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)
NSD at Peking University
• Investment on inland trans. infrastructure: €850 billion/year in 47 major countries, half of which in China (2% GDP in 2000 ↗ 5% in 2010)
• Blue: Expressway network 1999. Red: Expressway network 2010
What are the impacts of transportation infrastructure improvement on regional and aggregate economy

- Early work: first-order based measurement (Fogel, 1964) or reduced-form (Banerjee et al., 2012)
- GE in nature → necessitates a structural model

Recent progress:
- Market access approach: Donaldson and Hornbeck (2016), Alder (2016), Baum-Snow et al. (2018), ...
- Quantification via structural counterfactual: Donaldson (2018a), Allen and Arkolakis (2014, 2016), Fajgelbaum and Schaal (2017), ...
Key to both approaches: identify the trade cost elasticity

- travel distance $\xrightarrow{\text{trade cost elast.}}$ trade cost $\xrightarrow{\text{trade elast.}}$ trade flow $\xrightarrow{\text{GE}}$ emp./wage

- How existing work recovers trade cost elast.
  - (1) external measure of freight rates: Baum-Snow et al. (2018)
  - (2) estimate using price gaps of homogeneous goods: Asturias et al. (2018), Atkin and Donaldson (2015), Donaldson (2018b)
  - (3) estimate using shipment flows: Allen and Arkolakis (2014, 2016)

- Approach (1) rules out the non-monetary component of trade cost

- (2) and (3) both demanding in data → restricted to a small groups of products (thus one-sector models); trade cost elas. identified from cross-sectional variations in shipment flows
What we do

- A novel source of information to measure domestic shipment
  - export data from the Chinese customs 1999-2010
  - location of exporter, port of exit, volume and quantity \( \implies \) routing, price gap

- Combined with expressway expansion to estimate cost on expressway and regular roads
  - idea: A exports more through port 1 than port 2 \( \implies \tau_{A,1} < \tau_{A,2} \)
  - use over-time variations and an IV (Faber, 2014) to address various concerns
  - allow trade cost heterogeneity by weight-to-value ratio; discipline extent of heterogeneity using prices

- Parameterize a regional GE model
  - routing module from Allen and Arkolakis (2016)
    idiosyncratic trucker preference over routes \( \implies \) tractable for characterization of the welfare effects
  - Caliendo and Parro (2015) with sector heterogeneity in trade costs
Main findings

- **Transport costs parameters:**
  - ad valorem for each 100 kilometer on regular road (4.2%) and expressway (3.4%)
  - doubling weight-to-value ratio increases cost by 23%
Main findings

- Transport costs parameters:
  - ad valorem for each 100 kilometer on regular road (4.2%) and expressway (3.4%)
  - doubling weight-to-value ratio increases cost by 23%

- Evaluate the return to expressway expansion: 1999-2010
  - 50,000 kilometer expressways built; total cost $570 billion (10% of 2010 GDP)
  - welfare gains 5.1%, or 150% net return to investment
  - return smaller if shut down international trade (15% less), regional specialization (20%), sector heterogeneity in cost (5%), and intermediate linkages (75%)
    \[ \Rightarrow 0.89\% \text{ welfare gains in one-sector model, or 56\% negative return} \]
Main findings

- Transport costs parameters:
  - ad valorem for each 100 kilometer on regular road (4.2%) and expressway (3.4%)
  - doubling weight-to-value ratio increases cost by 23%

- Evaluate the return to expressway expansion: 1999-2010
  - 50,000 kilometer expressways built; total cost $570 billion (10% of 2010 GDP)
  - welfare gains 5.1%, or 150% net return to investment
  - return smaller if shut down international trade (15% less), regional specialization (20%), sector heterogeneity in cost (5%), and intermediate linkages (75%)
    \[\Rightarrow 0.89\% \text{ welfare gains in one-sector model, or 56\% negative return}\]

- The effects can be approximated accurately using a 2nd-order characterization
  - after the model is parameterized, no need for computing counterfactuals
  - apply to closed/open economy; accommodate mobile/immobile labor
**Literature**

- **Impacts of infrastructure projects on**
  - *difference*: a new way of estimating trade cost elasticity

- **Domestic trans. infra. promotes export**
  - using *country-level* (Lima and Venables, 2001) and *region-level* variations (Coșar and Demir, 2016 and Martincus et al., 2017)
  - *difference*: focus are impact on trade cost and welfare, rather than export per se

- **Chinese spatial economy.**
  - determine transport cost using *railway shipments* (account for only 10% of shipment; province level) or *regional input-output* table (imputed from railway)
  - *new*: parameterize a domestic trade cost matrix
Outline

• Data and Reduced-form Specification

• Model
  - Road network → trade cost
  - Multi-sector EK

• Quantification and Counterfactuals
  - Welfare gains of the expressway expansion 1999-2010
  - Welfare gains of mega expressway projects
  - Nonlinearity and second-order characterization of welfare gains
Data and Reduced-form Specifications
Data: transportation network (Baum-Snow et al., 2018)

- Left: expressways for 1999 and 2010
- Right: regular roads ('national' and 'provincial' roads) in 2007
- Find distance along the shortest path between \( o \) and \( d \),
  \[ \{ \text{dist}^t_{od} : t = 1999, 2010 \} \]
  - necessary to take a stand on relative costs of expressway and regular road
  - for now: 1 km on expressway equals to 0.5 km on regular road
  - later: pinned down in full quantification
Reduced-form specification: routing

\[
\ln(v_{(o,\text{RoW}),d}^t) = \beta_{od} + \beta_o^t + \beta_d^t + \gamma \cdot \text{dist}_{od}^t + \epsilon_{od}^t
\]

- \(v_{(o,\text{RoW}),d}^t\): value exported from city \(o\) via port \(d\) in year \(t\)
- \(\text{dist}_{od}^t\): shortest effective distance from \(o\) to \(d\): \(0.5 \times \text{dist}_{o\rightarrow d,H}^t + \text{dist}_{o\rightarrow d,L}^t\)
- \(\gamma\): composite of \(\kappa_L \times \theta^F\)
  - \(\kappa_L\): effective cost for regular roads; \(\theta^F\): elasticity of substitution between ports

Remarks
- limit case of the structural equation w/o. trader preference heterogeneity
- omitting \(\beta_{od}\) leads to biased \(\hat{\gamma}\)
- address endogeneity of road networks: (1) exclude major cities; (2) minimum-spanning tree IV; (3) sectoral-level specification

Minimum-spanning Tree

10/27
Reduced-form specification: routing

\[ \ln(\mathbf{v}_{(o,\text{RoW}),d}^t) = \beta_{od} + \beta_o^t + \beta_d^t + \gamma \cdot \text{dist}_{od}^t + \epsilon_{od}^t \]

- \( \mathbf{v}_{(o,\text{RoW}),d}^t \): value exported from city \( o \) via port \( d \) in year \( t \)
- \( \text{dist}_{od}^t \): shortest effective distance from \( o \) to \( d \): \( 0.5 \times \text{dist}_{o \rightarrow d,H}^t + \text{dist}_{o \rightarrow d,L}^t \)
- \( \gamma \): composite of \( \kappa_L \times \theta^F \)
  - \( \kappa_L \): effective cost for regular roads; \( \theta^F \): elasticity of substitution between ports
- Remarks
  - limit case of the structural equation w/o. trader preference heterogeneity
  - omitting \( \beta_{od} \) leads to biased \( \hat{\gamma} \)
  - address endogeneity of road networks: (1) exclude major cities; (2) minimum-spanning tree IV; (3) sectoral-level specification
### Expressway and Routing of Export Shipments

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) OLS</th>
<th>(3) OLS</th>
<th>(4) OLS</th>
<th>(5) 2SLS</th>
<th>(6) 2SLS</th>
<th>(7) OLS</th>
<th>(8) 2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>dist\textsubscript{od} \textsuperscript{t}</td>
<td>-0.341*** (0.011)</td>
<td>-0.384*** (0.011)</td>
<td>-0.157*** (0.037)</td>
<td>-0.174*** (0.045)</td>
<td>-0.170*** (0.058)</td>
<td></td>
<td>-0.088** (0.038)</td>
<td>-0.162** (0.068)</td>
</tr>
<tr>
<td>-on express</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.174*** (0.045)</td>
<td>-0.198*** (0.063)</td>
</tr>
<tr>
<td>-on regular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>o, d, t</td>
<td>ot, dt</td>
<td>od, ot, dt</td>
<td>od, ot, dt</td>
<td>od, ot, dt</td>
<td>od, ot, dt</td>
<td>od, ot, dt</td>
<td>od, ot, dt</td>
</tr>
<tr>
<td>Exclude Major Cities</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>3668</td>
<td>3660</td>
<td>2838</td>
<td>2068</td>
<td>2038</td>
<td>2068</td>
<td>2038</td>
<td>2068</td>
</tr>
<tr>
<td>R\textsuperscript{2}</td>
<td>0.646</td>
<td>0.709</td>
<td>0.906</td>
<td>0.897</td>
<td>0.020</td>
<td>0.897</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>First Stage K-P F</td>
<td>1400.799</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>170.204</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors are clustered at city-port level. * p < 0.10, ** p < 0.05, *** p < 0.01.
Summary and Motivation for a routing model

- Reduced-form elasticity of routing w.r.t. effective distance around 15%
  - Elasticity lower w.r.t. expressway distance
  - Using cross-sectional variations more than doubles the estimate
  - Needs to take a stand on the relative cost of express/national, for shortest path
  - Confounding with port choice elasticity and router preference heterogeneity
Summary and Motivation for a routing model

- Reduced-form elasticity of routing w.r.t. effective distance around 15%
  - Elasticity lower w.r.t. expressway distance
  - Using cross-sectional variations more than doubles the estimate
  - Needs to take a stand on the relative cost of express/national, for shortest path
  - Confounding with port choice elasticity and router preference heterogeneity

- Extend the routing problem and embed into a GE model
  - allow traders to have heterogeneous preference for routes $\implies$ both regular roads and expressways used; identify $\theta_F$, $\kappa_L$, $\kappa_H$
  - incorporates alternative modes
  - use the GE structure to infer the level of cost; counterfactuals
Routing Model
A Routing Model with Multiple Transport Modes

Figure: Routing on a Network: Four-Region Example

- iceberg cost $\mathcal{I} = \exp(\kappa \text{dist}_{kl})$

- Two direct (one-step) paths; trucker draws pref. shocks from Frechet for each

- if made choices among the two, the expected cost is:

$$\tau_{od,1} \propto \left( [\mathcal{I}^{L}]^{-\theta} + [\mathcal{I}^{H}]^{-\theta} \right)^{-\frac{1}{\theta}}$$
A Routing Model with Multiple Transport Modes

Figure: Routing on a Network: Four-Region Example

- Three two-step paths: \( o \rightarrow k \rightarrow d \), \( o \rightarrow k \rightarrow l \rightarrow d \), and \( o \rightarrow l \rightarrow d \)

- if made choices among routes \( \leq 2 \) steps, the expected cost is:

\[
\tau_{od,2} \propto \left( \tau_{od,1}^{-\theta} + (\nu_{ok}^L \nu_{kd}^H)^{-\theta} + (\nu_{ok}^L \nu_{kd}^L)^{-\theta} + (\nu_{ol}^L \nu_{ld}^H)^{-\theta} \right)^{-\frac{1}{\theta}}
\]
Define two adjacent matrices for regular and express

\[
L = \begin{pmatrix}
0 & \theta_{lo} & \theta_{od} & \theta_{ok} \\
\theta_{ol} & 0 & \theta_{od} & \theta_{ok} \\
\theta_{lo} & \theta_{lo} & 0 & \theta_{dl} \\
\theta_{ko} & \theta_{lo} & \theta_{dl} & 0 \\
\end{pmatrix}, \quad \quad H = \begin{pmatrix}
0 & 0 & \theta_{od} & 0 \\
0 & 0 & 0 & \theta_{od} \\
\theta_{do} & 0 & 0 & \theta_{dk} \\
0 & \theta_{ko} & \theta_{dk} & 0 \\
\end{pmatrix}
\]

Then \( \tau_{od,1} \propto [A_{od}]^{-\frac{1}{\theta}} \) and \( \tau_{od,2} \propto [A_{od} + A_{od}^2]^{-\frac{1}{\theta}} \), for \( A \equiv L + H \)
Define two adjacent matrices for regular and express

\[
L = \begin{pmatrix}
  0 & \ell_{ol}^L - \theta & \ell_{od}^L - \theta & \ell_{ok}^L - \theta \\
  \ell_{lo}^L - \theta & 0 & \ell_{ld}^L - \theta & 0 \\
  \ell_{do}^L - \theta & \ell_{dl}^L - \theta & 0 & \ell_{dk}^L - \theta \\
  \ell_{ko}^L - \theta & 0 & \ell_{kd}^L - \theta & 0 \\
\end{pmatrix}
\]

\[
H = \begin{pmatrix}
  0 & 0 & \ell_{od}^H - \theta & 0 \\
  0 & 0 & 0 & 0 \\
  \ell_{do}^H - \theta & 0 & 0 & \ell_{dk}^H - \theta \\
  0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

Then \( \tau_{od,1} \propto [A_{od}]^{-1/\theta} \) and \( \tau_{od,2} \propto [A_{od} + A_{od}^2]^{-1/\theta} \), for \( A \equiv L + H \)

Allowing all routes, average trade cost: \( \tau_{od} \equiv \tau_{od,\infty} \propto B_{od}^{-1/\theta} \), for \( B \equiv (I - A)^{-1} \).
Routing Model: Summary and Discussions

- Bilateral trade cost along road networks summarized by $\tau_{od} \propto B_{od}$ and $B = \tilde{B}(\kappa^L \theta, \kappa^H \theta)$.

- Routing reduces to shortest-path routing when $\theta \to \infty$.

- Extended with sectoral-heterogeneity, alternative transport modes, and port choices, we have the following

$$\pi(o, RoW), d = \frac{(\tau_{od} \cdot \tau_{d, RoW})^{-\theta_F}}{\sum_{\text{All ports } k} (\tau_{ok} \cdot \tau_{k, RoW})^{-\theta_F}},$$

$\pi(o, RoW), d$: share of export from $o$ via port $d$

$\theta_F$: elasticity of substitution across ports

$\tau_{d, RoW}$ and denominator: port and city fixed effects after log linearization.

- Use over-time variations in export routing to estimate $\kappa^L \theta, \kappa^H \theta, \theta_F$. 
Estimating the Routing Model

- Estimate the following with non-linear least square

\[
\min_{\theta^F, \kappa^H \theta, \kappa^L \theta, f} \sum_{o, d} \left[ \frac{\theta^F}{\theta} \log \left( \left[ \hat{B}^t (\kappa^H \theta, \kappa^L \theta)_{(od)} \right] \right) + f - \log (v^t_{(o, RoW), d}) \right]^2
\]

<table>
<thead>
<tr>
<th></th>
<th>Point estimates</th>
<th>S.d.</th>
<th>Median</th>
<th>p10</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: export routing data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa^L \theta$</td>
<td>4.68</td>
<td>1.90</td>
<td>4.67</td>
<td>4.26</td>
<td>6.18</td>
</tr>
<tr>
<td>$\kappa^H \theta$</td>
<td>3.78</td>
<td>1.08</td>
<td>3.77</td>
<td>3.36</td>
<td>4.83</td>
</tr>
<tr>
<td>$\theta^F / \theta$</td>
<td>0.06</td>
<td>0.03</td>
<td>0.05</td>
<td>0.03</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Notes: Standard errors and percentiles are based on 200 bootstrapped samples.

- Takeaways:
  - $\kappa^H / \kappa^L \approx 0.8$: the ad valorem cost of expressway is 20% lower than regular road
  - $\theta^F < \theta$: routing is much more substitutable than port choice
The rest of the model

- 323 regions (prefectures)+RoW, 25 sectors (2-digit). Regions differ by sector productivity, amenity and fixed land supply
- Mobile workers with Cobb-Douglas preference over housing and sectoral final goods
- Rent from land redistributed via lump-sum transfer
- Intermediate good production: combine labor and sector final goods
- Final good production: aggregate intermediate inputs within the sector across source regions a la Armington
Quantification and Counterfactuals
### Parameterize the rest of the model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Descriptions</th>
<th>Value</th>
<th>Targets/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameters calibrated externally</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta^i, \gamma^{ij}, \alpha^j$</td>
<td>IO structure and consumption share</td>
<td>-</td>
<td>2007 IO table for China</td>
</tr>
<tr>
<td>$L_d$</td>
<td>Total employment</td>
<td>-</td>
<td>2010 Population Census</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Trade elasticity</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>$\theta_M$</td>
<td>Elasticity of substitution across modes</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td><strong>Parameters calibrated in equilibrium</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>Routing elasticity</td>
<td>111.52</td>
<td></td>
</tr>
<tr>
<td>$\theta_F$</td>
<td>Port choice elasticity</td>
<td>6.35</td>
<td></td>
</tr>
<tr>
<td>$\kappa_H$</td>
<td>Expressway route cost</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td>$\kappa_L$</td>
<td>Regular route cost</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>$h_0$</td>
<td>Trade cost level</td>
<td>1.260</td>
<td>Average ground shipment distance: 177 km</td>
</tr>
<tr>
<td>$\bar{\kappa}$</td>
<td>Alternative mode cost</td>
<td>0.163</td>
<td>Share of non-road shipment: 0.24</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Cost-weight to value elasticity</td>
<td>0.3</td>
<td>estimated</td>
</tr>
<tr>
<td>$\tau^i_{RoW}, \tau^{i'}_{RoW}$</td>
<td>Export and import costs</td>
<td>-</td>
<td>Sectoral export and import</td>
</tr>
<tr>
<td>$T^i_d$</td>
<td>Region-sector productivity</td>
<td>-</td>
<td>City-sector sales in 2008 Economic Census</td>
</tr>
</tbody>
</table>

Joint estimate of $\kappa^H \theta = 3.78, \kappa^L \theta = 4.68, \theta_F = 0.06, \frac{\partial \log p}{\partial \text{dist}} = 0.06$
Model validation

**Figure:** Model Predicted Shipment Flows

Export and export growth by cities  Transhipment via cities
The Effects of the Expressway Expansion, 1999-2010

<table>
<thead>
<tr>
<th>Change in</th>
<th>Value</th>
<th>S.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate welfare</td>
<td>0.051</td>
<td>0.025</td>
</tr>
<tr>
<td>Log(Domestic trade)</td>
<td>0.136</td>
<td>0.052</td>
</tr>
<tr>
<td>Log(Exports)</td>
<td>0.097</td>
<td>0.080</td>
</tr>
<tr>
<td>Std Log(real wage) across regions</td>
<td>-0.017</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Note: Changes in model statistics are calculated by comparing the calibrated equilibrium and a counterfactual equilibrium with the 1999 expressway network.

- Numbers in perspective: between 1999 and 2010, aggregate TFP increased by 36% (Penn World Table)
- Expressway expansion explains about 14% of the TFP increase
The Effects of the Expressway Expansion, 1999-2010

<table>
<thead>
<tr>
<th>Change in</th>
<th>Value</th>
<th>S.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate welfare</td>
<td>0.051</td>
<td>0.025</td>
</tr>
<tr>
<td>Log(Domestic trade)</td>
<td>0.136</td>
<td>0.052</td>
</tr>
<tr>
<td>Log(Exports)</td>
<td>0.097</td>
<td>0.080</td>
</tr>
<tr>
<td>Std Log(real wage) across regions</td>
<td>-0.017</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Note: Changes in model statistics are calculated by comparing the calibrated equilibrium and a counterfactual equilibrium with the 1999 expressway network.

- Numbers in perspective: between 1999 and 2010, aggregate TFP increased by 36% (Penn World Table)
- Expressway expansion explains about 14% of the TFP increase
- Total cost: 10% of 2010 GDP. Assuming 10% depreciation rate (Bai and Qian, 2010), 10% return to capital (Bai et al., 2006) \( \Rightarrow \) 155% net return to investment
The role of sectors

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
<th>Model (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>International trade</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade cost heterogeneity</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional specialization</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate Input</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Welfare gains</td>
<td>5.10%</td>
<td>4.47%</td>
<td>4.29%</td>
<td>3.36%</td>
<td>0.89%</td>
</tr>
</tbody>
</table>

Each model recalibrated to match the same sales by city ($\{T_d^i\}$) and average shipment distance ($h_0$).

- baseline $\rightarrow$ (2): overlooks that expressways reduces import and export cost
## The role of sectors

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
<th>Model (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>International trade</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade cost heterogeneity</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional specialization</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate Input</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Welfare gains</td>
<td>5.10%</td>
<td>4.47%</td>
<td>4.29%</td>
<td>3.36%</td>
<td>0.89%</td>
</tr>
</tbody>
</table>

Each model recalibrated to match the same sales by city ($\{T^i_d\}$) and average shipment distance ($h_0$).

- baseline $\rightarrow$ (2): overlooks that expressways reduces import and export cost
- (2) $\rightarrow$ (3): matched to the same average shipment distance, model (2) infers higher shipment values, which to the first order determine the gains
The role of sectors

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
<th>Model (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>International trade</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Trade cost heterogeneity</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional specialization</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate Input</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Welfare gains</td>
<td>5.10%</td>
<td>4.47%</td>
<td>4.29%</td>
<td>3.36%</td>
<td>0.89%</td>
</tr>
</tbody>
</table>

Each model recalibrated to match the same sales by city ($\{T_d^i\}$) and average shipment distance ($h_0$).

- baseline $\rightarrow$ (2): overlooks that expressways reduces import and export cost
- (2) $\rightarrow$ (3): matched to the same average shipment distance, model (2) infers higher shipment values, which to the first order determine the gains
- (3) $\rightarrow$ (4): In the data and model (3) trade happens between large regions specializing in different sectors; model (4) predicts more trade between close partners.
# The role of sectors

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
<th>Model (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>International trade</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade cost heterogeneity</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional specialization</td>
<td>✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate Input</td>
<td>✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Welfare gains</td>
<td>5.10%</td>
<td>4.47%</td>
<td>4.29%</td>
<td>3.36%</td>
<td>0.89%</td>
</tr>
</tbody>
</table>

Each model recalibrated to match the same sales by city ($\{T_d^i\}$) and average shipment distance ($h_0$).

- baseline → (2): overlooks that expressways reduces import and export cost
- (2) → (3): matched to the same average shipment distance, model (2) infers higher shipment values, which to the first order determine the gains
- (3) → (4): In the data and model (3) trade happens between large regions specializing in different sectors; model (4) predicts more trade between close partners.  
  [Map](#)
- (4) → (5): inferred wrong sales/VA ratio
Evaluating mega projects

<table>
<thead>
<tr>
<th>ID</th>
<th>Length (km)</th>
<th>Cost as % GDP</th>
<th>Cost per km (million)</th>
<th>Welfare Gains (%)</th>
<th>Net return to investment</th>
<th>% Change in dom. trade</th>
<th>% Change in Export</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>1533.61</td>
<td>0.30</td>
<td>77.71</td>
<td>0.40</td>
<td>567.19%</td>
<td>1.16</td>
<td>0.56</td>
</tr>
<tr>
<td>G3</td>
<td>2513.38</td>
<td>0.54</td>
<td>85.53</td>
<td>0.49</td>
<td>354.10%</td>
<td>1.05</td>
<td>1.86</td>
</tr>
<tr>
<td>G10</td>
<td>891.73</td>
<td>0.15</td>
<td>67.25</td>
<td>0.02</td>
<td>-22.92%</td>
<td>0.09</td>
<td>0.02</td>
</tr>
<tr>
<td>G30</td>
<td>4356.49</td>
<td>0.85</td>
<td>78.04</td>
<td>0.39</td>
<td>129.32%</td>
<td>1.34</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>30012.46</td>
<td>6.16</td>
<td>3.47</td>
<td>8.76</td>
<td>6.48</td>
<td></td>
</tr>
</tbody>
</table>

Note: Each row corresponds to a counterfactual experiment by removing from the 2010 expressway network a mega expressway project referred by ‘ID’.
Welfare characterization

- **Goal:** calculate welfare gains from expressway expansion w/o solving equilibrium

- **Solution:** a formula that associates welfare gains from model statistics up to second order

\[
\Delta W = - \sum_{o,d,i} \left( \frac{X_{od}^i}{Y} - \frac{\Lambda_o^i}{Y} 1_{d=RoW} \right) \sum_{kl \in C} \pi_{od}^{\text{road}} \pi_{od}^{kl} \Delta \log(\epsilon_{kl}) + SO_R + \text{Residual}
\]

- \( \Delta \log(\epsilon_{kl}) \): change in route cost; known after estimating \((\kappa^L, \kappa^H)\)
- \( Y \): Domestic GDP
- \( X_{od}^i \): Domestic trade flows from \( o \) to \( d \) in sector \( i \)
- \( \Lambda_o^i \): foreign consumption exposure to export cost
- \( \pi_{od}^{\text{road}}, \pi_{od}^{kl} \): routing patterns
- \( SO_R \): second-order term that depends on routing patterns and routing elas. \( \theta \)
Welfare gains: The miss of FO and corrections of SO

Note: Each point corresponds to an experiment with one expressway segment removed. The sample segments are the top 100 busiest city pairs in the baseline equilibrium.

- The FO approach (cost-saving approach) misses welfare gains by 21% on average and most likely overestimates the gains
  - Road upgrading should be considered a “large” shock
  - FO ignores the rerouting behavior after downgrading an expressway
- Incorporating SO reduces the error by 2/3
Apply the welfare formula to large projects

- Consider removing all expressway segments where there are existing regular roads

\[ \Delta W = \text{FO effect} + \text{Own SO}_R + \text{Cross SO}_R + \text{Residual} \]

- Cross SO\(_R\) captures the cross-substitutes of road segments, and has an intuitive interpretation

- We publish the FO and the SO (Hessian) coefficients so policy makers can assess the welfare gains of expressway expansions without solving counterfactual
Conclusion
Conclusion

- Exploit over-time variations in city-to-port export to estimate the impact of transportation infrastructure on trade cost
  - construction of expressway reduces cost-distance elasticity by 20%

- Accommodating regional specialization / sectoral heterogeneity / intermediate input is important
  - neglecting these underestimate the gains and turns positive NPV into negative

- Our approach requires data on sectoral production and is computational intensive. For future work useful to think about ways to
  - circumvent parameterizing the full model and computing counterfactuals
    2nd-order characterization quite accurate, but requires full information on shipment and routing
  - reduce the data requirement while retaining accuracy
References


The minimum-spanning tree IV (Faber, 2014)

- **Red**: min-distance network connecting 55 major cities; **Blue**: 2010 expressway

- IV for $\text{dist}_{ij}^{2010}$: Effective length of shortest-path along the (Blue) network

- IV for $\text{dist}_{ij}^{1999}$: $\text{dist}_{ij}^{1999}$

- Identification: National Trunk Highway System exogenous to small cities
Change in shipment flows without regional specialization

Note: The numbers are the differences in shipment value/GDP between Model (3) (with specialization) and Model (4) (no specialization). Cold colors indicate that there is less shipment in Model (4) than in Model (3).
### Price-distance regression

**Table:** Price Distance Regression

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>dist_{od}</code></td>
<td>0.055***</td>
<td>0.061***</td>
<td>0.053***</td>
<td>0.058***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.022)</td>
<td>(0.012)</td>
<td>(0.021)</td>
</tr>
<tr>
<td><strong>IV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>dist_{od}</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>dci, oci</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude major cities</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Exclude differentiated goods</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1829372</td>
<td>232609</td>
<td>1829372</td>
<td>232609</td>
</tr>
<tr>
<td><code>R^2</code></td>
<td>0.323</td>
<td>0.340</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>First Stage KP-F statistic</td>
<td></td>
<td></td>
<td>1515.787</td>
<td>1156.297</td>
</tr>
</tbody>
</table>

Notes: `o, d, c, i` stand for origin city, port, destination country, and HS-8 product fixed effects, respectively. Standard errors are clustered at city-port level. * `p < 0.10`, ** `p < 0.05`, *** `p < 0.01`. 

Back
## Price regression: estimate trade cost-weight elasticity

### Table: Transport cost and weight-to-value ratio

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td>log price ratio</td>
<td>log price ratio</td>
<td>log price ratio</td>
<td>log price ratio</td>
<td>log price ratio</td>
<td>log price ratio</td>
<td>log price ratio</td>
</tr>
<tr>
<td>Heaviness- HS2 Category</td>
<td>0.163*** (0.056)</td>
<td>0.161*** (0.056)</td>
<td>0.278*** (0.086)</td>
<td>0.199** (0.089)</td>
<td>0.303*** (0.044)</td>
<td>0.362*** (0.050)</td>
<td>0.253*** (0.043)</td>
</tr>
<tr>
<td>Heaviness- HS4 Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>o, d, c</td>
<td>odc</td>
<td>fdc</td>
<td>fdc</td>
<td>fdc</td>
<td>fdc</td>
<td>fdc</td>
</tr>
<tr>
<td>Exclude major cities</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Exclude differentiated goods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1987140</td>
<td>1985946</td>
<td>1805563</td>
<td>190836</td>
<td>1805563</td>
<td>1126941</td>
<td>119077</td>
</tr>
<tr>
<td>R²</td>
<td>0.063</td>
<td>0.074</td>
<td>0.375</td>
<td>0.481</td>
<td>0.417</td>
<td>0.596</td>
<td>0.639</td>
</tr>
</tbody>
</table>

Notes: 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.
# Descriptive Statistics of Changes in Route Length

<table>
<thead>
<tr>
<th>Route-level variables</th>
<th>2000-2001</th>
<th></th>
<th>2010-2011</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>std</td>
<td>mean</td>
<td>std</td>
</tr>
<tr>
<td>Export</td>
<td>164</td>
<td>2050</td>
<td>989</td>
<td>12100</td>
</tr>
<tr>
<td>Total length</td>
<td>20.40</td>
<td>11.60</td>
<td>17.24</td>
<td>10.45</td>
</tr>
<tr>
<td>Length of expressway segments</td>
<td>12.93</td>
<td>7.71</td>
<td>16.88</td>
<td>9.72</td>
</tr>
<tr>
<td>Expressway share in total length</td>
<td>0.65</td>
<td>0.24</td>
<td>0.99</td>
<td>0.05</td>
</tr>
<tr>
<td>Effective length</td>
<td>13.93</td>
<td>9.00</td>
<td>8.80</td>
<td>5.77</td>
</tr>
</tbody>
</table>

Notes: Export value in current year million dollars. Length of routes in 100km.

\[
\Delta dist_{od}^t = \Delta dist_{o\rightarrow d}^L + 0.5 \times \Delta dist_{o\rightarrow d}^H = (\Delta dist_{o\rightarrow d}^L + \Delta dist_{o\rightarrow d}^H) - 513 \text{ km} = (\Delta dist_{o\rightarrow d}^L + \Delta dist_{o\rightarrow d}^H) - 316 \text{ km (60%)} - 0.5 \times \Delta dist_{o\rightarrow d}^H = -196 \text{ km (40%)}
\]
Export v.s Route Length, Cross-section and Over-time variations

Figure: Bin-scattered Plots of Log(Export) and Route Length

(a) Cross-section variations

(b) Over-time variations

Note: Panel (a) plots the log(export) of routes over the effective length of routes. Panel (b) plots the change in log(export) over the change in effective length from 1999 to 2010.
(a) Relative Change in Port Access

(b) Relative Change in log(Export)

**Figure:** Relative Change in Port Access and Export: North Minus South

Note: Port access is measured by average effective distance to ports within a port group. The change is from 1999 to 2010.

- Compared to northern cities, southern cities export relatively more via northern ports after the distance to northern ports get shorter.
### Growth v.s Rerouting; Results from Sectoral Data

<table>
<thead>
<tr>
<th></th>
<th>(1) Aggregate Data</th>
<th>(2) Sectoral Data</th>
<th>(3) Sectoral Data</th>
<th>(4) Sectoral Data</th>
<th>(5) Sectoral Data</th>
<th>(6) Sectoral Data</th>
<th>(7) Sectoral Data</th>
<th>(8) Sectoral Data</th>
<th>(9) Sectoral Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dist^t_{od}</strong></td>
<td>-0.185***</td>
<td>-0.166***</td>
<td>-0.157***</td>
<td>-0.373***</td>
<td>-0.138***</td>
<td>-0.157***</td>
<td>-0.137***</td>
<td>-0.137***</td>
<td>-0.137***</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.041)</td>
<td>(0.041)</td>
<td>(0.017)</td>
<td>(0.044)</td>
<td>(0.041)</td>
<td>(0.039)</td>
<td>(0.039)</td>
<td>(0.039)</td>
</tr>
<tr>
<td><strong>- on express</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>- on regular</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>log(city export through other routes)</strong></td>
<td>0.171**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.053**</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.026)</td>
</tr>
<tr>
<td><strong>Specification</strong></td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>og, ot, gt, od, p, dt</td>
<td>od, p, dt, dt, odi, oti, dti</td>
<td>od, p, dt, dt, odi, oti, dti</td>
<td>odi, oti, dti, odi, oti, dti</td>
<td>odi, oti, dti, odi, oti, dti</td>
<td>odi, oti, dti, odi, oti, dti</td>
<td>odi, oti, dti, odi, oti, dti</td>
<td>odi, oti, dti, odi, oti, dti</td>
<td>odi, oti, dti, odi, oti, dti</td>
</tr>
<tr>
<td>Exclude Major Cities</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1036</td>
<td>2082</td>
<td>2082</td>
<td>20946</td>
<td>11758</td>
<td>11758</td>
<td>6816</td>
<td>12920</td>
<td>12920</td>
</tr>
<tr>
<td>R²</td>
<td>0.933</td>
<td>0.870</td>
<td>0.869</td>
<td>0.593</td>
<td>0.896</td>
<td>0.896</td>
<td>0.927</td>
<td>0.850</td>
<td>0.850</td>
</tr>
</tbody>
</table>

Note: Column (1) replaces port fixed effects with port-group fixed effects. Column (2) replaces city fixed effects with province fixed effects.

- Controlling for different levels of fixed effects does not change estimates much; 
  ⇒ the variations in export mainly come from growth instead of rerouting
- Results are robust when using sectoral data
Robustness: Measuring in Weights

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distance</td>
<td>Total</td>
<td>Distance</td>
<td>Total</td>
<td>IV Reduced Form</td>
<td>2SLS</td>
<td>Distance</td>
<td>Total</td>
<td>IV Reduced Form</td>
</tr>
<tr>
<td><strong>Distance_total_w05</strong></td>
<td>-0.363***</td>
<td>-0.412***</td>
<td>-0.191***</td>
<td>-0.217***</td>
<td>-0.231***</td>
<td>-0.268***</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
<td>(0.052)</td>
<td>(0.067)</td>
<td>(0.079)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IV_distance_total_w05</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.268***</td>
<td>(0.079)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>(first) Distance</strong></td>
<td>-0.088**</td>
<td>-0.163**</td>
<td>-0.215***</td>
<td>-0.248***</td>
<td>-0.215***</td>
<td>-0.248***</td>
<td>(0.044)</td>
<td>(0.080)</td>
<td>(0.052)</td>
</tr>
<tr>
<td><strong>(first) Distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>(first) Distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>(first) IV_distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>(first) IV_distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>(first) IV_distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>Fixed Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude Major Cities</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.606</td>
<td>0.680</td>
<td>0.893</td>
<td>0.884</td>
<td>0.884</td>
<td>0.023</td>
<td>0.884</td>
<td>0.884</td>
<td>0.018</td>
</tr>
<tr>
<td>(Kleibergen-Paap F statistic):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1356.045</td>
<td></td>
<td></td>
<td></td>
<td>163.977</td>
</tr>
</tbody>
</table>

Notes: Export measured in kilograms. Standard errors are clustered at city-port level. * p < 0.10, ** p < 0.05, *** p < 0.01.

- Results are robust when measuring export in weights.
Identification of $\theta_F/\theta$

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

Figure: Model Prediction Varying $\theta$

Horizontal: change in shortest-path distance in regular-road equivalent distance  
Vertical: Model prediction $\Delta \log \left( \tilde{\mathcal{B}}^t (\kappa_H \theta, \kappa_L \theta)_{(od)} \right)$  
Read line: $\kappa_L \times \theta^F$
Identification of $\theta_F/\theta$

(a) $\theta = 0.75 \times \theta^*$

(b) $\theta = \theta^*$, the calibrated value

(c) $\theta = 1.5 \times \theta^*$

Figure: Model Prediction Varying $\theta$

Horizontal: change in shortest-path distance in regular-road equivalent distance

Vertical: Model prediction $\Delta \log \left( \hat{B}^t (\kappa^H \theta, \kappa^L \theta)_{(od)} \right)$

Read line: $\kappa_L \times \theta^F$

Key: Lowering $\theta$ preserves network structure other than shortest-path
Identification of $\theta_F/\theta$, cont’d

- Identification of $\theta$ (relatively to $\theta^F$) is then given by the importance of network structure other than shortest-path in accounting for the data.
- The figure reports the fraction of variations in data that is explained by the non-linear component when varying $\theta$.

$$
\log(v_{(o, RoW), d}) = \left\{ \begin{array}{ll}
\kappa^L \theta^F \text{dist}_{od} + \beta \cdot \left( \frac{\theta^F}{\theta} \log([\tilde{B}_{(o, d)}]) - \kappa^L \theta^F \text{dist}_{od} \right) & \text{linear} \\
\text{non-linear} 
\end{array} \right.
$$
Model validation: export level

**Figure:** City-level export in 2010: Model versus Data

Note: The figure plots model-implied city export against the data, controlling for city employment.
Table: Predicting Export Growth

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(export), model</td>
<td>0.465***</td>
<td>0.953***</td>
<td>0.871***</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.189)</td>
<td>(0.199)</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>t</td>
<td>oi, it</td>
<td>oi, it</td>
</tr>
<tr>
<td>Exclude major cities</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>8472</td>
<td>8472</td>
<td>6576</td>
</tr>
<tr>
<td>R²</td>
<td>0.333</td>
<td>0.878</td>
<td>0.860</td>
</tr>
<tr>
<td>F-statistic</td>
<td>92.706</td>
<td>25.332</td>
<td>19.223</td>
</tr>
</tbody>
</table>

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 the ‘Fixed Effects’ stand for time, city, and sector (two-digit) fixed effects, respectively. Standard errors (clustered by city) in parenthesis.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. **
Table: Correlation with Shipment

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(shipment), model</td>
<td>0.314***</td>
<td>0.177***</td>
<td>0.176***</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.035)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Log(employment)</td>
<td></td>
<td>0.594***</td>
<td>0.587***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.059)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>Observations</td>
<td>240</td>
<td>240</td>
<td>234</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>no</td>
<td>no</td>
<td>prov</td>
</tr>
<tr>
<td>R²</td>
<td>0.234</td>
<td>0.488</td>
<td>0.636</td>
</tr>
</tbody>
</table>

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