The effects of transportation infrastructure on cities: A review of the evidence

Lectures for EAERE Summer school, Venice, 2009

Matthew Turner

updated 10 July 2009

Outline

- Introduction
- Econometrics - estimating equations
- Econometrics - inference
- Stylized facts
- Papers
- Conclusion

Questions:

- What are the effects of roads and transit on land use, driving, and economic development?
- How big a road network should we build in Kuala Lampur, and what should it look like?

We would like to know the effects of roads, and also understand their implications for welfare.

Problems

With experimental data, there is no problem. With observed data, we confront the following problems:

- Roads and transit infrastructure are very durable, and there are many margins of adjustment to new infrastructure: migration, production, trade, modal choice
- Assignment of infrastructure is not random.
- Welfare is always hard (but we can use land rent) and anything that makes a city more attractive is probably good (caveat systems of cities vs partial equilibrium).

Econometric problem

- Cities indexed by $i \in \{1,..,N\}$.
- time $t \in \{1,...,T\}$.
- y dependent variable of interest, VMT, population, productivity, rent
- R explanatory variable of interest, Roads, Transit.
- Other controls X
- Unobserved characteristics ϵ
- $\Delta_t y = y_t - y_{t-1}$.

Our problem is to establish the nature of the causal relationship between R and y, if such exists.

Estimating equation

- Choice of dependent variable?
- Static equilibrium,

$$y_{it} = AR_{it} + BX_{it} + \epsilon_{it}, \tag{1}$$

or first difference equivalent,

$$\Delta y_{it} = A\Delta R_{it} + B\Delta X_{it} + \Delta\epsilon_{it}, \tag{2}$$

- Growth equation,

$$\Delta y_{it+1} = AR_{it} + BX_{it} + \epsilon_{it}. \tag{3}$$

Which is right? NB: Growth in systems of cities is a mess.

The inference problem

Suppose we observe a relationship between y and R. Why does this relationship occur?

- y causes R?
- R causes y?
- sorting of people or activities?

The welfare implications are completely different. Our problem is to distinguish between these hypotheses.

Estimation: Cross-section equilibrium

$$y_{it} = AR_{it} + BX_{it} + \epsilon_{it}. \tag{4}$$

If R is not assigned to cities at random, we have $cov(R,\epsilon) \neq 0$ and cross-sectional inference does not give us an unbiased estimate of A.

Estimation: Cross-section growth equation

$$\Delta y_{it+1} = AR_{it} + BX_{it} + \epsilon_{it}. \tag{5}$$

If $cov(R_t, \epsilon_t) \neq 0$ and $cov(R_t, R_{t+1}) \neq 0$ worry that $cov(R_t, \epsilon_{t+1}) \neq 0$. Again, cross-sectional inference does not give us an unbiased estimate of A.

Pre-existing conditions need not be exogenous.

Estimation: First difference/Fixed effect

$$y_{it} = AR_{it} + BX_{it} + \delta_i + \mu_{it}. \tag{6}$$

Taking first differences,

$$\Delta y_{it} = A\Delta R_{it} + B\Delta X_{it} + \Delta\mu_{it}, \tag{7}$$

Thus, to estimate A we now require that $cov(R,\mu) = 0$, not the stronger $cov(R,\delta + \mu)$.

Here it is OK if roads are assigned to cities on the basis of some constant but unobserved feature of cities. If roads are assigned to cities after they receive a productivity shock, we confound the effect of roads with the effect of this shock. Response: IV, ad hoc checks.

Estimation: IV

Estimate the system

$$R_i = CX_i + DZ_i + \mu_i$$
$$y_{it} = A\widehat{R_{it}} + BX_{it} + \epsilon_{it}.$$

If The validity of this procedure hinges on finding a magic variable Z, which has the following two properties:

$$cov(Z,R|X) \neq 0$$
$$cov(Z,\epsilon|X) = 0.$$

Issues: Exogeneity is hard, Over-id test, weak instruments.

Can do IV of FD.

Estimation: Other

- Structural?
- Calibration?

Stylized facts:Roads

- Roads increase the productivity of adjacent land, partly at the expense of more remote areas.
- Roads increase the population density of nearby land.
- Roads change the composition of production and population.
- Roads disportionately attract wealthier people.
- Roads decrease density in cities.
- Adjustments to new roads occurs over generational time scales.
- An increase in roads causes an exactly proportional increase in city VMT.

Stylized facts: Transit

- Transit increases the population of cities(?).
- Transit disportionately attracts poorer people to cities.
- An increase in transit does not affect city VMT.

Bogart 2009 - turnpike trusts

- County and parish level data for circa 1700 and 1800 in England.
- y is land rent.
- R presence of turnpike trust $\Longrightarrow$ private improved toll road
- Z parish on major trade route.
- estimating equation:

$$\Delta y_{it} = A\Delta \text{Turnpike Indicator}_{it} + B\Delta \text{Observed Controls}_{it} + \Delta \mu_{it}.$$

A turnpike increases land rent within 5km by about 20%.

Issue: turnpikes are assigned to places that are growing anyway – most turnpikes are built during the study period.

Bogart 2009 - continued

Estimate

$$\begin{aligned}\Delta y_{it} = A\Delta\ \widehat{\text{Turnpike Indicator}}_{it} + \\ B\Delta\text{Observed Controls}_{it} + \Delta\mu_{it}\end{aligned}$$

$$\begin{aligned}\Delta\widehat{\text{Turnpike Indicator}}_{i} = A\ \text{Old trade route}_{i} + \\ B\ \text{Observed Controls}_{i} + \gamma_{i}.\end{aligned}$$

This leads to a higher estimate of the effect of turnpikes on land rent.

Good:

- Measures land rent directly. If we know how much land is affected and how much a road costs we can calculate welfare!
- Makes a serious attempt at dealing with the endogeneity of roads.

Bad: Trade increased during the early parts of the industrial revolution. This means land rents should increase along trade routes independent of whether they had roads.

If so, 'on a trade route' could contribute land rents independent of its effect on roads thereby invalidating the instrument.

N.B.: It is very hard to sell IV.

Duranton and Turner 2008 - Growth and roads

- 275 US msa's in 1980,1990,2000.
- y is: population growth, employment growth, share poor.
- R lane km of 1980 major roads, 1983 count of large buses.
- Z km 1947 highway plan, km 1898 rail routes, km exploration routes (?), share of democratic voters in 1972.
- Estimating equations:

$$\Delta y = A\widehat{R} + BX + \epsilon$$

$$R = CZ + DX + \delta$$

Theory tells us that population and employment growth are directly related to land rent, so that these results also inform us about the effect of roads on welfare.

Results:

- 10% increase in msa road miles in 1980 causes a 2% increase in population 1980-2000
- 10% increase in msa road miles in 1980 causes a 1.3% increase in population 1980-1990, .7% 1990-2000. Total effect at constant decay rate 2.9%.
- 10% increase in 1898 rail associated with 80 year population growth of 3.5%.
- 10% increase in msa road miles in 1980 causes a 2% increase in employment 1980-2000
- 10% increase in large buses in 1980 causes a 0.7% increase in employment 1980-2000
- roads attract wealthier people, buses attract poorer.

- Building roads is a more expensive way to attract people than is buying buses.
- mechanism appear to be commuting rather than market access.
- IV estimates LARGER than OLS.
- Roads do not affect industrial composition.

Michaels 2008 – trade and specialization

- Non-msa (rural) counties in US, 1972,77,82,87,92.
- y earning in trucking and warehousing, retailing
- R indicator for presence of interstate highway in county
- Z 1947 highway plan, indicator for county is due east or west or south or north of the nearest big city (as opposed to southwest etc.)
- estimating equation:

$$y_{it} = A\widehat{R}_{it} + BX_{it} + \epsilon_{it}$$
$$R_{it} = CZ_i + DX_{it} + \delta_{it}$$

- IV estimates bigger than OLS.

- roads increase county earnings in trucking in the contemporaneous five year period, 1-16% depending on the period and specification.

  This suggests that roads either increase the ability of these rural counties to trade with the outside, or cause them to specialize in moving goods.

- Addition of a highway causes an 8-10% increase in retail earnings in the period in which roads were added.

  This suggests that people in these counties are spending more, which in turn suggests that they are earning more, causing a large increase in real wages over the course of just a few years.

- Small effect on the returns to skill in manufacturing and on the share of skilled workers in manufacturing. (he was looking in the wrong place?)

Chandra and Thompson 2000 - roads and productivity

- Non-msa (rural) counties in US, 1969-93.
- y earnings in 1 digit SIC codes
- counties are either: road, adjacent, control.
- R_{it}^T indicator for presence of T year old interstate highway in county i or adjacent county to i at time t.
- estimating equation:

$$y_i t = AR_{it}^T + \delta_i + \gamma_t + B \text{ Other industry earning in state}_t + \epsilon_{it}$$

- Estimate for 1 digit SIC codes for Road and control, Road, Adjacent and Control, Adjacent and control.

- $cov(R^T, \epsilon|.) \neq 0 \Longrightarrow$ serial correlation of errors, not found. N.B.: another solution to endogeneity is to control for everything – maybe they've done it.
- Chandra and Thompson 2000 and Michaels 2008 get same estimated effect on retail. Both are right, or wrong in the same way.

Chandra and Thompson 2000 - continued

Results:

- Overall economic activity in road or adjacent counties increases by about 4% by 25 years after the road is constructed.
- Adjacent counties lose about 3% of earnings 25 years after a road is built, while road counties gain 8-10%.
- Construction and farming are harmed in road counties. Retail, services, business services and government all grow, manufacturing unchanged.
- manufacturing and business services increases in adjacent counties and most other industries decrease.
- In adjacent counties it takes several years before government earnings increase, but business services increase almost immediately and almost monotonically over 25 years.

NB: This suggests that skill bias that Michaels looked for may be present. Skill bias is also broadly consistent with Duranton and Turner 2008a finding that roads disportionately attract wealthy people.

Fernald 1999 - roads and productivity

- Annual US national data productivity, by industry 1947-95.
- y_{jt} productivity in industry j year t.
- R_t dollar value of roads in year t.
- Estimating equation (loose),

$$\Delta y_{jt} = A \Delta R_t + \epsilon_{jt}$$

for each j (simultaneously) and for aggregate. To deal with endogeneity use cross equation restriction. Also allow A to vary before and after 1973.

Fernald 1999 - cont'd

- Interstate highway investment has a high return in national productivity during the period before 1973, about 10% per year, and an almost zero return after. Why?
  - highway network was mostly complete after 1973, so marginal improvements to the network were no longer as productive
  - the interstate highway network became progressively more crowded after 1972, and thus (probably) more expensive to use.

Fernald finds no productivity effect during the period that Chandra and Thompson, Michaels, and Duranton and Turner find an effect! Why?

- Chandra and Thompson, Baum-Snow, and Duranton and Turner find near zero effects of roads in OLS. Maybe Fernald doesn't fix endogeneity?
- All other papers consider longer periods than the single years considered in Fernald. Maybe Fernald needs to wait longer to see an effect?
- Fernald uses national data. All others use city or county level data. If city or county level road investments only move economic activity from one place to another, they do not affect aggregate productivity. Chandra and Thompson provide some support for this.

- The fact that Fernald's 1973 cut-off coincides with an oil shock does and a recession suggests that the decrease in road investments may have been caused by the decrease in economic activity, rather than the converse. (my favorite)

The extent to which aggregate effects differ from marginal, city level effects merits further investigation.

Baum-Snow 2007

Census tract exercise:

- Each MSA census tract in 1970 and 1990.
- y population density;
- R distance to the nearest limited access highway
- X distance to the CBD, msa
- estimating equation

$$y = AR + BX + \epsilon$$

  and first difference.

- Population density decreases by 1-2% per mile from highway.
- no control for endogeneity. Estimates too small?
- Do highways increase average density or decrease it?

Baum-Snow 2007

MSA exercise:

- US MSA's in 1950, 1970 and 1990.
- y CBD population density;
- R Number of highway rays to CBD.
- Z rays in 1947 highways plan
- X Demographics for 1950-1990.
- estimating equation

$$\begin{aligned} y_{it} &= A\widehat{R}_{it} + BX_{it} + \delta_i + \epsilon_{it} \\ R_{it} &= CZ_i + DX_{it} + \gamma_{it} \end{aligned}$$

Baum-Snow 2007 cont'd

- IV estimates bigger than OLS.
- 9% per ray decrease in CBD density from 1950-1990.
- Mean rays between 2 and 3, so highways caused 18-27% decline in central city population density – most of observed decrease.
- 'too much' migration out of CBD since congestion is unpriced?

Burchfield et al 2006

- Satellite data on MSA scatterdness of development in 1976 and 1992
- y, scatterdness of new development between 1976 and 1992, all development in 1992.
- R, 1902 streetcars per capita, 1980 roads per km^2.
- estimating equation

$$\Delta y = AR + BX + \epsilon$$

- Roads don't matter (endogeneity ?)
- Streetcars $\Longrightarrow$ more density. What does this mean?

N.B.: Baum-Snow and Burchfield are some of our only evidence on infrastructure and urban form.

Bento et al 2005

- 114 urbanized areas circa 1990, 1990 NHTS.
- y, modal choice, car ownership, vmt for individual (really mean individual)
- R miles of interstate per hectare, count of large buses
- X individual demographics
- estimating equation (non-linearities)

$$\begin{aligned}\text{Individual Travel Behavior}_{ij} = & A\text{Roads,Large buses}_j \\ & + B\text{individual controls} + \epsilon_{ij}\end{aligned}$$

where i indexes individuals and j indexes urbanized areas.

- results: road density is positively associated with the probability of riding a bus, increases to vmt, and increases to the likelihood of car ownership.
- No causal inference should be made. It is possible that (1) infrastructure provision is determined by some unobserved feature that also affects transportation behavior, or (2) that individuals sort on the basis of infrastructure.
- These are some of the only estimates available.

Duranton Turner 2008b - Fundamental Law...

- 275 US MSA's, 1980,1990,2000
- y city vmt, interstate, urban interstate, urban arterial, total HH vmt, truck vkt
- R lane km interstate, urban interstate, urban arterial, large buses
- Z km 1947 planned interstate, 1898 rail routes, 1820-35 exploration routes, share 1972 democratic vote.
- estimating equation,

$$\text{vmt}_i = A\text{lane miles}_i + B\text{controls} + \epsilon_i$$

as cross section, first difference, IV. The fundamental inference problem that they must resolve is that infrastructure is not assigned to cities at random.

Duranton Turner 2008b - cont'd

- Road elasticity of vmt is positive, and almost never statistically different from one. That is, vmt increases in direct proportion to road miles.
- This appears to results from traffic creation, not from substitution between road networks.
- New traffic appears to has three main sources.
  - Changes to individual driving behavior.
  - Migration to road rich cities
  - changes to trucking
- The number of large buses has no effect on vmt. Transit increases capacity in exactly the same way as does increases in lane miles.

Stylized facts:Roads

- Roads increase the productivity of adjacent land, partly at the expense of more remote areas.
- Roads increase the population density nearby land.
- Roads change the composition of production and population.
- Roads disportionately attract wealthier people.
- Roads decrease density in cities.
- Adjustments to new roads occurs over generational time scales.
- An increase in roads causes an exactly proportional increase in city VMT.

Stylized facts: Transit

- Transit increases the population of cities.(?)
- Transit disportionately attracts poorer people to cities.
- An increase in transit does not affect city VMT.

Conclusion

- We can assess many of the long run affects of road.
- We have a very limited ability to do welfare analysis.
- Areas for future research
  - Understanding the effects of public transit
  - Understanding effects of built environment on travel behavior
  - Understanding the effects of development outside of the US.