Distributional Approaches to Understand Patterns of Urban Differentiation

Andres Gomez-Lievano

PhD candidate in Applied Mathematics for the Life and Social Sciences

Mathematical, Computational & Modeling Sciences Center,

School of Human Evolution and Social Change,

Arizona State University

Collaborators:

Luís M.A. Bettencourt (SFI)

José Lobo (ASU)

Rachata Muneepeerakul (ASU)

HyeJin Youn (INET, Oxford, SFI)

Deborah Strumsky (SFI)

Kevin Stolarick (MPI)

Vladislav Vysotsky (PDMI, ASU)

The Big Question

- What are the obstacles to economic development?

The Big Question

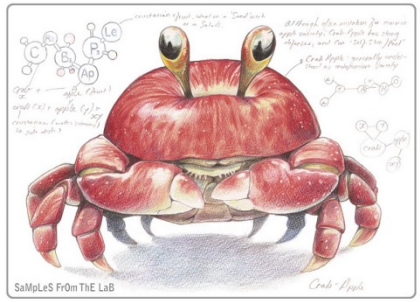
- What are the obstacles to economic development?

- Socio-economic processes in cities

The Big Question

- What are the obstacles to economic development?

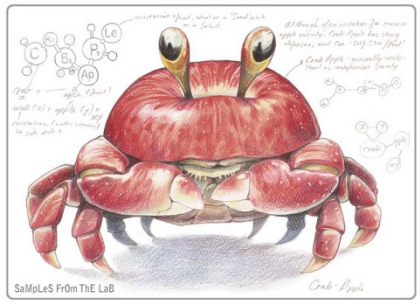
- Socio-economic processes in cities

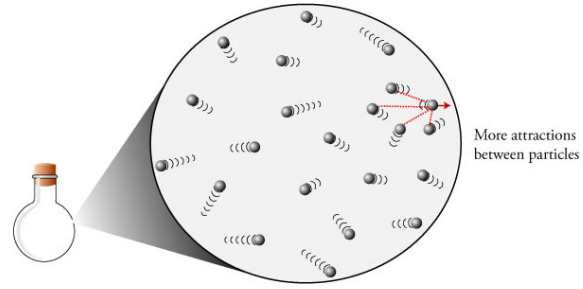


The Big Question

- What are the obstacles to economic development?

- Socio-economic processes in cities

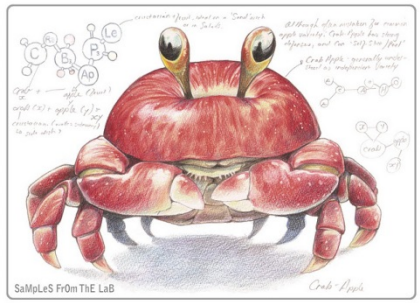


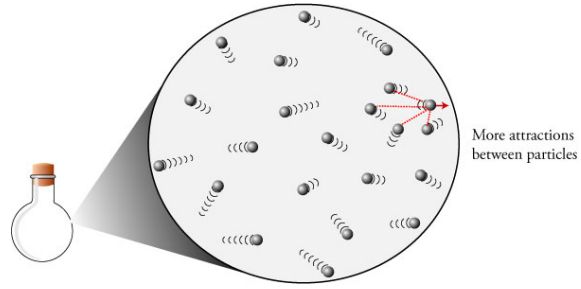


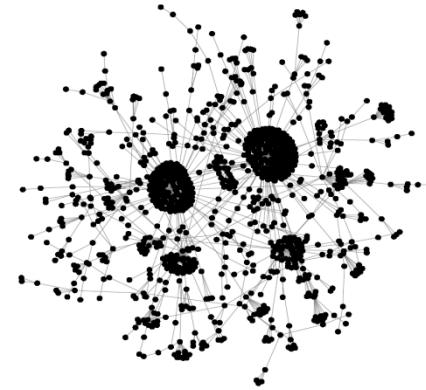
The Big Question

- What are the obstacles to economic development?

- Socio-economic processes in cities



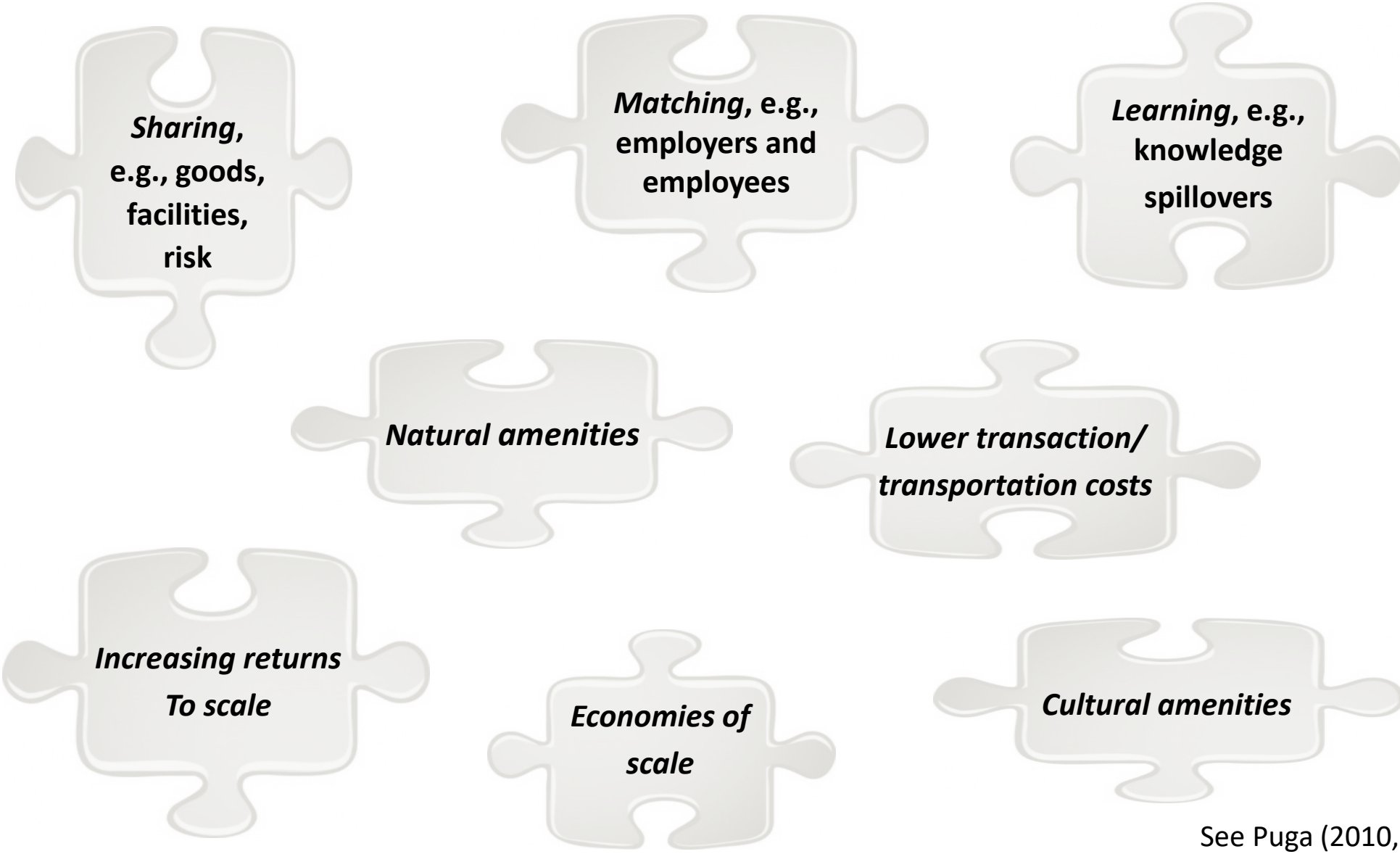




Diversity of: Shapes, Colors, *and* Sizes.

Artist**: Ahmed Farid**. Title: **Urbanization** (Oil on canvas, 2012)
From: http://almasargallery.com/ahmed-farid-urban-diversity-11-may-30-may-2013-solo-exhibition

Diversity and the Puzzle of Cities



See Puga (2010, *JRS*)

Explaining Productivity Differentials

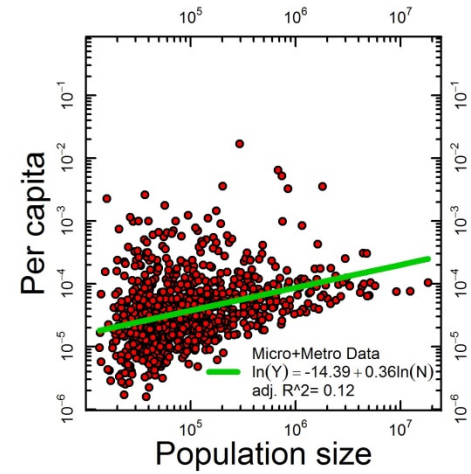
- **Classical models:**

$$\frac{Y}{N} = A\left(\frac{K}{N}\right)^{\kappa}\left(\frac{L}{N}\right)^{\lambda}\left(\frac{H}{N}\right)^{\eta}\left(\frac{M}{N}\right)^{\mu}\cdots\left(\frac{P}{N}\right)^{\pi}, \quad \kappa+\lambda+\ldots+\pi=1$$

Explaining Productivity Differentials

- **Classical models:**

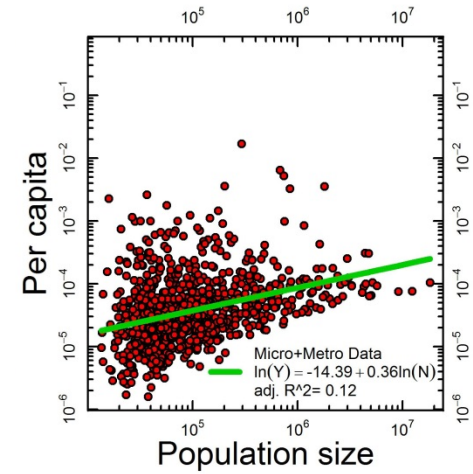
$$\frac{Y}{N} = A\left(\frac{K}{N}\right)^{\kappa}\left(\frac{L}{N}\right)^{\lambda}\left(\frac{H}{N}\right)^{\eta}\left(\frac{M}{N}\right)^{\mu}\cdots\left(\frac{P}{N}\right)^{\pi}, \quad \kappa+\lambda+\ldots+\pi=1$$



Explaining Productivity Differentials

- **Classical models:**

$$\frac{Y}{N} = A\left(\frac{K}{N}\right)^{\kappa}\left(\frac{L}{N}\right)^{\lambda}\left(\frac{H}{N}\right)^{\eta}\left(\frac{M}{N}\right)^{\mu}\cdots\left(\frac{P}{N}\right)^{\pi}, \quad \kappa+\lambda+\ldots+\pi=1$$



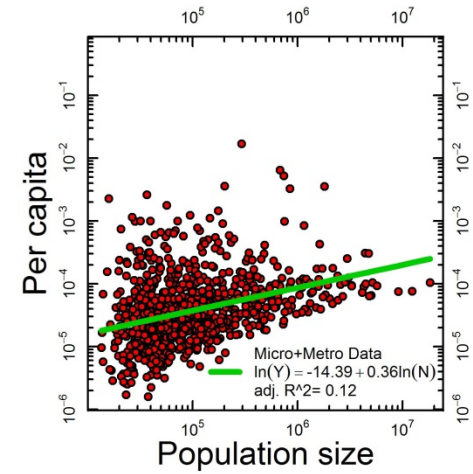
- **Definition:**

$$\frac{Y}{N} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

Explaining Productivity Differentials

- **Classical models:**

$$\frac{Y}{N} = A\left(\frac{K}{N}\right)^{\kappa}\left(\frac{L}{N}\right)^{\lambda}\left(\frac{H}{N}\right)^{\eta}\left(\frac{M}{N}\right)^{\mu}\cdots\left(\frac{P}{N}\right)^{\pi}, \quad \kappa+\lambda+\ldots+\pi=1$$

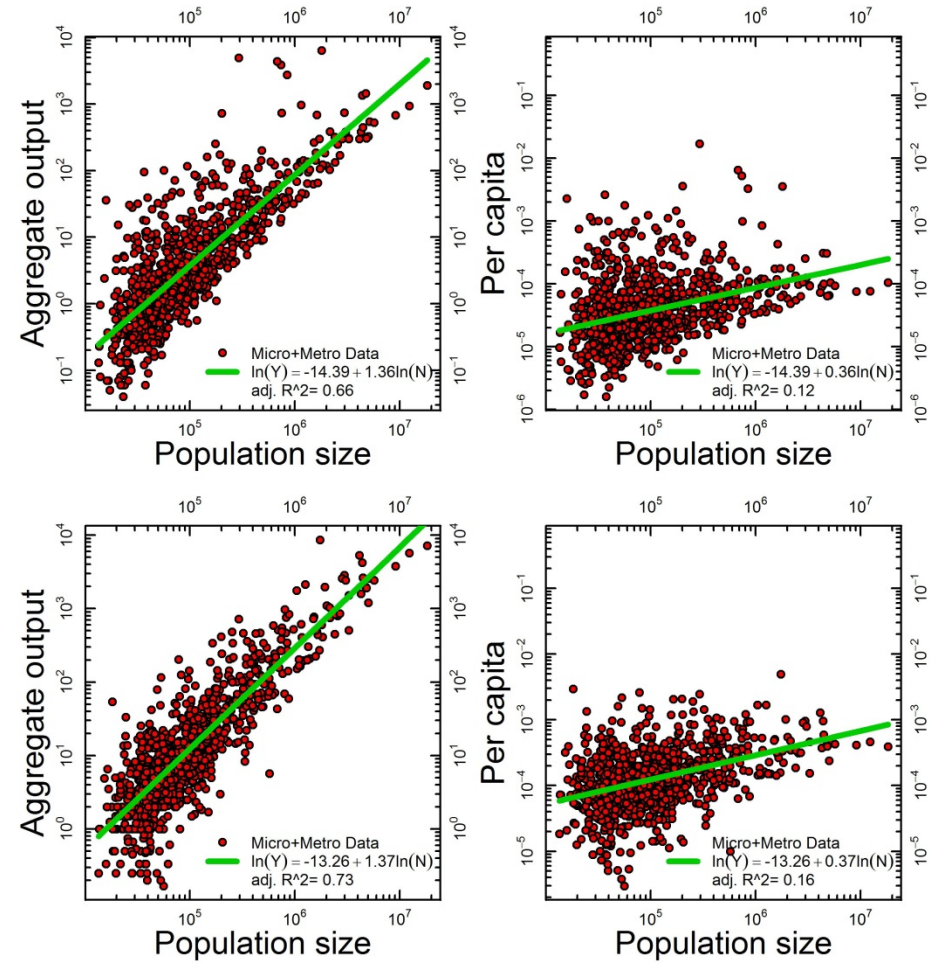


- **Definition:**

$$\frac{Y}{N} = \frac{1}{N}\sum_{i=1}^{N} y_i \approx \mathrm{E}[y_i]$$

Explaining Productivity Differentials

$$\frac{Y}{N} = \frac{1}{N}\sum_{i=1}^{N} y_i \approx \mathrm{E}[y_i]$$

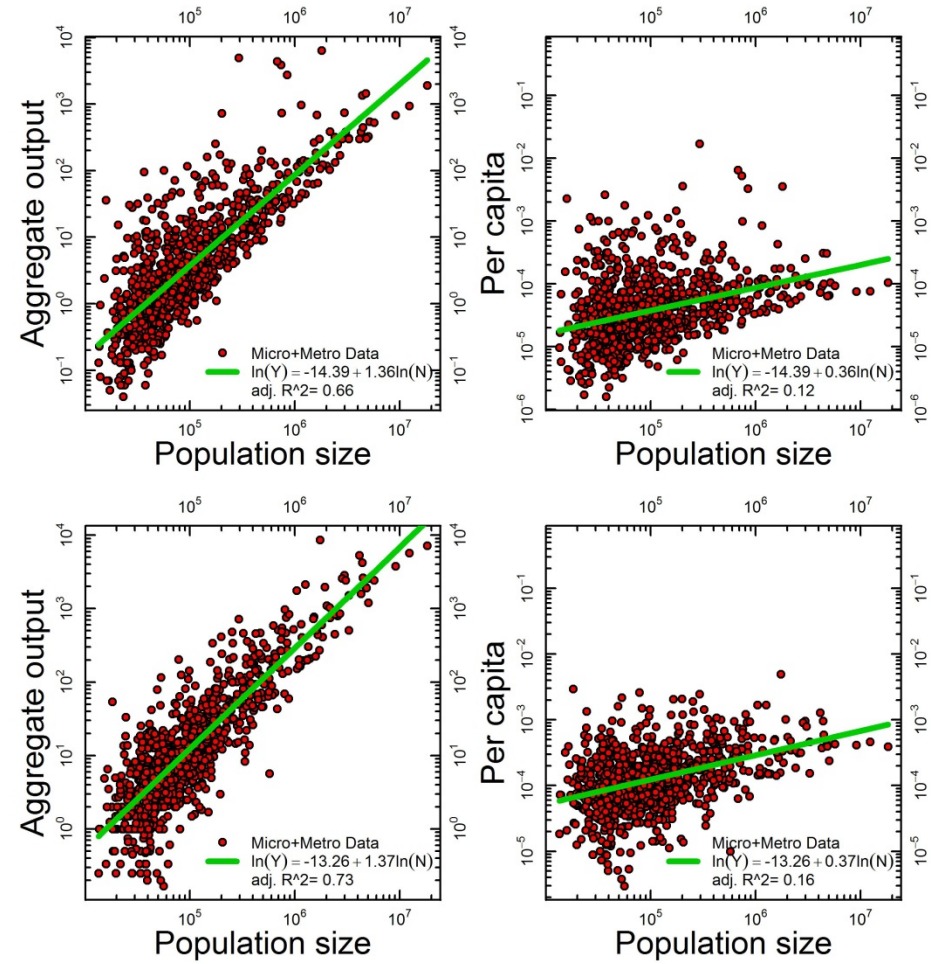


Simple Null Model

Simulated data!

- Real populations,
- but $y_i \sim \mathcal{LN}\left(\mu, \sigma^2\right)$ for **all** cities.

$$\frac{Y}{N} = \frac{1}{N}\sum_{i=1}^{N} y_i$$



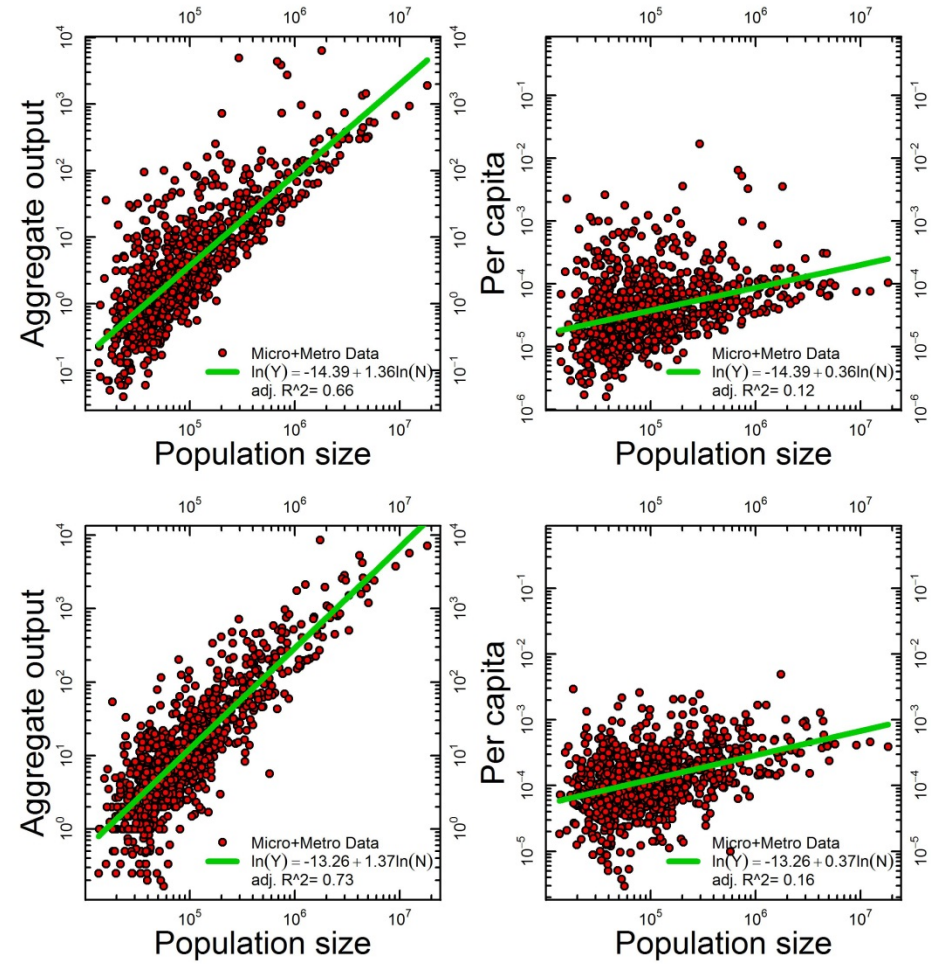
Simple Null Model

Simulated data!

- Real populations,
- but $y_i \sim \mathcal{LN}\left(\mu, \sigma^2\right)$ for **all** cities.

$$\frac{Y}{N} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

$$Y = \sum_{i=1}^{N} y_i \propto N^{\beta}, \text{ where } \beta > 1$$



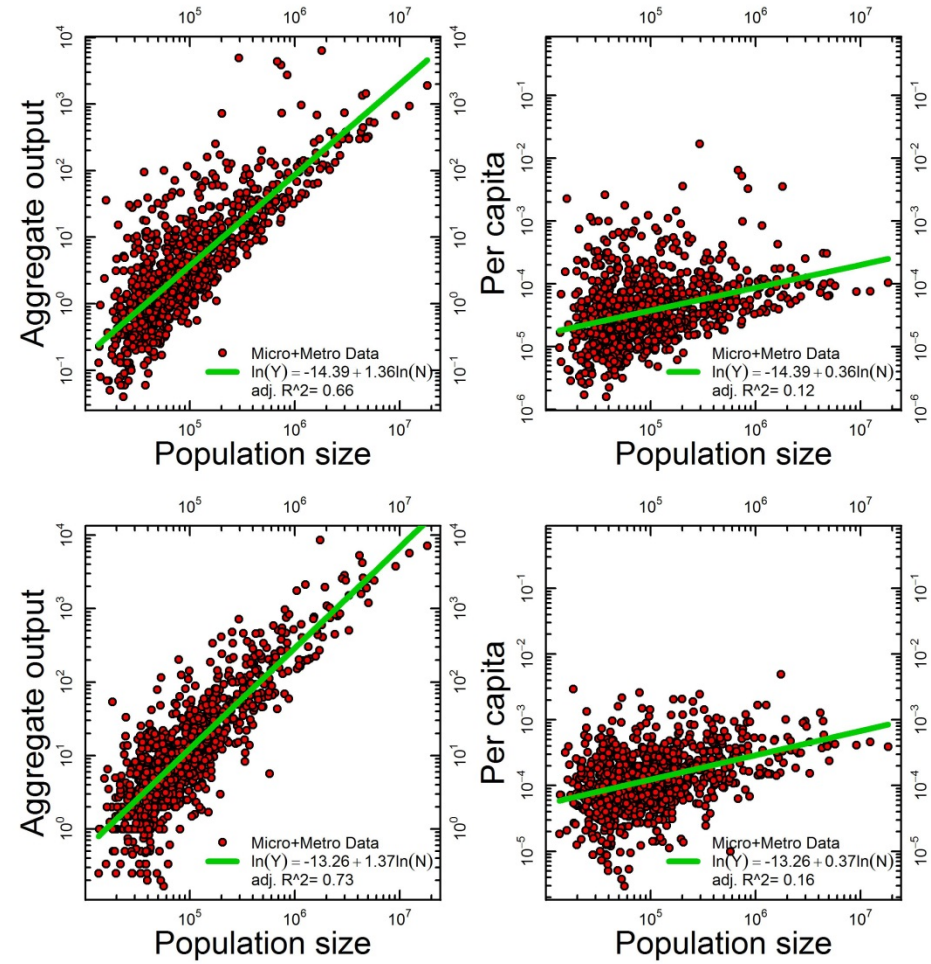
Simple Null Model

Simulated data!

- Real populations,
- but $y_i \sim \mathcal{LN}\left(\mu, \sigma^2\right)$ for **all** cities.

$$Y = \frac{1}{N} \sum_{i=1}^{N} y_i$$

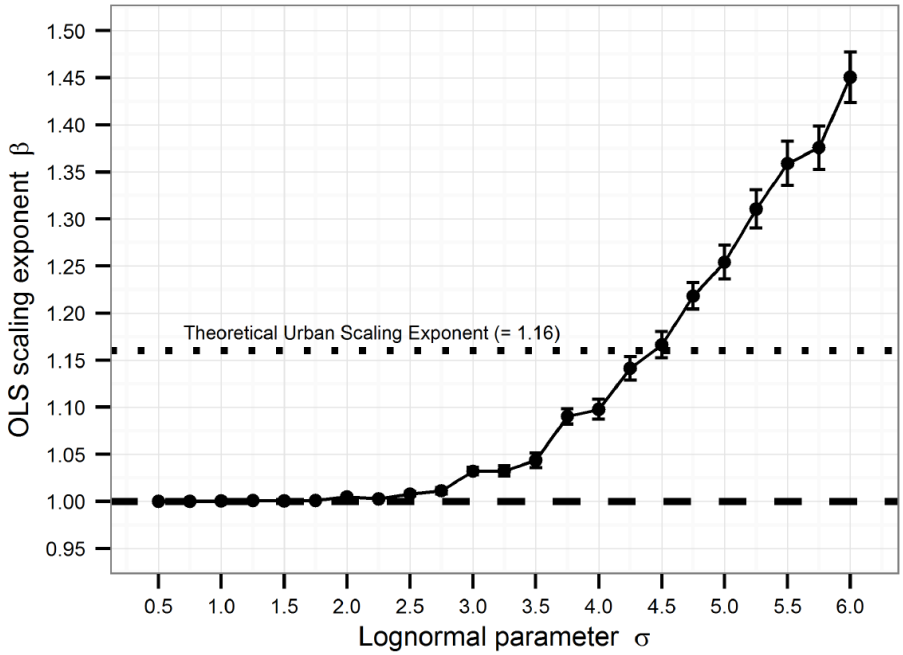
$$Y = \sum_{i=1}^{N} y_i \propto N^{\beta}, \text{ where } \beta > 1$$

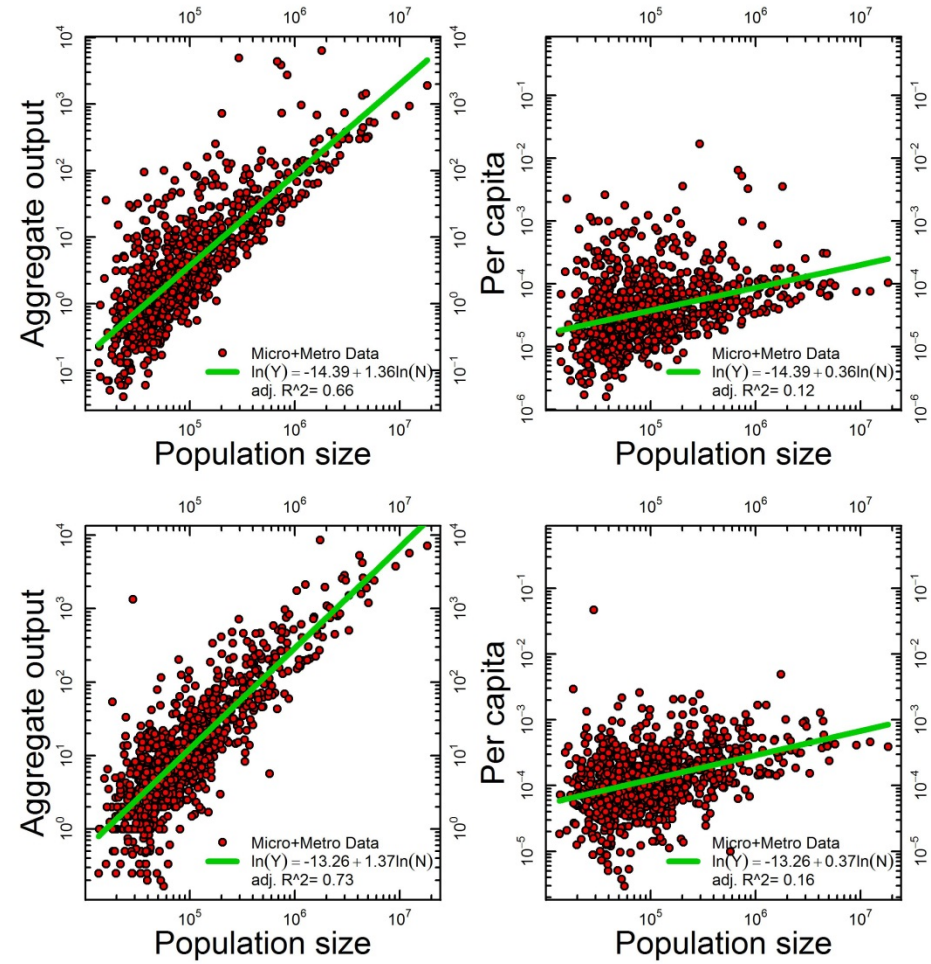


Simple Null Model

Simulated data!

- Real populations,
- but $y_i \sim \mathcal{LN}\left(\mu, \sigma^2\right)$

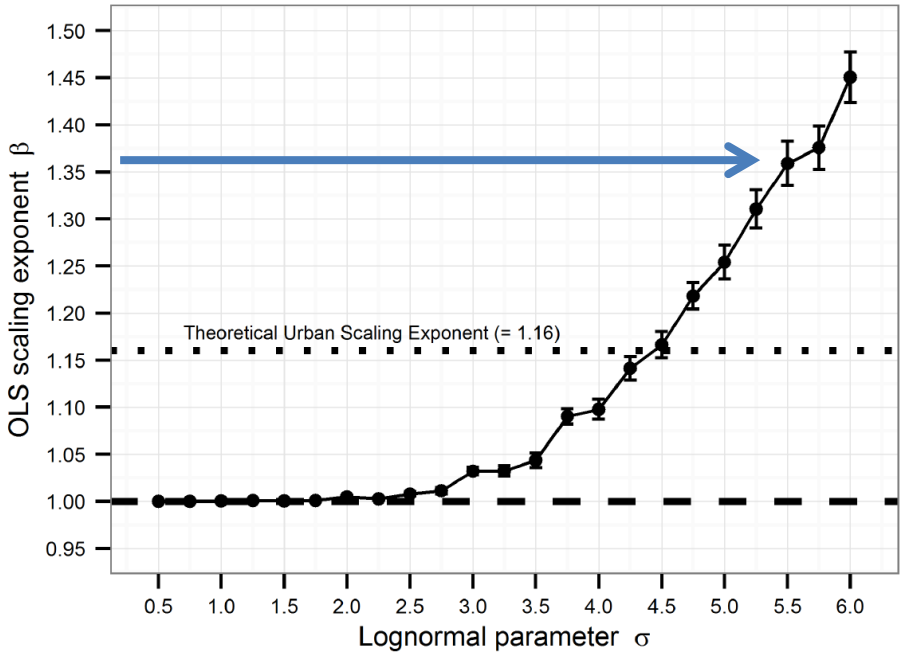


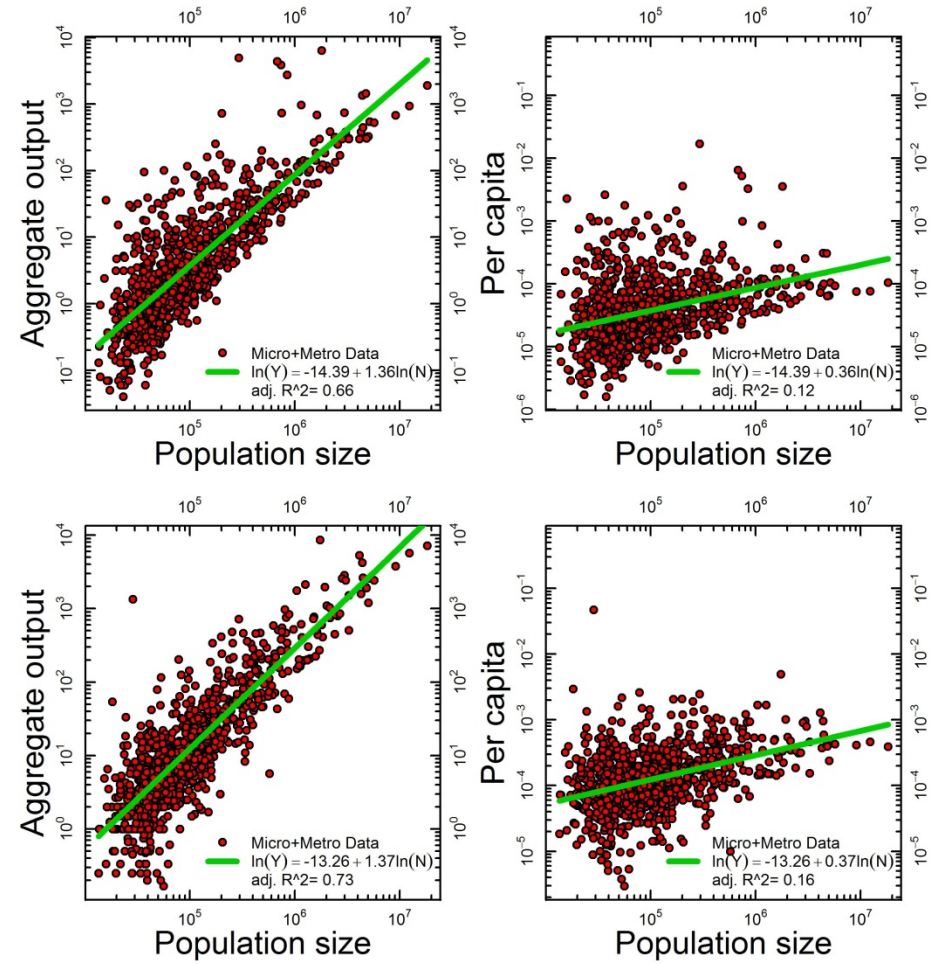


Simple Null Model

Simulated data!

- Real populations,
- but $y_i \sim \mathcal{LN}\left(\mu, \sigma^2\right)$

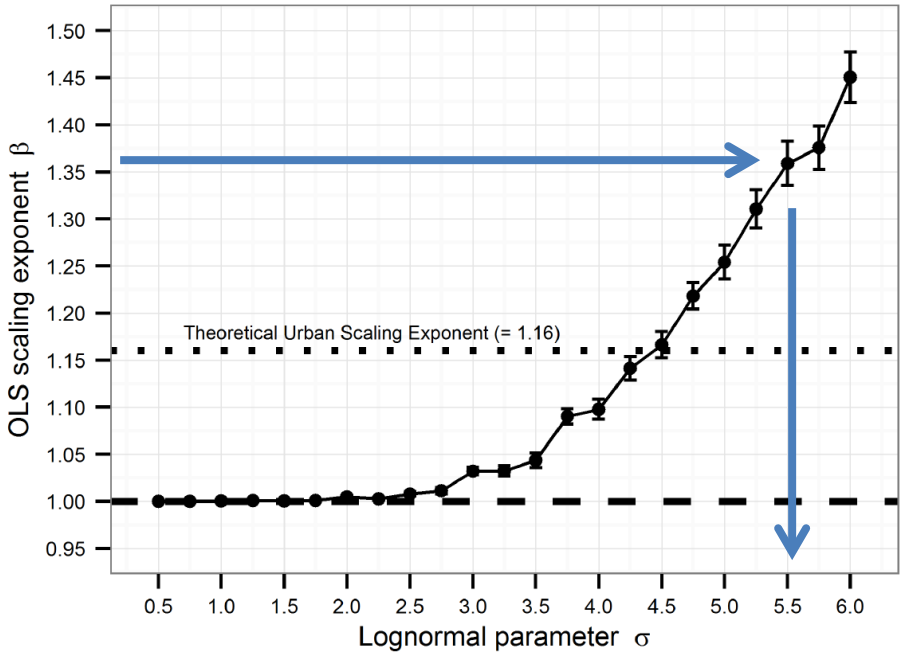


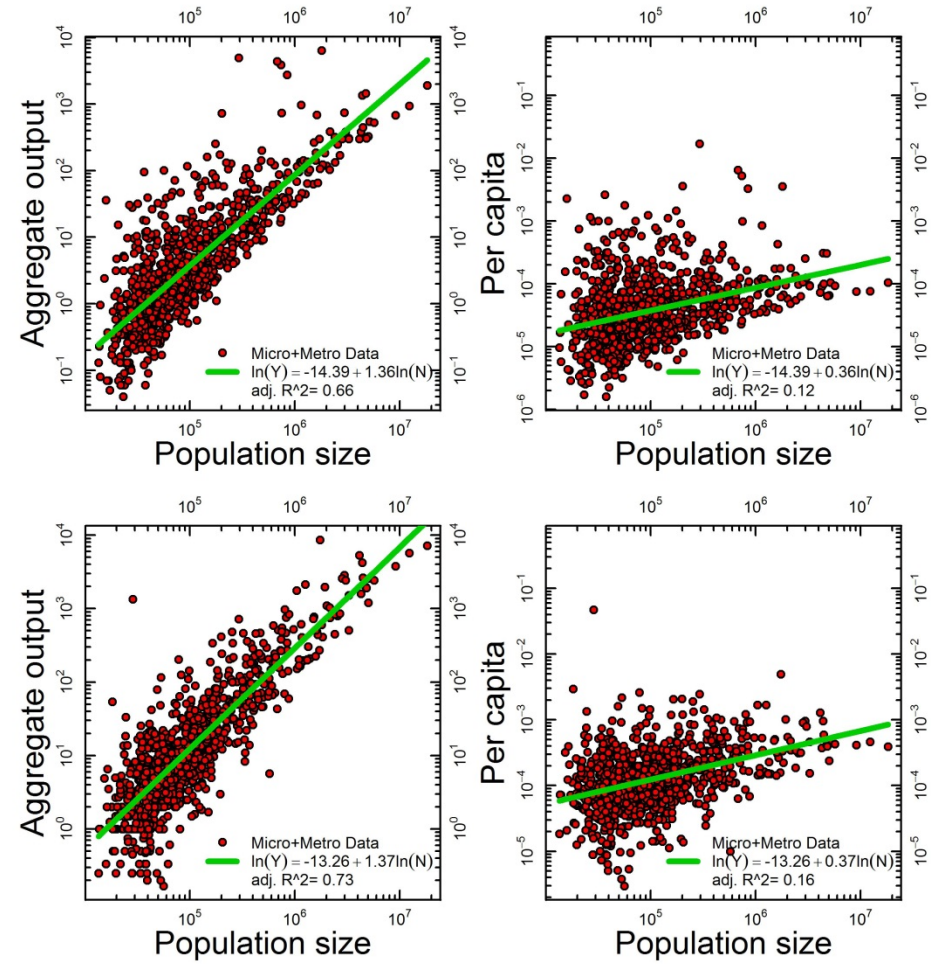


Simple Null Model

Simulated data!

- Real populations,
- but $y_i \sim \mathcal{LN}\left(\mu, \sigma^2\right)$



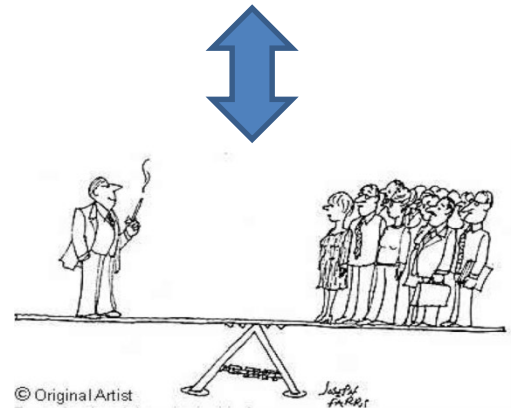


The interplay between city **sizes** and their internal **heterogeneity** determines**:**

- how we should **measure** productivity,
- and the **statistical properties** of the **aggregate output.**

Facing the Heterogeneity

Are cities **finite** systems that **violate** the *Law of Large Numbers*?



Gomez-Lievano, Vysotsky, and Lobo (2014, *in preparation*)

Consensus

Consensus

1. **Knowledge** as major economic driver.

   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)

Consensus

1. **Knowledge** as major economic driver.
   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)
2. Diversity of **skills**, (complementary to specialization):

Consensus

1. **Knowledge** as major economic driver.
   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)
2. Diversity of **skills**, (complementary to specialization):
   - Marshall (1890) vs. Jacobs (1969)

Consensus

1. **Knowledge** as major economic driver.
   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)
2. Diversity of **skills**, (complementary to specialization):

   - Marshall (1890) vs. Jacobs (1969)
   - Glaeser (1992); Faggio, Silva & Strange (2013-Dec, Working Paper)

Consensus

1. **Knowledge** as major economic driver.
   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)
2. Diversity of **skills**, (complementary to specialization):

   - Marshall (1890) vs. Jacobs (1969)
   - Glaeser (1992); Faggio, Silva & Strange (2013-Dec, Working Paper)
   - Quigley (1998), Beaudry & Schiffauerova (2009)

Consensus

1. **Knowledge** as major economic driver.
   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)
2. Diversity of **skills**, (complementary to specialization):
   - Marshall (1890) vs. Jacobs (1969)
   - Glaeser (1992); Faggio, Silva & Strange (2013-Dec, Working Paper)
   - Quigley (1998), Beaudry & Schiffauerova (2009)
3. Population **size** matters:

Consensus

1. **Knowledge** as major economic driver.
   - Romer (1986, 1990), Lucas (1988), Glaeser et al. (1995), Rauch (1993), Jones & Romer (2010)
2. Diversity of **skills**, (complementary to specialization):
   - Marshall (1890) vs. Jacobs (1969)
   - Glaeser (1992); Faggio, Silva & Strange (2013-Dec, Working Paper)
   - Quigley (1998), Beaudry & Schiffauerova (2009)
3. Population **size** matters:

   - Sveikauskaus (1975); Ó hUallacháin (1999); Bettencourt et al. (2007a, 2007b, 2010); Bettencourt (2013).

Measurement of Knowledge

Stock of knowledge

Measurement of Knowledge

<u>Stock</u> of knowledge

- Educational attainment

Measurement of Knowledge

Stock of knowledge

- Educational attainment
- Embodied in technology

Measurement of Knowledge

<u>Stock</u> of knowledge

- Educational attainment
- Embodied in technology
- "Total Factor Productivity" (*A*)

Measurement of Knowledge

<u>Stock</u> of knowledge

- Educational attainment
- Embodied in technology
- "Total Factor Productivity" (*A*)

Knowledge is spatially localized

- E.g., Jaffe et al. (1993)

Measurement of Knowledge

<u>Stock</u> of knowledge

- Educational attainment
- Embodied in technology
- "Total Factor Productivity" (*A*)

Knowledge is spatially localized

- E.g., Jaffe et al. (1993)
- "tacit knowledge"

Measurement of Knowledge

<u>Stock</u> of knowledge

- Educational attainment
- Embodied in technology
- "Total Factor Productivity" (*A*)

Knowledge is spatially localized

- E.g., Jaffe et al. (1993)
- "tacit knowledge"
- People move with *less difficulty* than their ideas
  - (See Breschi & Lissoni, 2001a,b, 2004, 2009; Klepper & Sleeper, 2005; Klepper, 2010; Hausmann, Neffke, Otto, 2013)

Conceptually, the economic output of an entire city (a *system-wide* property) should not be understood using the average properties of individuals (e.g., years of schooling)

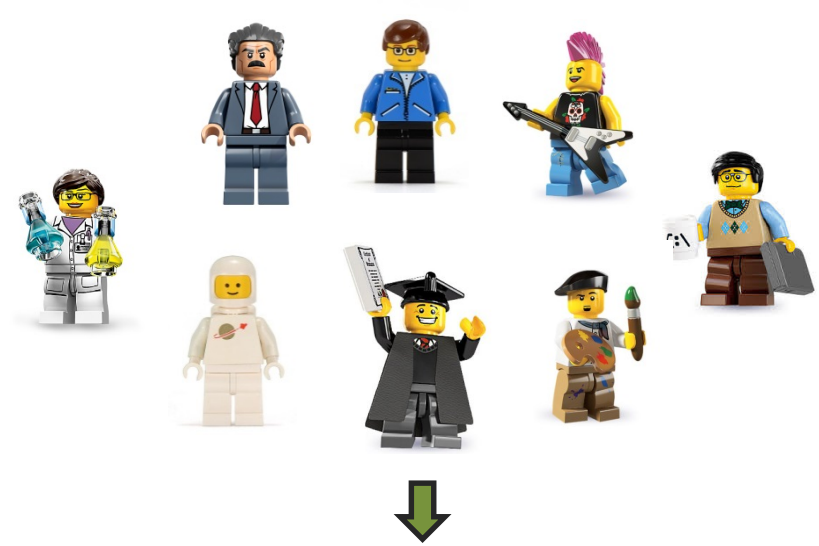
What citizens *do*

- Creative Class (Florida, 2004)

- Inventors

Obstacles to Economic Development

- Are the counts of ‘creatives’ and inventors ***constrained*** in urban areas?

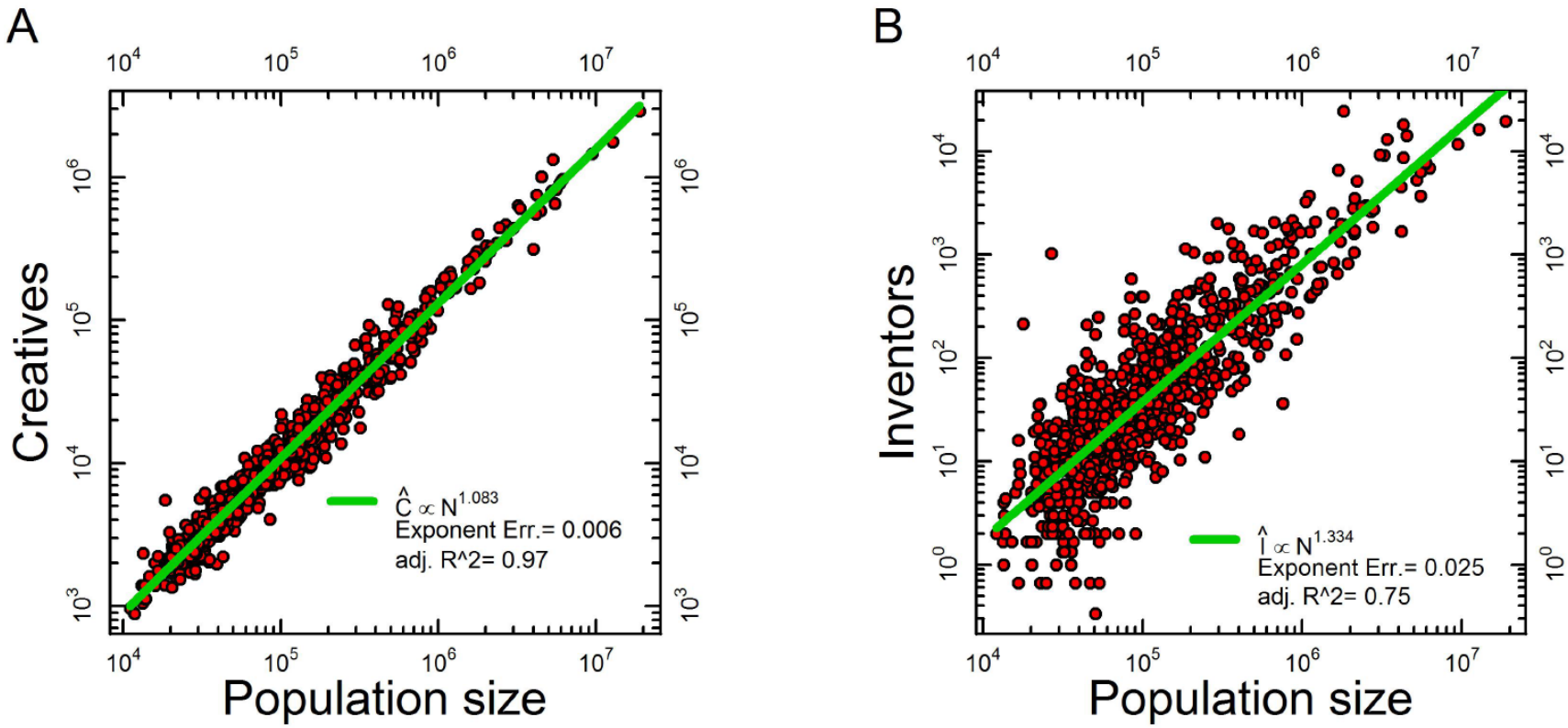


Gomez-Lievano, Bettencourt, Stolarick, Strumsky, Lobo (2014, *in submission*)

Assumption

The ability of a city to “learn” is constrained by its **capacity to attract** creative and inventive individuals (Florida, 1995, *Futures*).

More than 70% explained by urban population

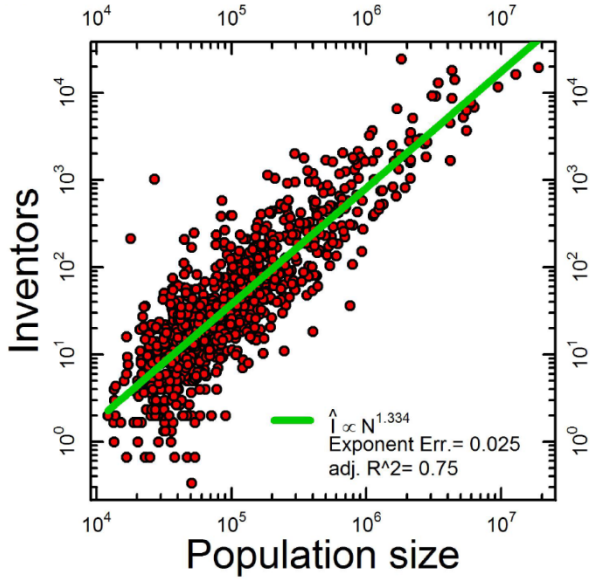


$$E(Y|N) = Y_0 N^{\beta}$$

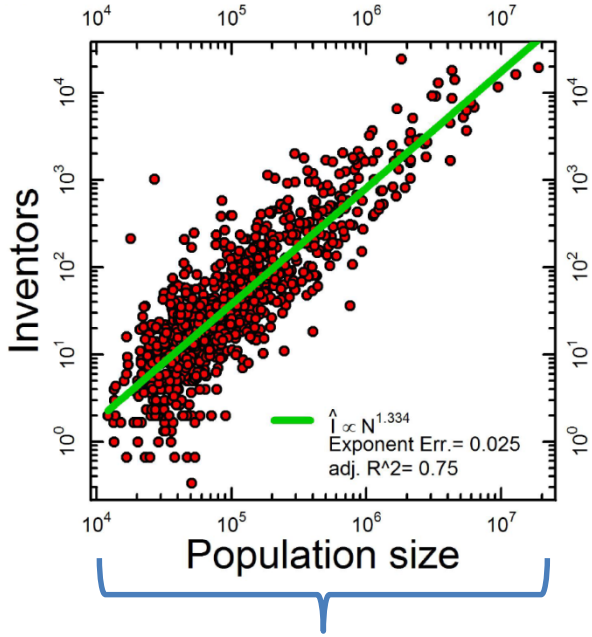
(Rauch, 1993; Glaeser & Saiz, 2004; Bettencourt et al., 2007)

Methodology & Results: $P(Y,N)=P(Y|N)P(N)$

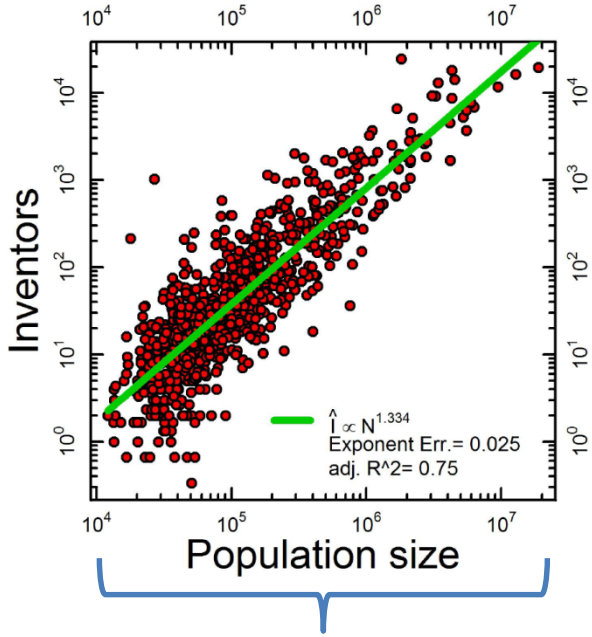
Methodology & Results: $P(Y,N)=P(Y|N)P(N)$

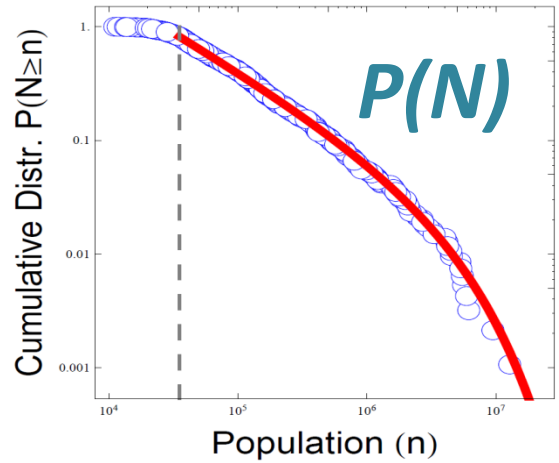


Methodology & Results: $P(Y,N)=P(Y|N)P(N)$



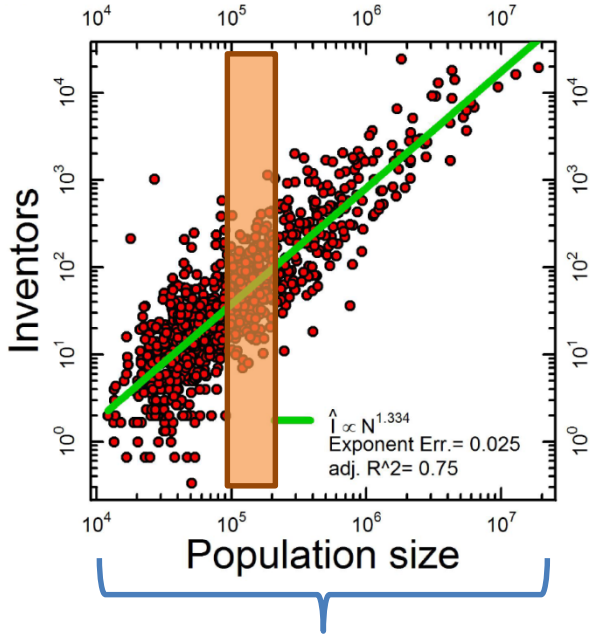
Methodology & Results: P(Y,N)=P(Y|N)P(N)

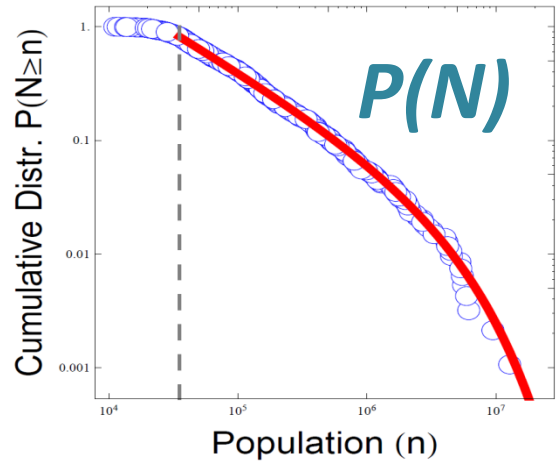




$$p_N(n;\alpha,\beta,\nu,n_{\min}) = C\frac{\mathrm{e}^{-\left(\frac{n}{\nu}\right)^{\beta}}}{\left(\frac{n}{\nu}\right)^{\alpha}}, \quad n \geq n_{\min}$$

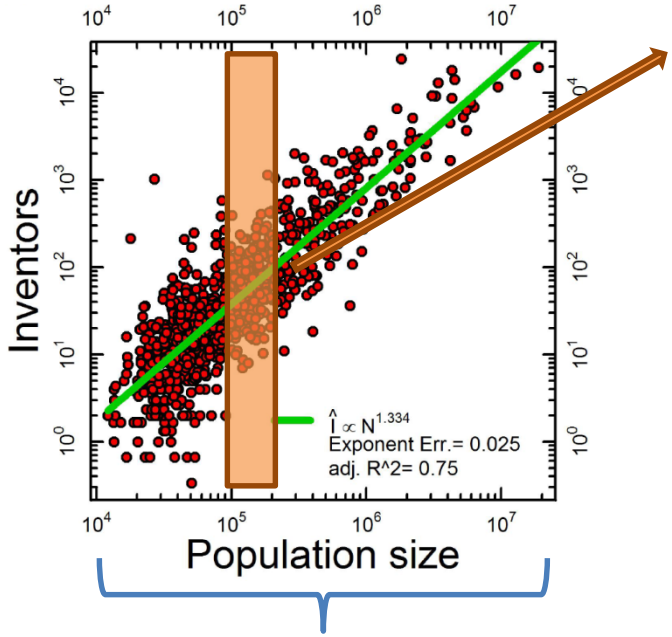
Methodology & Results: P(Y,N)=P(Y|N)P(N)



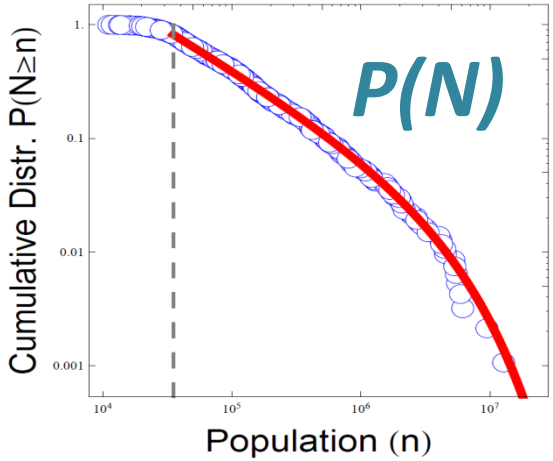


$$p_N(n;\alpha,\beta,\nu,n_{\min}) = C\frac{\mathrm{e}^{-\left(\frac{n}{\nu}\right)^{\beta}}}{\left(\frac{n}{\nu}\right)^{\alpha}}, \quad n \geq n_{\min}$$

Methodology & Results: P(Y,N)=P(Y|N)P(N)



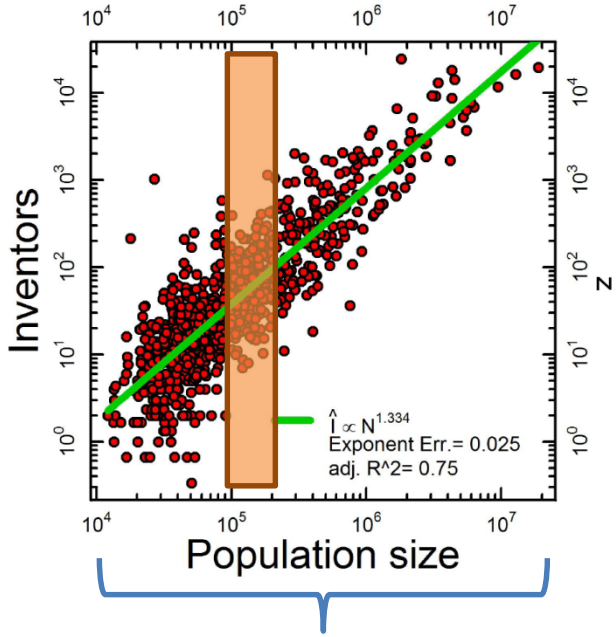
P(Y|N)



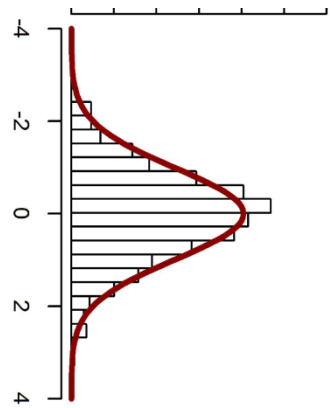
P(N)

$$p_N(n;\alpha,\beta,\nu,n_{\min}) = C\frac{\mathrm{e}^{-\left(\frac{n}{\nu}\right)^{\beta}}}{\left(\frac{n}{\nu}\right)^{\alpha}}, \quad n \geq n_{\min}$$

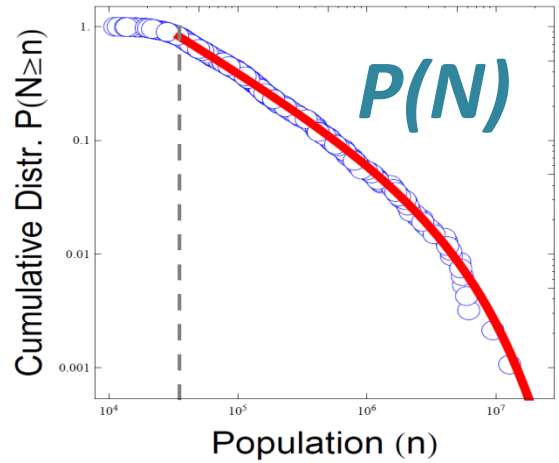
Methodology & Results: P(Y,N)=P(Y|N)P(N)



P(Y|N)



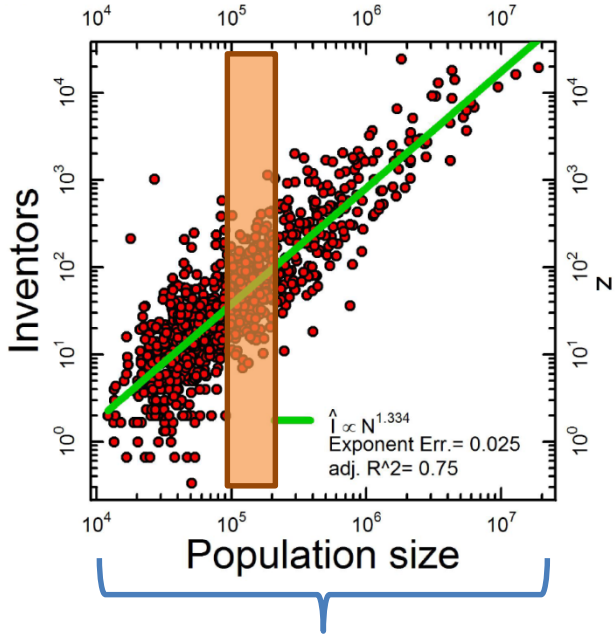
$$p_{Y|N}(y \mid n\,;\ \mu_n, \sigma_n) = \frac{1}{y\sqrt{2\pi\sigma_n^2}} \mathrm{e}^{-\frac{1}{2\sigma_n^2}(\ln(y)-\mu_n)^2}$$



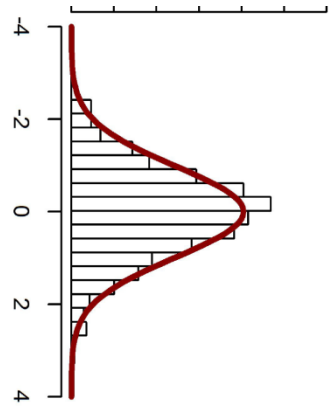
P(N)

$$p_N(n;\alpha,\beta,\nu,n_{\min}) = C\frac{\mathrm{e}^{-\left(\frac{n}{\nu}\right)^{\beta}}}{\left(\frac{n}{\nu}\right)^{\alpha}}, \quad n \geq n_{\min}$$

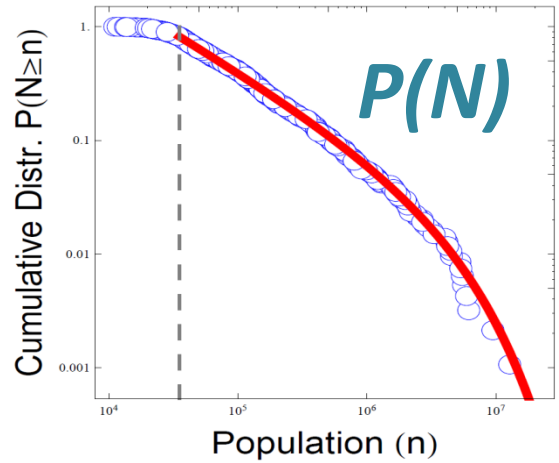
Methodology & Results: P(Y,N)=P(Y|N)P(N)



P(Y|N)



$$p_{Y|N}(y \mid n\,;\ \mu_n, \sigma_n) = \frac{1}{y\sqrt{2\pi\sigma_n^2}} \mathrm{e}^{-\frac{1}{2\sigma_n^2}(\ln(y)-\mu_n)^2}$$



P(N)

$$p_N(n;\alpha,\beta,\nu,n_{\min}) = C\frac{\mathrm{e}^{-\left(\frac{n}{\nu}\right)^{\beta}}}{\left(\frac{n}{\nu}\right)^{\alpha}}, \quad n \geq n_{\min}$$

For a similar study regarding homicides in Brazil, Colombia, and Mexico:
Gomez-Lievano, Youn, Bettencourt (2012, *PLoS ONE*)

Findings: a distributional account

Findings: a distributional account

- Estimated cutoff for population sizes around and **above** 11million.

Findings: a distributional account

- Estimated cutoff for population sizes around and **above** 11million.

- Constraints to growth (e.g., Black & Henderson, 2003).

Findings: a distributional account

- Estimated cutoff for population sizes around and **above** 11million.

- Constraints to growth (e.g., Black & Henderson, 2003).
- **Why** *P(N) ~ PL with EC* **?**
  - Extensive literature (recently: Berry & Okulicz-Kozaryn, 2012; Hsu, 2012; Ioannides & Skouras, 2013; Gonzalez-Val et al., 2013)

Findings: a distributional account

- Estimated cutoff for population sizes around and **above** 11million.

- **Lognormal** deviations, with log-variance *weakly* varying with ***N***

- Constraints to growth (e.g., Black & Henderson, 2003).
- **Why** *P(N) ~ PL with EC* **?**
  - Extensive literature (recently: Berry & Okulicz-Kozaryn, 2012; Hsu, 2012; Ioannides & Skouras, 2013; Gonzalez-Val et al., 2013)

Findings: a distributional account

- Estimated cutoff for population sizes around and **above** 11million.

- **Lognormal** deviations, with log-variance *weakly* varying with ***N***

- Constraints to growth (e.g., Black & Henderson, 2003).
- **Why** *P(N) ~ PL with EC* **?**
  - Extensive literature (recently: Berry & Okulicz-Kozaryn, 2012; Hsu, 2012; Ioannides & Skouras, 2013; Gonzalez-Val et al., 2013)
- Divergence (e.g., Ó hUallacháin, 1999; Berry & Glaeser, 2005).

Findings: a distributional account

- Estimated cutoff for population sizes around and **above** 11million.

- **Lognormal** deviations, with log-variance *weakly* varying with ***N***

- Constraints to growth (e.g., Black & Henderson, 2003).
- **Why** *P(N) ~ PL with EC* **?**
  - Extensive literature (recently: Berry & Okulicz-Kozaryn, 2012; Hsu, 2012; Ioannides & Skouras, 2013; Gonzalez-Val et al., 2013)

- Divergence (e.g., Ó hUallacháin, 1999; Berry & Glaeser, 2005).
- **Why** *P(Y|N) ~ Lognormal* **?**

Why is this important?

Why is this important?

- Whereas **normally** distributed RVs often arise from random *additive* processes

Why is this important?

- Whereas **normally** distributed RVs often arise from random *additive* processes

$X_1 + X_2 + X_3 + \ldots =$ "X_1 **OR** X_2 **OR** X_3 **OR** ..."

Why is this important?

- Whereas **normally** distributed RVs often arise from random *additive* processes

$X_1 + X_2 + X_3 + \ldots =$ "X_1 **OR** X_2 **OR** X_3 **OR** ..."

- … **lognormal** RVs arise, *often*, from random **multiplicative** processes:

Why is this important?

- Whereas **normally** distributed RVs often arise from random *additive* processes

$X_1 + X_2 + X_3 + \ldots =$ "X_1 **OR** X_2 **OR** X_3 **OR** ..."

- … **lognormal** RVs arise, *often*, from random **multiplicative** processes:

$X_1 * X_2 * X_3 * \ldots =$ "X_1 **AND** X_2 **AND** X_3 **AND** ..."

Why is this important?

- Whereas **normally** distributed RVs often arise from random *additive* processes

$X_1 + X_2 + X_3 + \ldots =$ "X_1 **OR** X_2 **OR** X_3 **OR** ..."

- … **lognormal** RVs arise, *often*, from random **multiplicative** processes:

$X_1 * X_2 * X_3 * \ldots =$ "X_1 **AND** X_2 **AND** X_3 **AND** ..."

- Jaynes, 2003; Frank, 2009; Frank and Smith, 2011

Conjecture

Conjecture

Differences between cities are *structural*

Conjecture

Differences between cities are *structural*

and come from a large number of effects acting multiplicatively

Conjecture

Differences between cities are *structural*

and come from a large number of effects acting multiplicatively

- Hidalgo et al. (2007), Hidalgo & Hausmann (2009), Neffke et al. (2013, Working Paper), Neffke & Henning (2013);

Conjecture

Differences between cities are *structural*

and come from a large number of effects acting multiplicatively

- Hidalgo et al. (2007), Hidalgo & Hausmann (2009), Neffke et al. (2013, Working Paper), Neffke & Henning (2013);
- Muneepeerakul, Lobo, Shutters, Gomez-Lievano, Qubbaj (2013);

Conjecture

Differences between cities are *structural*

and come from a large number of effects acting multiplicatively

- Hidalgo et al. (2007), Hidalgo & Hausmann (2009), Neffke et al. (2013, Working Paper), Neffke & Henning (2013);
- Muneepeerakul, Lobo, Shutters, Gomez-Lievano, Qubbaj (2013);
  - Hausmann & Hidalgo (2011)

“Urban dynamics, as can be seen, are the ultimate noisy social science problem”

– Storper et al. (2012, *JRS*, p. 4)

- *“Size”, “Heterogeneity” and “Structure”*
- *Policies and Distributions*

“**Urban dynamics, as can be seen, are the ultimate noisy social science problem**”

– Storper et al. (2012, *JRS*, p. 4)

THANKS!

Question

Question

Does the acquisition of skills (“lego pieces”)**,**

Question

Does the acquisition of skills (“lego pieces”)**,**
constrained by the space of products *and* the existing set of skills in the city,

Question

Does the acquisition of skills (“lego pieces”)**,**
constrained by the space of products *and* the existing set of skills in the city,
conditioned on the population size (# of lego pieces)**,**

Question

Does the acquisition of skills (“lego pieces”)**,**
constrained by the space of products *and* the existing set of skills in the city,
conditioned on the population size (# of lego pieces)**,**
generate lognormal distributions in output?

Playing with a model of the world

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.
- Persons are carriers of skills.

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.
- Persons are carriers of skills.
- 1 person = "1 pbyte"

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.
- Persons are carriers of skills.
- 1 person = "1 pbyte"
- A pbyte "counts" for economic development only when it is different from the others.

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.
- Persons are carriers of skills.
- 1 person = "1 pbyte"
- A pbyte "counts" for economic development only when it is different from the others.
- Products require different pbytes.

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.
- Persons are carriers of skills.
- 1 person = "1 pbyte"
- A pbyte "counts" for economic development only when it is different from the others.
- Products require different pbytes.
- Countries/Cities that possess many different pbytes will be able to produce many different, and more complex products

Playing with a model of the world

Hausmann & Hidalgo (2011)'s conceptual framework:

- Skills are represent non-tradable capabilities.
- Persons are carriers of skills.
- 1 person = "1 pbyte"
- A pbyte "counts" for economic development only when it is different from the others.
- Products require different pbytes.
- Countries/Cities that possess many different pbytes will be able to produce many different, and more complex products

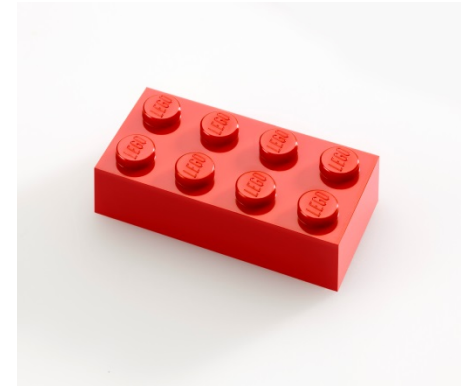
I expand H&H's model in **two** ways:

1. I will model a **process of acquisition** of skills (pbytes).

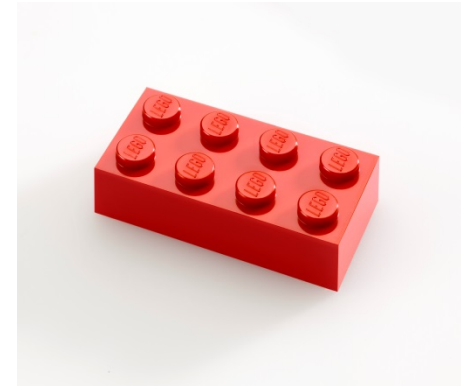
2. I will model *what* cities produce, as a **step** to know ***how much*** they produce.


C:\

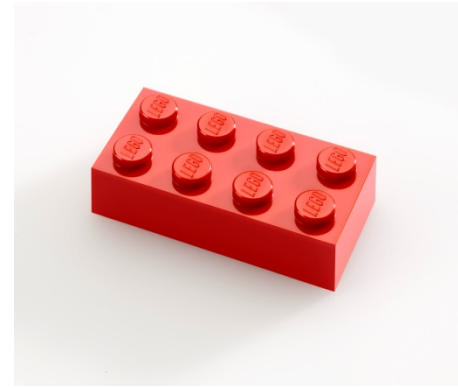
=

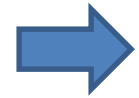


=



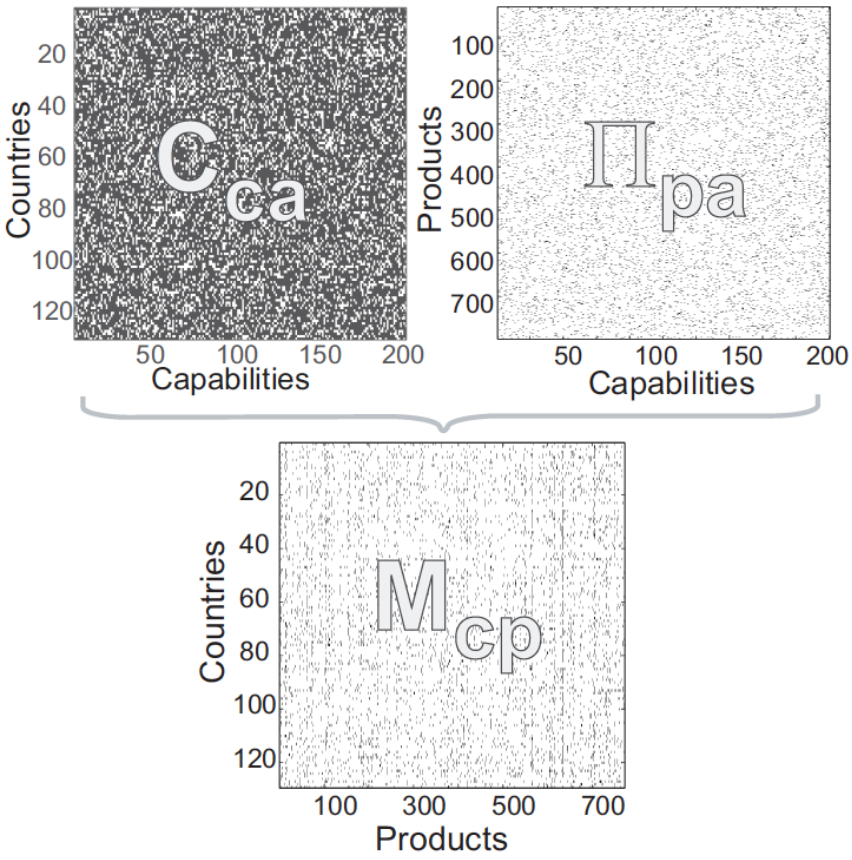
=





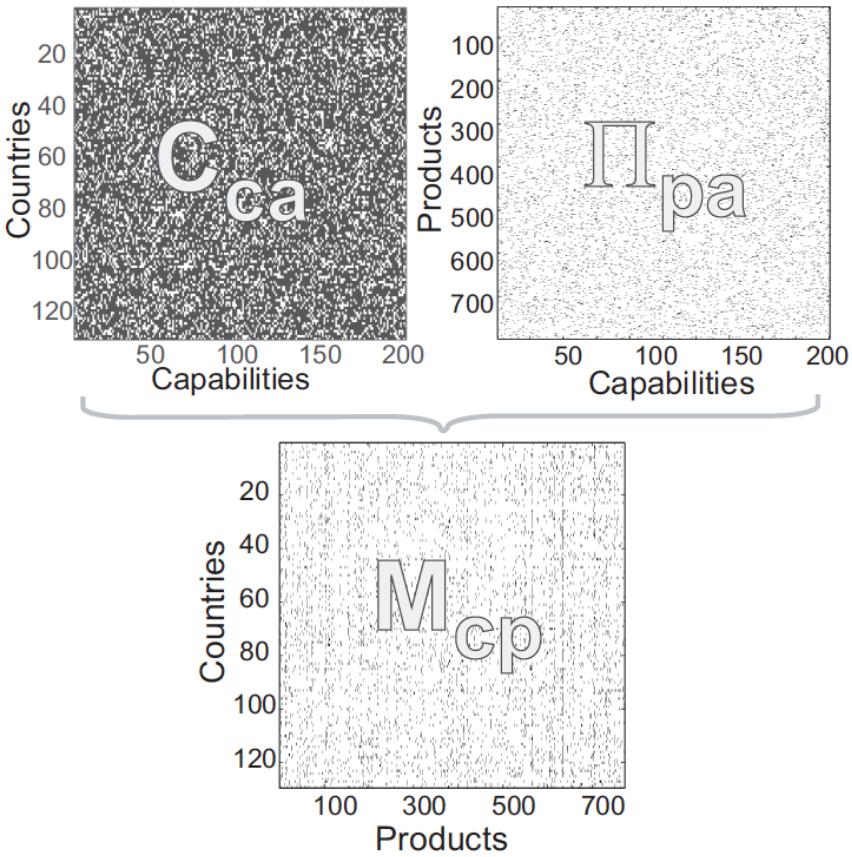
Lego analogy inspired by:
TEDxBoston - César A. Hidalgo - Global Product Space

The Model



From: Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. PNAS, 106(25):10570-10575.

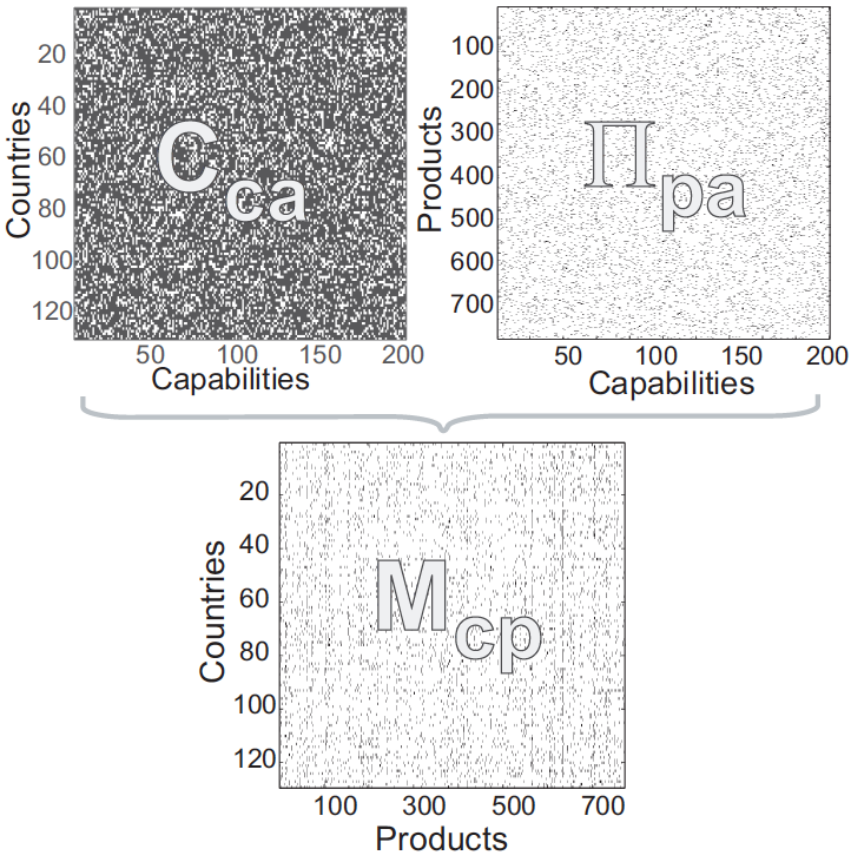
The Model



$$X^t = C^t \odot P$$

From: Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. PNAS, 106(25):10570-10575.

The Model

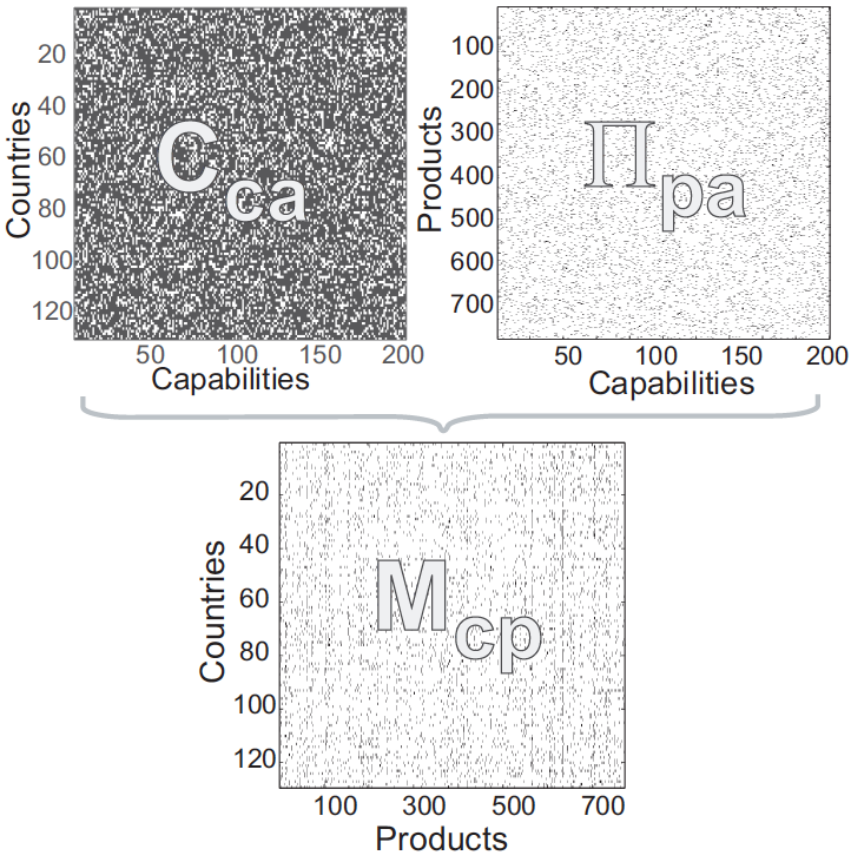


$$\boldsymbol{X^t = C^t \odot P}$$

$$X_{cp} = F_{\odot}(C_{ci}, C_{cj}, C_{ck})$$

From: Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. PNAS, 106(25):10570-10575.

The Model



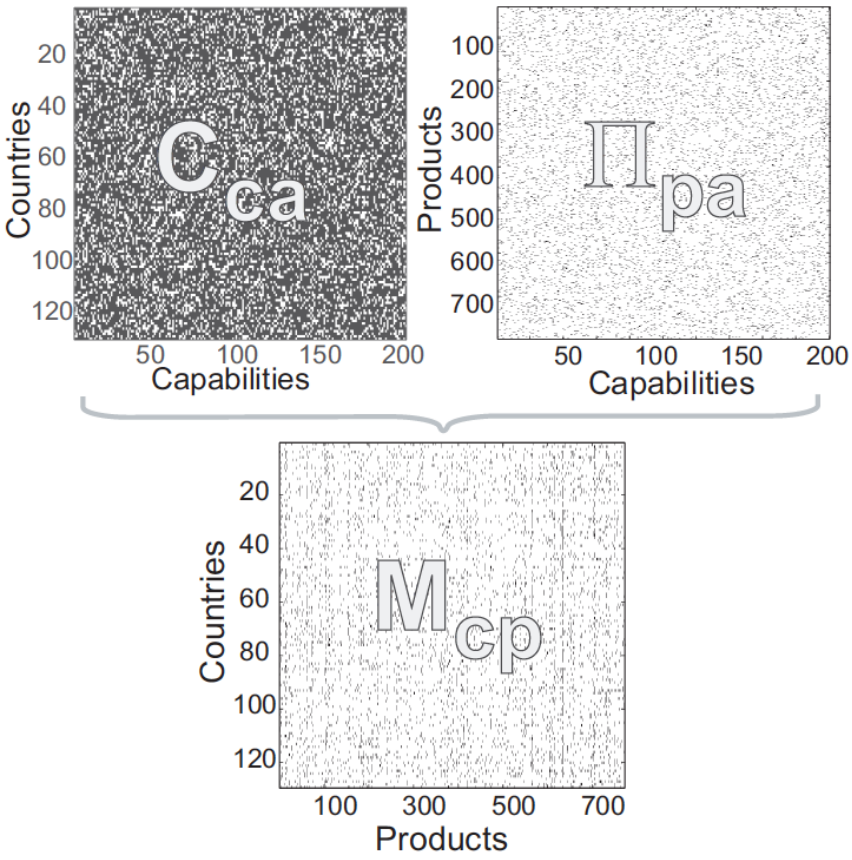
From: Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. PNAS, 106(25):10570-10575.

$$\boldsymbol{X^t} = \boldsymbol{C^t} \odot \boldsymbol{P}$$

$$X_{cp} = F_{\odot}\left(C_{ci}, C_{cj}, C_{ck}\right)$$

$$pop_c = \sum_{a=1}^{N_a} C_{ca}$$

The Model



From: Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. PNAS, 106(25):10570-10575.

$$\boldsymbol{X^t} = \boldsymbol{C^t} \odot \boldsymbol{P}$$

$$X_{cp} = F_{\odot}\left(C_{ci}, C_{cj}, C_{ck}\right)$$

$$pop_c = \sum_{a=1}^{N_a} C_{ca}$$

$$y_c^t = \sum_{p=1}^{N_p} X_{cp}^t$$

Model (currently):

1. The matrix ***C*** starts empty.
2. The matrix ***P*** is Bernoulli-filled (with prob. q), and fixed.
3. For each city (i.e., row) in ***C***,
   i. a random skill is added,
   ii. and lost with a certain probability.
4. We calculate the matrix ***X*** of output

Will the model reproduced data?

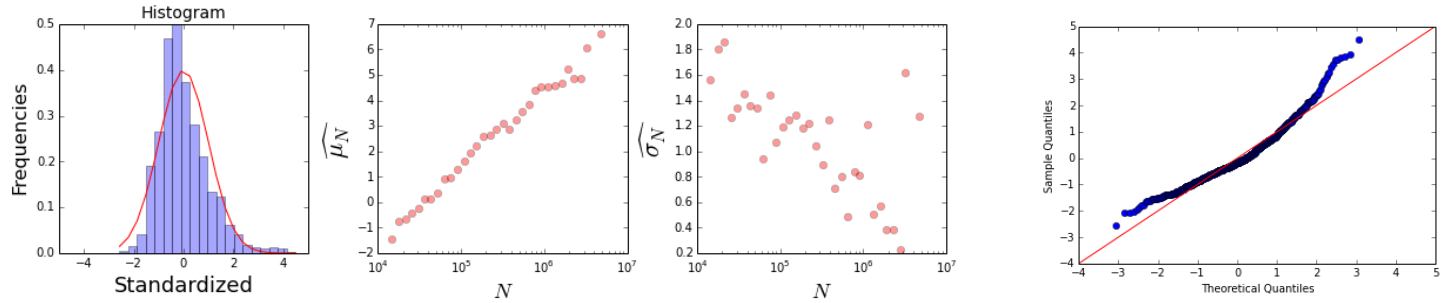
$$y_c^t = \sum_{p=1}^{N_p} X_{cp}^t$$

$$pop_c = \sum_{a=1}^{N_a} C_{ca}$$

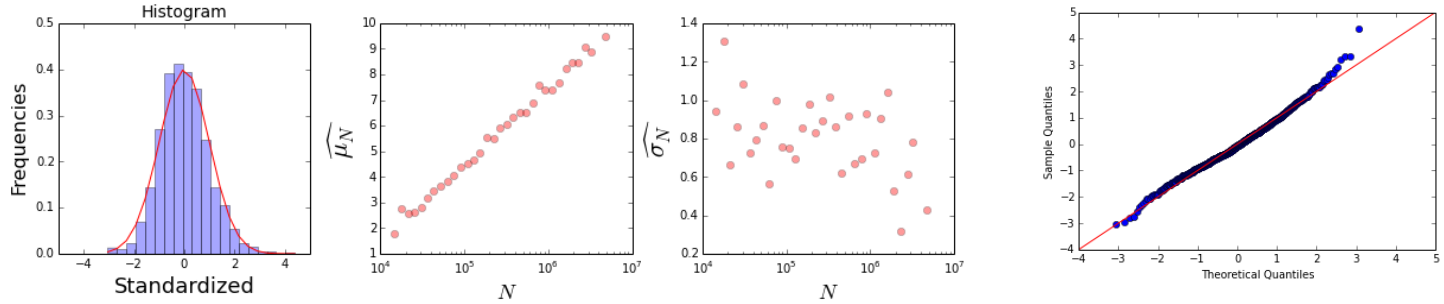
$$\Pr\left(y_c^t \mid pop_c\right) \sim \text{Lognormal}$$

Profiles of P(Y|N)

Null Model:



Inventors:



H&H Model

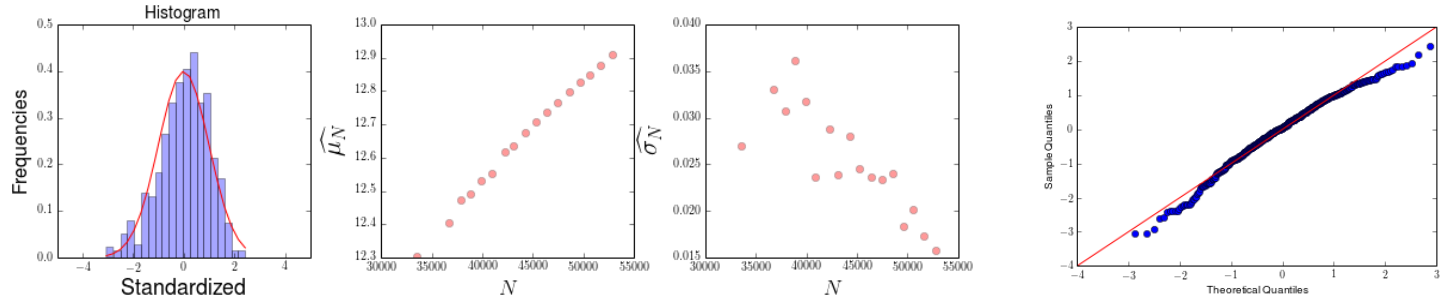


Table 1: US urban areas, sorted by population size, that have counts, in inventors or creatives, that are outside a $z = 3.07$ *sigma* interval around the log-meann, i.e., $y \notin [e^{\mu(n)-z\sigma(n)}, e^{\mu(n)+z\sigma(n)}]$. The numbers shown in the Population and Inventors columns are estimated averages from 2008-2010, although all values representing counts have been rounded to the nearest integer. The value $z = 3.07$ corresponds to a log-deviation such that $1 - \Phi(z) < 1/938$. The random variables and their values are for inventors I and i, for creatives C and c, respectively

Name of Urban Area	Population (n)	Inventors (i)	Most probable I	$P(I \geq i\|n)$	Creatives (c)	Most probable C	$P(C \geq c\|n)$
Los Alamos, NM (Micro)	17,899	213	1	0.	5,502	1,697	0.
Mountain Home, ID (Micro)	26,926	1,027	2	0.	2,801	2,751	0.562
Clewiston, FL (Micro)	39,109	8	4	0.626	1,976	3,848	0.999
Clovis, NM (Micro)	47,009	1	5	0.999	4,504	4,461	0.577
Eagle Pass, TX (Micro)	53,392	1	7	1.	4,328	5,144	0.834
Palm Coast, FL (Metro)	94,755	44	15	0.426	4,040	8,958	1.
Lake Havasu City-Kingman, AZ (Metro)	200,447	43	43	0.82	10,940	21,603	0.999
Merced, CA (Metro)	253,198	37	59	0.924	13,650	27,633	0.999
Ocala, FL (Metro)	330,780	69	87	0.873	17,620	37,218	1.
Durham-Chapel Hill, NC (Metro)	498,511	1,692	155	0.028	128,900	57,900	0.001
McAllen-Edinburg-Mission, TX (Metro)	758,064	36	279	0.999	60,400	87,582	0.963
San Jose-Sunnyvale-Santa Clara, CA (Metro)	1,818,864	24,531	952	0.	396,820	240,275	0.033

Example

C	a1	a2	a3	a4	a5
city 1	1	1	0	0	1
city 2	0	1	0	1	1

P	a1	a2	a3	a4	a5
p1	0	1	0	0	1
p2	0	0	0	1	0

M	p1	p2
city 1	1	0
city 2	1	1

- City 1 has skills {1,2,5}
- City 2 has skills {2,4,5}

- Product 1 requires {2,5}
- Product 2 requires {4}

- City 1 only produces Product 1
- City 2 produces Both

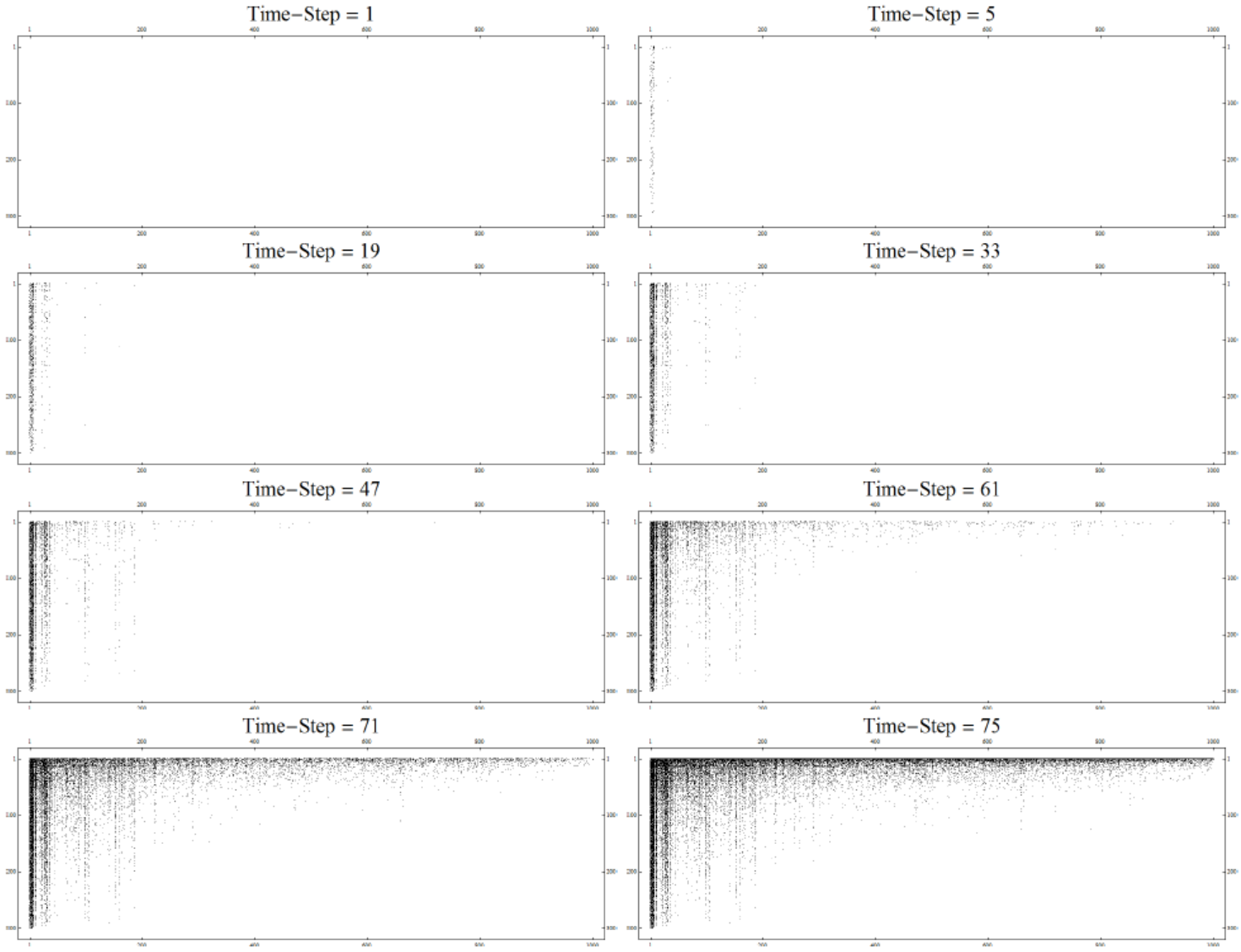
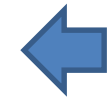


Figure 5: **Model 1: Evolution of the M^{t} matrix of cities versus products.** The model assumes a fixed set of products uniquely defined by their set of capabilities. In each time-step, each city acquires a new random capability previously missing. If the capability allows the city to produce a new product, the capability is fixed, otherwise it is lost (until it acquires it again by chance in the future). Here, $N_c = 300$, $N_p = 1000$, $N_a = 50$, $r = 0$, and $q = 0.13$. The M_{cp} elements are sorted in the same order throughout all steps from most diversified city to the least, and from the most ubiquitous product to the least, as they appear in the last time-step.

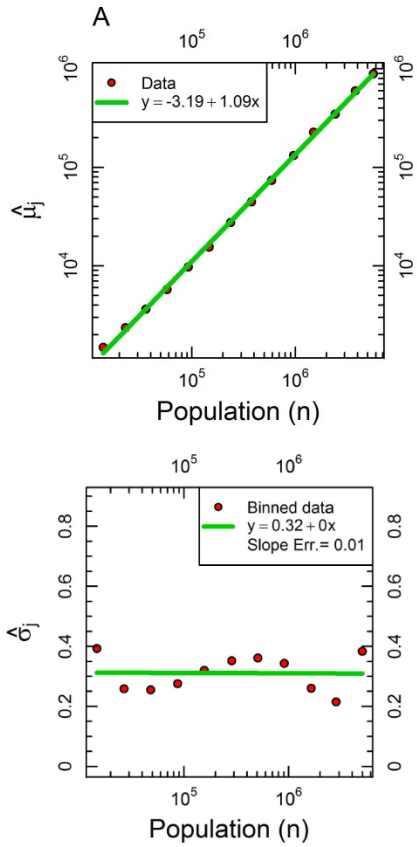
The following relationship has to hold if we want a confidence *(1-α)* that *Y/N* is close to the mean of the process

$$(\alpha\epsilon^2)N^r \geq c(r)\left(N\mathrm{E}\left[y_1^r\right] + N^{r/2}\sqrt{\mathrm{Var}[y_1]}\right)$$

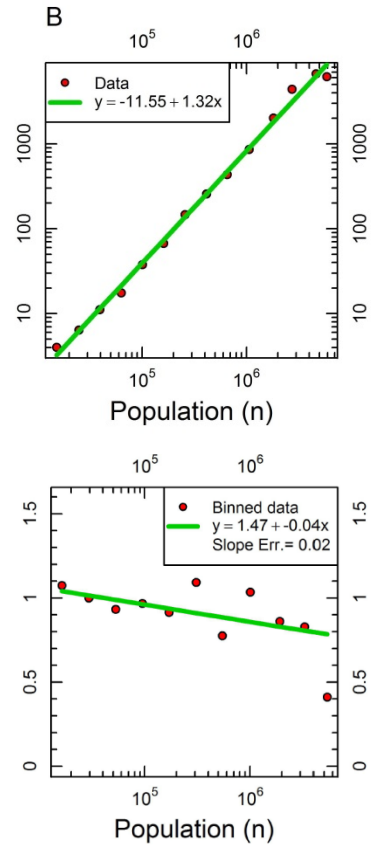


Population size dependence

Creatives

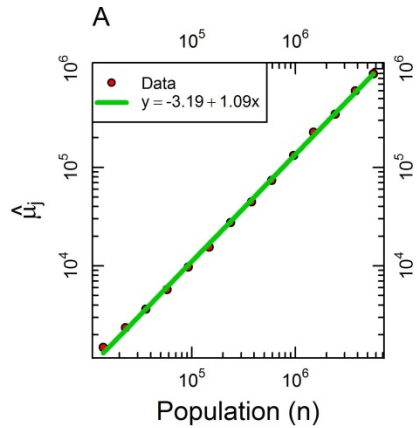


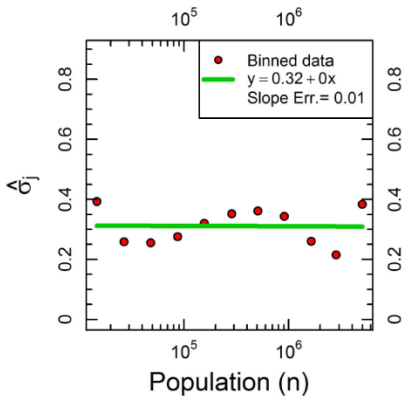
Inventors



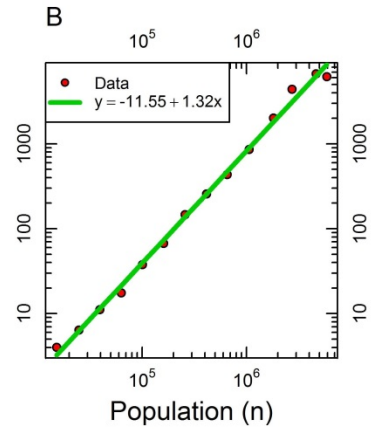
Population size dependence

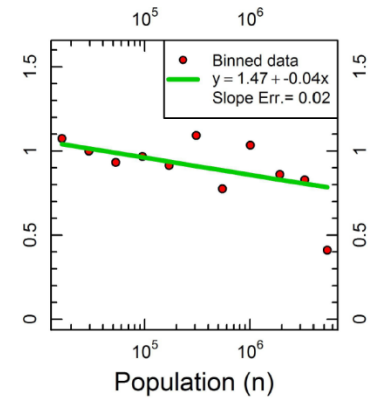
Creatives





Inventors

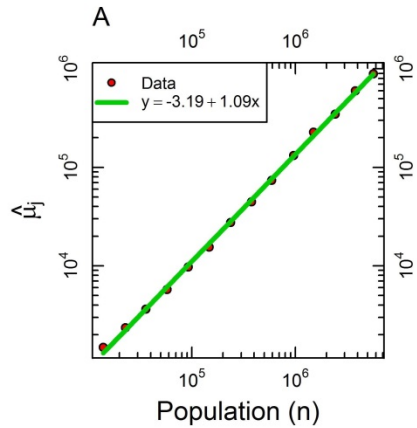


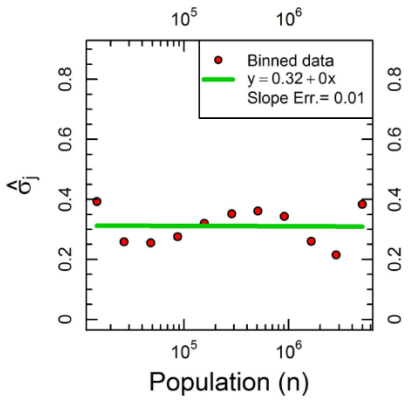


- Note this way we have completely characterized

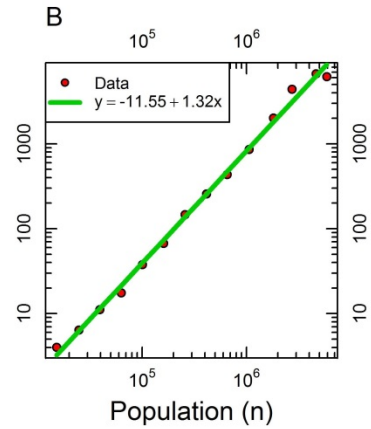
Population size dependence

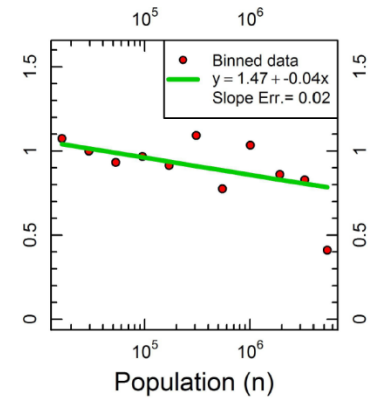
Creatives





Inventors





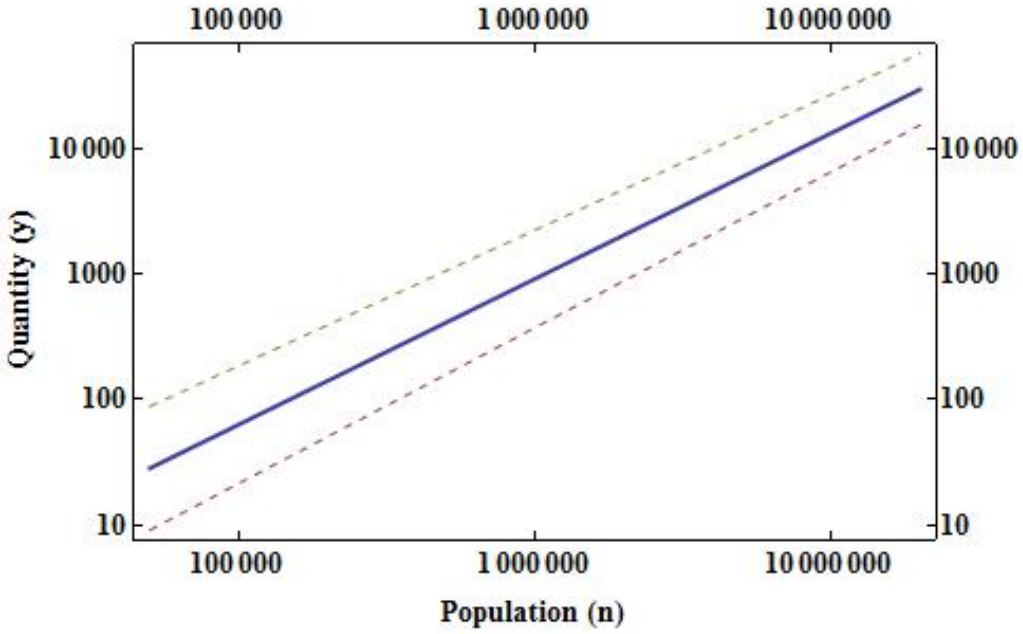
- Note this way we have completely characterized

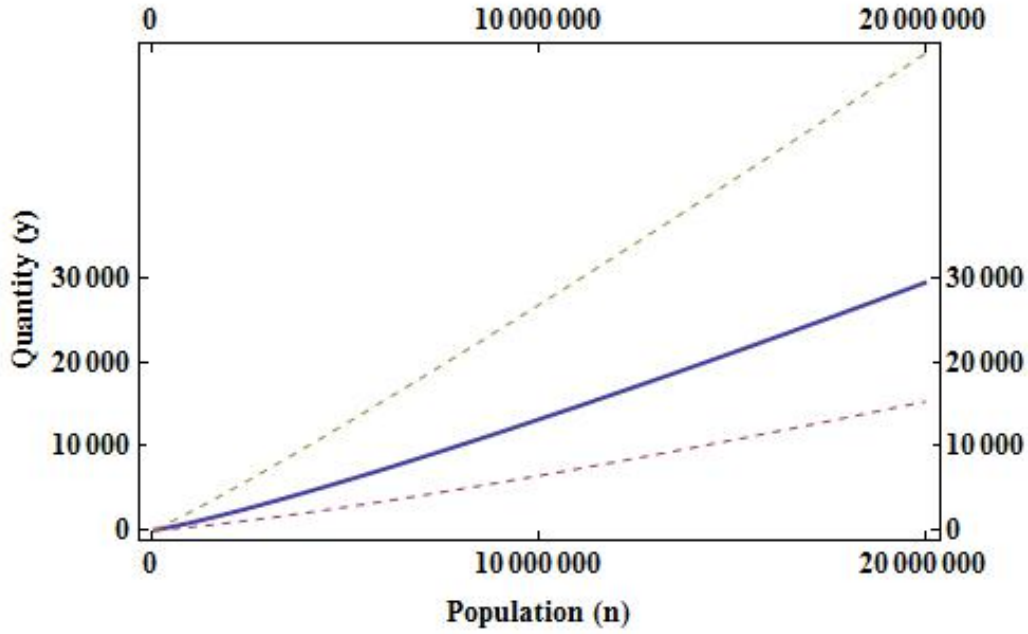
$$P(Y,N)=P(Y|N)P(N)$$

Logarithmic convergence

vs.

Absolute divergence of cities





OPEN ACCESS Freely available online

PLOS ONE

Urban Economies and Occupation Space: Can They Get "There" from "Here"?

Rachata Muneepeerakul[1,2]*, José Lobo[1], Shade T. Shutters[1,3], Andrés Goméz-Liévano[2,4], Murad R. Qubbaj[1]

1 School of Sustainability, Arizona State University, Tempe, Arizona, United States of America, **2** Mathematical, Computational, and Modeling Sciences Center, Arizona State University, Tempe, Arizona, United States of America, **3** Center for Social Dynamics and Complexity, Arizona State University, Tempe, Arizona, United States of America, **4** School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona, United States of America

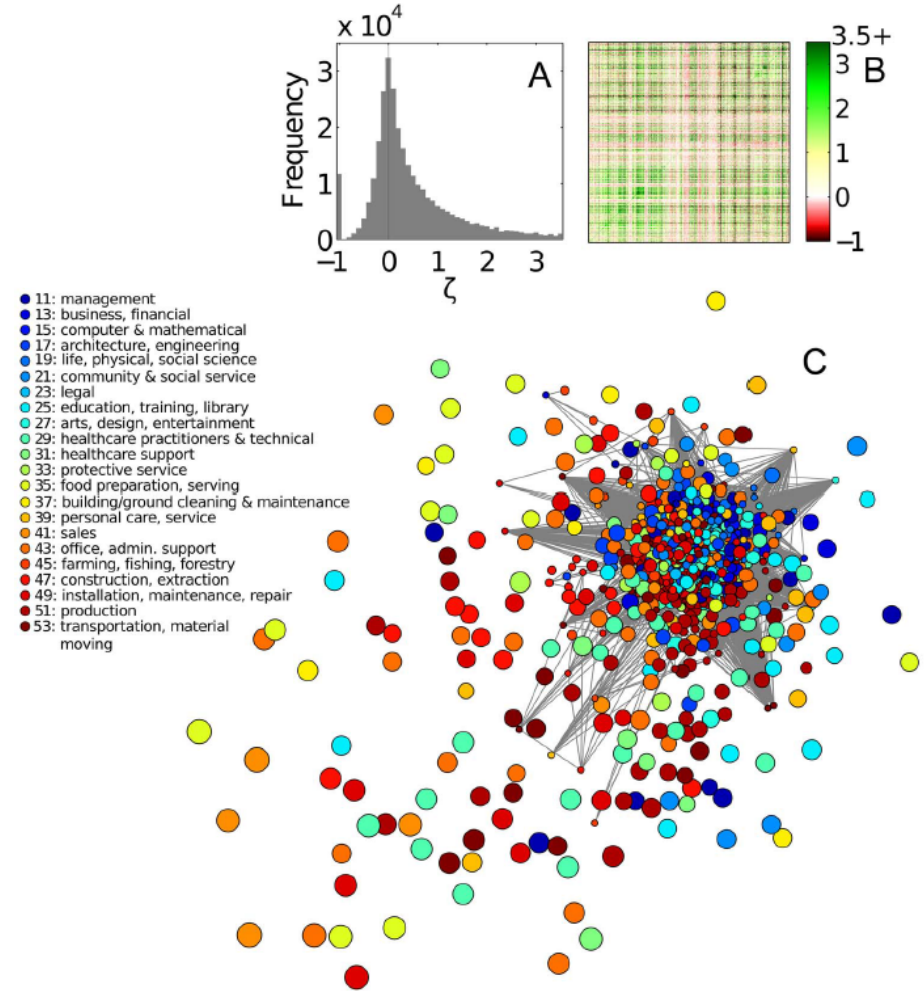
Abstract

Much of the socioeconomic life in the United States occurs in its urban areas. While an urban economy is defined to a large extent by its network of occupational specializations, an examination of this important network is absent from the considerable body of work on the determinants of urban economic performance. Here we develop a structure-based analysis addressing how the network of interdependencies among occupational specializations affects the ease with which urban economies can transform themselves. While most occupational specializations exhibit positive relationships between one another, many exhibit negative ones, and the balance between the two partially explains the productivity of an urban economy. The current set of occupational specializations of an urban economy and its location in the occupation space constrain its future development paths. Important tradeoffs exist between different alternatives for altering an occupational specialization pattern, both at a single occupation and an entire occupational portfolio levels.

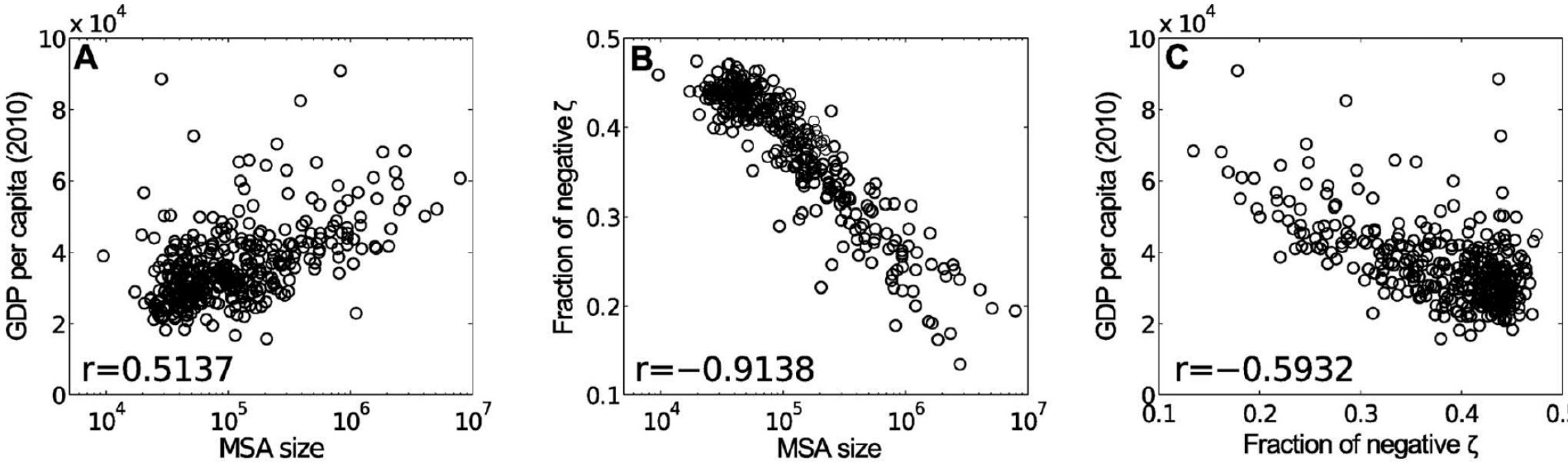
Occupational "interdependencies"

- Definition:

$$\zeta_{ij} = \frac{P[LQ_i^{(M)} > 1, LQ_j^{(M)} > 1]}{P[LQ_i^{(M')} > 1] P[LQ_j^{(M'')} > 1]} - 1$$



“Harmony” between occupations



Scaling relationships

(sensu Stat. Mech.*)*

$$
\begin{aligned}
X_1|N &\sim \mathcal{LN}(\mu_1, \sigma_1^2) \\
X_2|N &\sim \mathcal{LN}(\mu_2, \sigma_2^2) \\
X_2 &= f(X_1) \\
&\Rightarrow X_2 = aX_1^{\beta}, \beta = \frac{\sigma_2}{\sigma_1}
\end{aligned}
$$

$$
\begin{aligned}
X_1 &\sim \mathcal{P}(\alpha) \\
&\Rightarrow X_2 \sim \mathcal{P}(\tau),\ \tau = \frac{\alpha}{\beta}
\end{aligned}
$$