The Wealth of Cities

David Comerford[§] David Pugh[†]

[§]University of Strathclyde
[†]King Abdullah University of Science and Technology

ERSA 2019
Lyon, France

Introduction

- Well known that larger places are more productive
- This issue has been investigated from an empirical point of view (Bettencourt et al (2007), __ (2010), Bettencourt & West (2010), & Bettencourt (2013): many socio-economic indicators display “super-linear” scaling) and from economic theory point of view (e.g. economic geography of Fujita, Krugman & Venables, 1999)
- Super linear scaling is essentially increasing returns to scale (IRS)
- The economic theory approach links to the economics of trade (how large are the gains from trade?), & to macroeconomics (how much do macroeconomies multiply small shocks to produce observed fluctuations?) and provides micro-foundations for IRS
- Zhelobodko et al (2012) propose an increasing “relative love for variety (RLV)” which generates pro-competitive effects i.e. a reduction in IRS
- In this paper, we use both theory and empirics to investigate this issue

Data

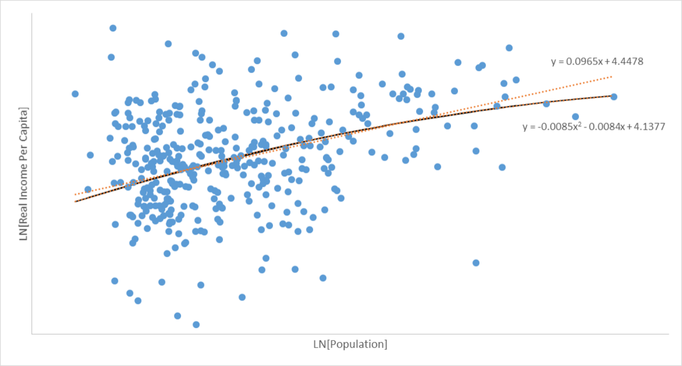


Figure: Metropolitan Statistical Areas in the US, from US Bureau of Economic Analysis (BEA): `https://www.bea.gov/data/gdp/gdp-metropolitan-area`

- We use MSA GDP, adjusted by MSA specific price deflator, to generate real economic size, E_h, for each location h, for 381 MSAs
- Divided by MSA population generates real GDP per capita, y_h

Data

- Upward trend: 10% increase in population, L_h, associated with, on average, around a 1% increase in real income per capita, y_h.
- But this is maybe not scale invariant, with returns to scale apparently reducing as cities become larger
- This is purely empirical and very naive (e.g. no consideration of endogeneity etc). Empirical model explains very little of the variation in the data
- Let's discipline using a theoretical model
- To do so we augment our dataset with MSA level human capital data, H_h, from: `https://www.anderson.ucla.edu/centers/ucla-anderson-forecast/projects-and-partnerships/city-human-capital-index`
- Note $\exists$ +ve correl between log human capital and log population (correl coeff = 25%)

Islands Model

- All models in this paper are variations upon Krugman (1980)
- First model features no trade, and so cities are effectively "Islands"
- Consumers in city h have CES preferences with elasticity of substitution θ_h
- Firms are monopolistically competitive. Pay fixed cost to create firm with marginal cost of production $1/\phi$ (in units of effective labour). ϕ is common across cities
- Free entry, so firms are created until profits driven to zero
- Goods and labour markets clear: wages that are paid to the entire effective labour supply are fully spent buying firms' output
- Cities differ only in their effective labour supply, S_h, and (potentially) in their elasticity of substitution, θ_h

Islands Model: Detail

- Demand function given CES preferences: $q(p) = p^{-\theta_h} E_h$
- Monopolistic competition $\Rightarrow$ price is constant markup over marginal cost $p = \frac{w_h}{\phi}\frac{\theta_h}{\theta_h - 1}$
- Income equals expenditure: $w_h S_h = y_h L_h$
- Zero profit condition: $w_h = \frac{y_h L_h}{S_h} = \frac{1}{\theta_h}\left(\phi\left(\theta_h - 1\right)\right)^{\frac{\theta_h - 1}{\theta_h}} E_h^{\frac{1}{\theta_h}}$

Islands Model

- Such a model produces a real income per capita in city h of:

$$y_h = \frac{1}{\theta_h}\left(\phi\left(\theta_h - 1\right)\right)^{\frac{\theta_h - 1}{\theta_h}} \frac{S_h}{L_h} E_h^{\frac{1}{\theta_h}} \tag{1}$$

- Income p.c. depends on effective labour p.c., $\frac{S_h}{L_h}$, & market size, E_h
- If we assume that:
  - $\theta_h = \theta, \forall h$ (i.e. const $\theta =$ standard Krugman trade model w. no trade)
  - Effective labour, $S_h \sim H_h^{\beta} L_h$, where H_h is data on human capital
  - E_h, is an exogenous data series w.r.t. estimation of y_h (!!)
- then we can estimate

$$\ln y_h = \gamma + \beta lnH_h + \frac{1}{\theta} \ln E_h$$

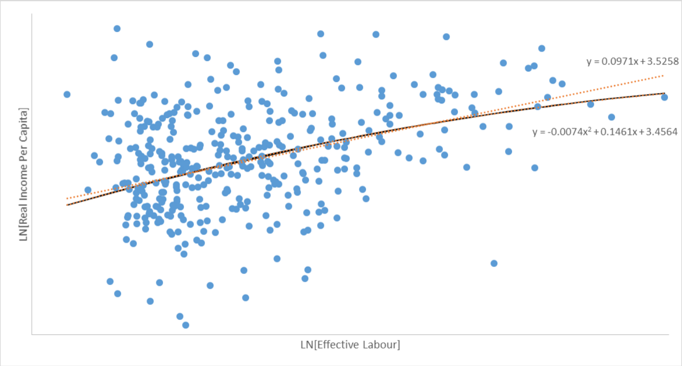
with

$$\gamma = \ln\left[\frac{1}{\theta}\left(\phi\left(\theta - 1\right)\right)^{\frac{\theta - 1}{\theta}}\right] \Rightarrow \phi$$

Islands Model: Results

- Clearly this ignores endogeneity, but estimates are s.t. $\ln \hat{y}_h = \hat{\gamma} + \hat{\beta} \ln H_h + \frac{1}{\hat{\theta}} \ln E_h$ does minimise $\sum_h (\ln y_h - \ln \hat{y}_h)^2$
- Results are: $\hat{\theta} = 15.18$; and $\hat{\beta} = 1.91$
- A fully internally consistent model (rearr. Eq (1)) is described by: $E_h = y_h L_h \sim S_h^{\frac{\theta}{\theta-1}}$, so value for $\hat{\theta}$ translates to exponent of 1.071
- If this model was true DGP, then increase in city population of 10% causes increase in income per capita of around 0.7%
- This is high for θ / low for degree of increasing returns to scale
- Approaches used in international trade suggest values for θ in range 8 (Eaton & Kortum, 2002) to 4 (Simonovska & Waugh, 2014)
- Approaches used in international macroeconomics produces even lower values: $\theta \in (2, 3)$ (Backus et al, 1994)
- Expect our approach to produce higher θ estimates since use of MSA specific price indices (incorporate congestion effects esp. property rental prices) dampens real incomes in highly prosperous cities

Islands Model: Results



- New x-axis: $\ln S_h = \ln L_h + \hat{\beta} \ln H_h$
- slope here is $\sim$ the same as the slope vs population
- 10% increase in S_h, (either from inc in L_h or H_h) associated with, on average, around a 1% increase in real income per capita, y_h.
- scale (in)variance similar: coeff on $(\ln S_h)^2$ is 0.74% c.f. 0.85% on $(\ln L_h)^2$

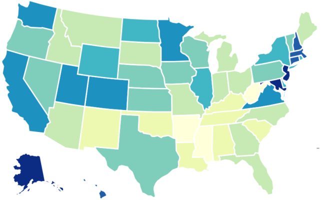
Islands Model: Variable θ_h

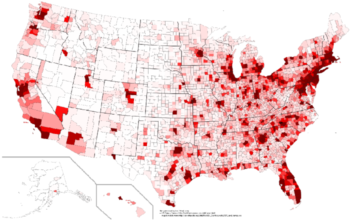
- Now postulate: $\theta_h = \hat{\theta} + \epsilon \left(\ln E_h - \ln \bar{E} \right)$
- Test model fit with parameters $\{\hat{\theta}, \hat{\beta}, \hat{\phi}\}$ from previous, and now also with small +ve or -ve ϵ. Is fit better with +ve or -ve?
- Hypothesis that degree of returns to scale reduce as city size increases consistent with $\epsilon > 0$ (Zhelobodko et al (2012)'s "Increasing RLV")
- Note also an impact of variable θ_h upon the "constant" term

$$\gamma_h = \ln \left[\frac{1}{\theta_h} \left(\hat{\phi} \left(\theta_h - 1 \right) \right)^{\frac{\theta_h - 1}{\theta_h}} \right]$$

  which is a decreasing (increasing) function of E_h for +ve (-ve) ϵ
- This accentuates impact on returns to scale from $\frac{1}{\theta_h} \ln E_h$ term
- Using goalseek to minimise squared errors, $\sum_h \left(\ln y_h - \ln \hat{y}_h \right)^2$, keeping $\{\hat{\theta}, \hat{\beta}, \hat{\phi}\}$ from previous estimation, get $\hat{\epsilon} = 0.50 > 0$
- If however we allow other parameters to vary, get $\hat{\epsilon} = 1.20$, plus new $\hat{\theta} = 13.37$, and essentially no change in $\hat{\beta}$ or $\hat{\phi}$

Connectivity Model: Generalised Krugman Trade Model





- Some correlation between population density and income per capita
- Theory suggests market size should correlate with productivity
- Places can be part of a large market either by being large, or by being well connected
- Generalise the model to include trade

Connectivity Model: Detail

- Demand function for a good consumed in location j:
  $q(p) = \left(\frac{p}{P_j}\right)^{-\theta_j} E_j$, where P_j is the price index in location j
- Monopolistic competition $\Rightarrow$ price chosen by a firm in h to sell into market j is constant markup over marginal cost $p_{hj} = w_h \frac{\delta_{hj}}{\phi} \frac{\theta_j}{\theta_j - 1}$
- $\delta_{hj} > 1, \forall j \neq h$ is the relative extra marginal cost experienced by a firm in h selling into j rather than into h (i.e. $\delta_{hh} = 1, \forall h$)
- Income equals expenditure in h: $w_h S_h = y_h L_h$
- Zero profit condition for firms in h:

$$\phi = \sum_j \left(\frac{1}{\theta_j - 1}\right) \left(\frac{\theta_j}{\theta_j - 1} \frac{1}{\phi}\right)^{-\theta_j} \left(\frac{y_h L_h}{S_h}\right)^{-\theta_j} \left(\frac{P_h}{P_j}\right)^{-\theta_j} \delta_{hj}^{1-\theta_j} E_j$$

- Define $M_h \equiv$ measure of firms in h, then labour market clearing in h:

$$S_h = \frac{M_h}{\phi} \sum_j \left(\frac{\theta_j}{\theta_j - 1}\right) \left(\frac{\theta_j}{\theta_j - 1} \frac{1}{\phi}\right)^{-\theta_j} \left(\frac{y_h L_h}{S_h}\right)^{-\theta_j} \left(\frac{P_h}{P_j}\right)^{-\theta_j} \delta_{hj}^{1-\theta_j} E_j$$

Connectivity Model: Detail

- Exports from h to j:

$$
\begin{aligned}
X_{hj} &= M_h q_{hj} p_{hj} \\
&= M_h P_h \left(\frac{\theta_j}{\theta_j - 1}\frac{1}{\phi}\right)^{1-\theta_j} \left(\frac{y_h L_h}{S_h}\right)^{1-\theta_j} \left(\frac{P_h}{P_j}\right)^{-\theta_j} \delta_{hj}^{1-\theta_j} E_j
\end{aligned}
$$

- Balanced trade in h, $\sum_j X_{hj} = \sum_j X_{jh}$, i.e.

$$
\sum_j M_h P_h \left(\frac{\theta_j}{\theta_j - 1}\frac{1}{\phi}\right)^{1-\theta_j} \left(\frac{y_h L_h}{S_h}\right)^{1-\theta_j} \left(\frac{P_h}{P_j}\right)^{-\theta_j} \delta_{hj}^{1-\theta_j} E_j
$$

$$
= \sum_j M_j P_j \left(\frac{\theta_h}{\theta_h - 1}\frac{1}{\phi}\right)^{1-\theta_h} \left(\frac{y_j L_j}{S_j}\right)^{1-\theta_h} \left(\frac{P_j}{P_h}\right)^{-\theta_h} \delta_{jh}^{1-\theta_h} E_h
$$

Connectivity Model: Constant θ

- Firstly simplify: assume constant $\theta \Rightarrow$ standard Krugman trade model
- Define effective demand experienced by city h as:

$$D_h = P_h^{-\theta} \sum_j P_j^{\theta} \delta_{hj}^{1-\theta} E_j$$

- Then get zero profit condition in the same form as the island model:

$$y_h = \frac{1}{\theta} \left(\phi \left(\theta - 1 \right) \right)^{\frac{\theta-1}{\theta}} \frac{S_h}{L_h} D_h^{\frac{1}{\theta}} \tag{2}$$

- Eliminating Ms using the labour market clearing condition, we also have an additional condition on Ps from the balanced trade condition:

$$P_h^{\theta} = \left(\frac{\theta^{\frac{\theta}{\theta-1}}}{\phi \left(\theta - 1 \right)} \right)^{\theta} \left[\sum_j S_j^{\theta} E_j^{1-\theta} \left(P_j^{\theta} \right)^{\frac{1-\theta}{\theta}} \delta_{jh}^{1-\theta} \right]^{\frac{\theta}{1-\theta}} \tag{3}$$

Connectivity Model: Constant θ

- Use latitude & longitude of each MSA to construct a (symmetric) matrix of bilateral distances, $d_{hj} = d_{jh}$
- Calibrate model to full dataset using following algorithm:
  - Suppose $\delta_{hj}^{1-\theta} = \exp(-d \times d_{hj})$
  - Choose some value for $d > 0$
  - Initially assume $\hat{P}_h^\theta = 1, \forall h$
  - Evaluate $D_h = \hat{P}_h^{-\theta} \sum_j \hat{P}_j^\theta \exp(-d \times d_{hj}) E_j$
  - Eq (2): obtain $\hat{\phi}, \hat{\beta}, \hat{\theta}$ by estimating $\ln y_h = \gamma + \beta \ln H_h + \frac{1}{\theta} \ln D_h$
  - Use Eq (3) to evaluate

$$P_h^\theta = \left(\frac{\hat{\theta}^{\frac{\hat{\theta}}{\hat{\theta}-1}}}{\hat{\phi}\left(\hat{\theta}-1\right)} \right)^{\hat{\theta}} \left[\sum_j \left(L_j H_j^{\hat{\beta}} \right)^{\hat{\theta}} E_j^{1-\hat{\theta}} \left(\hat{P}_j^\theta \right)^{\frac{1-\hat{\theta}}{\hat{\theta}}} \exp(-d \times d_{hj}) \right]^{\frac{\hat{\theta}}{1-\hat{\theta}}}$$

  - Iteratively adjust $\hat{P}_h^\theta$ (& hence D_h, $\hat{\theta}$, etc) to minimise $\sum_h \left(\frac{P_h^\theta}{\hat{P}_h^\theta} - 1 \right)^2$
  - Then repeat for different $d > 0$ until minimise $\sum_h (\ln y_h - \ln \hat{y}_h)^2$

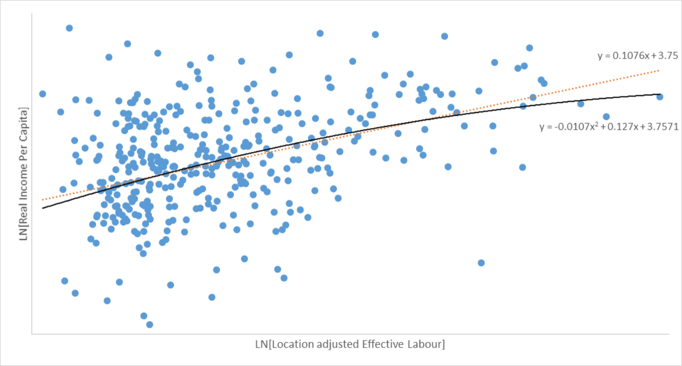
Connectivity Model: Constant θ, Results

- Produces results: $\{\hat{d} = 0.07, \hat{\theta} = 10.10, \hat{\beta} = 1.40\}$ - stronger returns to scale, and lower returns to human capital than Islands Model
- Implies mean value for $\min_j\{\delta_{hj}\} \sim 2$ i.e. firms' costs supplying nearest neighbours $\sim 2\times$ costs supplying their own markets
- Now define

$$C_h = \left(\frac{D_h}{E_h}\right)^{\frac{1}{\theta}}$$

  Then a fully internally consistent model (rearr. Eq (2)) is described by: $E_h = y_h L_h \sim (S_h C_h)^{\frac{\theta}{\theta-1}}$ i.e. C_h is like a connectivity adjustment to the factors of production available in a location
- Value for $\hat{\theta}$ translates to exponent of 1.110. If this model was true DGP, then increase in city population of 10% causes increase in income per capita of around 1.1% (c.f. 0.7% from Islands)
- Note $\exists$ -ve correl between log connectivity adj and log population (correl coeff $= -46\%$)

Connectivity Model: Constant θ, Results



- New x-axis: $\ln(S_h C_h)$
- slope here is $\sim$ the same as the slope vs population (slightly steeper)
- 10% increase in $S_h C_h$, (from inc in L_h, H_h, or C_h) associated with, on average, around a 1% increase in real income per capita, y_h.
- larger dependence on scale: coeff on $(\ln S_h C_h)^2$ is 1.07% c.f. 0.85% on $(\ln L_h)^2$

Connectivity Model: Variable θ_h

- Now postulate: $\theta_h = \hat{\theta} + \epsilon \left(\ln \hat{D}_h - \ln \bar{D} \right)$ using $\hat{D}_h$ from constant θ model calibration
- Test model fit with parameters $\{\hat{\theta}, \hat{\beta}, \hat{\phi}\}$ from previous, using prices from constant θ model calibration, $\hat{P}_h$, and now also with small $+$ve or -ve ϵ. Is fit better with $+$ve or -ve?
- Hypothesis that degree of returns to scale reduce as economic size increases consistent with $\epsilon > 0$ (Increasing RLV)
- We use Zero profit condition for firms in h:

$$\phi = \sum_j \left(\frac{1}{\theta_j - 1} \right) \left(\frac{\theta_j}{\theta_j - 1} \frac{1}{\hat{\phi}} \right)^{-\theta_j} \left(\frac{\hat{y}_h}{H_h^{\hat{\beta}}} \right)^{-\theta_j} \left(\frac{\hat{P}_h}{\hat{P}_j} \right)^{-\theta_j} \hat{\delta}_{hj}^{1-\theta_j} E_j$$

- We adjust $\hat{y}_h$ until the RHS of this equation equals $\hat{\phi}$
- We then evaluate $\sum_h (\ln y_h - \ln \hat{y}_h)^2$. We do this, following the results from the Islands Model, for $\epsilon = +0.50, 0, -0.50$
- The best fit is $\hat{\epsilon} = +0.50$, but note that this is not any real sort of calibration as we have not re-optimised on $\hat{P}$s or $\hat{\theta}$ etc

Quantitative Implications

- Suppose a full calibration of the model produced $\hat{\theta} = 10.10$ and $\hat{\epsilon} = 1.20$
- Consider then the impact on the MSAs around New York and Kansas City:
  - New York has a population of approximately 20m, while Kansas City has a population of around 2m
  - They have approximately equal human capital index scores and real incomes per capita
  - The model predicts that New York's real income per capita is around 22% higher because its real economic size is around 229% higher
- Obviously there is a lot of noise in real data, but we can suppose that the "errors" are approximately constant when we perform policy experiments
  - i.e. the model is better at predicting real income per capita changes even if it doesn't do such a good job of fitting real income per capita levels

Quantitative Implications

- Consider 2 scenarios:
  - Adding $\sim 100,000$ population to each city
  - Adding 5% population to each city ($\sim 1,000,000$ to New York, and $100,000$ to Kansas City)
- In the Constant θ model,
  - Scenario 1 causes an increase in real income per capita of 0.05% for New York, and 0.48% for Kansas City
  - Scenario 2 causes an increase in real income per capita of 0.48% for New York, and 0.48% for Kansas City
- In the Variable θ model,
  - Scenario 1 causes an increase in real income per capita of 0.04% for New York, and 0.50% for Kansas City
  - Scenario 2 causes an increase in real income per capita of 0.39% for New York, and 0.50% for Kansas City

Quantitative Implications

- Scenario 1 causes much bigger increase in Kansas City income per capita: same population increase causes a much bigger % change than for New York. But size of change is 12.2 times bigger in Variable θ model, compared with 9.5 times bigger in Constant θ model
- Scenario 2 causes same change in income per capita for both cities in Constant θ model, but $1.3\times$ bigger change for Kansas City in Variable θ model
- Maybe quantitatively significant for comparing some population increasing investment in New York and Kansas City since their real income per capita levels are similar to start with
- However, if considering some (approx equal cost) investment that causes equal population increase in two cities for which model fits well, then even $0.50\% - 0.04\% = 0.46\%$ relative increase in inc p.c. levels applied to whole of smaller city population, does not counteract $\sim 20\%$ difference in inc p.c. levels
  - Investment makes more sense in larger city, as in standard case

Conclusions

- Calibrated an extended Krugman trade model, with variable elasticity of substitution, to US MSA real income p.c. data
- Produces estimates of θ that are consistent with other estimates produced in the literature (higher, as expected since the data we use incorporates congestion effects)
- Best fit model exhibits increasing RLV (Zhelobodko et al, 2012): IRS are exhausted in the limit of large and highly connected cities
- While only quantitatively important in a few cases, does imply that centralisation has limits
- Need to finalise calibration and quantify significance of model in explaining the data
- Data issues: better connectivity data? Use travel time (from e.g. Google Maps) rather than latitude and longitude?
- Model issues: Counterfactual predictions wrt price indices? Should we extend model to include non-tradables which suffer congestion, and then match to price index data?