Railroads of the Raj

Dave Donaldson

Samuel Kortum (5/2/2009)
Contributions

- Building railroads raised real income in India by 20%.

- Introduces a powerful method to calculate transport costs.

- Reformulates Ricardian trade model for agriculture.

- New standard for connecting vast amounts of historical data to a structural model.

- Template for modelling global effects of climate change?
Transport Costs and How They Fell

- Pricing equation for a good shipped from origin $o$ to destination $d$:
  \[ P^o_{dt} = P^o_{ot}T^o_{odt}. \]

- Iceberg specification works for Armington, Monopolistic competition, or Ricardian model.

- Salt was “branded” according to 8 distinct origins $o$ in India.

- Collects annual data from 1861-1930 on price $P^o_{dt}$ of $o$ salt in each of 124 districts ($o$ is not one of districts).
• Data on transport network $R_t$ from 8 origins to 128 destinations, rail, road, river, and coast.

• Assume constant coefficients for cost per unit distance relative to rail: $\alpha_{rail} = 1$, $\alpha_{road}$, $\alpha_{river}$, and $\alpha_{coast}$.

• Use Dijkstra’s algorithm to calculate $LCR(R_t, \alpha)_{odt}$: rail distance from $o$ to $d$ at date $t$ that mimics actual best combination of transport.

• Estimates: $\alpha_{rail} = 1$, $\alpha_{road} = 7.3$, $\alpha_{river} = 3.4$, $\alpha_{coast} = 3.2$, and an elasticity $\delta = 0.25$ of iceberg cost with respect to $LCR$.

• I’d like to interpret $\delta$ as rail cost per mile per Dollar shipped:

$$\ln P_{dt}^o = \beta_{ot}^o + \ln [1 + \delta * LCR(R_t, \alpha)_{odt}].$$
Incredible Detail of Data

- For 239 districts in India, 1861-1930 (first track laid in India, 1853).

- Daily rainfall data by district (India’s economy dominated by non-irrigated agriculture).

- Bilateral trade between 45 trade blocks and 23 foreign countries for each of 85 commodities.

- Yields and land under cultivation by district for each of 17 principle crops: barley, chick pea, corn, cotton, indigo, sorghum, opium, rice, sesame, sugar, lentils, wheat, .....
Model 1

- Utility of the representative agent in district $d$

$$\ln U_d = \sum_{k=1}^{K} \mu_k \left[ \frac{\sigma_k}{\sigma_k - 1} \right] \ln \int_{0}^{1} \left[ C_d^k(j) \right]^{(\sigma_k - 1)/\sigma_k} dj.$$ 

- Ricardian land yield for variety $j \in [0, 1]$ of good $k$:

$$\Pr[Z_d^k(j) \leq z] = e^{-A_d^k z^{-\theta_k}}$$

- Price index, assuming $\sigma_k < \theta_k + 1$:

$$P_d^k = \chi^k \left[ \sum_{o=1}^{D} A_o^k(r_o T_{od}^k)^{-\theta_k} \right]^{-1/\theta_k}.$$
• Trade share

\[ \pi^k_{od} = \frac{X^k_{od}}{X^k_d} = \lambda^k_3 A^k_0 \left( \frac{r^k_{od} T^k_{od}}{P^k_d} \right)^{-\theta_k}. \]
Key Relationships

• Bilateral trade equation to estimate $\theta_k$ and relate $A^k_o$ to rainfall:

$$\ln X^k_{od} = \ln \lambda^k_3 + \ln A^k_o - \theta_k \ln T^k_{od} + \theta_k \ln P^k_d + \ln X^k_d.$$ 

• Back out land rent with good $k$ as numeraire:

$$\frac{r_o}{P^k_o} = \left( \frac{\pi^k_{oo}}{\lambda^k_3 A^k_o} \right)^{-1/\theta_k}.$$ 

• Gives real income per unit of land:

$$W_o = \frac{r_o}{P_o} = \prod_{k=1}^{K} \left( \frac{r_o}{P^k_o} \right)^{\mu_k}.$$
Hence:

\[ \ln W_o = \Omega + \sum_{k=1}^{K} \frac{\mu_k}{\theta_k} \ln A_o^k - \sum_{k=1}^{K} \frac{\mu_k}{\theta_k} \ln \pi_{oo}^k \]

- Gains from: intra-good trade, inter-good trade, inter-temporal trade.
Interpretation

- It's a structural model, so we can think through what it means.

- Random yields at the level of varieties $j$ smooths out the Ricardian model in a useful way.

- Best interpretation: seed varieties (not variation in land quality or the weather).

- But do consumers really care what variety of chick pea they're eating?
• I’d expect $\sigma_k = 20$, but then $\theta_k$ must also be large for a well defined price index for chick peas.

• Problem: the estimate of $\theta_k$ are more like 3 to 5 (could use more detail in the paper).
Solution I

- Go for the case with very high $\sigma_k$ and $\theta_k$.

- Almost all the gains from trade come from inter-good trade, as in a traditional Ricardian model.

- The continuum of varieties is more of a device to make the model behave nicely.

- Figure out why estimates of $\theta_k$ are downward biased.
Solution II

- Aggregate! Why run bilateral trade equations with 85 commodities?.

- The point of varieties is to not fuss with all the detail: (i) grains, (ii) other.

- Aggregation will eliminate a huge amount of measurement error.

- Now we’d expect to see much lower estimates of $\theta_k$, perhaps similar to the current ones.

- Much of the gains from trade will come from intra-good trade.