We got the skills to pay the bills: exploring the link between occupation diversity and innovation

Andrés Gomez-Lievano, Charlie Brummitt, Gareth Haslam, and Nic Goudemand
1. Cities as the cauldron of innovation

2. Go deep or go broad? How does specialization or diversity affect innovation in US cities

3. Model(l)ing the growth of occupation diversity
1. Cities as the cauldron of innovation

**People**

Diversity + Large Population

**Places**

Increased interactions

**Products**

Innovation

Wealth and Growth

Jacobs (1961 and 1984)
Hypothesis:

The above-average patent production is due to:

the greater diversity of (creative) occupations

* Creative jobs are defined according to the US Economic Research Service to be those involving developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions. Based on Florida (2002)
2. Go deep or go broad?

Some definitions:

• **City - Metropolitan Statistical Area (MSA)**
  A geographical region with a relatively high population density, as defined by US Office of Management and Budget.

• **Dataset**
Occupation distributions in USA cities

fraction of employees

smallest city

largest city

ranks of the occupations
Occupation distributions in USA cities

fraction of employees

- cashiers, sales
- smallest city
- shampooers
- largest city

ranks of the occupations
2. Go deep or go broad?

Diversity measures:

- Shannon entropy:
  \[ H(X) = -\sum_{i=1}^{m} p_i \log p_i \]

  How many questions you need to ask to determine someone’s occupation (on average)

- Herfindahl Index (HHI):
  \[ H = \sum_{i=1}^{m} p_i^2 \]

  A measure of specialization.
  \( 1/H = \text{diversity} \)
OMG! Power! Law! (kindof)

Figure 2. Patents scale superlinearly with population size. The scaling exponent $\hat{\beta} \approx 1.40$ is close to the one predicted $\beta_p \approx 1.53$. 

$H = \sum_i s_i^2$
How entropic is your city?
Universality class of formless scatter plots

\[ \frac{\text{inventors}}{\text{population}} \]

\[ \log_2 (\text{occupations}) \]

\[ 0.001 \quad 0.002 \quad 0.005 \quad 0.010 \quad 0.020 \]
Universality class of formless scatter plots

- San Jose
- Corvallis
- NYC (college towns)
- farm towns

Graph: inventors vs. population

- X-axis: entropy (log₂ (occupations))
- Y-axis: inventors / population

- Data points clustered in different regions:
  - San Jose
  - Corvallis
  - NYC (college towns)
  - farm towns

- Universality class of formless scatter plots
2. Go deep or go broad?

\[ q_{ij} = \begin{cases} 
1, & \text{if } Q_i = Q_j \\
< 1, & \text{if } Q_i < Q_j \\
> 1, & \text{if } Q_i > Q_j.
\end{cases} \]

\[ Y \propto \sum_{ij} q_{ij}. \quad Y \propto N^2 \left\langle \frac{1}{Q} \right\rangle. \quad Y \sim N^{1.53}. \]

(a) Population dependence of no. of occupations, \( m \), and
(b) average no. of people in each occupation, \( Q \).
Does **diversity** explain **over-**, **under-performance**?
Nope.
Pretty plots
3. Modelling Occupation Growth

\[ y = 0.0934x + 1.4495 \]

\[ R^2 = 0.9971 \]
4. Conclusions

1. Population largely determines patents
2. Occupation diversity doesn’t explain above-average innovation
3. Restricting to creative types doesn’t help
4. Model: Viability of occupation doesn’t lead to power law distribution of occupations
5. Future: data on skills and abilities, such as night vision and hand-eye coordination
Acknowledgements

• Luís Bettencourt
• Geoffrey West
• Hyejin Youn
• Dan Wu
• Jianfeng Xu
• Power laws