) v_{k,t+1}^s(z_{t+1}) \right] dG(z_{t+1}|z_t) \end{aligned}$$

Rewrite the value function:

$$v_{kt}^s(z_t) = \max_{l \in \mathbb{F}_{kt}^s(z_t)} \left[u(\delta^s w_{kt} z_t - \lambda_{klt}) - \tau_{klt} + \kappa \epsilon_l + \beta W_{kl,t+1}^s(z_t) \right] - \bar{\epsilon}_{kt}^s(z_t)$$

The probability of patient i at location k seeking treatment at location l :

$$\mu_{klt}^s(z_t) = \begin{cases} \frac{\exp[u(\delta^s w_{kt} z_t - \mathbb{1}(l \neq k) \cdot \lambda) - \tau_{klt} + \beta W_{kl,t+1}^s(z_t)]}{\sum_{l' \in \mathbb{F}_{kt}^s(z_t)} \exp[u(\delta^s w_{kt} z_t - \mathbb{1}(l' \neq k) \cdot \lambda) - \tau_{kl't} + \beta W_{kl',t+1}^s(z_t)]}^{1/\kappa} & \text{if } l \in \mathbb{F}_{kt}^s(z_t) \\ 0 & \text{otherwise} \end{cases}$$

Aggregation and Equilibrium

The law of motion of healthy population:

$$L_{kt}^0(z_t) = \left(1 - \sum_{s=1}^S \pi_{kt}^{0,s}\right) \int_0^\infty L_{k,t-1}^0(z_{t-1}) g(z_t | z_{t-1}) dG(z_{t-1})$$

remain healthy

$$+ \int_0^\infty \sum_{s=1}^S \sum_{l=1}^K \pi_{l,t-1}^{s,0} \mu_{kl,t-1}^s(z_{t-1}) L_{k,t-1}^s(z_{t-1}) g(z_t | z_{t-1}) dG(z_{t-1})$$
$$+ \int_0^\infty \sum_{s=1}^S \sum_{l=1}^K \pi_{l,t-1}^{s,-1} \mu_{kl,t-1}^s(z_{t-1}) L_{k,t-1}^s(z_{t-1}) g(z_t | z_{t-1}) dG(z_{t-1})$$

recovery

replacement

Aggregation and Equilibrium

The law of motion of ill population:

$$L_{kt}^s(z_t) = \int_0^\infty \pi_{kt}^{0,s} L_{k,t-1}^0(z_{t-1}) g(z_t | z_{t-1}) dG(z_{t-1}) +$$

$$\int_0^\infty \sum_{l=1}^K \mu_{kl,t-1}^s \left(1 - \pi_{l,t-1}^{s,-1}(z_{t-1}) - \pi_{l,t-1}^{s,0}(z_{t-1}) \right) L_{k,t-1}^s(z_{t-1}) g(z_t | z_{t-1}) dG(z_{t-1})$$

↑ remaining stock

The number of patients seeking treatment in location l is:

$$p_{lt} = \int_0^\infty \sum_{s=1}^S \sum_{k=1}^K \mu_{klt}^s(z_t) L_{kt}^s(z_t) dG(z_t).$$

Equilibrium

Sequential Equilibrium: Given a sequence of location fundamentals, $\{w_{kt}, m_{kt}\}$, population distribution $\{L_k\}$, and travel costs $\{\tau_{klt}\}$, a sequential equilibrium is the vectors of value functions $\{v_{kt}^s\}$ and population distribution $\{L_{kt}^s\}$ such that:

1. Individuals maximize their expected utility by choosing locations for medical treatments, subject to distribution of patients, $\{p_{lt}\}_{l=1}^K$.
2. Rational expectation holds so that the law of motion and the patient distribution described above are consistent with individual choices.

Steady State: With constant location fundamentals and travel costs, $\{L_k, w_k, m_k, \tau_{kl}\}$, the steady state of the economy consists of constant value functions and population distribution $\{v_k^s, L_k^s\}$ that satisfy the definition of a sequential equilibrium.

Outline

Data

Stylized Facts

Model

Estimation

Quantitative Results

Functional Form Assumptions: Treatment Outcome

- Assume a multinomial structure to map the outcome probabilities to resources and patient volume:

$$\pi_{lt}^{s,0} = \frac{\exp(f^{s,0}(\cdot))}{1 + \exp(f^{s,0}(\cdot)) + \exp(f^{s,-1}(\cdot))}, \text{ and } \pi_{lt}^{s,-1} = \frac{\exp(f^{s,-1}(\cdot))}{1 + \exp(f^{s,0}(\cdot)) + \exp(f^{s,-1}(\cdot))}$$

where

$$f^{s,0}(m_{lt}, p_{lt}) = \gamma_1^{sH} + \gamma_2^{sH} \log(m_{lt}) + \gamma_3^{sH} \log(p_{lt}),$$

$$f^{s,-1}(m_{lt}, p_{lt}) = \gamma_1^{sD} + \gamma_2^{sD} \log(m_{lt}) + \gamma_3^{sD} \log(p_{lt}).$$

- 6S parameters ($\{\gamma(\cdot)\}$) to estimate.
- $\gamma_2^{sH} > 0$ and $\gamma_2^{sD} < 0 \implies$ resources improve outcome.
- $\gamma_3^{s(\cdot)} \leq 0 \implies$ congestion or agglomeration.

Functional Form Assumptions

- Assume that the travel cost τ_{klt} as:

$$\tau_{klt} = \begin{cases} \exp(\eta_0 + \eta_1 D_{klt}), & \text{if } l \neq k \\ 0 & \text{if } l = k \end{cases}$$

- D_{klt} is observed travel time. Two parameters (η_0 and η_1) to estimate.
- Utility function, CRRA, with $\psi = 2$:

$$u(c) = \frac{c^{1-\psi} - 1}{1 - \psi}.$$

Step 1: Calibration

The first group of parameters are calibrated outside of the model.

- Time frequency: 30 days.
- $\beta = 0.98^{1/12} = 0.9983$: assume an annual discount rate of 2%.
- $\lambda = 0.0005$: match the average cardinality of choice sets in the data ($\sum_{l=1}^K \sum_{k=1}^K \mathbb{1}(p_{kl} > 0)/K$).
- $\delta^s = 0.67$: share of working days in the hospital for patients.
- $\rho_z = 0.998$ (Fan et al (2010)).
- $\sigma_\nu = \sqrt{(1 - \rho_z^2)} \times sd(\log z) = \sqrt{(1 - \rho_z^2)} \times 1.6272$ (dispersion of income in data).
- w_k : observed wage rates.
- $S = 2$: general and severe.
- $\pi^{0,s}$: $\pi^{0,1} = 0.00354/24$, and $\pi^{0,2} = 0.00223/24$
- m_k : the number of tier-3 hospital beds in a prefecture.

Step 2: Indirect Inference

Remaining parameters:

1. (Inverse) distance elasticity, κ
2. Travel cost function, $\{\eta_0, \eta_1\}$
3. The treatment outcome functions $\left\{ \gamma_{(\cdot)}^{s,0}, \gamma_{(\cdot)}^{s,-1} \right\}_{s=1}^S$

We estimate these parameters using indirect inference.

Indirect Inference: Auxiliary Regressions

$$\log \mu_{klt}^s(z_t) = \frac{\beta}{\kappa} W_{kl,t+1}^s(z_t) + \underbrace{\frac{1}{\kappa} u(\delta^s w_{kt} z_t - \mathbb{1}(l \neq k) \cdot \lambda)}_{\text{extensive margin}} - \underbrace{\frac{1}{\kappa} \exp(\eta_0 + \eta_1 D_{klt})}_{\text{intensive margin}} - \Phi_{kt}^s(z_t).$$

1. Correlation between travel probability (μ_{klt}) and individual income conditional on disease severity help identify κ :

$$\mathbb{1}(l \neq k)_i = \alpha_1^{\text{out}} + \alpha_2^{\text{out}} \log(\text{Income}_{ik}) + \alpha_3^{\text{out}} s_{ik} + \alpha_4^{\text{out}} \log(\text{Income}_{ik}) \times s_{ik} + \nu_{ik}$$

2. Correlation between μ_{klt} and D_{klt} help identify η_0 and η_1 conditional on κ (gravity equation):

$$\log \mu_{kl} = \alpha_1^{\text{pair}} + \alpha_2^{\text{pair}} \log(\text{Travel Time}_{kl}) + \alpha_3^{\text{pair}} \frac{m_l}{m_k} + \alpha_4^{\text{pair}} \frac{p_l}{p_k} + \text{FE}_l + \text{FE}_k + \nu_{kl}.$$

Indirect Inference: Auxiliary Regressions

3. The auxiliary regressions below help identify treatment-related parameters:

$$\left\{ \gamma_{(\cdot)}^{sH}, \gamma_{(\cdot)}^{sD} \right\}_{s=1}^S$$

$$\mathbb{1}(s' = 0 | s, l)_i = \alpha_1^{sH} + \alpha_2^{sH} \log(m_l) + \alpha_3^{sH} \log(p_l) + \nu_{il}^s$$

$$\mathbb{1}(s' = -1 | s, l)_i = \alpha_1^{sD} + \alpha_2^{sD} \log(m_l) + \alpha_3^{sD} \log(p_l) + \nu_{il}^s.$$

20 moment conditions $\mathbf{A} = \{\alpha_{(\cdot)}^{\text{out}}, \alpha_{(\cdot)}^{\text{pair}}, \alpha_{(\cdot)}^{sH}, \alpha_{(\cdot)}^{sD}\}$, 15 parameters
 $\Theta = \{\kappa, \eta_0, \eta_1, \gamma_{(\cdot)}^{sH}, \gamma_{(\cdot)}^{sD}\}.$

The indirect inference estimator solves the following minimization problem:

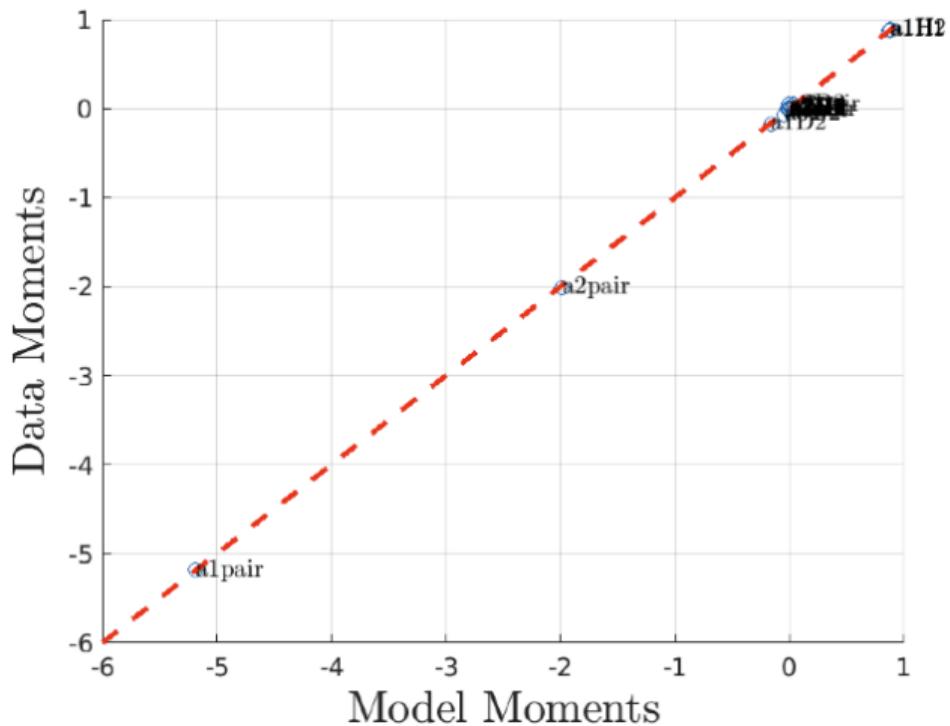
$$\min_{\Theta} \left[\mathbf{A} - \frac{1}{R} \sum_{r=1}^R \tilde{\mathbf{A}}_r(\Theta) \right]' \mathbf{W} \left[\mathbf{A} - \frac{1}{R} \sum_{r=1}^R \tilde{\mathbf{A}}_r(\Theta) \right],$$

where \mathbf{W} is the diagonal weighting matrix in which the i th diagonal equals the inverse of the squared standard errors of the i th elements in \mathbf{A} .

Estimation Results

name	value	s.e.	note
κ	0.249***	0.030	inverse travel elasticity
η_0	0.062	0.052	travel cost function, intercept
η_1	0.144***	0.031	travel cost function, slope
γ_1^{1H}	1.735***	0.063	recovery rate, intercept, $s = 1$
γ_2^{1H}	0.133***	0.022	recovery rate, slope on m_l , $s = 1$
γ_3^{1H}	-0.288***	0.050	recovery rate, slope on p_l , $s = 1$
γ_1^{2H}	0.958***	0.071	recovery rate, intercept, $s = 2$
γ_2^{2H}	0.102***	0.027	recovery rate, slope on m_l , $s = 2$
γ_3^{2H}	-0.314***	0.076	recovery rate, slope on p_l , $s = 2$
γ_1^{1D}	-3.088**	1.543	mortality rate, intercept, $s = 1$
γ_2^{1D}	-2.488***	0.868	mortality rate, slope on m_l , $s = 1$
γ_3^{1D}	0.749	0.942	mortality rate, slope on p_l , $s = 1$
γ_1^{2D}	-4.113***	0.374	mortality rate, intercept, $s = 2$
γ_2^{2D}	-1.044*	0.612	mortality rate, slope on m_l , $s = 2$
γ_3^{2D}	-1.102***	0.150	mortality rate, slope on p_l , $s = 2$

Model Fit



Outline

Data

Stylized Facts

Model

Estimation

Quantitative Results

The Impacts of Transportation Networks

Counterfactual analysis: all prefectures, transportation networks (2010 → 2018)

Require new data inputs:

1. Prefecture-level population (Chinese Population Census)
2. Prefecture-level average income (China City Statistical Yearbooks)
3. Prefecture-level medical resources (China health Commission)
4. Transportation networks between prefectures (Ma and Tang (2024))

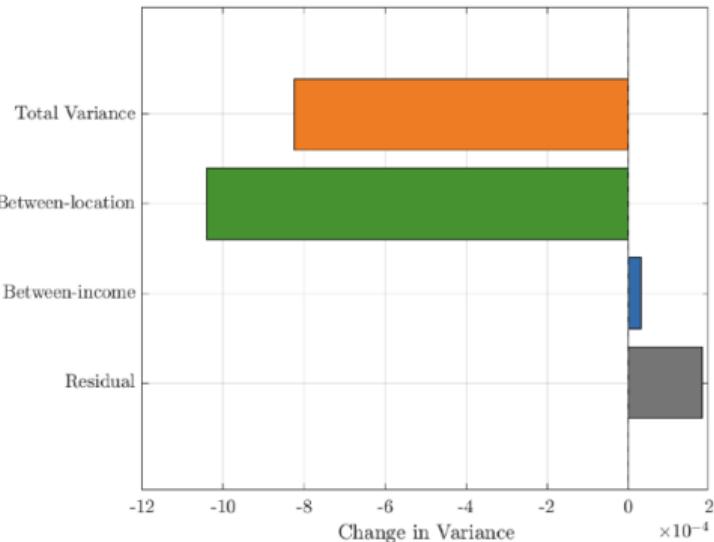
Two sets of analyses:

- 2010 base: all other variables fixed at 2010, transportation (2010 → 2018)
- 2018 base: all other variables fixed at 2018, transportation (2018 → 2010)

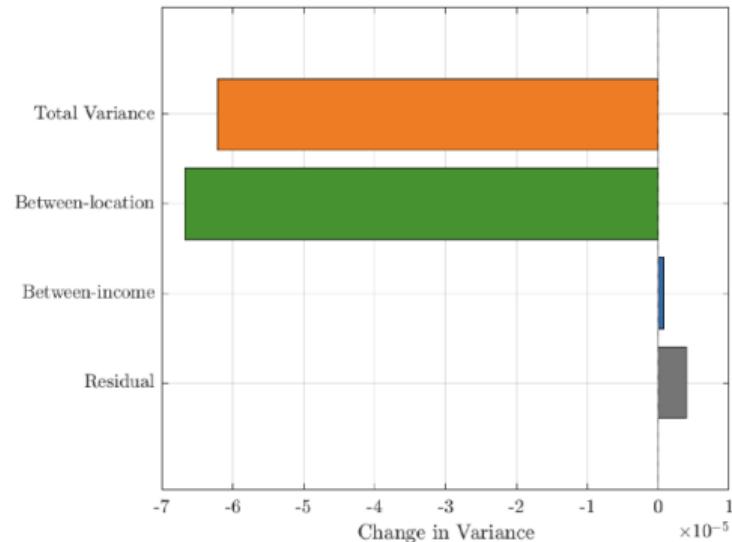
Aggregate Impacts of Transportation on Mortality

	%Δ Mortality (1)	Δ Mortality (thousands)				Δ VSL (billion ¥) (6)
		Total (2)	Induced-traveler (3)	Never-traveler (4)	Always-traveler (5)	
Reduce τ , 2010 base	-3.08	-9.88	-9.01	-0.85	-0.01	47.01
Reduce τ , 2018 base	-2.01	-1.82	-2.17	0.35	0.00	8.65

Changes in Variance of Expected Mortality Rates

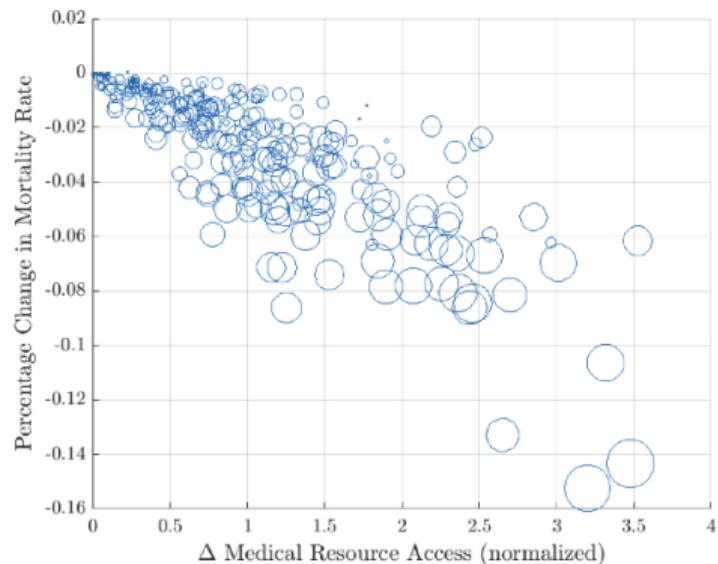


(a) 2010 Base

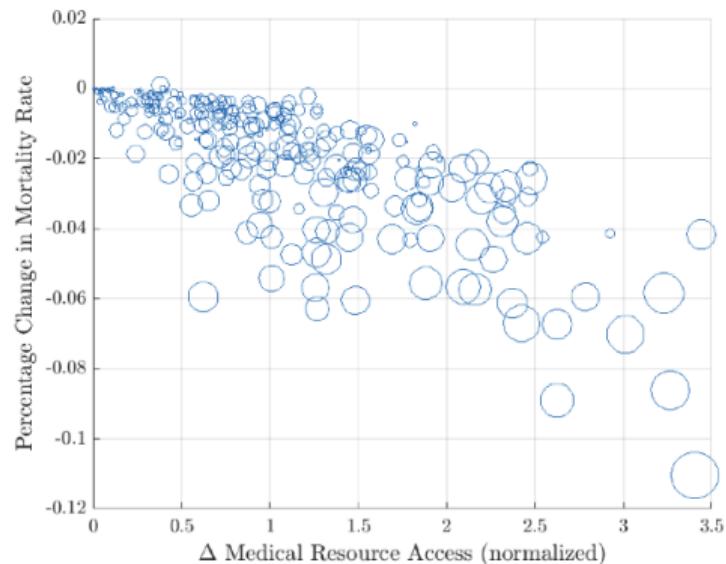


(b) 2018 Base

Impacts of Transportation Networks on Mortality Rates Across Cities

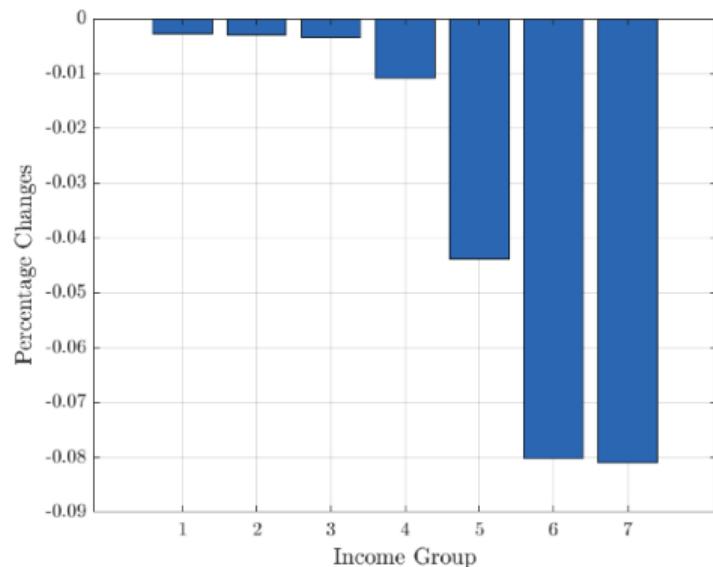


(a) 2010 Base

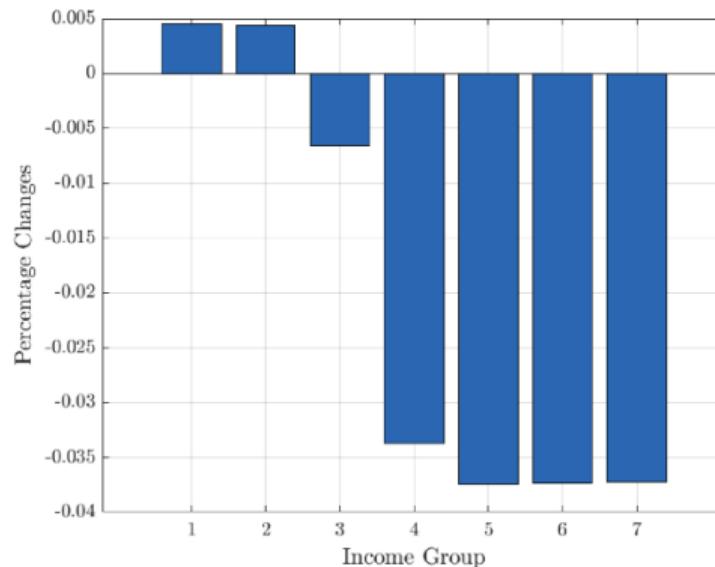


(b) 2018 Base

Impacts of Transportation Networks on Mortality Rates Across Income Groups



(a) 2010 Base



(b) 2018 Base

Conclusions and Implications

- Improvements in transportation infrastructure reduces mortality, primarily by enabling faster access to high-quality medical facilities from underserved areas.
 - Better transport reduces spatial inequality in access to care, but the benefits are distributed unevenly across income groups
-
- Implication: without financial mechanisms to alleviate the costs of traveling for care, improvements in connectivity may reinforce existing socioeconomic inequalities.

Backup Slides

Background: Inequality of Healthcare Resources in China

- Central planning period (1949–1978):
 - Urban: public health insurance schemes secured high-quality, heavily subsidized healthcare resources for urban populations.
 - Rural: rural commune system focused on promoting basic and preventive healthcare through various public health campaigns; bare-foot doctors and local clinics
- Economic reforms since 1978 exacerbated healthcare inequality across regions.
 - With fiscal decentralization, local governments became increasingly responsible for funding healthcare.
- Since the 2009 healthcare reform, China has expanded public insurance coverage and increased investment in healthcare infrastructure. However, regional disparities persist due to uneven implementation.

Back