Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis

Simon Alder\textsuperscript{1}  Illenin Kondo\textsuperscript{2}

\textsuperscript{1}University of North Carolina at Chapel Hill
\textsuperscript{2}University of Notre Dame

Annual Bank Conference on Development Economics
World Bank
June 25-26, 2018
Outline

Introduction

Literature

Data

Model and Network Design

Analysis

Robustness

Conclusion
Motivation

- China substantially expanded and modernized its national highway network starting in the 1990s
  - Approximately 35,000 km of new high speed four-lane highways
  - Massive and rapid buildup in the late 1990s and early 2000s
  - Estimated cost of US$ 120 billion until 2007 (Faber, 2014)

- Clearly specified policy
  - Connect cities with at least 500,000 people + provincial capitals
  - Network nodes are specified, but not the individual connections
    - Which cities are directly linked
    - How the paths are chosen

- How was the policy implemented?
  - Was the road network implemented efficiently?
  - What can explain potential distortions?
China’s National Expressway Network

The orange lines show the NTHS that connects the targeted cities (shown in red). The background shows the slope of the terrain of mainland China as a proxy for road construction costs.
This Paper

We focus on political distortions in infrastructure allocation

Questions

- Are the birthplaces of top Chinese leaders favored in the actual highway network relative to the optimal network benchmark?
- What are the aggregate costs of the political distortions?

Our approach

1. Approximate the income-maximizing transport network in a general equilibrium trade model
2. Combine transport network with CV data on Chinese politicians
3. Estimate effect of politicians’ birthplace on network distortions at county level
4. Evaluate aggregate welfare effects using general equilibrium model
Literature overview (selection)

- China’s National Expressway Network
  - Faber (2014), Lu and Wang (2016), Roberts et al. (2012)

- Transport network design in general equilibrium trade models

- Misallocation of infrastructure, ethnic favoritism, and birthplaces of politicians
## Outline

- Introduction
- Literature
- Data
- Model and Network Design
- Analysis
- Robustness
- Conclusion
Data

▶ Geography
  ▶ Location of targeted cities (500k+ pop, provincial capitals)
  ▶ Slope of the terrain

▶ Transport infrastructure
  ▶ Maps of the actual highway network + other existing infrastructure

▶ Income
  ▶ Income of each city proxied by luminosity within a 30km radius

▶ Politicians
  ▶ CVs of Chinese politicians (China Vitae)
  ▶ Match birthplaces to counties in the census
  ▶ Code county as ‘birthplace’ if a politician i) was born in that county, and ii) was in a political office during the planning phase
    ▶ Politburo, Provincial Governors, expanding
    ▶ Planning and implementation phase 1995–2001 (14th & 15th CCP)
    ▶ Placebo if in office after 2008 (17th CCP)
Model

Setup

- Inter-regional trade model based on Eaton and Kortum (2002)
- Donaldson and Hornbeck (2016) derive ‘Market Access’ as a measure for the effect of transport infrastructure
- Trade from location $i$ to location $j$ is

$$X_{ij} = T_i (q_i^\alpha w_i^\gamma)^{-\theta} \times Y_j \times \tau_{ij}^{-\theta} \times \kappa_1 MA_j^{-1}$$

- Origin’s productivity and factor costs
- Destination’s income
- Bilateral trade costs
- Destination’s market access

- Summing the gravity equation over destinations $j$ yields total expenditure (income) of origin $i$:

$$Y_i = \sum_{j} X_{ij} = \kappa_1 T_i (q_i^\alpha w_i^\gamma)^{-\theta} \sum_{j} \left[ Y_j \tau_{ij}^{-\theta} MA_j^{-1} \right]$$

- Market Access of $i$
Model

Real Income

- Rewriting the above income gravity equation using the price index (market access) and factor shares yields

\[ Y_i' = \left( \kappa_2 T_i \right)^{\frac{1}{1+\theta(\alpha+\gamma)}} \left( \frac{\alpha}{L_i} \right)^{-\frac{\theta \alpha}{1+\theta(\alpha+\gamma)}} \left( \frac{\gamma}{H_i} \right)^{-\frac{\theta \gamma}{1+\theta(\alpha+\gamma)}} MA_i^{\frac{1+\theta(1+\alpha+\gamma)}{1+\theta(\alpha+\gamma)} \theta} \]  

\[ MA_i = \rho^{\frac{1+\theta}{\theta}} \sum_j \tau_{ij}^{-\theta} MA_j^{\frac{-(1+\theta)}{\theta}} Y_j'. \]

- \( \kappa_2 \) contains the interest rate, which is equalized across locations.

(\( \rightsquigarrow \) Details)
Network Design Algorithm

- Heuristic iterative algorithm based on Alder (2017) to approximate the (net) income-maximizing network
- The net effect of adding a link $l$ is

$$W_l = \Delta Y_l - (r + m)\lambda C_l,$$

where $r$ is the rental rate of capital, $m$ is the annual maintenance cost, $\lambda C_l$ are the road construction costs

⇒ Gastner (2005) and Gastner and Newman (2006) combined with general equilibrium model

1. Start from the fully connected network
2. Remove each link individually and compute the corresponding trade costs and income net of road construction costs
3. Remove links with lowest (most negative) net effect
4. Check if there are links that would be beneficial to add
5. Back to step 2 until no further improvements are possible to connect all cities
Network design algorithm
Network design algorithm
Network design algorithm
The maps shows the approximation of the optimal highway network in China based on the heuristic algorithm with the constraint that all targeted cities are connected. The background shows the slope of the terrain of mainland China. The nodes show the location of all cities with a population of at least 500,000 and all provincial capitals.
The map shows the approximation of the optimal highway network in China (thicker orange line) together with the actual NTHS (thinner black line). The background shows the slope of the terrain of mainland China. The nodes show the location of all cities with a population of at least 500,000 and all provincial capitals.
Optimal and Actual Chinese Highway Network

- Optimal network has similar structure as actual
  
  ⇒ but smaller (costs 86% of NTHS)

- Replacing the actual network with the approximation of the optimal network would imply a 1.4% higher aggregate net income annually

- Next, we will test whether the deviations between the two networks can be explained by the birthplaces of politicians
Outline

Introduction

Literature

Data

Model and Network Design

Analysis

Robustness

Conclusion
Reduced form estimation

- Effect of birthplaces on proximity to actual highway network (relative to optimal benchmark)

\[ \Delta_{opt-act} \text{distance}_i = \alpha + \beta \text{political access}_i + \gamma X_i \]

- Dependent variable: deviation between a county’s distance to the optimal and actual network
- Independent variable: Political access in 1995–2001

\[ \text{Political Access}_i = \sum_{t=t_1}^{t_2} \# \text{politicians born in } i \text{ and ‘in power’ at } t \frac{t_2 - t_1 + 1}{t_2 - t_1 + 1} \]

- Placebo: politicians in office after 2008
- Controls and robustness
  - controls: initial light density, area, distance to optimal network, distance to railroads, distance to ports, distance to trunk network, province FE
  - robustness: minimum spanning tree or optimal network with the same cost as actual network
## Effects of Birthplaces on Roads

<table>
<thead>
<tr>
<th></th>
<th>BCT $\Delta \text{NTHS}_{\text{BCT-act}}^i$ (1)</th>
<th>OPT $\Delta \text{NTHS}_{\text{OPT-act}}^i$ (2)</th>
<th>MST $\Delta \text{NTHS}_{\text{MST-act}}^i$ (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Political Access$_{1995-2001}$</td>
<td>0.539** (0.224)</td>
<td>0.514** (0.246)</td>
<td>0.548** (0.228)</td>
</tr>
<tr>
<td>Political Access$_{2008-2012}$ (placebo)</td>
<td>0.100 (0.183)</td>
<td>0.097 (0.183)</td>
<td>0.094 (0.169)</td>
</tr>
<tr>
<td>Dist Optim (NTHS budget)</td>
<td>0.789*** (0.043)</td>
<td>0.789*** (0.043)</td>
<td></td>
</tr>
<tr>
<td>Dist Optim</td>
<td>0.767*** (0.051)</td>
<td>0.767*** (0.051)</td>
<td></td>
</tr>
<tr>
<td>Dist MST</td>
<td></td>
<td>0.662*** (0.040)</td>
<td>0.662*** (0.040)</td>
</tr>
<tr>
<td>Observations</td>
<td>2175</td>
<td>2175</td>
<td>2175</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.479</td>
<td>0.479</td>
<td>0.467</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in columns 1 and 2 is the difference between each county’s distance to the optimal network (with the same cost as the actual network) and the actual network. The dependent variable in columns 3 and 4 is the difference between each county’s distance to the (unconstrained) optimal network and the actual network. The dependent variable in columns 5 and 6 is the difference between each county’s distance to the minimum spanning tree (MST) and the actual network. The main explanatory variable is an index for politicians’ place of birth. All regressions control for initial light density, county area, and province fixed effects. Standard errors are clustered at the province level.
Quantifying the distortions

- Reduced form regressions suggest that top incumbent politician birthplaces are closer to the actual network relative to the optimal network benchmark.

- Optimal network implies 1.4% higher aggregate net income relative to actual NTHS network.
  - But only part of the deviation explained by birthplaces.

- What are the aggregate costs of the distortion caused by birthplaces?
  - We use the general equilibrium model to compare aggregate welfare of counterfactual transport networks.
Actual vs optimal network and birthplaces

Birthplaces that are predicted to be connected

The orange lines show the optimal network among the 62 officially targeted cities. The black lines show the actual network. The green dots show the cities that are predicted to be connected by the reduced form regression based on politicians’ birthplaces. The background shows the slope of the terrain of mainland China.
Quantifying the distortions

Consider network consisting of 102 potential nodes

- 62 cities that are targeted by the official policy
  - >500k & provincial capitals must always be connected
- 20 cities that are predicted to be connected
  - Because they are birthplaces
- 20 ‘counterfactual cities’ that could be connected
  - Closer to the optimal network and large in terms of light

We then compare welfare from two networks that maximize income under the constraint that

1. 62 targeted cities are connected (optimal network so far)
2. 62 targeted cities + 20 predicted birthplaces are connected

- + any of the 102 may be connected if worthwhile

Result suggests 0.2% lower income annually when forcing to connect birthplaces

- Note that overall, we found 1.4% higher annual income difference between actual and optimal
- Also estimate direct effect on local light growth
Outline

Introduction

Literature

Data

Model and Network Design

Analysis

Robustness

Conclusion
Robustness

We considered the following robustness checks:

1. Including exports to international markets in the network design
2. Different parameter value that determines the elasticity of income with respect to market access
3. Dropping the constraint that all 62 cities have to be connected
4. Imposing the constraint that the optimal network has the same cost as the actual network (see regression table)
5. Using minimum spanning tree (least-cost) network as benchmark (see regression table)

(⇒ Details)
Outline

Introduction

Literature

Data

Model and Network Design

Analysis

Robustness

Conclusion
Conclusion

- We compare the actual and approximately optimal highway network in China using a general equilibrium framework.
  - Implementing policy of connecting “intermediate sized cities”

- Actual and optimal networks have similar structure.
  - But precise connections and paths differ

- Difference between actual and optimal network can partly be explained by the birthplace of politicians.
  - Using time variation to address unobserved heterogeneity.
  - Birthplaces matter only if politician was in office during planning and implementation phase of highway network.

- Non-negligible welfare and growth effects of political distortions to large transportation infrastructure networks.
Model

Setup

- Inter-regional trade model based on Eaton and Kortum (2002)
- Donaldson and Hornbeck (2016) derive ‘Market Access’ as a measure for the effect of transport infrastructure

Setup

- Geography
  - Many trading locations
  - Iceberg trade costs
- Technology
  - Cobb-Douglas production with capital (mobile), labor (fixed), and land (fixed)
  - Probabilistic productivity of each district in each variety (comparative advantage)
- Preferences
  - CES utility over a continuum of differentiated goods
Model
Gravity

- Eaton and Kortum (2002): trade from location $i$ to location $j$ is

$$X_{ij} = \underbrace{T_i (q_i^\alpha w_i^\gamma)^{-\theta}}_{\text{Origin's productivity and factor costs}} \times \underbrace{Y_j}_{\text{Destination's income}} \times \underbrace{\tau_{ij}^{-\theta}}_{\text{Bilateral trade costs}} \times \underbrace{\kappa_1 MA_j^{-1}}_{\text{Destination's market access}}$$

where

$$MA_j \equiv \kappa_1 \sum_l T_l (q_l^\alpha w_l^\gamma)^{-\theta} \tau_{jl}^{-\theta} = (P_j)^{-\theta}$$

- Summing the gravity equation over destinations $j$ yields total expenditure (income) of origin $i$:

$$Y_i = \sum_j X_{ij} = \kappa_1 T_i (q_i^\alpha w_i^\gamma)^{-\theta} \sum_j \left[ Y_j \tau_{ij}^{-\theta} MA_j^{-1} \right]$$

Market Access of $i$
Rewriting the above income gravity equation using the price index \((\text{market access})\) and factor shares yields

\[
Y_i' = \left( \kappa_2 T_i \right)^{\frac{1}{1+\theta(\alpha+\gamma)}} \left( \frac{\alpha}{L_i} \right)^{-\frac{\theta \alpha}{1+\theta(\alpha+\gamma)}} \left( \frac{\gamma}{H_i} \right)^{-\frac{\theta \gamma}{1+\theta(\alpha+\gamma)}} (MA_i)^{\frac{1+\theta(1+\alpha+\gamma)}{(1+\theta(\alpha+\gamma))\theta}} \tag{3}
\]

\[
MA_i = \rho^{\frac{1+\theta}{\theta}} \sum_j \tau_{ij}^{-\theta} MA_j^{-\frac{(1+\theta)}{\theta}} Y_j'. \tag{4}
\]

\(\kappa_2\) contains the interest rate, which is equalized across locations.
Travel Times

- Dijkstra’s algorithm is also used to find the shortest path through the terrain and infrastructure network.
- The driving time is then mapped into an iceberg trade cost based on

\[ \tau_{ij} = 1 + \omega t_{ij}^\chi, \]  

(5)
Geography

Road construction costs and least-cost path

The map shows the optimal path through a road construction cost surface based on the slope of the terrain. Darker pixels represent steeper slope.
The maps shows the light density in 1992 and the 30 km buffers around the targeted cities in mainland China.
## Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.2</td>
<td>Elasticity of income with respect to market access</td>
</tr>
<tr>
<td>$\theta$</td>
<td>8</td>
<td>Trade elasticity</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>– (enters through $\beta$)</td>
<td>Land share in the production function</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>– (enters through $\beta$)</td>
<td>Labor share in the production function</td>
</tr>
<tr>
<td>$\rho$</td>
<td>1</td>
<td>Scalar for $FMA = \rho CMA$</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Calibrated to match median iceberg trade cost of 1.25</td>
<td>Scaling of travel time</td>
</tr>
<tr>
<td>$\chi$</td>
<td>0.8</td>
<td>Concavity of trade cost in travel time</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Calibrated to match ratio of USD cost of NTHS to costs based on topography</td>
<td>Scalar to map road construction cost based on topography to USD costs</td>
</tr>
<tr>
<td>$r$</td>
<td>0.05</td>
<td>Annual cost of capital</td>
</tr>
<tr>
<td>$m$</td>
<td>0.12</td>
<td>Annual maintenance costs</td>
</tr>
</tbody>
</table>
Effects of Birthplaces on Light Growth

- We also estimate the effect of birthplaces on light growth through the network design.

- Trace out the growth effects of politically-driven NTHS access across ‘peripheral’ locations using the following 2SLS equation:

\[
\Delta \log \text{light}_{2002-2007}^i = \theta_0 + \theta_1 \Delta \hat{\text{NTHS}}_{\text{opt-act}}^i + \theta_2 Z^i + \eta^i,
\]

where the first-stage is based on the political bias regression

\[
\Delta \hat{\text{NTHS}}_{\text{opt-act}}^i = \hat{\alpha}_0 + \hat{\alpha}_1 \text{Political Access}_{1995-2001}^i + \hat{\alpha}_2 X^i.
\]
## Effects of Birthplaces on Light Growth

\[ \Delta \log \text{light}_{2002-2007} \]

with instruments for

\[ \Delta \text{NTHS}_i^{\text{bct-act}} \]

\[ \Delta \text{NTHS}_i^{\text{opt-act}} \]

\[ \Delta \text{NTHS}_i^{\text{mst-act}} \]

<table>
<thead>
<tr>
<th></th>
<th>( \Delta \text{NTHS}_i^{\text{bct-act}} )</th>
<th>( \Delta \text{NTHS}_i^{\text{opt-act}} )</th>
<th>( \Delta \text{NTHS}_i^{\text{mst-act}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dist Optim (NTHS budget)</td>
<td>0.647*** (0.134)</td>
<td>3.003*** (0.540)</td>
<td></td>
</tr>
<tr>
<td>Dist Optim</td>
<td>0.755*** (0.133)</td>
<td>2.924*** (0.510)</td>
<td></td>
</tr>
<tr>
<td>Dist MST</td>
<td>0.656*** (0.180)</td>
<td>2.349*** (0.394)</td>
<td></td>
</tr>
<tr>
<td>Political Access(_i^{1995-2001})</td>
<td>-0.080 (2.633)</td>
<td>-0.086 (2.659)</td>
<td>0.161 (2.504)</td>
</tr>
</tbody>
</table>

Observations 1650 1650 1650 1650 1650 1650

Adjusted \(R^2\) 0.174 – 0.176 – 0.171 –

Kleibergen-Paap Wald F – 9.42 – 9.95 – 15.36

Standard errors in parentheses

Note: Sample restricted using distance to target cities.

\* \( p < 0.10 \), \** \( p < 0.05 \), \*** \( p < 0.01 \)
The maps shows the approximation of the optimal highway network in China based on the heuristic algorithm with the constraint that all targeted cities are connected. The background shows the slope of the terrain of mainland China. The nodes show the location of all cities with a population of at least 500,000 and all provincial capitals.
The maps shows the least-cost network, i.e. the minimum spanning tree that is computed with the Kruskal algorithm. The background shows the slope of the terrain of mainland China. The nodes show the location of all cities with a population of at least 500,000 and all provincial capitals.