# Modeling cities using ideas from Economic Complexity and Cultural Evolution

SALURBAL, November 14, 2017

**Andres Gomez-Lievano, PhD**

Center for International Development,
Harvard University

Contact info:
Email: andres_gomez@hks.harvard.edu
Twitter: @GomezLievano

# The question

The **differences** in levels of wealth, innovation, disease prevalence, crime, etc., are often as large, or larger, across cities **within** a given country, than between countries.

# The question

The **differences** in levels of wealth, innovation, disease prevalence, crime, etc., are often as large, or larger, across cities **within** a given country, than between countries.

**How come?**

# Traditional approach

**Y:** dependent variable of interest

**X:** We hypothesize an explanatory variable

# Traditional approach

**Y:** dependent variable of interest

**X:** We hypothesize an explanatory variable

- Number of crimes ← Poverty, unemployment

# Traditional approach

**Y:** dependent variable of interest

**X:** We hypothesize an explanatory variable

- Number of crimes ← Poverty, unemployment
- Prevalence of disease ← Provision of public health

# Traditional approach

**Y:** dependent variable of interest

**X:** We hypothesize an explanatory variable

- Number of crimes ← Poverty, unemployment
- Prevalence of disease ← Provision of public health
- Income ← Education level

# Traditional approach

**Y:** dependent variable of interest

**X:** We hypothesize an explanatory variable

- Number of crimes ← Poverty, unemployment
- Prevalence of disease ← Provision of public health
- Income ← Education level
- Innovation output ← Number of universities, government funding

# Traditional approach

**Y:** dependent variable of interest

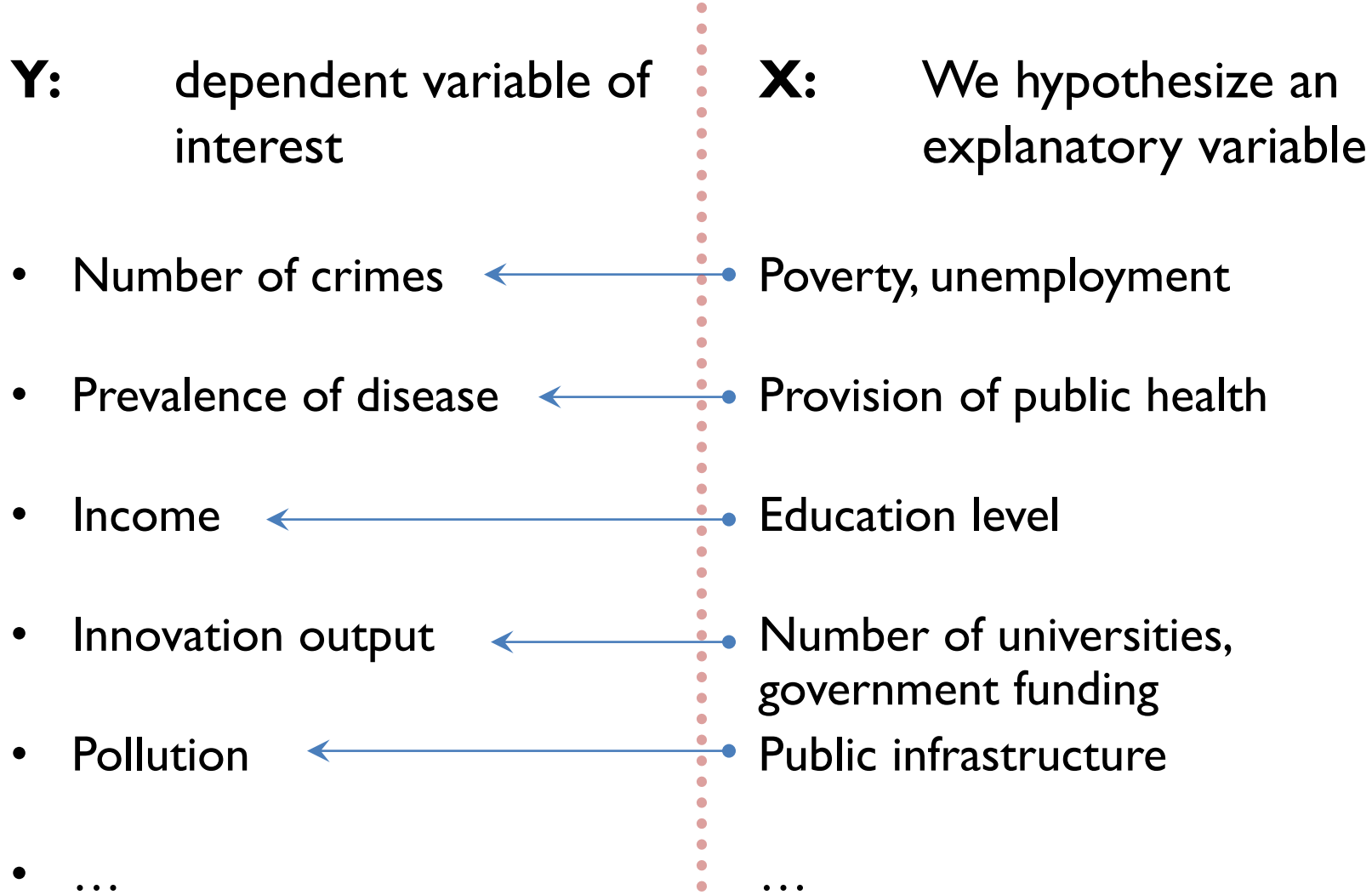
**X:** We hypothesize an explanatory variable

- Number of crimes ← Poverty, unemployment
- Prevalence of disease ← Provision of public health
- Income ← Education level
- Innovation output ← Number of universities, government funding
- Pollution ← Public infrastructure

# Traditional approach

**Y:** dependent variable of interest

**X:** We hypothesize an explanatory variable



This is great, but it is important to remind ourselves that this is not the only approach.

# Thought experiment

Theory: $Y = f_1(X)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$

Theory: $Y = f_2(X, W)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$

Theory: $Y = f_3(X, W, Z)$ → Empirics: …

# Thought experiment

Theory: $Y = f_1(X)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$

Theory: $Y = f_2(X, W)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$

Theory: $Y = f_3(X, W, Z)$ → Empirics: …

- Where are we going?

# Thought experiment

Theory: $Y = f_1(X)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$

Theory: $Y = f_2(X, W)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$

Theory: $Y = f_3(X, W, Z)$ → Empirics: …

- Where are we going?
- What happens if we think in the limit?

# Thought experiment

Theory: $Y = f_1(X)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$

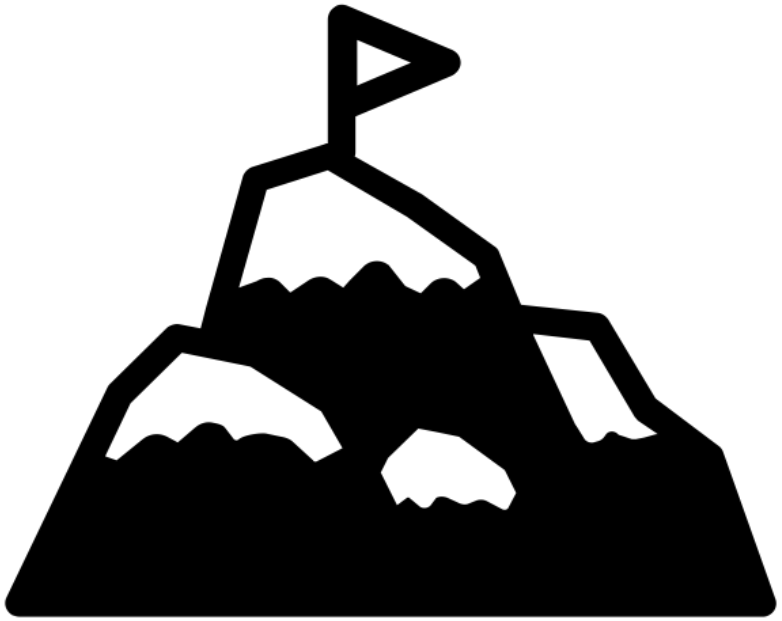
Theory: $Y = f_2(X, W)$ → Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$

Theory: $Y = f_3(X, W, Z)$ → Empirics: …

- Where are we going?
- What happens if we think in the limit?

$$Y = f(X_1, X_2, X_3, X_4, \ldots \longrightarrow \infty)$$

*TRUTH*



Icons:

Created by Peter van Driel
from Noun Project

Created by Martyn Jasinski
from Noun Project

*TRUTH*

## Traditional approach

| Y: dependent variable of interest | X: We hypothesize an explanatory variable |
|---|---|
| • Number of crimes | Poverty, unemployment |
| • Prevalence of disease | Provision of public health |
| • Income | Education level |
| • Innovation output | Number of universities, government funding |
| • Pollution | Income |
| • … | … |

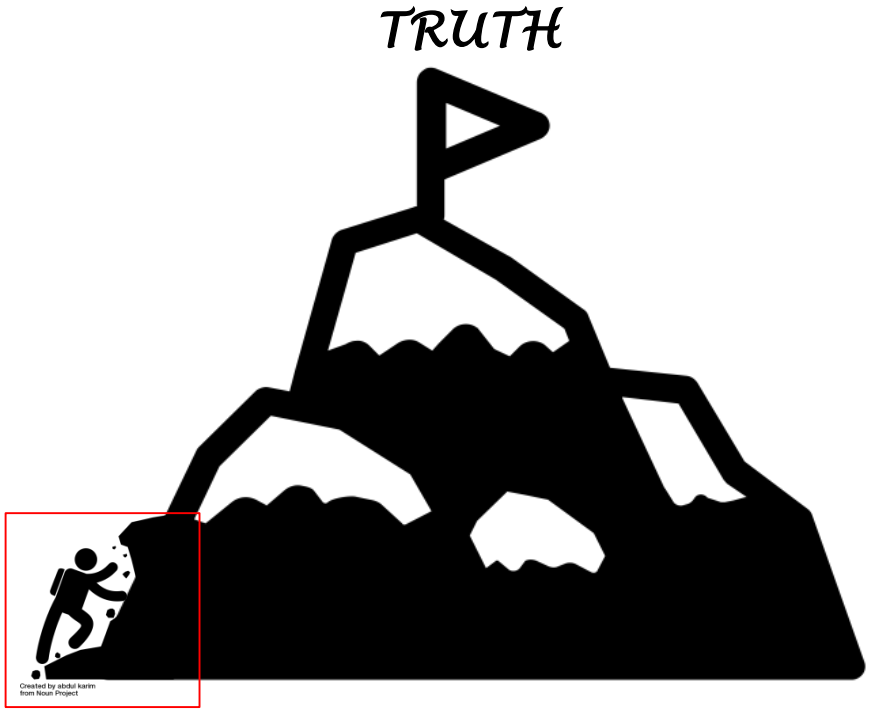
Icons:

Created by Peter van Driel from Noun Project

Created by Martyn Jasinski from Noun Project

## Traditional approach

| Y: dependent variable of interest | X: We hypothesize an explanatory variable |
|---|---|
| • Number of crimes | Poverty, unemployment |
| • Prevalence of disease | Provision of public health |
| • Income | Education level |
| • Innovation output | Number of universities, government funding |
| • Pollution | Income |
| • … | … |



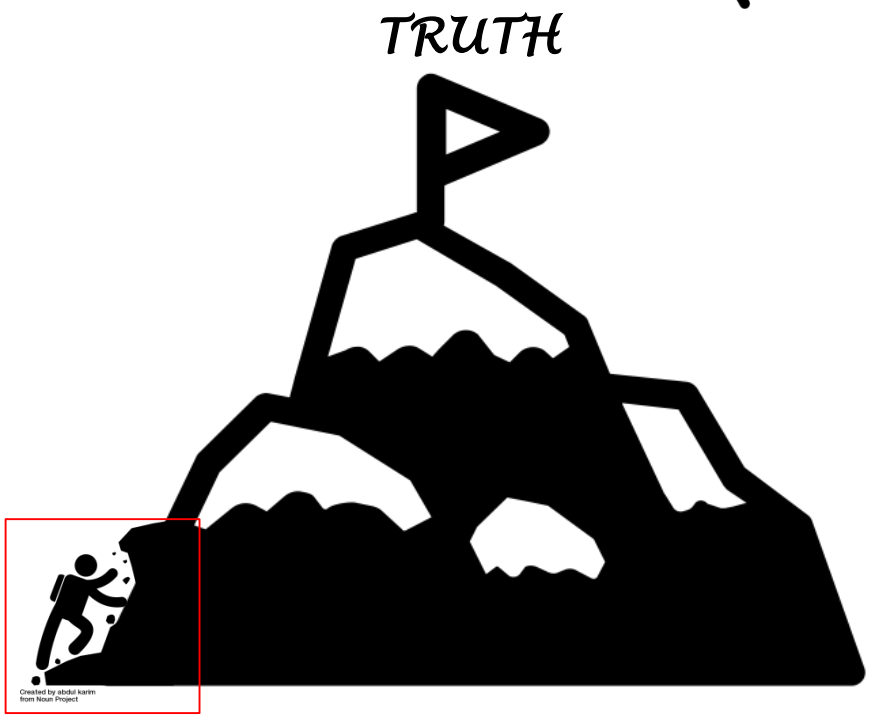

Explaining ***Y*** one variable at a time…

Icons:

Created by Peter van Driel from Noun Project

Created by Martyn Jasinski from Noun Project

Explicitly assume ***Y*** is the consequence of ***many, many, many explanatory variables***.




## Traditional approach

| Y: dependent variable of interest | X: We hypothesize an explanatory variable |
|---|---|
| Number of crimes | Poverty, unemployment |
| Prevalence of disease | Provision of public health |
| Income | Education level |
| Innovation output | Number of universities, government funding |
| Pollution | Income |
| … | … |

Explaining ***Y*** one variable at a time…

Icons:

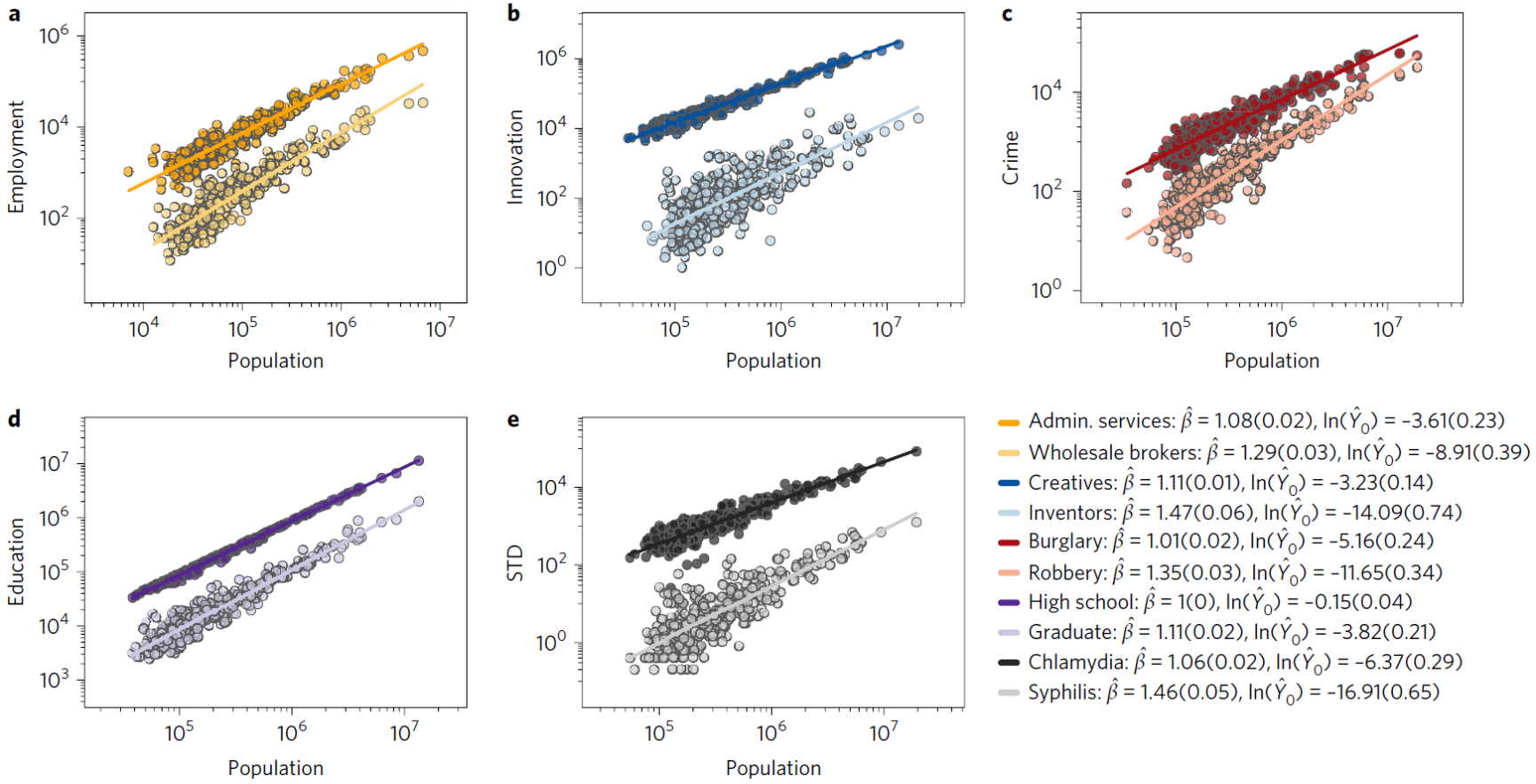
Created by Peter van Driel from Noun Project

Created by Martyn Jasinski from Noun Project

# The "view from above"

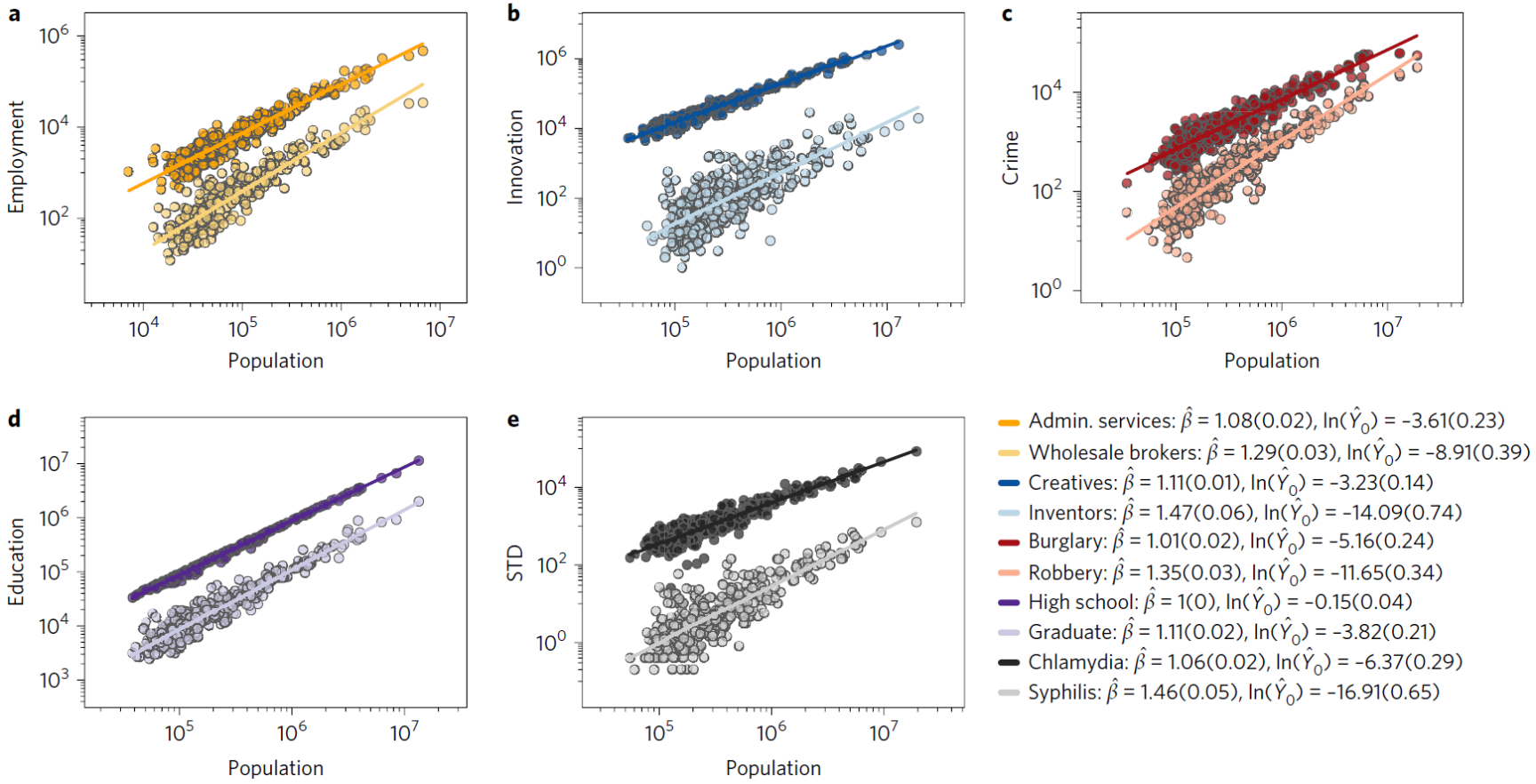
- The study of probability **distributions**,
  *not* of means.
- The study of the **processes** that give rise to the distributions,
  *not* of specific pathways/mechanisms.
- The study of **scale** and how phenomena in a system change with it,
  *not* just looking at a single city and a single scale.

# Four facts (among others) to try explaining



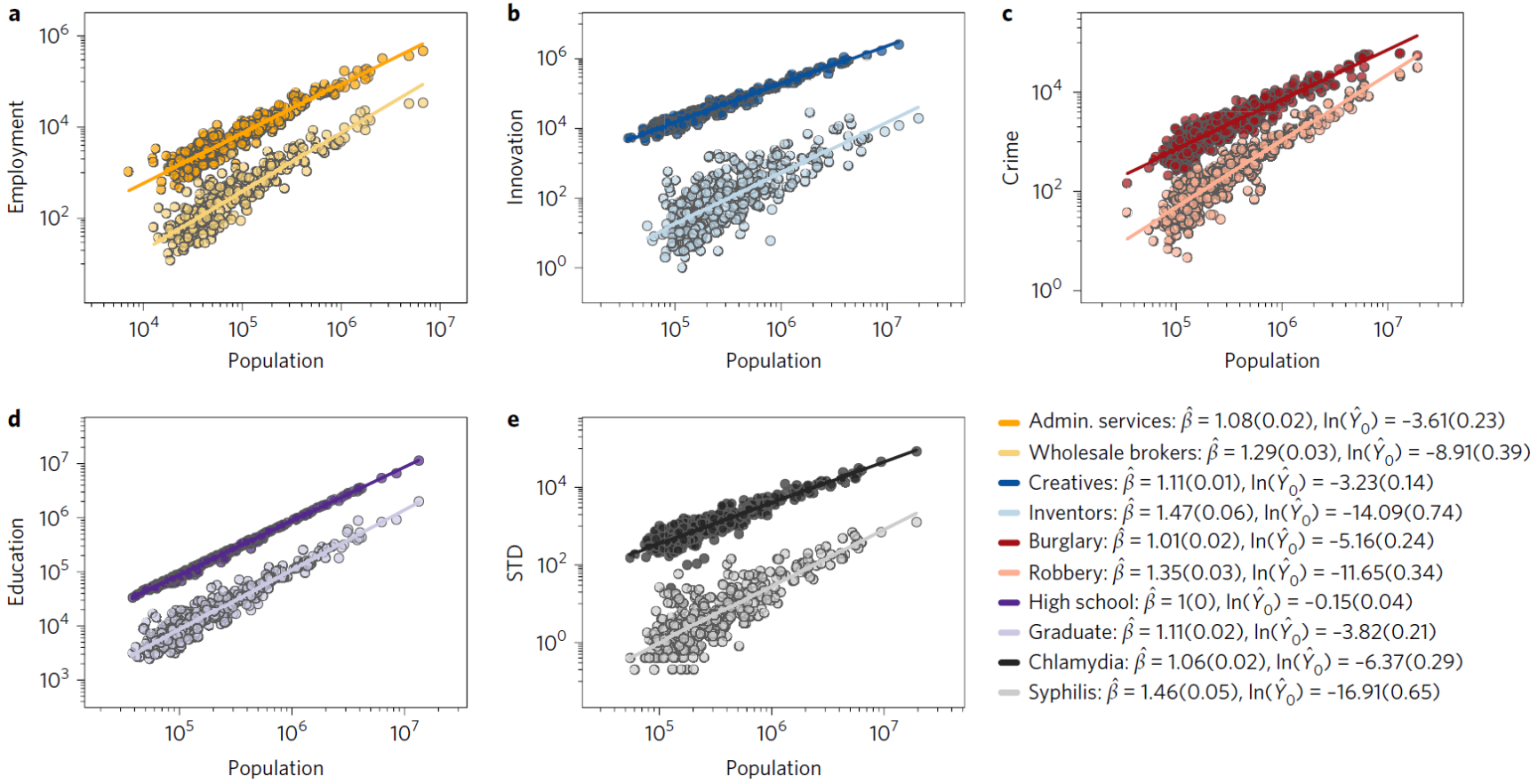
1. log-log → straight line

# Four facts (among others) to try explaining



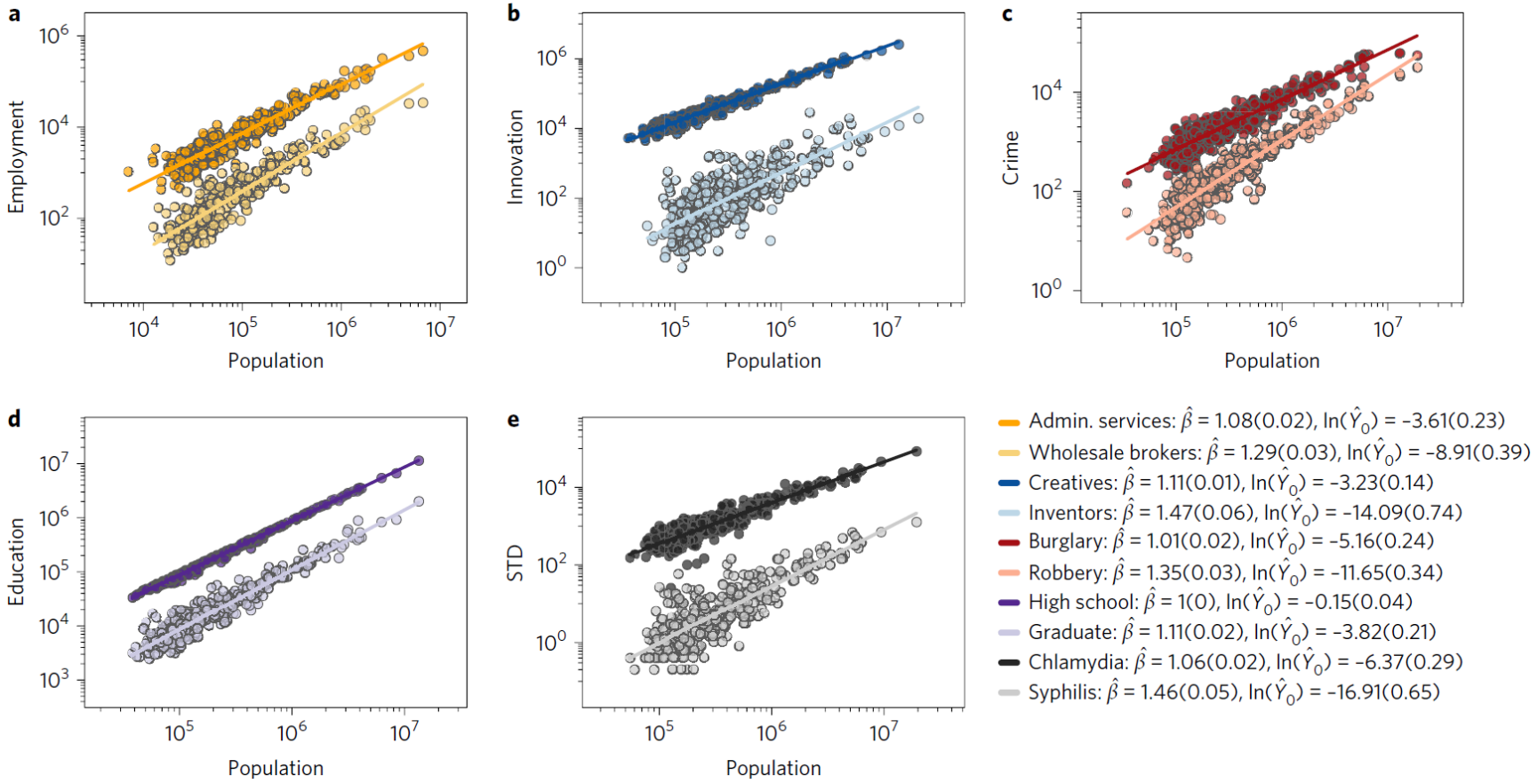
# Four facts (among others) to try explaining

1. log-log → straight line
2. each phenomenon with its own slope,



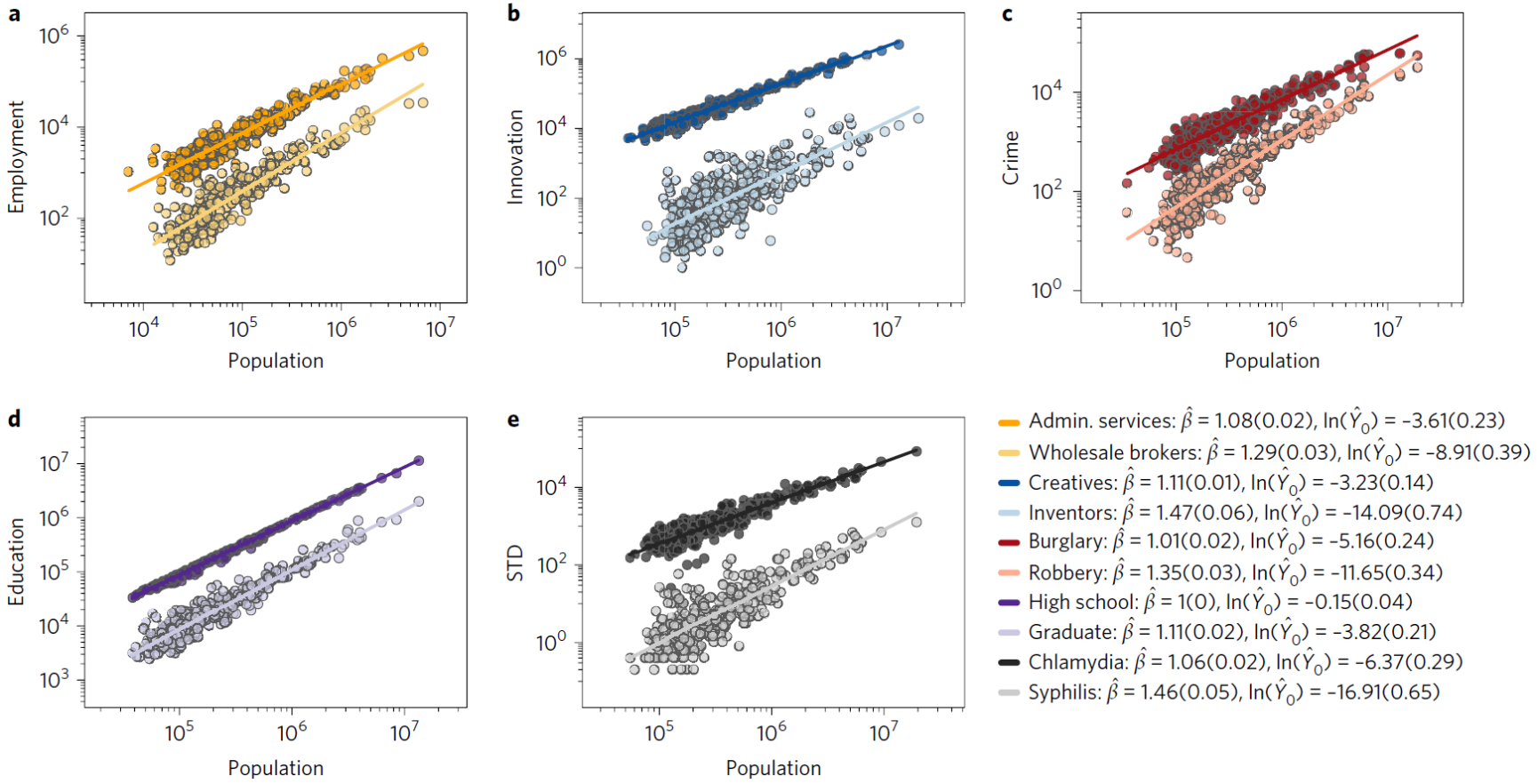
# Four facts (among others) to try explaining

1. log-log → straight line
2. each phenomenon with its own slope,
3. with its own intercept, and



# Four facts (among others) to try explaining

1. log-log → straight line
2. each phenomenon with its own slope,
3. with its own intercept, and
4. with its own dispersion.



# The theory:

(work together with O. Patterson-Lomba and R. Hausmann)

# The theory:

## (work together with O. Patterson-Lomba and R. Hausmann)

1. Most of urban phenomena are the conjunction of complementary factors.

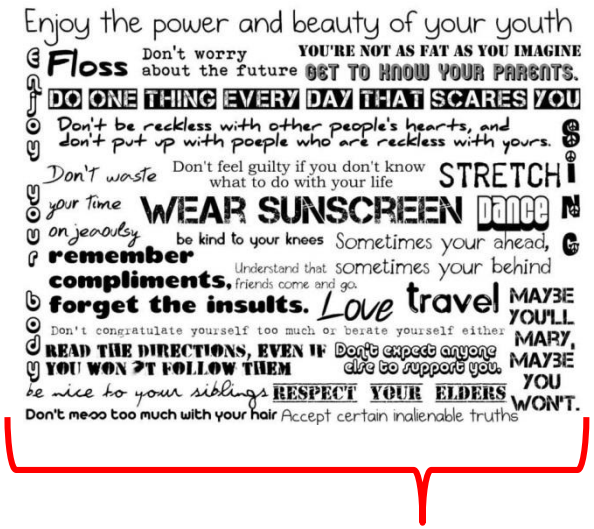
Being happy:

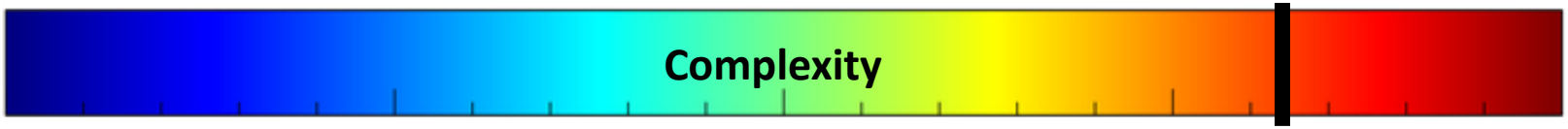
# The theory:

## (work together with O. Patterson-Lomba and R. Hausmann)

1. Most of urban phenomena are the conjunction of complementary factors.

Being happy:



# The theory:

## (work together with O. Patterson-Lomba and R. Hausmann)

1. Most of urban phenomena are the conjunction of complementary factors.

2. Cities accumulate these factors through a stochastic evolutionary process of variation and selection, and cities are the vessels were these factors combine.

# The theory:

## (work together with O. Patterson-Lomba and R. Hausmann)

1. Most of urban phenomena are the conjunction of complementary factors.
2. Cities accumulate these factors through a stochastic evolutionary process of variation and selection, and cities are the vessels were these factors combine.
3. And each person in the city is different in the factors they are susceptible to.

# The theory:

(work together with O. Patterson-Lomba and R. Hausmann)

1. Most of urban phenomena are the conjunction of complementary factors.

2. Cities accumulate these factors through a stochastic evolutionary process of variation and selection, and cities are the vessels were these factors combine.

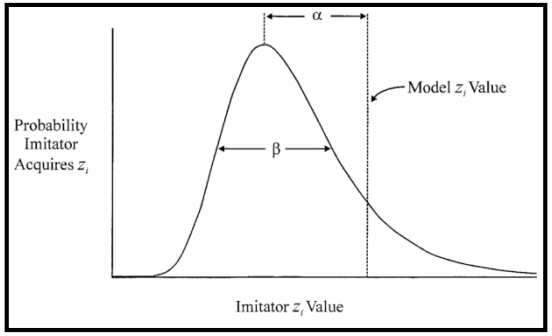
3. And each person in the city is different in the factors they are susceptible to.

$$\Pr\{X_{i,c,f} = 1\} = e^{-M_f q_i (1 - r_c)}$$

## Models of economic complexity


&

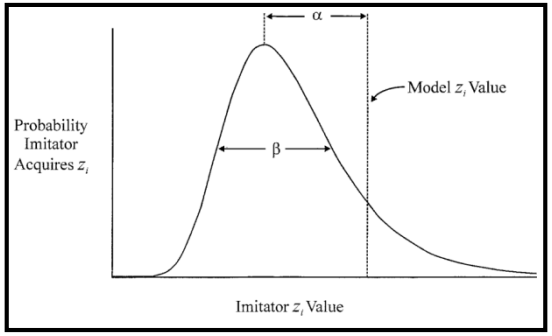
## Models of cultural evolution



Models of economic complexity


&

Models of cultural evolution



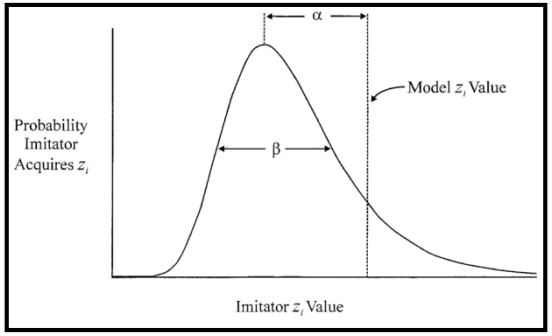

$$\mathbb{E}[Y] = e^{-Mq} N e^{Mq\, r(N)}$$

$$r(N) = a + b\ln(N)$$

Models of economic complexity


&

Models of cultural evolution




$$\mathbb{E}[Y] = e^{-Mq} N e^{Mq\, r(N)}$$

$$r(N) = a + b\ln(N)$$

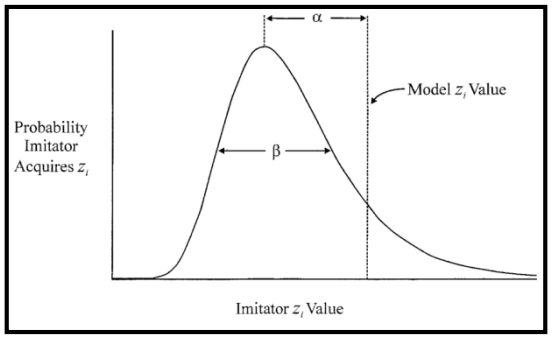
- Urban Scaling: $\mathbb{E}[Y] = e^{-Mq(1-a)} N^{1+Mqb}$

$$= Y_0\, N^{\beta}$$

Models of economic complexity


&

Models of cultural evolution

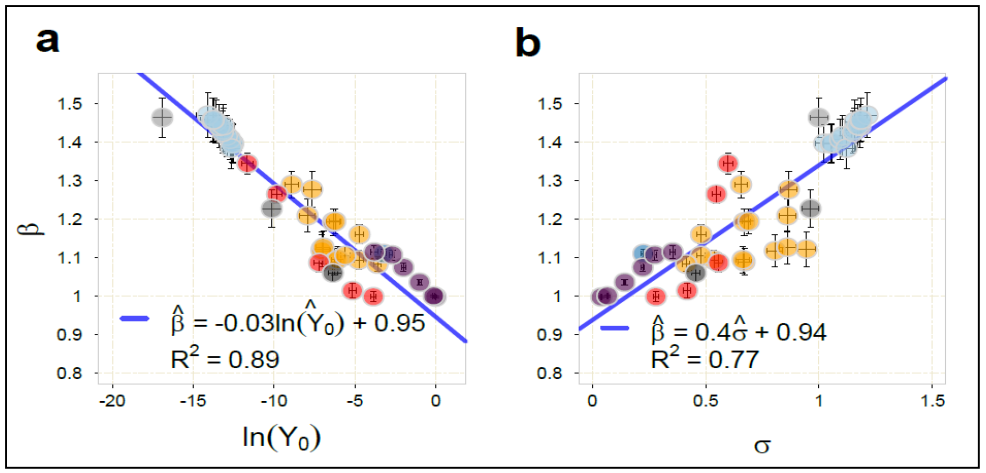



$$\mathbb{E}[Y] = e^{-Mq} N e^{Mq\, r(N)}$$

$$r(N) = a + b\ln(N)$$

- Urban Scaling: $\mathbb{E}[Y] = e^{-Mq(1-a)} N^{1+Mqb}$

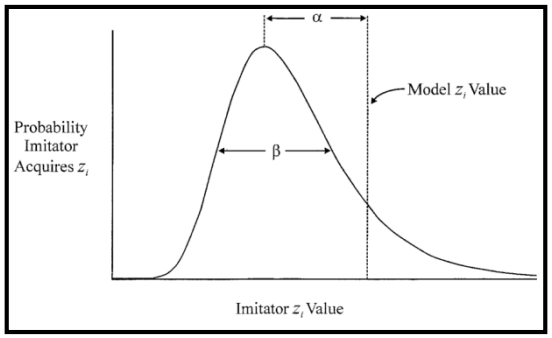
$$= Y_0\, N^{\beta}$$



Models of economic complexity


&

Models of cultural evolution

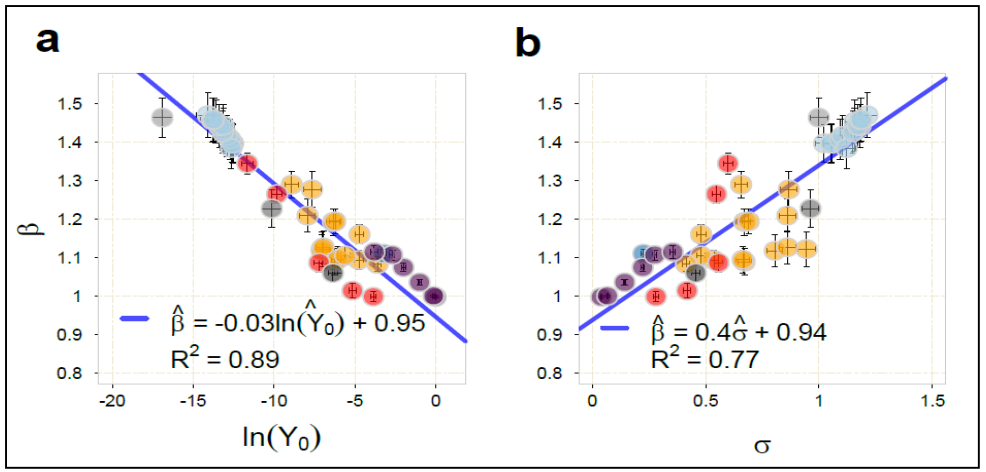


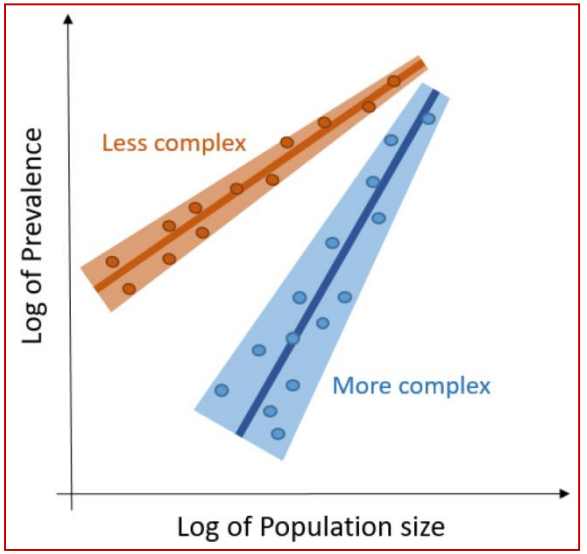

$$\mathbb{E}[Y] = e^{-Mq} N e^{Mq\, r(N)}$$

$$r(N) = a + b\ln(N)$$

- Urban Scaling: $\mathbb{E}[Y] = e^{-Mq(1-a)} N^{1+Mqb}$

$$= Y_0\, N^{\beta}$$

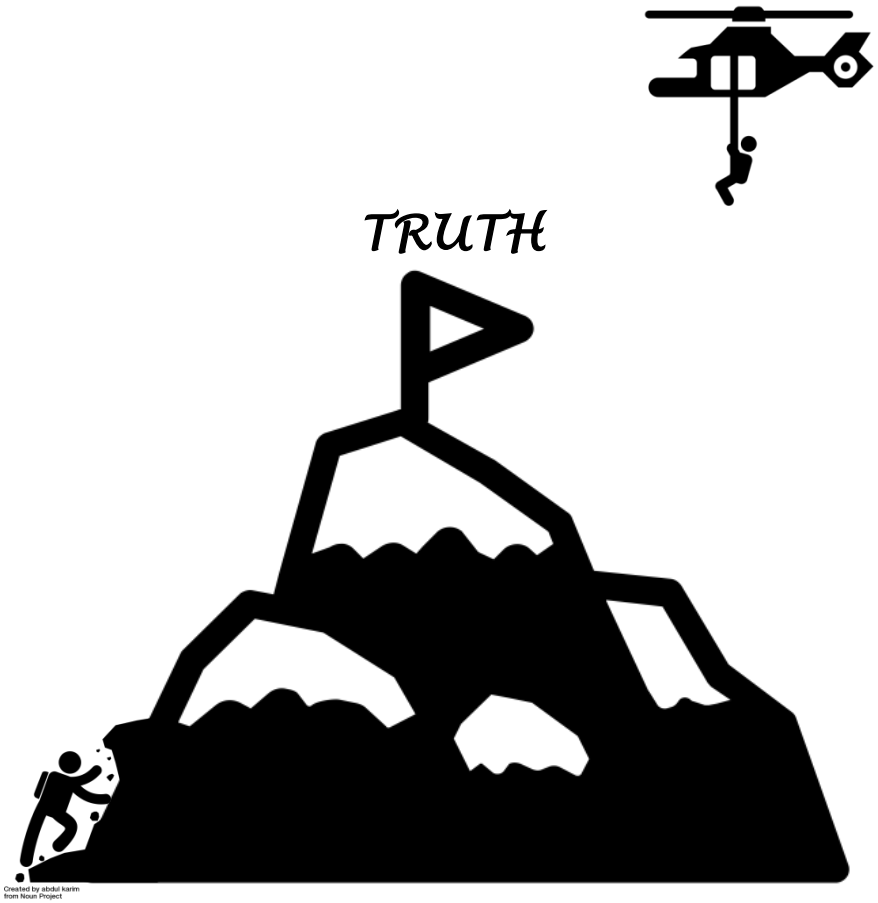


# Limitations/Future Directions

- Too simple. E.g., there are no dynamics. Hence, no sense of time-scales for how phenomena change in time.
- Cities are internally very heterogeneous, and segregated. How to include *spatial* heterogeneity?
- Too discrete (person *in* or *out*).

# Takeaways


TRUTH
Created by abdul karim
from Noun Project

# The importance of emergent phenomena

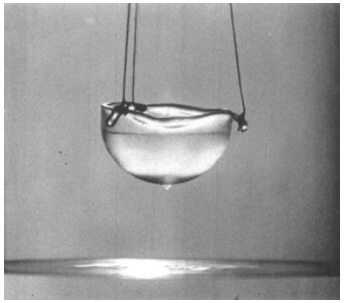
To understand a systemic phenomenon, consider

***how many*** parts and
how they ***interact***,

than

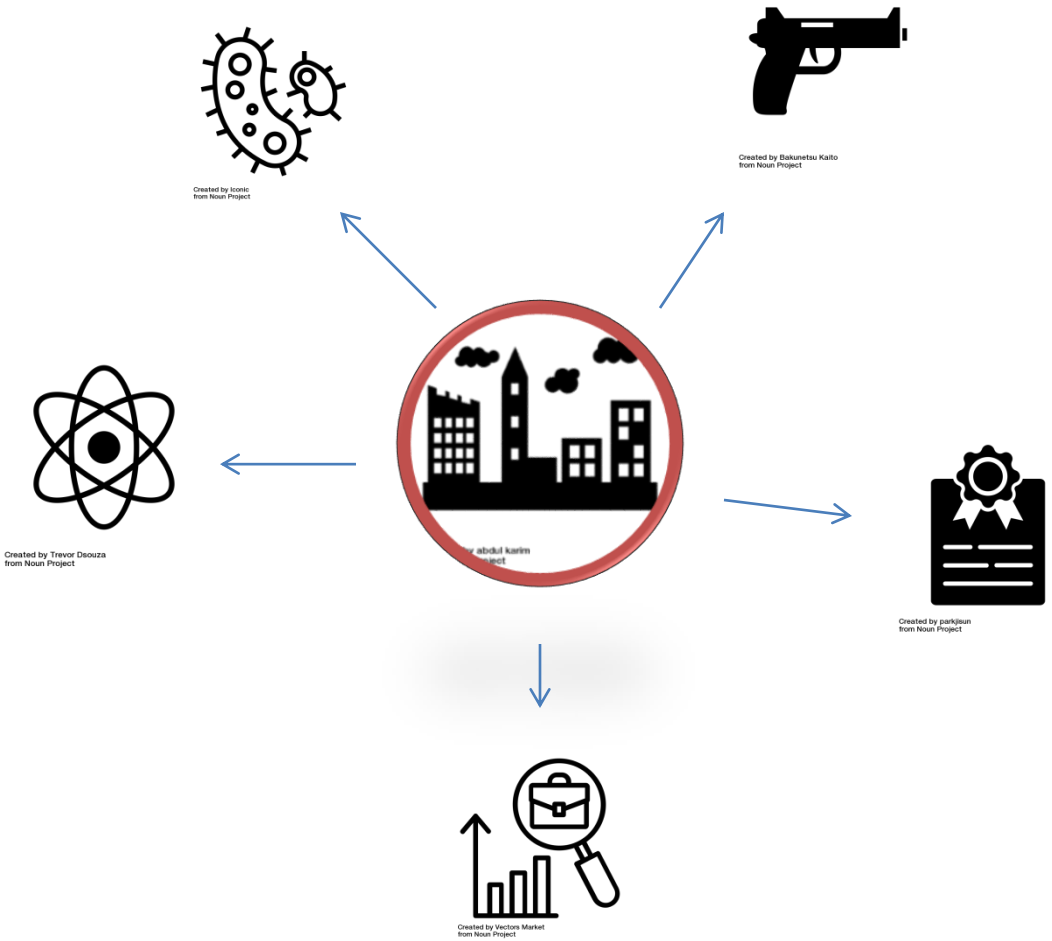
***what are*** the parts.

The fact that we understand the liquidity of water *means* that we understand liquidity in other materials.



# Urban “outcomes” as emergent phenomena

(→ my research ☺ )



# mentors and colleagues











# THANK YOU

Contact info:
Email: andres_gomez@hks.harvard.edu
Twitter: @GomezLievano

**References:**


- Gomez-Lievano, Patterson-Lomba, and Hausmann (2016, Nature Hum Behav);

Henrich (2004, Am. Ant.)

Hidalgo, Klinger, Barabási, and Hausmann (2007, Science);

Hidalgo and Hausmann (2009, PNAS);

Bettencourt, Lobo, Helbing, Kühnert, and West (2007, PNAS);

Gomez-Lievano, Youn, and Bettencourt (2012, PLoS ONE);

Gomez-Lievano, Bettencourt, Stolarick, Strumsky, and Lobo (2017, in preparation);

Muneepeerakul, Lobo, Shutters, Gomez-Lievano, and Qubbaj (2013, PLoS ONE);

Towers, Gomez-Lievano, Khan, Mubayi, and Castillo-Chavez (2015, PLoS ONE);

Patterson-Lomba, Goldstein, Gomez-Lievano, Castillo-Chavez, and Towers (2015, Sex Transm Infect);