Commuting, Labor, & Housing Market Effects of Mass Transportation: Welfare and Identification

Christopher Severen
Federal Reserve Bank of Philadelphia

Econometric Society Session at the AEA/ASSA Meetings

January 4, 2019

Disclaimer: This presentation represents preliminary research that is being circulated for discussion purposes. The views expressed in this paper are solely those of the author and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the author.
Question: What are the welfare effects of urban rail infrastructure in car-oriented Los Angeles?

1. Establish the causal effect of rail infrastructure on commuting flows between locations connected by LA Metro

2. Develop and estimate parameters of (relatively) simple quantitative, spatial GE model of internal city structure
   - Novel identification for elasticities (common in urban EG lit)
   - Disentangle commuting effect of transit from other margins

3. Quantify welfare effects of rail infrastructure in Los Angeles
   - After 10 years: Clean answer, costs > benefits
   - After 25 years: More speculative, still costs > benefits

+ Assess common methods/assumptions in urban economic geography
  - Differences between commuting and trade
Why transit infrastructure?

Rail systems seen as beneficial, but expensive policy option

- Light rail is 10-20x cost of roadway, subway is 30-100x
- Large US cities on a transit building spree!
- US cities not dense, less monocentric (Anas, Arnott, Small 1998)

Households care: high commuting costs limit residential/job access

- Households spend 10-15% of income & 220 hrs/yr commuting
- Increasing congestion (commutes times up 230% since 1985)

Transportation infrastructure in the aggregate

- Trade between cities & growth (Fogel 1964; Donaldson 2018)
- Commuting within cities, urban form, neighborhood growth (Bento et al. 2003; Gibbons & Machin 2005; Gonzalez-Navarro & Turner 2016)
- Ease congestion, but Fundamental Law of Congestion –
- So transit improves cities by allowing more people to fit in
Transit and commuting in LA

**Setting:** Los Angeles in 1990 and 2000
- No rail → 47 stations on 4 lines by 2000
- Historically automobile-oriented

**Data:** Census Transportation Planning Package (1990, 2000)
- Develop **panel** of all **bilateral commuting flows** for LA (tract-tract)
- Median wage at place (tract) of work

Other data sources
- Housing and population variables (IPUMS)
- LEHD LODES for more recent commuting (not directly comparable to CTPP)

Standardize to 1990 geographies to minimize information loss
Identification of commuting effect

$$\ln(N_{ijt}) = \delta_{it} + \delta_{jt} + \delta_{ij} + \lambda T_{ijt} + \varepsilon_{ijt}$$

Gravity style diff-in-diff isolates commuting effect of transit

- Origin-by-year, Destination-by-year fixed effects
  - Capture non-commuting effects of transit (i.e. amenities)
  - Control for standard concerns about location selection
- Pair-fixed effects capture time-invariant chars., like distance

Identification: parallel counterfactual trends in commuting flows between treated pairs & control pairs

- Use proposed subway plan (1925); streetcar routes (PER)
- Lower bound by adjacencies (Dube, Lester, & Reich 2010), adapted for flows
Identification of commuting effect

Arguments:

- Staggered rollout based on political expediency
- Inconsequential units along routes
- Geologic shock: Ross Dress for Less blew up!
- Similar evolution of built environment (Brooks & Lutz 2016)

Pre-trends? Cannot directly test, but

- Mostly (not perfectly) parallel pre-trends in residential commuting (tracts, not tract pairs – NCDB)
## Effects of stations on commuting flows; 1990–2000

### Subway Plan (All) Sample

<table>
<thead>
<tr>
<th>Model Description</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O &amp; D contain station</td>
<td>0.127**</td>
<td>0.147**</td>
<td>0.152**</td>
<td>0.162**</td>
<td>0.146**</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.046)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>O &amp; D &lt;250m from station</td>
<td>0.115*</td>
<td>0.122*</td>
<td>0.101*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.050)</td>
<td>(0.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O &amp; D &lt;500m from station</td>
<td>0.054</td>
<td>0.018</td>
<td>0.023</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.044)</td>
<td>(0.042)</td>
<td>(0.042)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>74046</td>
<td>74046</td>
<td>74046</td>
<td>74040</td>
<td>74040</td>
</tr>
<tr>
<td>Tract Pair FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>POW-X-Yr FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>RES-X-Yr FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sbcty-X-Sbcty-X-Yr FE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Highway Control</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Y</td>
</tr>
</tbody>
</table>

Metro **increases commuting by 15% (10%)** between connected tracts

- Consistent across various strategies & functional form (PPML)
Toward welfare? Model summary

To translate into welfare, turn to **quantitative, spatial GE model**
- HH dual location choice (similar to Ahlfeldt et al. 2015)
- **Bonus 1!** Generates reduced form commuting flow equation
- **Bonus 2!** Can test for other margins of effects from subway
Toward welfare? Model summary

To translate into welfare, turn to **quantitative, spatial GE model**

- **HH dual location choice** (similar to Ahlfeldt et al. 2015)
- **Bonus 1!** Generates reduced form commuting flow equation
- **Bonus 2!** Can test for other margins of effects from subway

Generate within city Rosen-Roback model with commuting

- Simple, competitive labor and housing markets in each location
- Frechet-distributed idiosyncratic preference over locations pairs

**Labor**

\[
\ln(W_{jt}) = \tilde{\alpha} \ln \left( \sum_r N_{rjt} \right) + \ln(A_{jt})
\]

- Wage
- Productivity

**Housing**

\[
\ln(Q_{it}) = \psi \ln \left( \sum_s N_{ist} \right) + \ln(C_{it})
\]

- H. Price
- H. Eff.
Toward welfare? Model summary

To translate into welfare, turn to **quantitative, spatial GE model**

- HH dual location choice (similar to Ahlfeldt et al. 2015)
- **Bonus 1!** Generates reduced form commuting flow equation
- **Bonus 2!** Can test for other margins of effects from subway

Generate within city Rosen-Roback model with commuting

- Simple, competitive labor and housing markets in each location
- Frechet-distributed idiosyncratic preference over locations pairs

**Labor**

\[
\ln(W_{jt}) = \tilde{\alpha} \ln (\sum_r N_{rjt}) + \ln(A_{jt})
\]

**Productivity**

**Commut.**

\[
\ln(N_{ijt}) = \epsilon \ln(W_{jt}) + \epsilon \zeta \ln(Q_{it}) + \epsilon \kappa \tau_{ijt} + \ln(B_{it}E_{jt}D_{ijt})
\]

**Flow**

**Housing**

\[
\ln(Q_{it}) = \psi \ln (\sum_s N_{ist}) + \ln(C_{it})
\]

**H. Price**

**H. Eff.**

**Amenities**
Identification

$\epsilon$ is key: Location preference homogeneity $\equiv$ Local labor supply elast.

- Existing estimates use cross-sectional variation or calibrate (ARSW 2015; Monte, Redding, & Rossi-Hansberg 2018; Allen, Arkolakis, & Li 2018)
- Model (or parameter) dependent identification

Two special ingredients here:
1. Panel of average wage at place of work
2. Employment by industry at place of work

⇒ Tract-of-work based shift-share instrument (local Bartik variant)

Standard approach takes wage as D-by-year FE; performs poorly
(\epsilon) is key: Location preference homogeneity $\equiv$ Local labor supply elast.

- Existing estimates use cross-sectional variation or calibrate (ARSW 2015; Monte, Redding, & Rossi-Hansberg 2018; Allen, Arkolakis, & Li 2018)
- Model (or parameter) dependent identification

Two special ingredients here:

1. Panel of average wage at place of work
2. Employment by industry at place of work

$\Rightarrow$ Tract-of-work based shift-share instrument (local Bartik variant)

- Standard approach takes wage as D-by-year FE; performs poorly
- Here, much less reliance on Frechet commuting behavior
IV identification of $\epsilon$: labor supply & pref. homogeneity

$$E[\Delta z_{jt} \cdot \Delta \ln(E_{jt})] = 0, \forall j$$

- Changes in non-pecuniary workplace amenity orthogonal to shock

$\hat{\epsilon} = 1.83 \times (\pm 0.78)$

- Locations decisions heterogeneous and sticky even within cities
- Conditions on residential structure of city and commuting

- Smaller than cross-sectional trade-style estimates ($\sim 6.7$)
- More in line with LS elasticity (Falch 2010; Suarez Serrato & Zidar 2014)
IV identification of $\epsilon$: labor supply & pref. homogeneity

$$E[\Delta z_{jt} \cdot \Delta \ln(E_{jt})] = 0, \forall j$$

- Changes in non-pecuniary workplace amenity orthogonal to shock

Flexible assumption, as compared with literature

- Weaker assumption than standard city-level Bartik
- Permits variation in workplace amenities (unlike ARSW)
- Does not require correct travel costs (unlike AAL, ARSW, MRR-H)
- Does not condition on model components being correct (unlike MRR-H)
IV identification of $\varepsilon$: labor supply & pref. homogeneity

$$\mathbb{E}[\Delta z_{jt} \cdot \Delta \ln(E_{jt})] = 0, \forall j$$

- Changes in non-pecuniary workplace amenity orthogonal to shock

Flexible assumption, as compared with literature

- Weaker assumption than standard city-level Bartik
- Permits variation in workplace amenities (unlike ARSW)
- Does not require correct travel costs (unlike AAL, ARSW, MRR-H)
- Does not condition on model components being correct (unlike MRR-H)

$\hat{\varepsilon} = 1.83^* \pm 0.78$

- Locations decisions heterogeneous and sticky even within cities
  - Conditions on residential structure of city and commuting
- Smaller than cross-sectional trade-style estimates ($\sim 6.7$)
- More in line with LS elasticity (Falch 2010; Suarez Serrato & Zidar 2014)
**Bonus!** Interact with geography to plausibly identify all elasticities!
Other elasticities IV estimates

\[ \mathbb{E}[\Delta z_{jt} \cdot \Delta \ln(C_{it})] = 0, \forall i \neq j: \text{housing supply} \]
- Local adaptation of Saiz (2010); Guerrieri, Hartley, Hurst (2013)

\[ \frac{1}{\hat{\psi}} \approx 0.62^{**} (\pm 0.17): \text{Housing supply elasticity} \]
- Less elastic than longer run median across US cities (Saiz 2010)
- CA has inelastic housing supply (Quigley & Raphael 2005)

Reasonable estimates for other elasticities

\[ \mathbb{E}[\Delta z_{j't} \cdot \Delta \ln(B_{it}D_{ijt})] = 0, \forall ij' \neq ij: \text{housing demand} \]

\[ \mathbb{E}[\Delta z_{j't} \cdot \Delta \ln(A_{jt})] = 0, \forall j' \neq j: \text{labor demand} \]
Welfare effects in GE

Model suggests structural interpretation of residuals:

- Define Proximity_{it} ∈ [0, 1], so 0 → station > 500m away
- Estimate the effect of transit on these primitives, e.g.

\[
\hat{\ln(B_{it})} = \lambda^B \text{Proximity}_{it} + \delta_i + \varepsilon^B_{it}
\]

No evidence of non-commuting effects
Welfare effects in GE

Model suggests structural interpretation of residuals:

▶ Define Proximity\(_{it} \in [0, 1]\), so 0 → station > 500m away
▶ Estimate the effect of transit on these primitives, e.g.

\[
\hat{\ln(B_{it})} = \lambda^B \text{Proximity}_{it} + \delta_i + \varepsilon^B_{it}
\]

No evidence of non-commuting effects

Some evidence of small decreases in automobile commute times for routes ≤2km from stations, no effect ∈ (2, 4] km out
▶ Can include in welfare calcs (given η) to provide bounds

Use structural effects \{\lambda^A, \lambda^B, \lambda^C, \lambda^D, \lambda^E\} and elasticities (LS, LD, HS, HD) to simulate counterfactual model and determine welfare
▶ \(\epsilon\) key to valuation
Welfare effects of system in 2000 (in $2016)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.680</td>
<td>0.680</td>
<td>0.680</td>
<td>0.680</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>1.830</td>
<td>1.830</td>
<td>1.830</td>
<td>1.830</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.650</td>
<td>0.650</td>
<td>0.650</td>
<td>0.650</td>
</tr>
<tr>
<td>$\psi$</td>
<td>1.693</td>
<td>1.693</td>
<td>2.290</td>
<td>2.290</td>
</tr>
<tr>
<td>$\epsilon_k$</td>
<td>-</td>
<td>-0.020</td>
<td>-</td>
<td>-0.020</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in fundamentals</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda^D$, O &amp; D contain station</td>
<td>0.146</td>
<td>0.146</td>
<td>0.146</td>
<td>0.146</td>
</tr>
<tr>
<td>$\lambda^D$, O &amp; D &lt;250m from station</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
</tr>
<tr>
<td>$\lambda^\tau$, O &amp; D &lt;2km from station</td>
<td>-</td>
<td>-0.033</td>
<td>-</td>
<td>-0.033</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Closed Economy</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual $\Delta$ in welfare</td>
<td>0.051%</td>
<td>0.069%</td>
<td>0.051%</td>
<td>0.069%</td>
</tr>
<tr>
<td>(in millions of $$2016$)</td>
<td>108.9</td>
<td>145.7</td>
<td>108.9</td>
<td>145.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Open Economy</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population $\Delta$</td>
<td>0.109%</td>
<td>0.146%</td>
<td>0.106%</td>
<td>0.141%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic impact</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Op. subsidy + capital cost (6%, 30yy)</td>
<td>-$797 mil.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Op. subsidy + capital cost (5%, 50yr)</td>
<td>-$641 mil.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Op. subsidy + capital cost (5%, $\infty$)</td>
<td>-$597 mil.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Op. subsidy + capital cost (2.5%, $\infty$)</td>
<td>-$380 mil.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operation subsidy only</td>
<td>-$162 mil.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Welfare effects of system, other margins

Benefits $<$ Costs

Other margins?

► If Fundamental Law of Congestion doesn’t take hold (or slow):
  - Air pollution benefits, roughly $\sim$182 million per year (using generous estimates from Gendron-Carrier et al. 2018)
  - Congestion benefits already incorporated, though I find/use smaller $\kappa$ than (Anderson 2014)

► Non-rail or non-commuter benefits
  - Equity: Unemployed/injured? Elderly/school
  - Better bus integration and service
  - Other trips?

► Dynamics: adjustment might be slower than decadal
  - Challenge to measuring: (i) data changed, (ii) network grew
  - Best attempt at this, assuming all changes due to habituation, yields additional $\sim$70 million per year by 2015
Evaluating some assumptions in the new urban EG literature

1. How well do standard implementations of “market access” reflect observed commuting behavior?
   - Modeled commuting ignores persistent, pair-specific factors
   - Market access terms indicate smoother geography and interventions

2. Are measures of gravity reasonable?

3. How do model-derived wages accord with observed wages?
   - Not very well
   - Model-derived wages actually reflect population
Market Access

How does market access compare with direct commuting flow measure?

- Define relative change in accessibility from residential places

\[
\Delta CF = \frac{\sum_s N_{is}}{\sum_s (1 - \lambda D T_{is}) N_{is}} - 1, \quad \Delta MA = \frac{\sum_s e^{-\tilde{\kappa} \tau_{is}} Y_s}{\sum_s e^{-\tilde{\kappa} \tau_{is}} (1 - \lambda D T_{is}) Y_s} - 1
\]

Benefits of CF terms:
1. No need to know \(\tilde{\kappa}\) or \(\tau\)
2. Preserves heterogeneity; idiosyncratic factors (besides distance) determine commuting
3. “Observed” accessibility

Benefits of MA terms:
1. Can scrape/model travel time data
2. Smooths spatial economy, like spatial weights? (Sp. E/metrics)
3. “Potential” accessibility
Market Access

How does market access compare with direct commuting flow measure?

Define relative change in accessibility from residential places

$$\Delta CF = \frac{\sum_s N_{is}}{\sum_s (1 - \lambda^D T_{is}) N_{is}} - 1 \quad \Delta MA = \frac{\sum_s e^{-\tilde{\kappa} \tau_{is}} Y_s}{\sum_s e^{-\tilde{\kappa} \tau_{is}} (1 - \lambda^D T_{is}) Y_s} - 1$$

Benefits of CF terms:

1. No need to know $\tilde{\kappa}$ or $\tau$
2. Preserves heterogeneity; idiosyncratic factors (besides distance) determine commuting
3. “Observed” accessibility

Benefits of MA terms:

1. Can scrape/model travel time data
2. Smooths spatial economy, like spatial weights? (Sp. E/metrics)
3. “Potential” accessibility
Gravity and $\epsilon \kappa$

Consider panel gravity equation:

$$\ln(N_{ijt}) = \theta_{it} + \omega_{jt} + \delta_{ij} - \tilde{\kappa} \tau_{ijt} + u_{ijt}$$

1. $R^2$ without $\delta_{ij}$ is 0.20, $R^2$ with $\delta_{ij}$ 0.80
   - Time-invariant characteristics of pairs $\gg$ changes in travel time

2. Now run (1) excluding $\tau_{ijt}$ and estimate $\delta_{ij}$

   $$\hat{\delta}_{ij} = \alpha - \tilde{\kappa} \tau_{ij} + u_{ij}$$

   - $R^2 \approx 0.20$, travel time $\ll$ time-invariant determinants of flows
Gravity and \( \epsilon K \)

Consider panel gravity equation:

\[
\ln(N_{ijt}) = \theta_{it} + \omega_{jt} + \delta_{ij} - \tilde{\kappa} \tau_{ijt} + u_{ijt}
\]

1. \( R^2 \) without \( \delta_{ij} \) is 0.20, \( R^2 \) with \( \delta_{ij} \) 0.80
   - Time-invariant characteristics of pairs \( \gg \) changes in travel time

2. Now run (1) excluding \( \tau_{ijt} \) and estimate \( \delta_{ij} \)

\[
\hat{\delta}_{ij} = \alpha - \tilde{\kappa} \tau_{ij} + u_{ij}
\]

- \( R^2 \approx 0.20 \), travel time \( \ll \) time-invariant determinants of flows

Different estimates of \(-\tilde{\kappa}\):

<table>
<thead>
<tr>
<th></th>
<th>LA 1-yr</th>
<th>LA Panel</th>
<th>LA 2-step</th>
<th>ARSW Gravity</th>
<th>ARSW GMM</th>
<th>MRR-H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.053</td>
<td>0.000</td>
<td>-0.024</td>
<td>-0.077</td>
<td>-0.099</td>
<td>(-0.1)</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.002)</td>
<td></td>
</tr>
</tbody>
</table>
Model wages vs. observed wages

(Step 1) \[ \ln(N_{ij}) = \omega_j + \theta_i - \epsilon \kappa \tau_{ij} + d_{ij} \]

(Step 2: Standard) \[ \omega_j = \epsilon w_j \]

(Step 2: Here) \[ \omega_{jt} = \epsilon w_{jt} + e_{jt} \]
\[ 0 = \mathbb{E}[\Delta z_{jt} \cdot \Delta \ln(E_{jt})] \]
Model wages vs. observed wages

(Step 1) \[ \ln(N_{ij}) = \omega_j + \theta_i - \epsilon \kappa \tau_{ij} + d_{ij} \]

(Step 2: Standard) \[ \omega_j = \epsilon w_j \]

(Step 2: Here) \[ \omega_{jt} = \epsilon w_{jt} + e_{jt} \]

\[ 0 = \mathbb{E}[\Delta z_{jt} \cdot \Delta \ln(E_{jt})] \]
Summary

Develop new data sources to estimate effects of LA Metro:

▶ **Positive** effect on commuting between connected tracts
▶ Little adjustment on other margins

Carefully identify elasticities that populate econ. geo. model

▶ New identification strategy based on tract-level shift-share instrument
▶ Local stickiness, limited mobility even within city
▶ Permits more retention of unmodeled heterogeneity

Calculate welfare benefits of LA Metro

▶ Significant benefits, but **costs are larger**
▶ Even after 25 years...

Critically examine urban EG modeling
Thank you