

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](https://twitter.com/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

- Number of crimes

- Prevalence of disease

X: We hypothesize an explanatory variable

- Poverty, unemployment

- Provision of public health



Traditional approach

Y: dependent variable of interest

- Number of crimes

- Prevalence of disease

- Income

X: We hypothesize an explanatory variable

- Poverty, unemployment

- Provision of public health

- Education level



Traditional approach

Y: dependent variable of interest

- Number of crimes

- Prevalence of disease

- Income

- Innovation output

X: We hypothesize an explanatory variable

- Poverty, unemployment

- Provision of public health

- Education level

- Number of universities, government funding

Traditional approach

Y: dependent variable of interest

- Number of crimes

- Prevalence of disease

- Income

- Innovation output

- Pollution

X: We hypothesize an explanatory variable

- Poverty, unemployment

- Provision of public health

- Education level

- Number of universities, government funding

- Public infrastructure

Traditional approach

Y: dependent variable of interest

- Number of crimes

- Prevalence of disease

- Income

- Innovation output

- Pollution

- ...

X: We hypothesize an explanatory variable

- Poverty, unemployment

- Provision of public health

- Education level

- Number of universities, government funding

- Public infrastructure

- ...

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

Thought experiment

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

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

Theory: $Y = f_3(X, W, Z)$ \longrightarrow Empirics: ...

Thought experiment

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

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

Theory: $Y = f_3(X, W, Z)$ \longrightarrow Empirics: ...

- Where are we going?

Thought experiment

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

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

Theory: $Y = f_3(X, W, Z)$ \longrightarrow Empirics: ...

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

Thought experiment

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

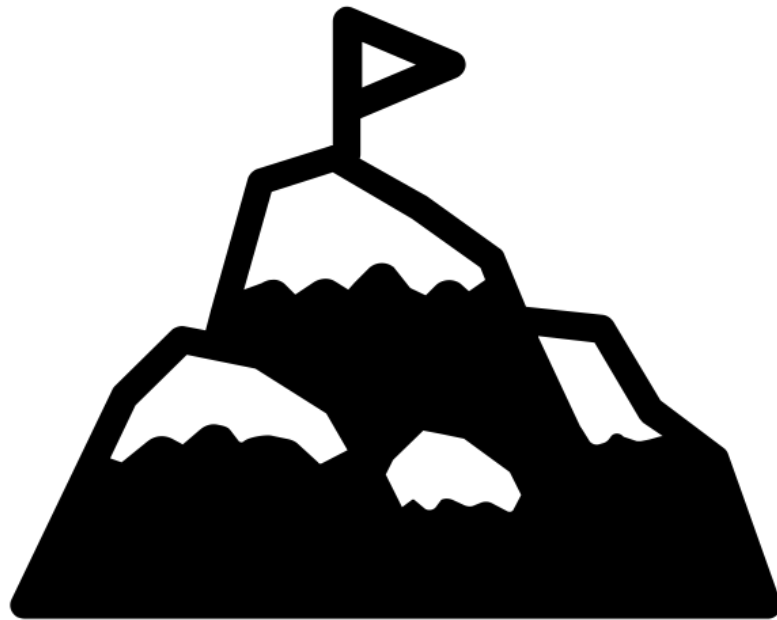
Theory: $Y = f_2(X, W)$ \longrightarrow Empirics: $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$

Theory: $Y = f_3(X, W, Z)$ \longrightarrow Empirics: ...

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

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

TRUTH

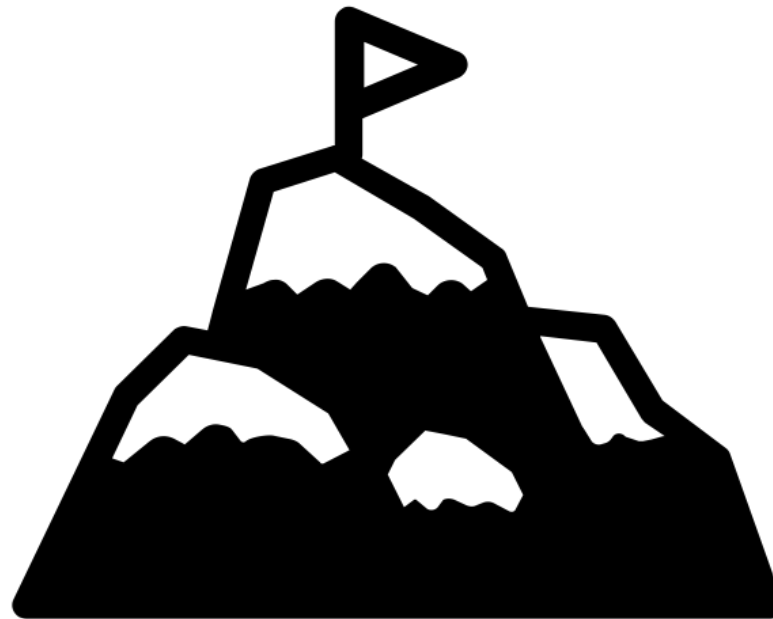


Icons:

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

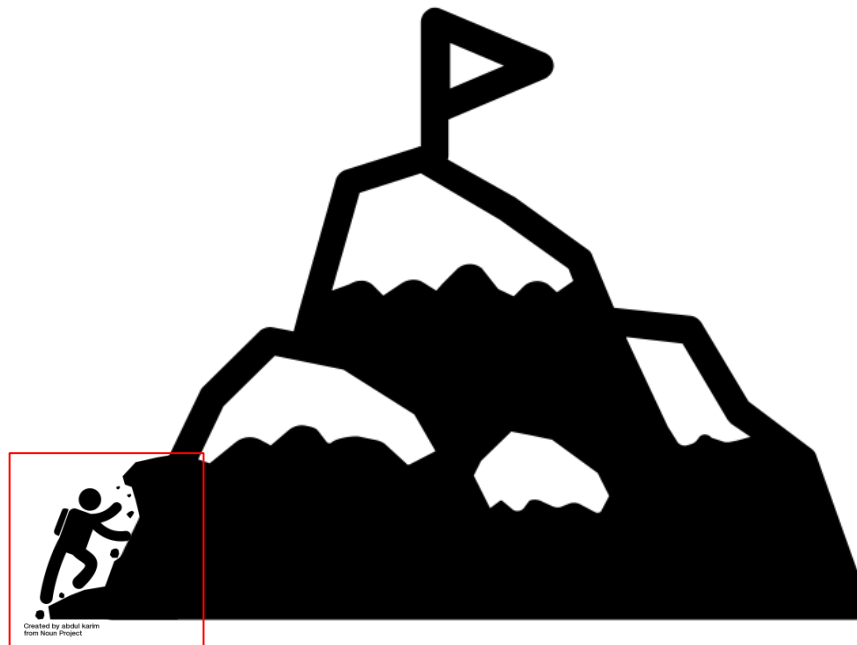
- | | | |
|-------------------------|---|--|
| • Number of crimes | ← | • Poverty, unemployment |
| • Prevalence of disease | ← | • Provision of public health |
| • Income | ← | • Education level |
| • Innovation output | ← | • Number of universities, government funding |
| • Pollution | ← | • Income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

Icons:

Created by Peter van Driel
from Noun Project

Created by Martyn Jasinski
from Noun Project

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 |
| • ... | | • ... |

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

Icons:

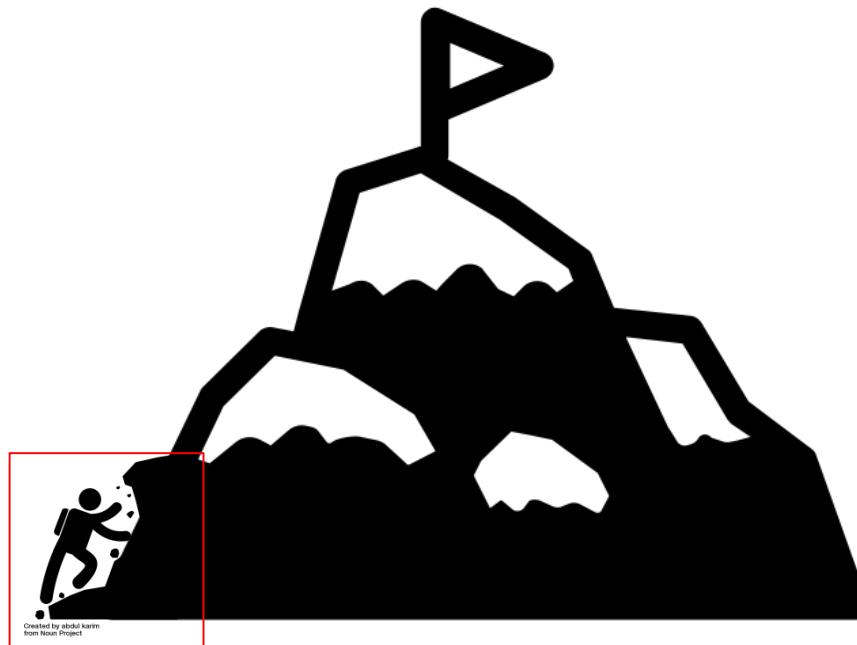
Created by Peter van Driel
from Noun Project

Created by Martyn Jasinski
from Noun Project



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

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 |
| • ... | | • ... |

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

Icons:

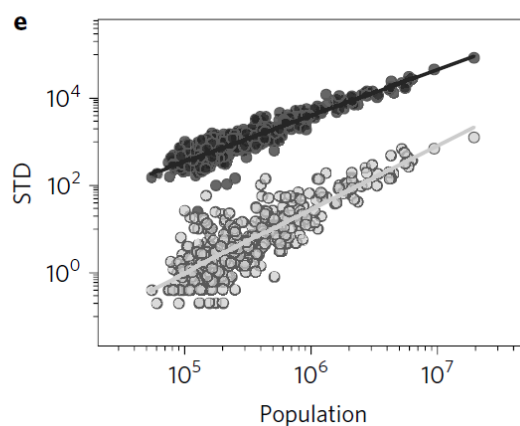
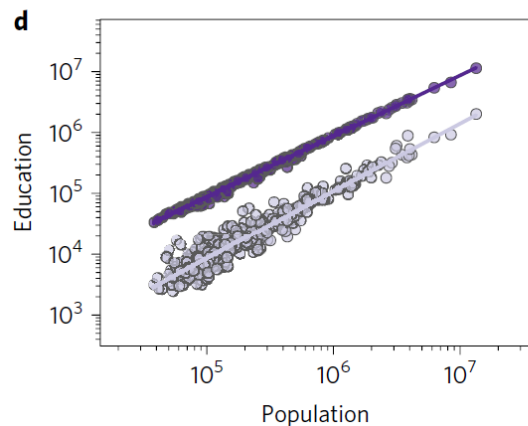
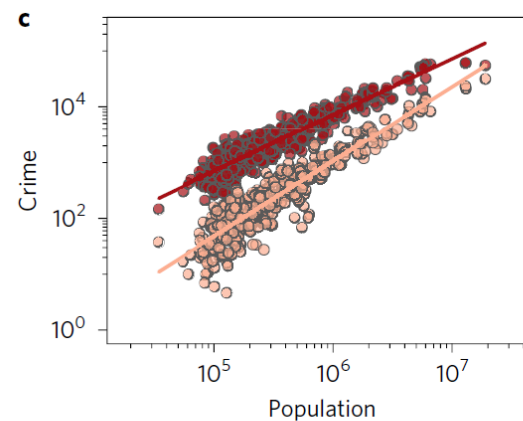
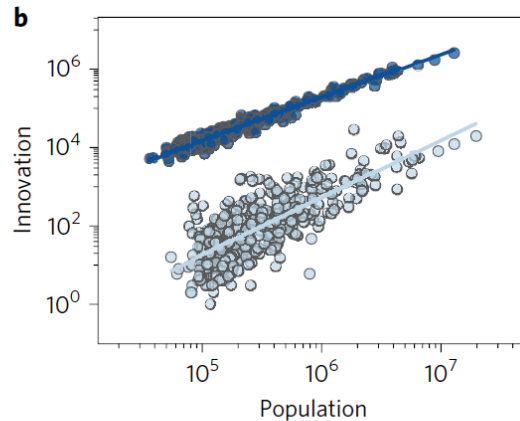
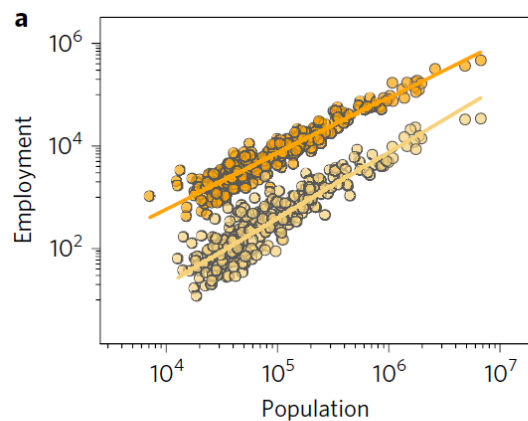
Created by Peter van Driel
from Noun Project

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.

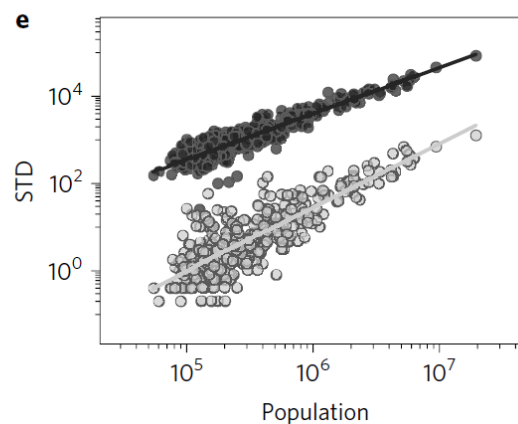
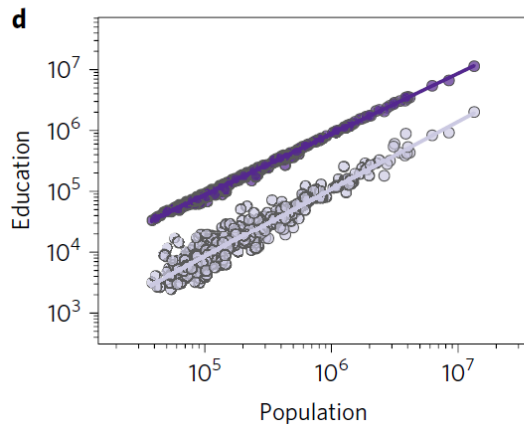
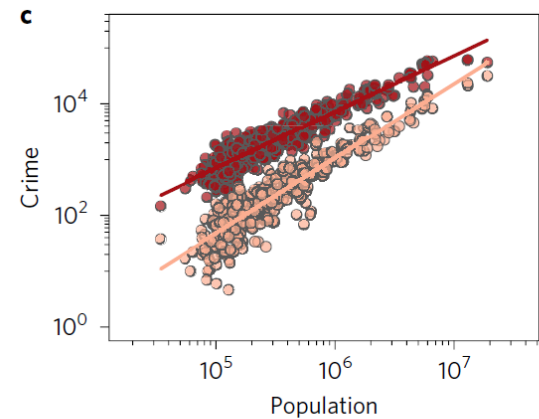
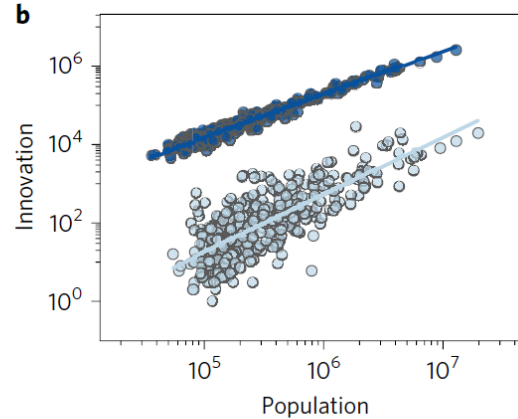
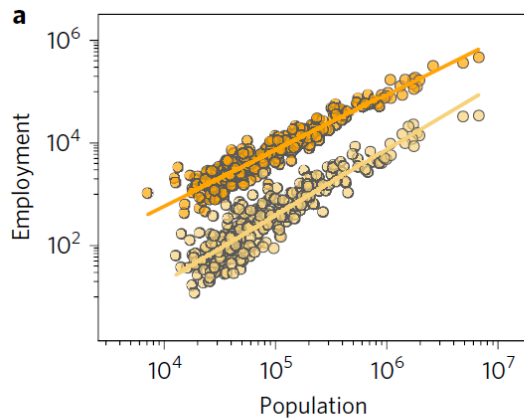
Four facts (among others) to try explaining



- Admin. services: $\hat{\beta} = 1.08(0.02)$, $\ln(\hat{Y}_0) = -3.61(0.23)$
- Wholesale brokers: $\hat{\beta} = 1.29(0.03)$, $\ln(\hat{Y}_0) = -8.91(0.39)$
- Creatives: $\hat{\beta} = 1.11(0.01)$, $\ln(\hat{Y}_0) = -3.23(0.14)$
- Inventors: $\hat{\beta} = 1.47(0.06)$, $\ln(\hat{Y}_0) = -14.09(0.74)$
- Burglary: $\hat{\beta} = 1.01(0.02)$, $\ln(\hat{Y}_0) = -5.16(0.24)$
- Robbery: $\hat{\beta} = 1.35(0.03)$, $\ln(\hat{Y}_0) = -11.65(0.34)$
- High school: $\hat{\beta} = 1(0)$, $\ln(\hat{Y}_0) = -0.15(0.04)$
- Graduate: $\hat{\beta} = 1.11(0.02)$, $\ln(\hat{Y}_0) = -3.82(0.21)$
- Chlamydia: $\hat{\beta} = 1.06(0.02)$, $\ln(\hat{Y}_0) = -6.37(0.29)$
- Syphilis: $\hat{\beta} = 1.46(0.05)$, $\ln(\hat{Y}_0) = -16.91(0.65)$

I. log-log \rightarrow straight line

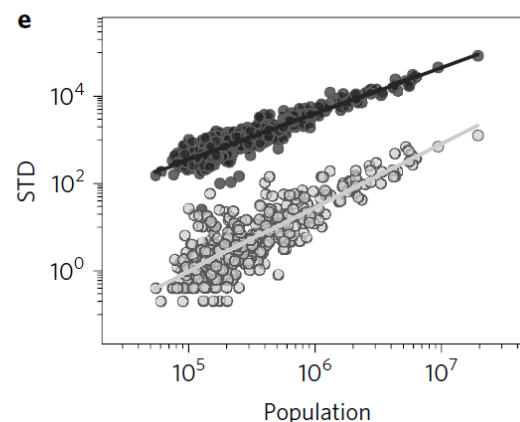
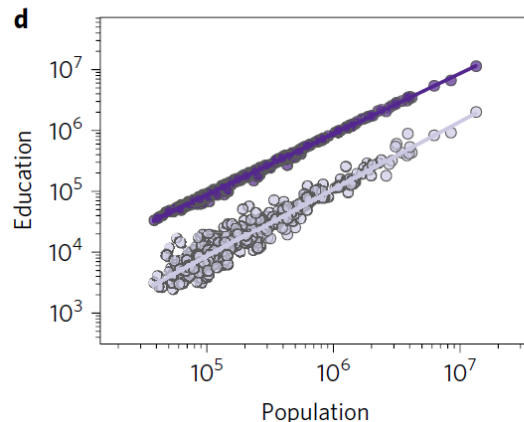
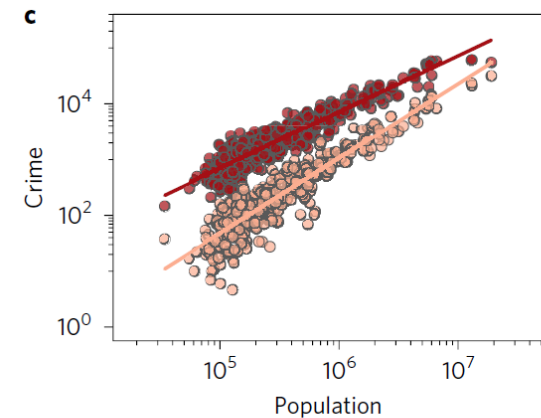
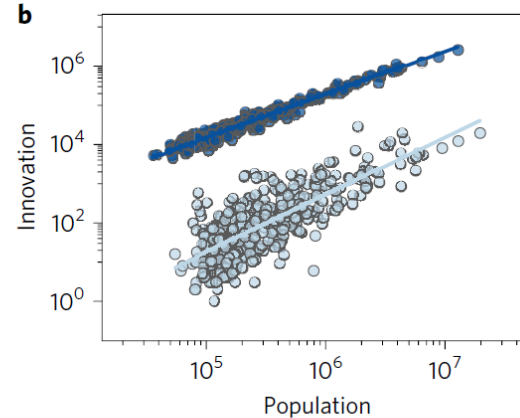
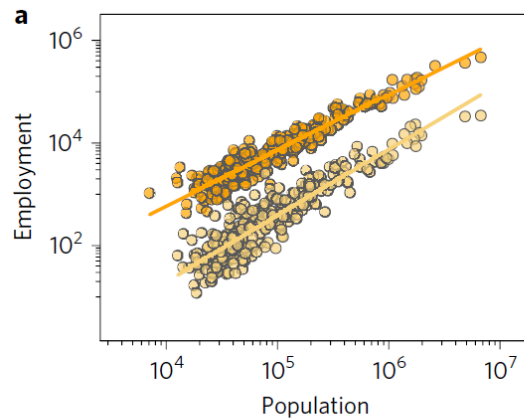
Four facts (among others) to try explaining



- Admin. services: $\hat{\beta} = 1.08(0.02)$, $\ln(\hat{Y}_0) = -3.61(0.23)$
- Wholesale brokers: $\hat{\beta} = 1.29(0.03)$, $\ln(\hat{Y}_0) = -8.91(0.39)$
- Creatives: $\hat{\beta} = 1.11(0.01)$, $\ln(\hat{Y}_0) = -3.23(0.14)$
- Inventors: $\hat{\beta} = 1.47(0.06)$, $\ln(\hat{Y}_0) = -14.09(0.74)$
- Burglary: $\hat{\beta} = 1.01(0.02)$, $\ln(\hat{Y}_0) = -5.16(0.24)$
- Robbery: $\hat{\beta} = 1.35(0.03)$, $\ln(\hat{Y}_0) = -11.65(0.34)$
- High school: $\hat{\beta} = 1(0)$, $\ln(\hat{Y}_0) = -0.15(0.04)$
- Graduate: $\hat{\beta} = 1.11(0.02)$, $\ln(\hat{Y}_0) = -3.82(0.21)$
- Chlamydia: $\hat{\beta} = 1.06(0.02)$, $\ln(\hat{Y}_0) = -6.37(0.29)$
- Syphilis: $\hat{\beta} = 1.46(0.05)$, $\ln(\hat{Y}_0) = -16.91(0.65)$

Four facts (among others) to try explaining

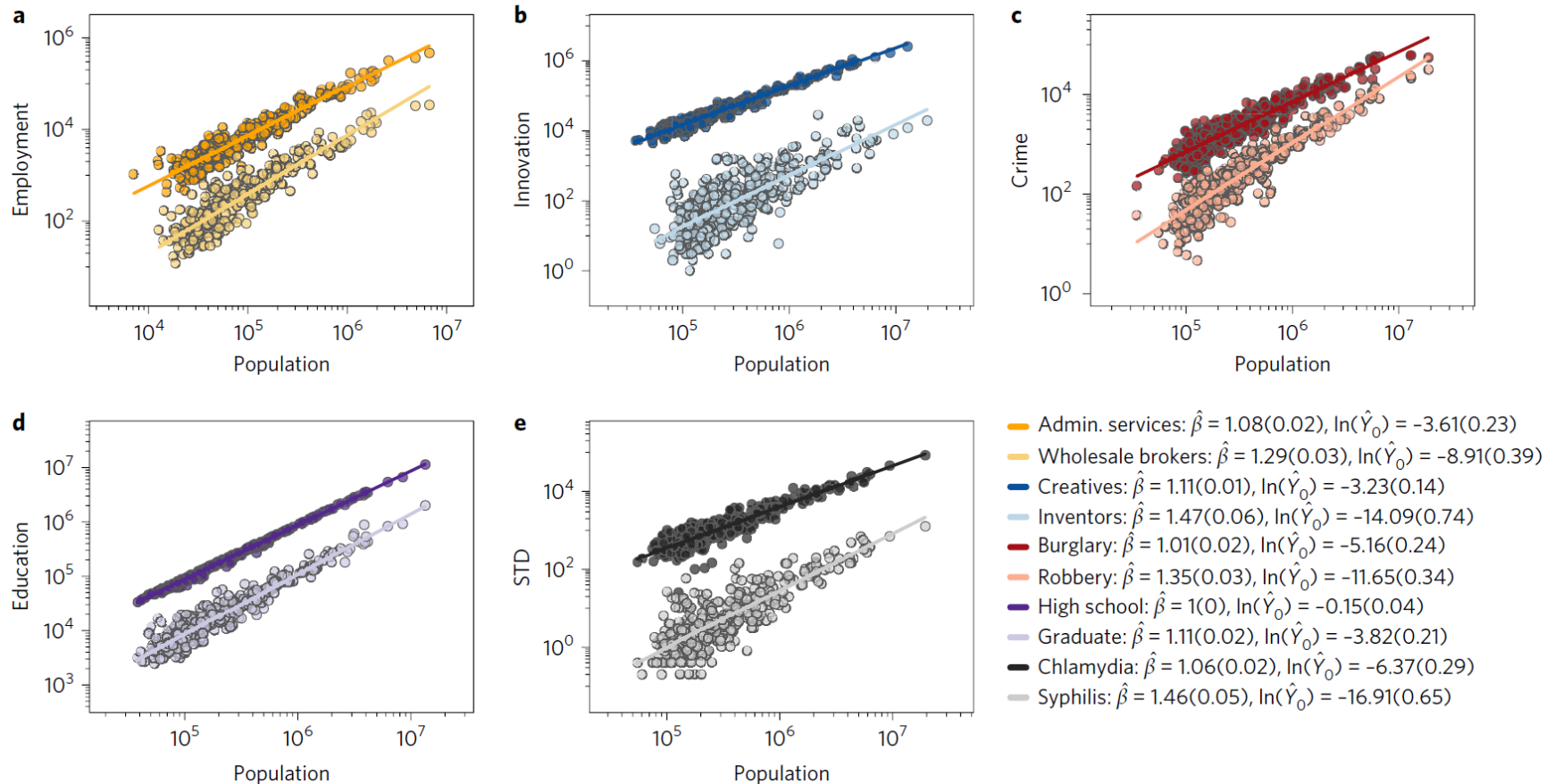
1. log-log \rightarrow straight line
2. each phenomenon with its own slope,



- Admin. services: $\hat{\beta} = 1.08(0.02)$, $\ln(\hat{Y}_0) = -3.61(0.23)$
- Wholesale brokers: $\hat{\beta} = 1.29(0.03)$, $\ln(\hat{Y}_0) = -8.91(0.39)$
- Creatives: $\hat{\beta} = 1.11(0.01)$, $\ln(\hat{Y}_0) = -3.23(0.14)$
- Inventors: $\hat{\beta} = 1.47(0.06)$, $\ln(\hat{Y}_0) = -14.09(0.74)$
- Burglary: $\hat{\beta} = 1.01(0.02)$, $\ln(\hat{Y}_0) = -5.16(0.24)$
- Robbery: $\hat{\beta} = 1.35(0.03)$, $\ln(\hat{Y}_0) = -11.65(0.34)$
- High school: $\hat{\beta} = 1(0)$, $\ln(\hat{Y}_0) = -0.15(0.04)$
- Graduate: $\hat{\beta} = 1.11(0.02)$, $\ln(\hat{Y}_0) = -3.82(0.21)$
- Chlamydia: $\hat{\beta} = 1.06(0.02)$, $\ln(\hat{Y}_0) = -6.37(0.29)$
- Syphilis: $\hat{\beta} = 1.46(0.05)$, $\ln(\hat{Y}_0) = -16.91(0.65)$

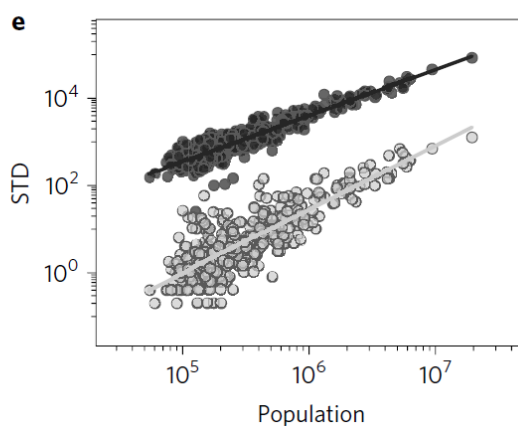
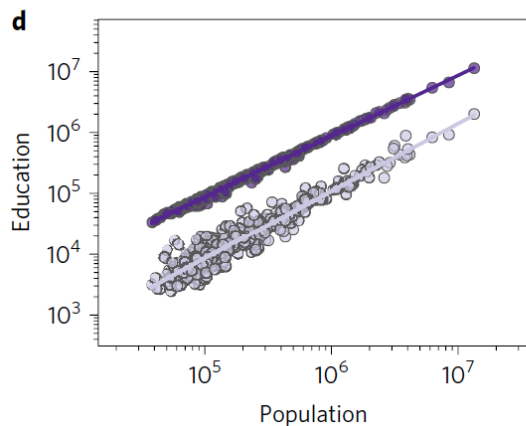
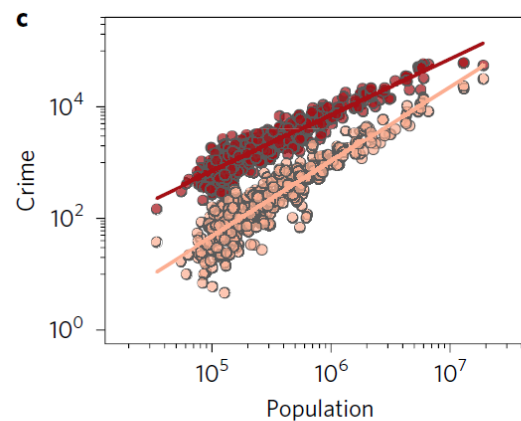
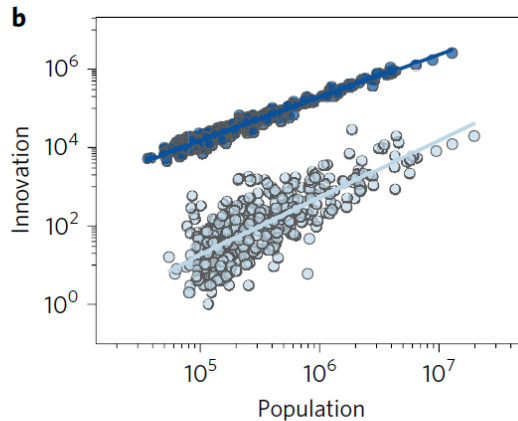
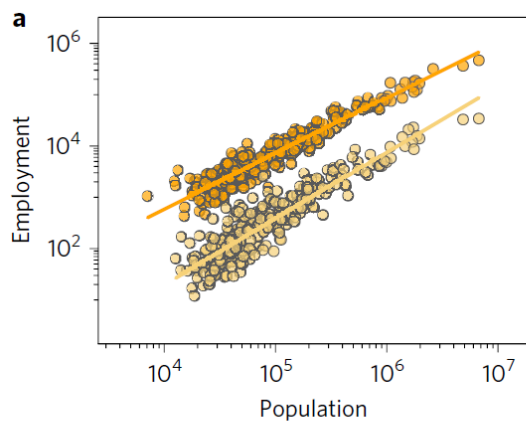
Four facts (among others) to try explaining

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



- Admin. services: $\hat{\beta} = 1.08(0.02)$, $\ln(\hat{Y}_0) = -3.61(0.23)$
- Wholesale brokers: $\hat{\beta} = 1.29(0.03)$, $\ln(\hat{Y}_0) = -8.91(0.39)$
- Creatives: $\hat{\beta} = 1.11(0.01)$, $\ln(\hat{Y}_0) = -3.23(0.14)$
- Inventors: $\hat{\beta} = 1.47(0.06)$, $\ln(\hat{Y}_0) = -14.09(0.74)$
- Burglary: $\hat{\beta} = 1.01(0.02)$, $\ln(\hat{Y}_0) = -5.16(0.24)$
- Robbery: $\hat{\beta} = 1.35(0.03)$, $\ln(\hat{Y}_0) = -11.65(0.34)$
- High school: $\hat{\beta} = 1(0)$, $\ln(\hat{Y}_0) = -0.15(0.04)$
- Graduate: $\hat{\beta} = 1.11(0.02)$, $\ln(\hat{Y}_0) = -3.82(0.21)$
- Chlamydia: $\hat{\beta} = 1.06(0.02)$, $\ln(\hat{Y}_0) = -6.37(0.29)$
- Syphilis: $\hat{\beta} = 1.46(0.05)$, $\ln(\hat{Y}_0) = -16.91(0.65)$

The theory:

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

The theory:

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

- I. Most of urban phenomena are the conjunction of complementary factors.

Being happy:

Enjoy the power and beauty of your youth
E **Floss** Don't worry about the future **YOU'RE NOT AS FAT AS YOU IMAGINE**
A **DO ONE THING EVERY DAY THAT SCARES YOU** **GET TO KNOW YOUR PARENTS.**
J Don't be reckless with other people's hearts, and
O don't put up with people who are reckless with yours. S
Y Don't waste Don't feel guilty if you don't know what to do with your life **STRETCH**
O your time **WEAR SUNSCREEN** **DANCE** N
U on jealousy be kind to your knees Sometimes your ahead, C
F **remember** Understand that sometimes your behind
C **compliments.** friends come and go.
B **forget the insults.** Love travel MAYBE
O Don't congratulate yourself too much or berate yourself either YOU'LL
D **READ THE DIRECTIONS, EVEN IF** Don't expect anyone MARY,
Y **YOU WON'T FOLLOW THEM** die to support you. MAYBE
be nice to your siblings **RESPECT YOUR ELDERS** YOU
Don't mess too much with your hair Accept certain inalienable truths **WON'T.**

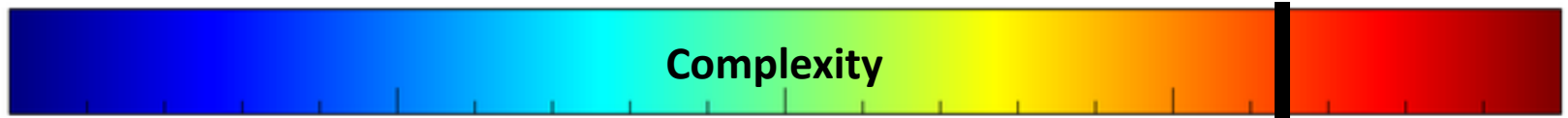
The theory:

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

- I. Most of urban phenomena are the conjunction of complementary factors.

Being happy:

Enjoy the power and beauty of your youth
Don't worry about the future **Floss** **YOU'RE NOT AS FAT AS YOU IMAGINE**
DO ONE THING EVERY DAY THAT SCARES YOU
Don't be reckless with other people's hearts, and don't put up with people who are reckless with yours.
Don't waste your time Don't feel guilty if you don't know what to do with your life **STRETCH**
Don't be jealous **WEAR SUNSCREEN** **DANCE**
Remember compliments. Understand that sometimes your behind friends come and go.
Forget the insults. Love travel MAYBE YOU'LL MARY, MAYBE YOU WON'T.
Don't congratulate yourself too much or berate yourself either
Read the directions, even if you won't follow them Don't expect anyone else to support you.
Be nice to your siblings **RESPECT YOUR ELDER**
Don't mess too much with your hair Accept certain inalienable truths



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

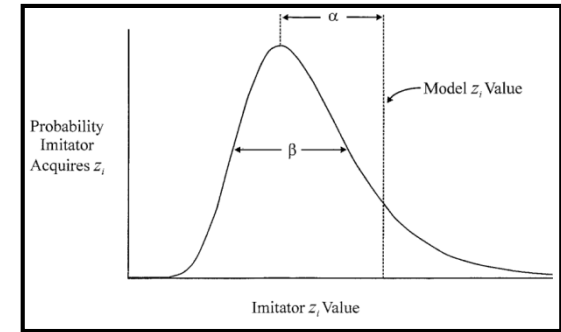
$$\Pr\{X_{i,c,f} = 1\} = e^{-M_f q_i (1-r_c)}$$

Models of economic complexity



&

Models of cultural evolution

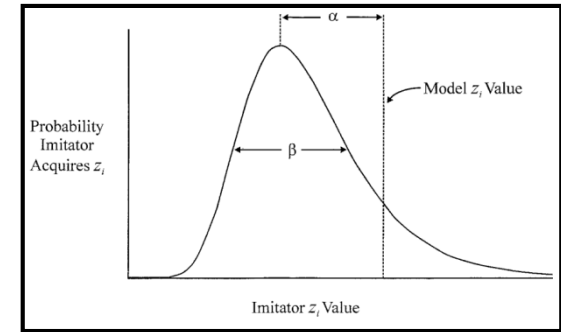


Models of economic complexity



&

Models of cultural evolution



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

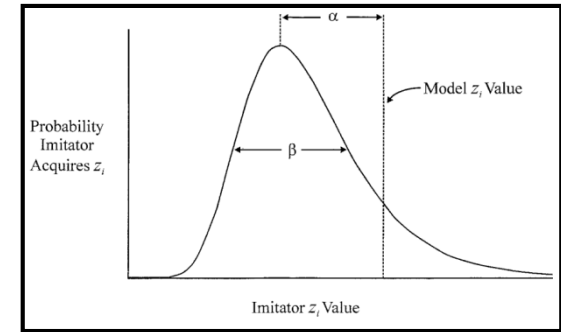
$$r(N) = a + b \ln(N)$$

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

$$= Y_0 N^\beta$$

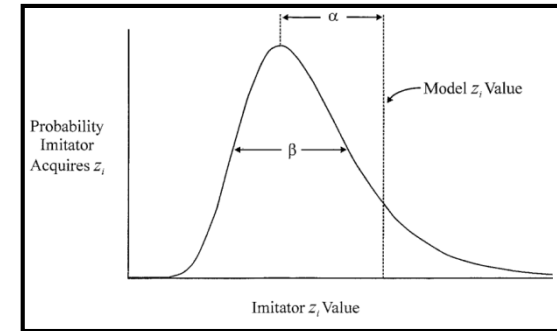


Models of economic complexity



&

Models of cultural evolution

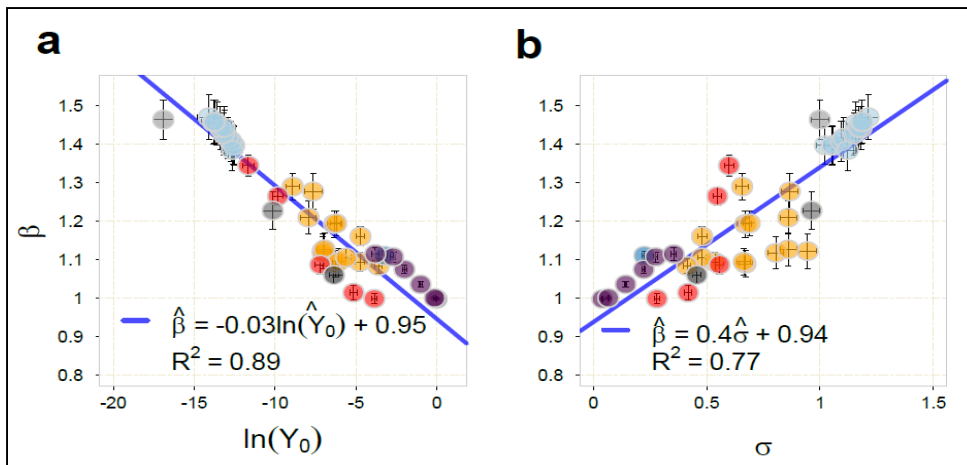


$$\mathbb{E}[Y] = e^{-Mq} N e^{Mq} r(N)$$

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

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

$$= Y_0 N^\beta$$

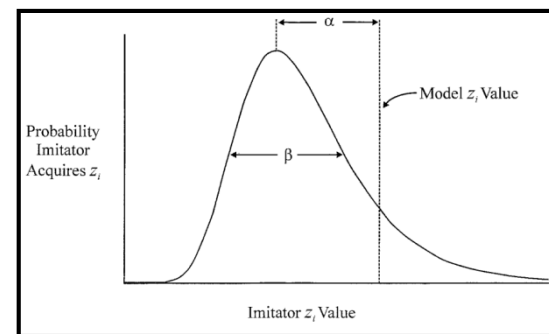


Models of economic complexity



&

Models of cultural evolution



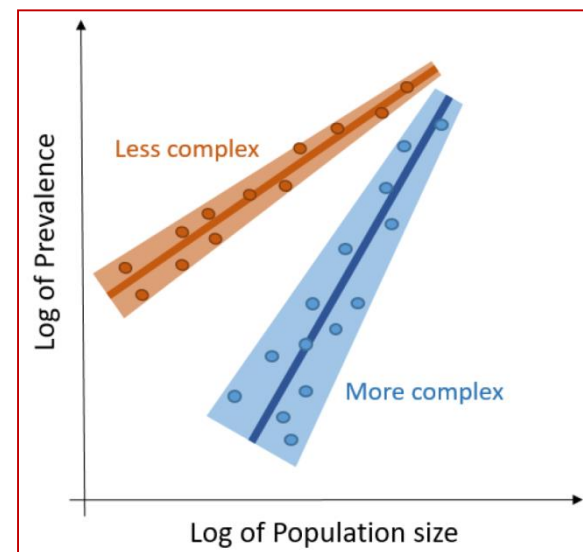
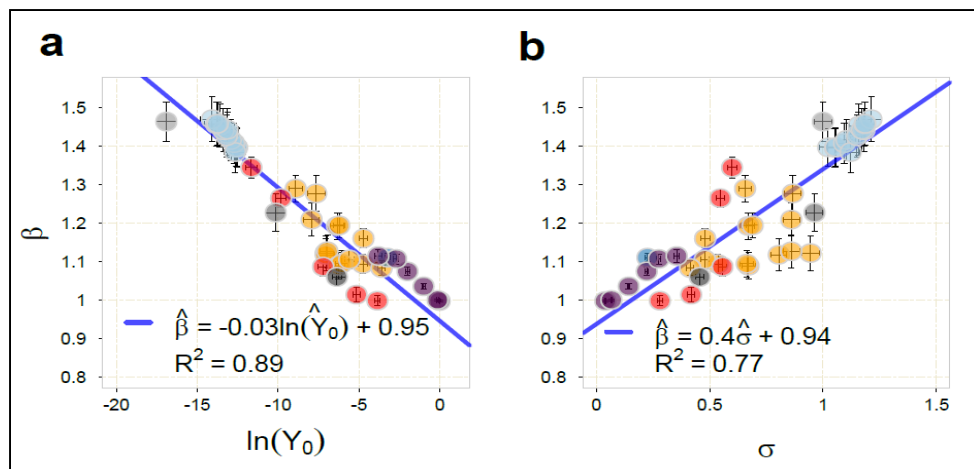
$$\mathbb{E}[Y] = e^{-Mq} N e^{Mq} r(N)$$

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

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



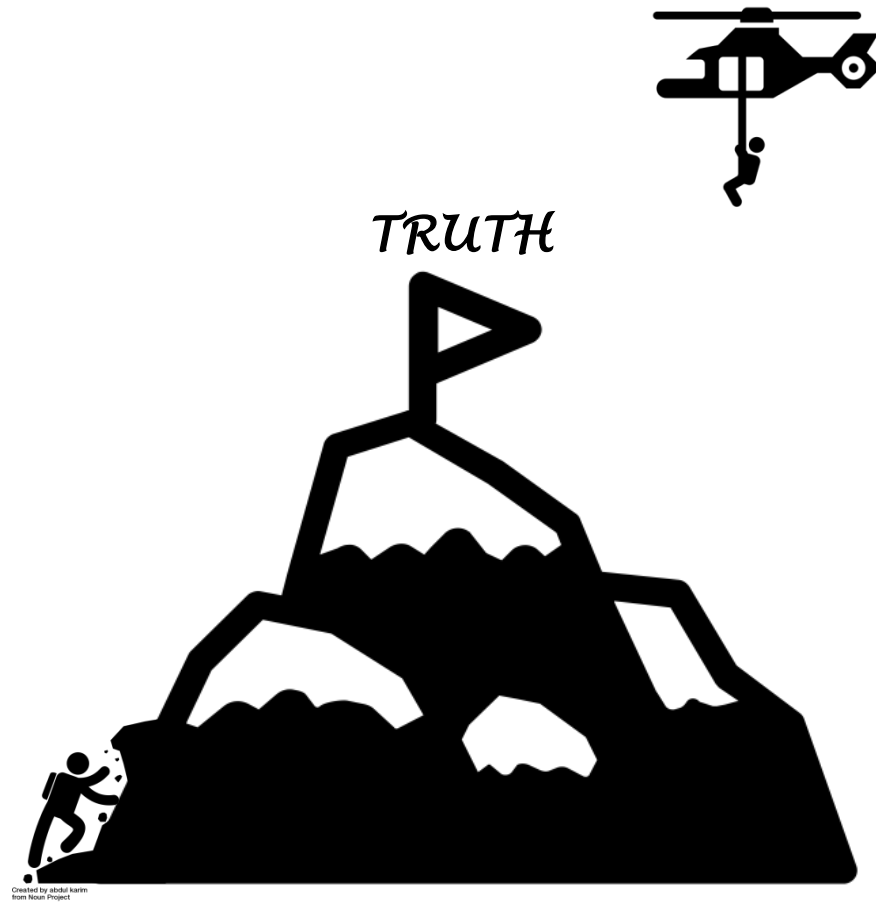
$$= Y_0 N^\beta$$



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



Created by Abdul Karim
from Noun Project

The importance of emergent phenomena



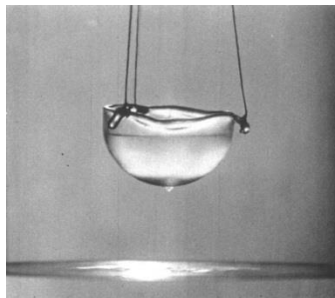
To understand a systemic phenomenon, consider

how many parts and
how they *interact*,

than

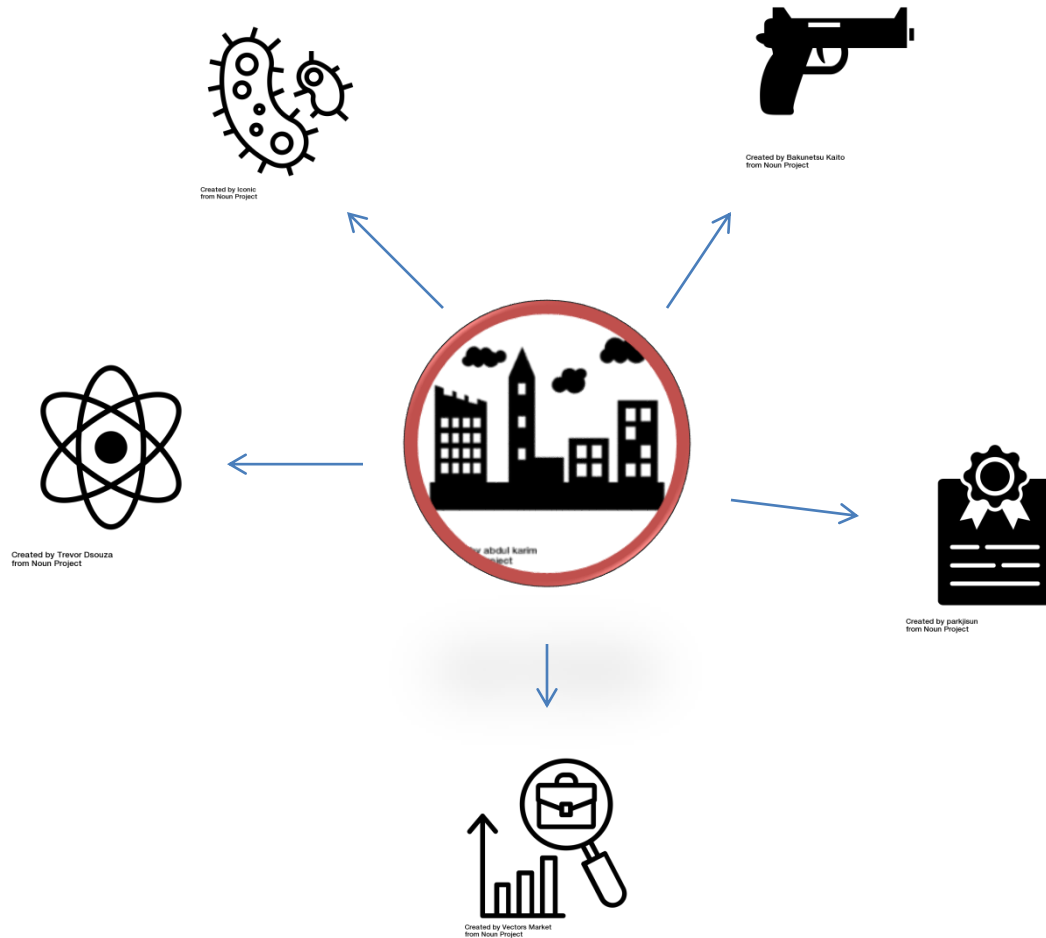
what are the parts.

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

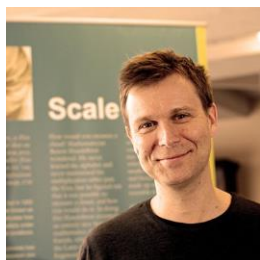


Urban “outcomes” as emergent phenomena

(→ 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);