Slides at:
dakotamurray.me/talk/2021-copenhagen/

Embedding mobility

Dakota Murray
@dakotasmurray
dakota.s.murray@gmail.com

SCHOOL OF
INFORMATICS, COMPUTING, AND ENGINEERING

Mobility drives economies, cultural exchange, epidemics

Hanson, R., Mouton, C. A., Grissom, A. R., & Godges, J. P. (2020). *COVID-19 Air Traffic Visualization: Decisionmakers Should Base Travel Restrictions on Infection Rates Per Capita and Air Traffic Levels*.

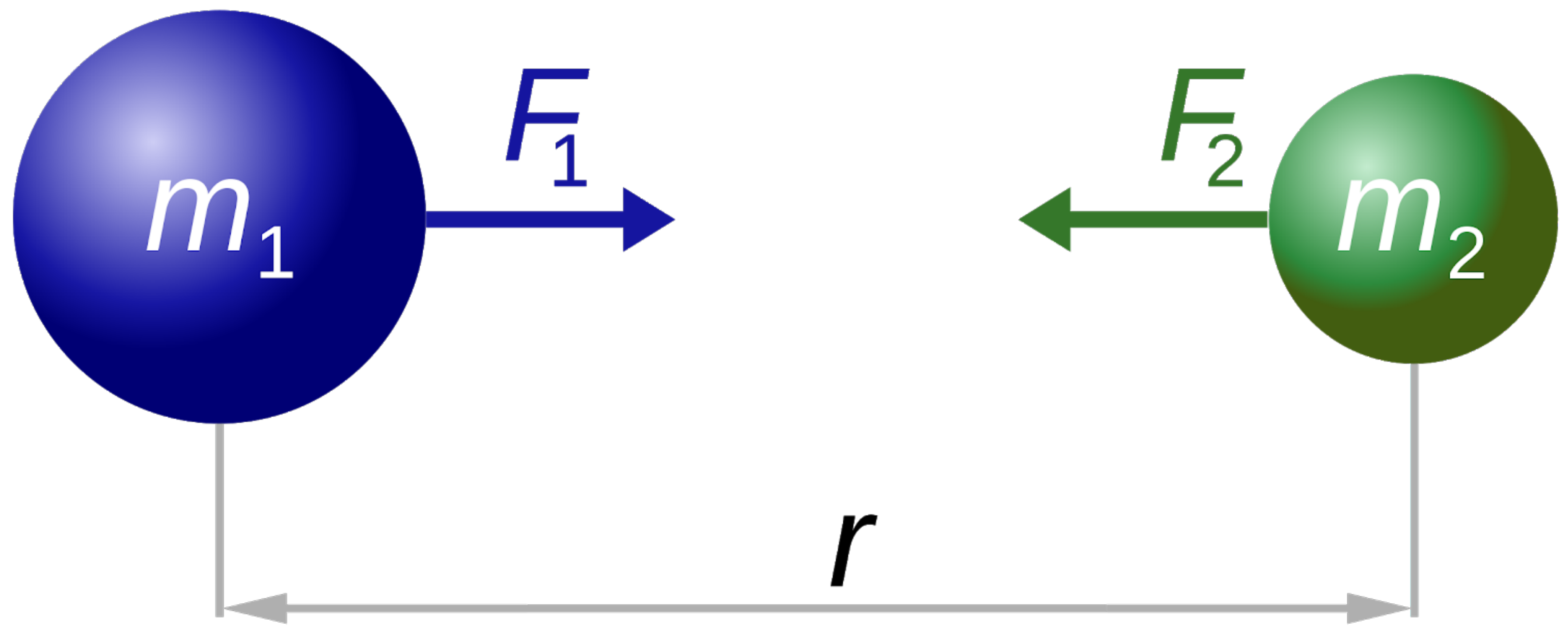
Understanding mobility is essential to understanding a range of human phenomenon

Understanding mobility is essential to understanding a range of human phenomenon

How do we study it?

Gravity model

- Powerful, intuitive, and ubiquitous
- Models mobility akin to newton's law of gravity
- Function of sizes and distance



$$F_1 = F_2 = G\frac{m_1 \times m_2}{r^2}$$

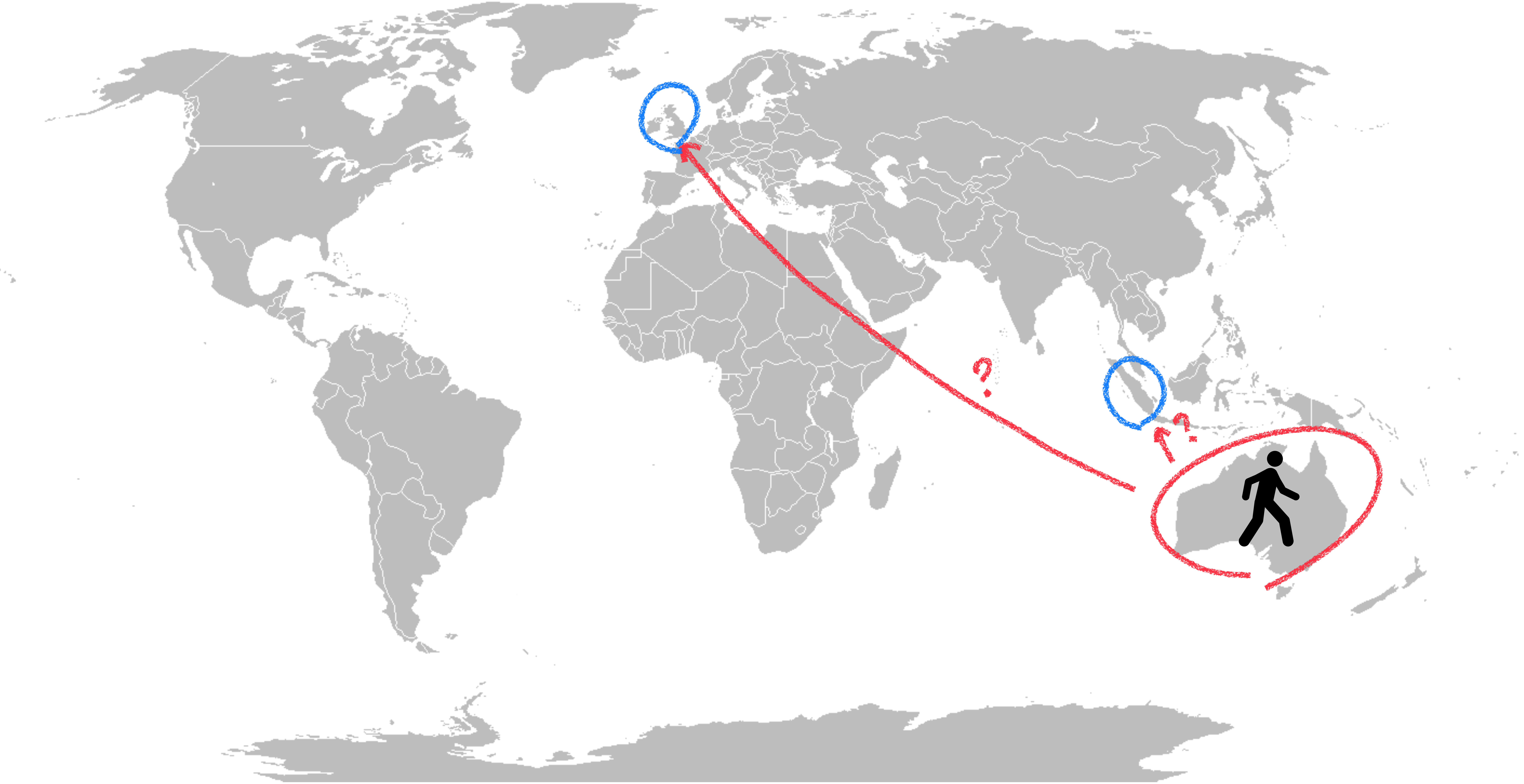
Wikipedia user Dennis Nilsson

$$\hat{T}_{ij} = C m_i m_j f(r_{ij})$$

Flux *Population* *Decay*

The gravity model is effective…

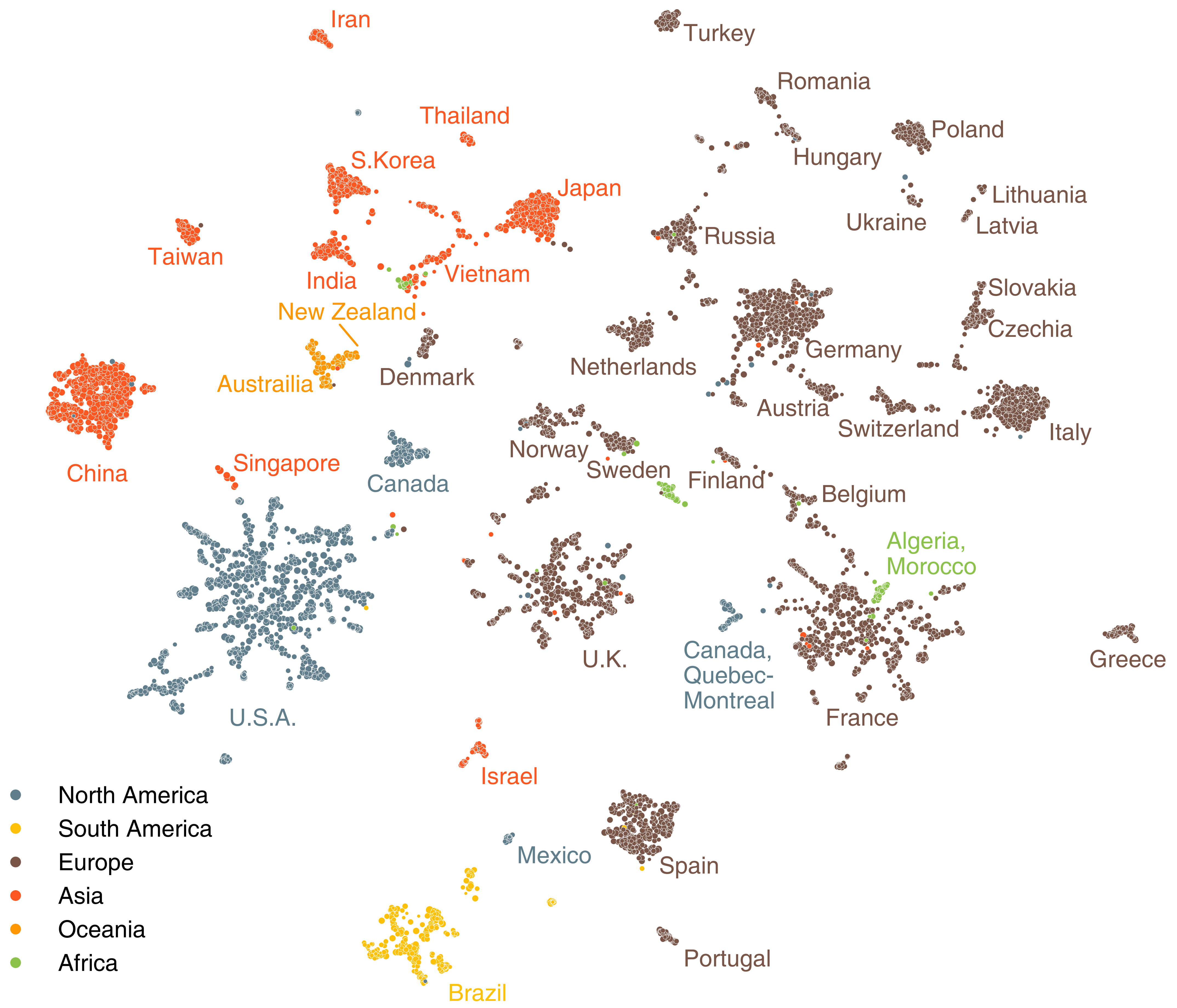
But geographic distance is not always appropriate



Geography matters, but…

Cultural, linguistic, economic, and political distance are also important!

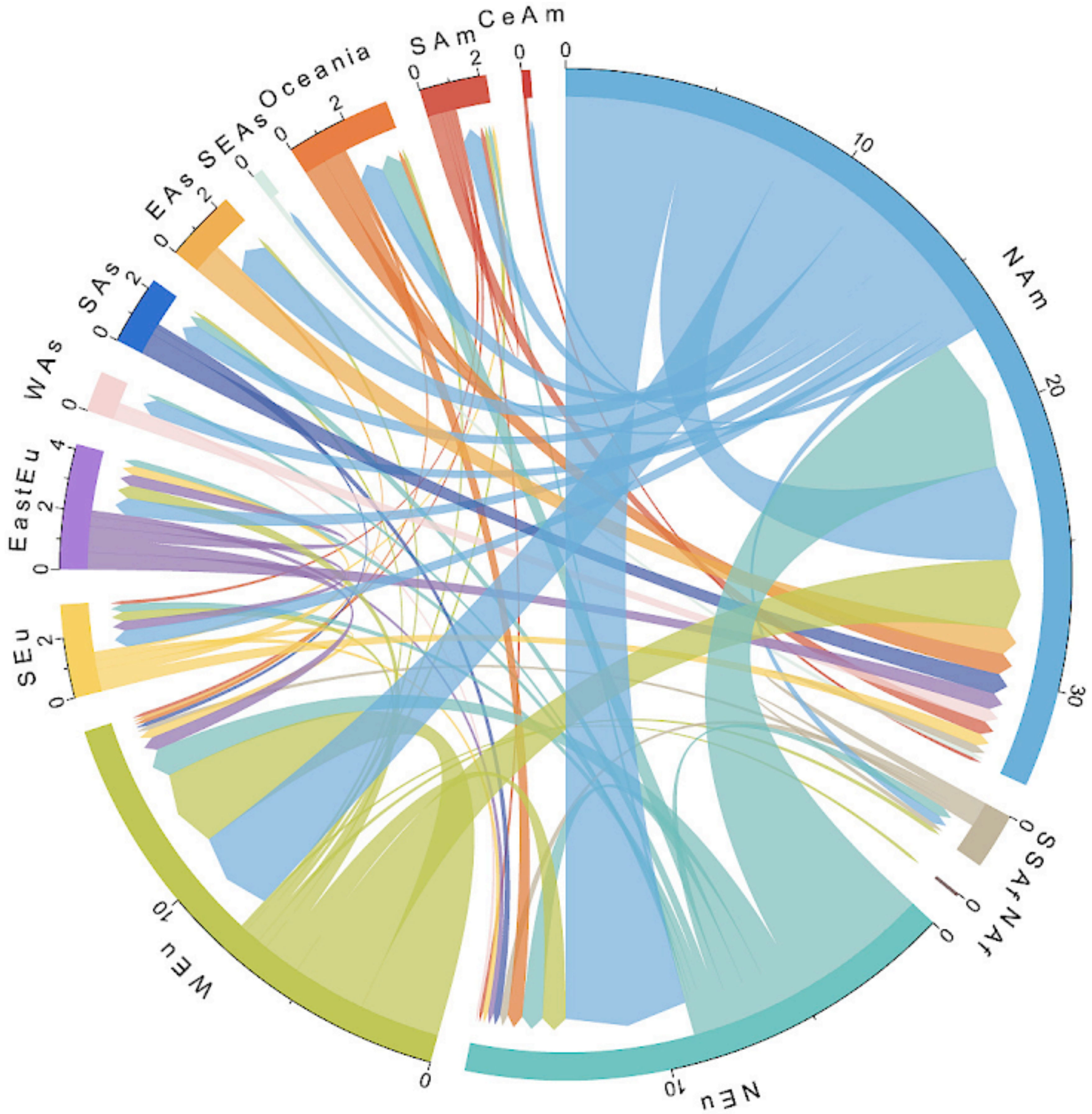
Can we instead learn an "embedding" that captures latent distance between places?

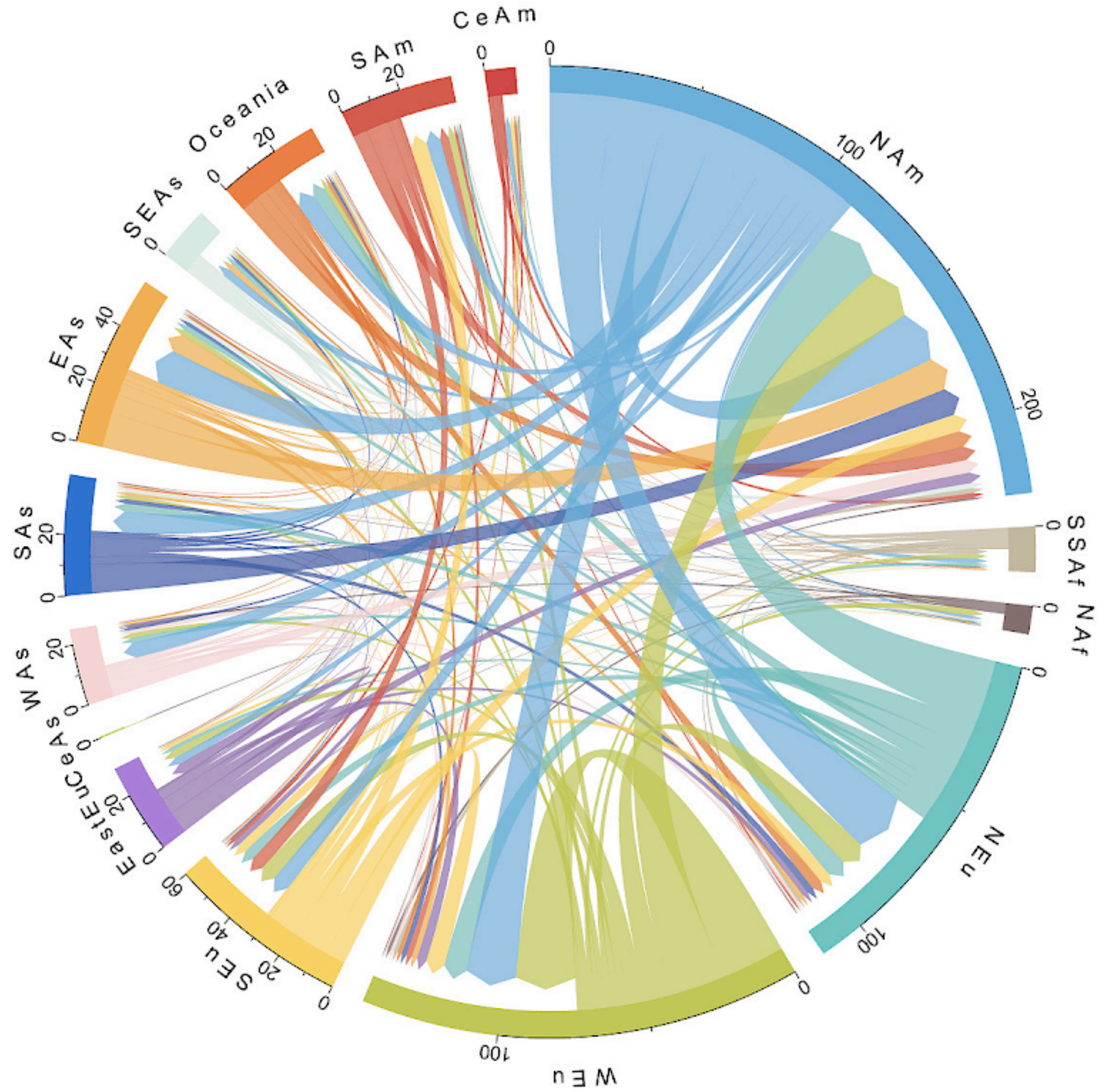


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The case of scientific mobility

Drives collaboration, innovation, and the diffusion of ideas





Czaika, M., & Orazbayev, S. (2018). The globalisation of scientific mobility, 1970–2014. *Applied Geography*, *96*, 1–10.

Neural embedding

"The quick brown fox …"

Word2vec
(or other word embedding methods)

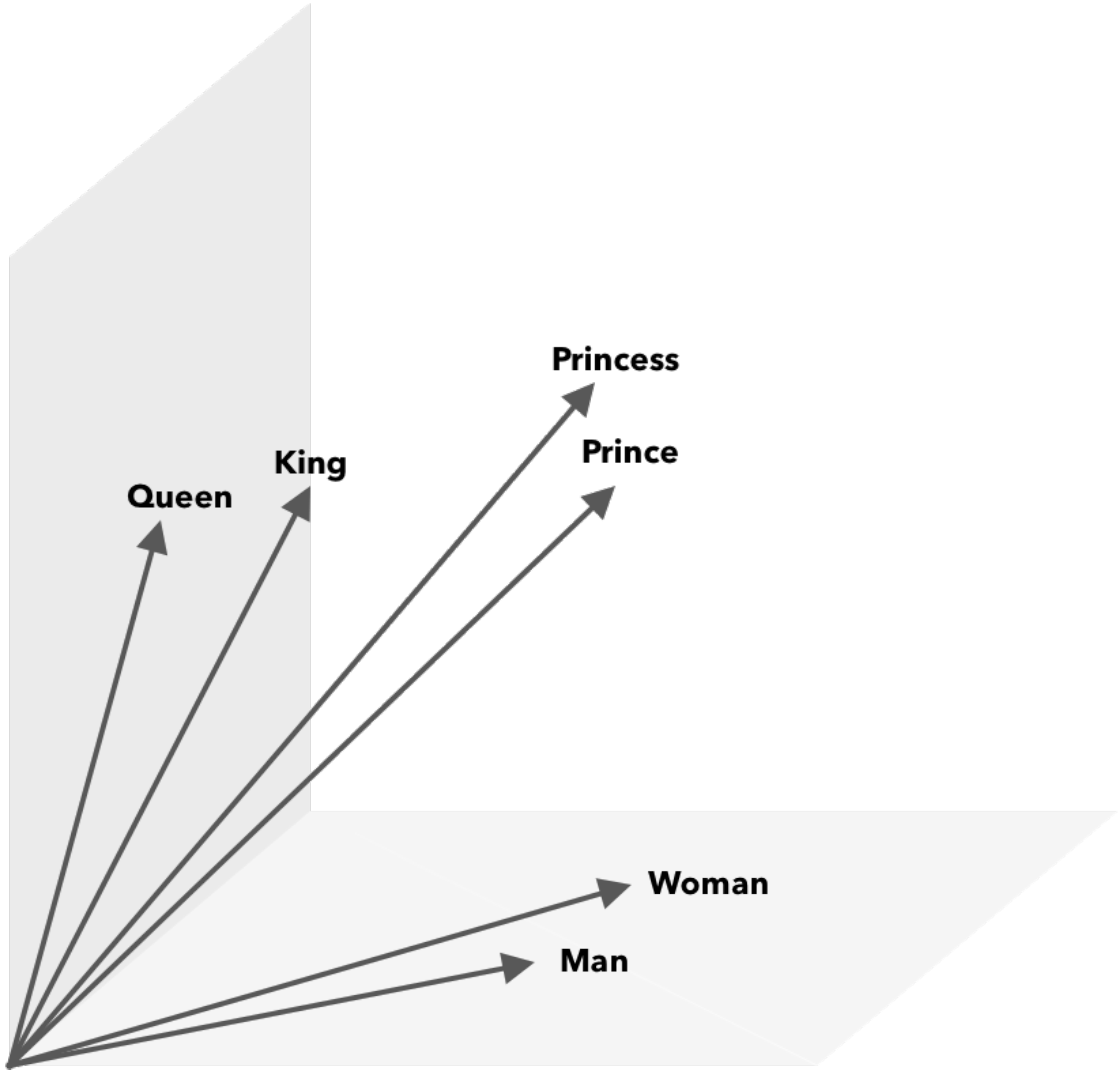
"quick": [0.51, 0.12, 0.69, …]
"brown": [0.11, 0.92, 0.29, …]

Dense vectors

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The geometry of the vector space encodes semantic relationships

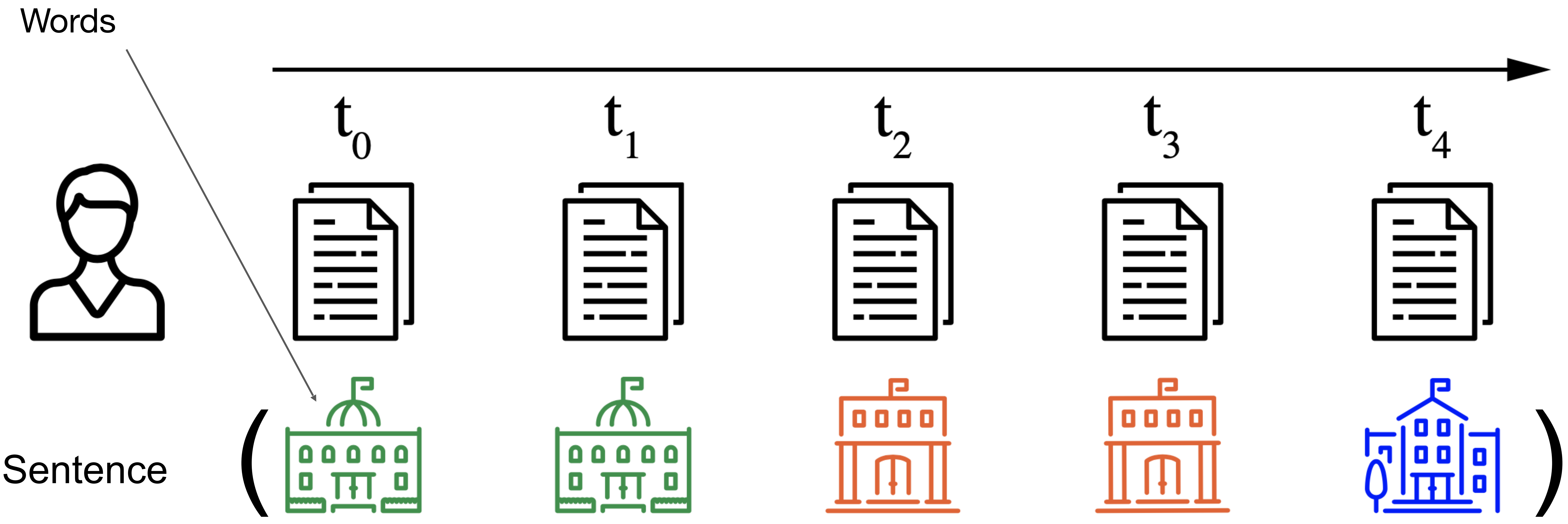
Using cosine distance



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

They don’t have to be “real” words or sentences

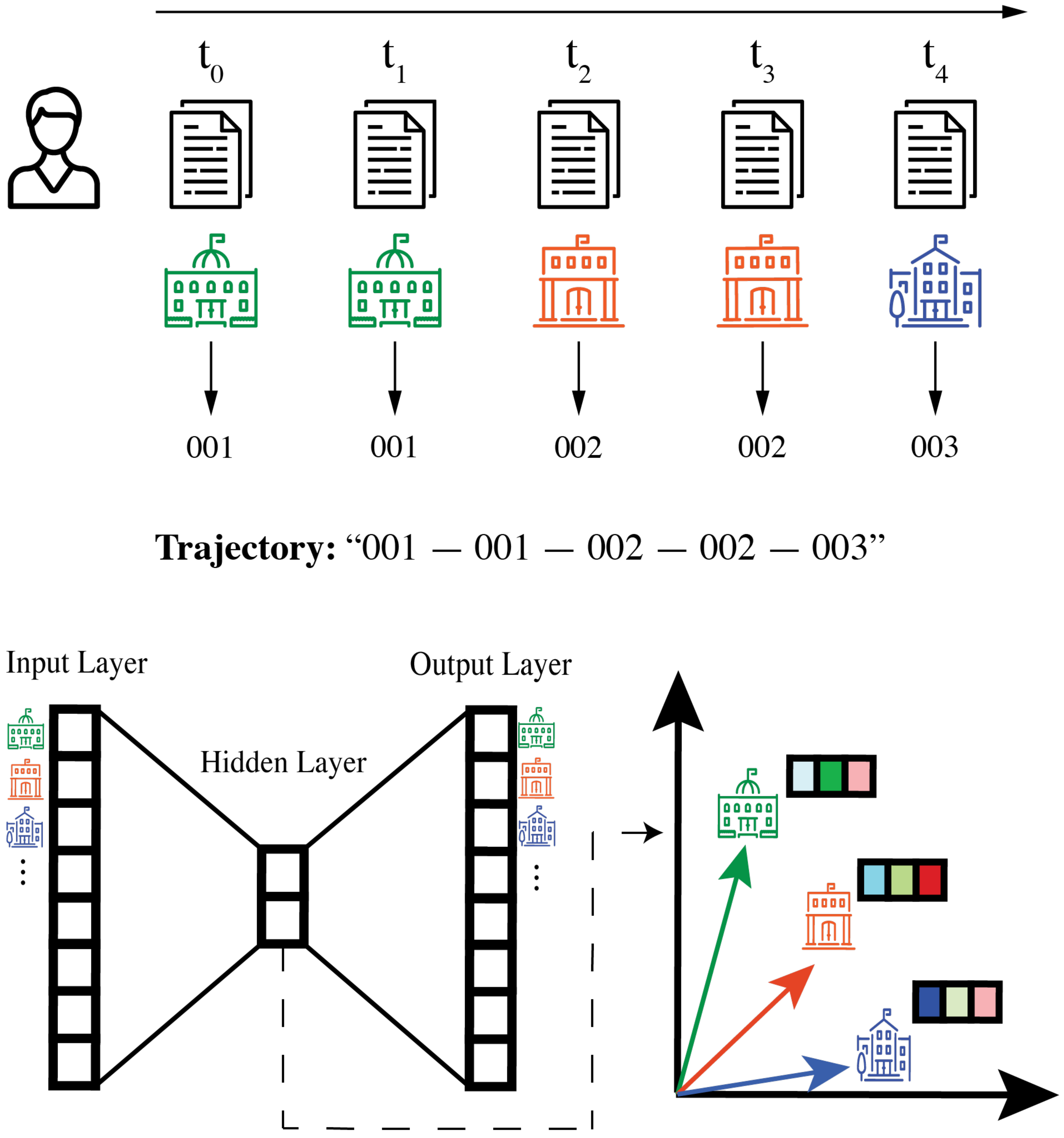
Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use **trajectories of scientists** as **sentences** and **organizations** as **words**.



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Modelling mobility

- Career trajectories of 3 million scientists derived from publications
- Give as input to *word2vec*
- Can measure embedding distance between any pair of organizations

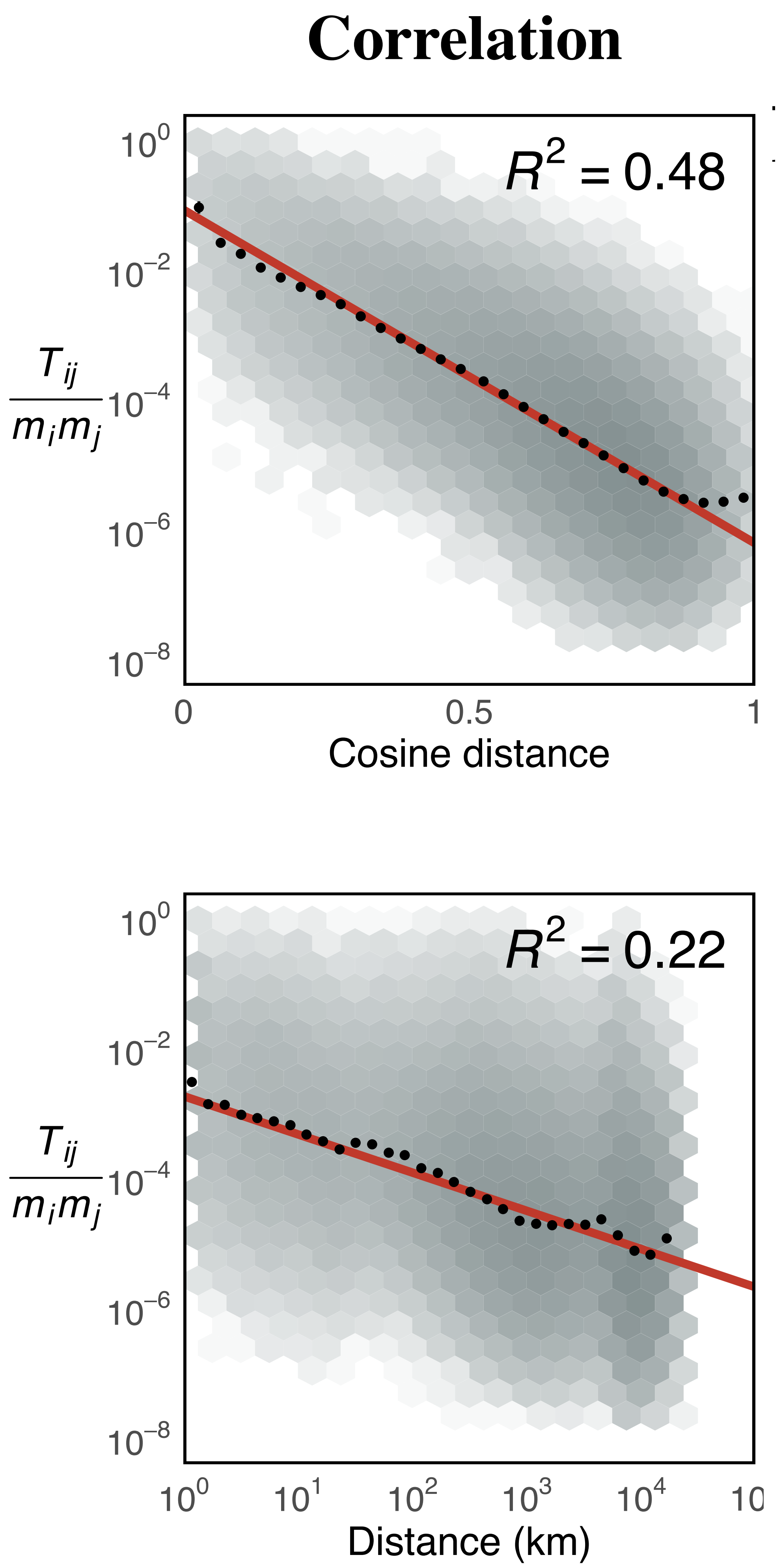


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Embedding Distance

Geographic Distance



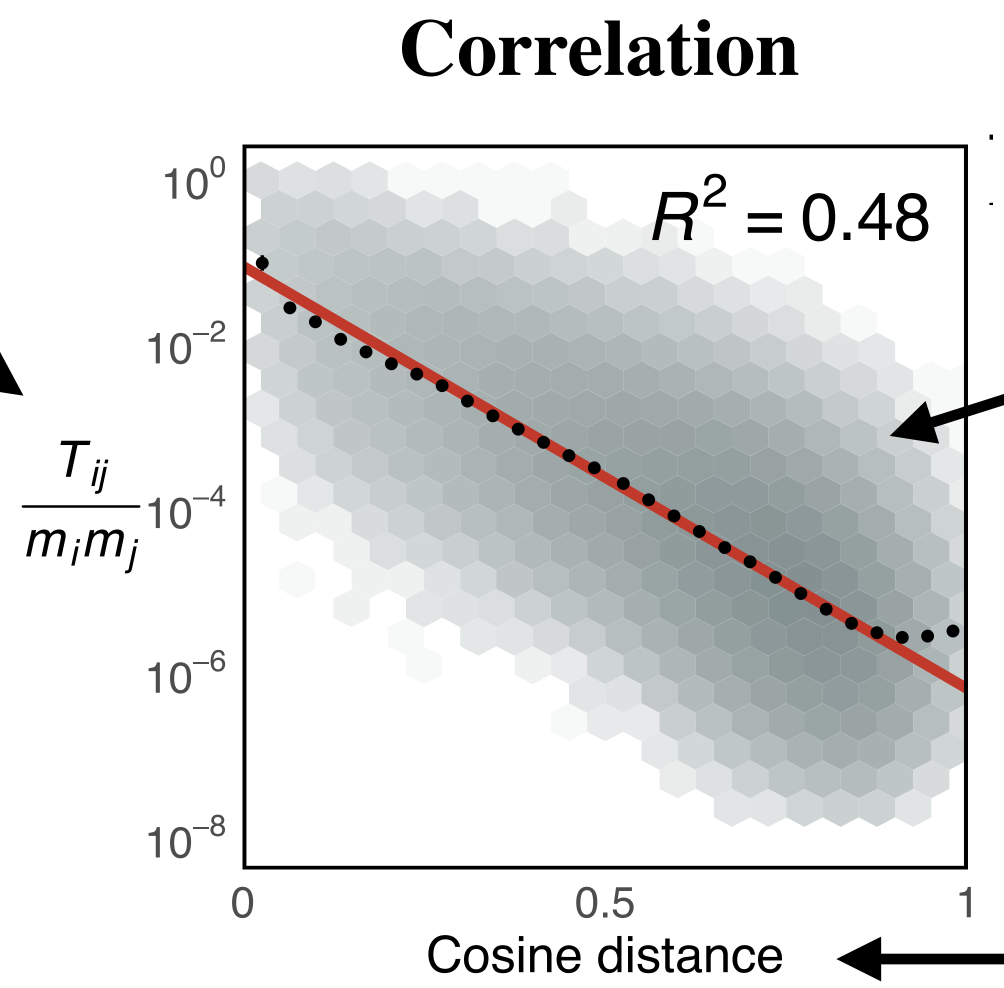
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Flux, given the organizations sizes

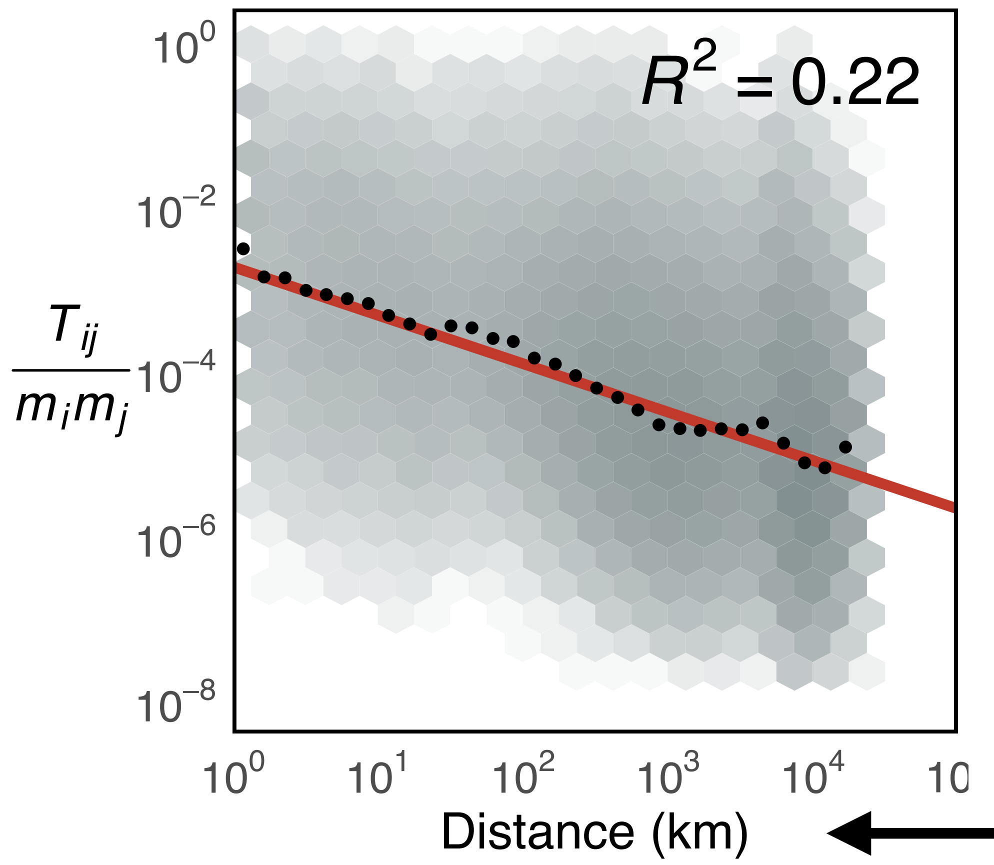
Every point is a pair of organizations (binned)

Embedding Distance



Cosine distance between organization vectors

Geographic Distance



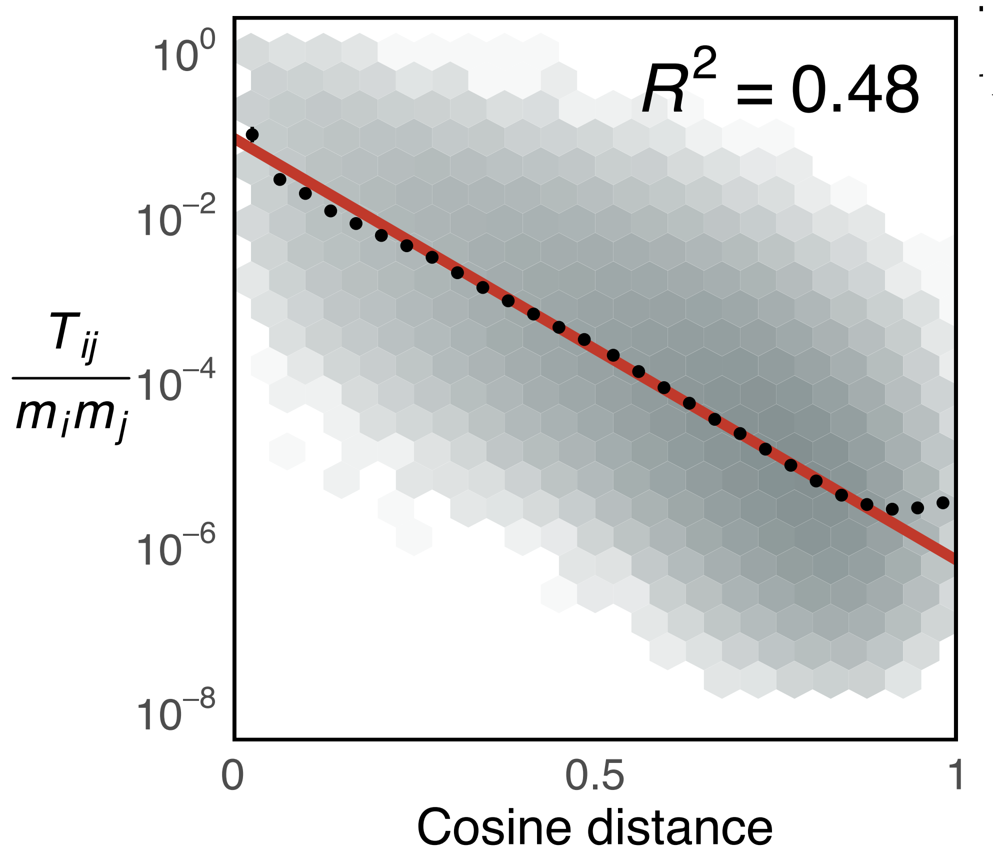
Geographic distance between organizations

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

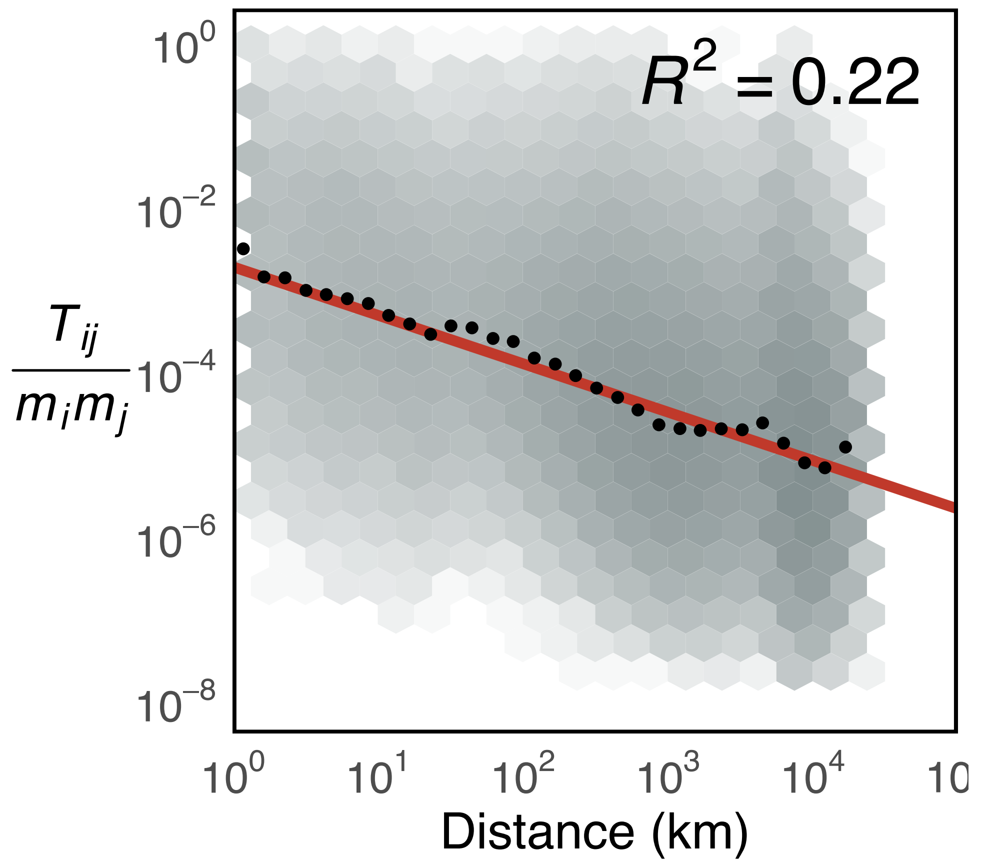
Correlation

Embedding Distance



Flux more strongly correlates with embedding distance than geographic distance

Geographic Distance

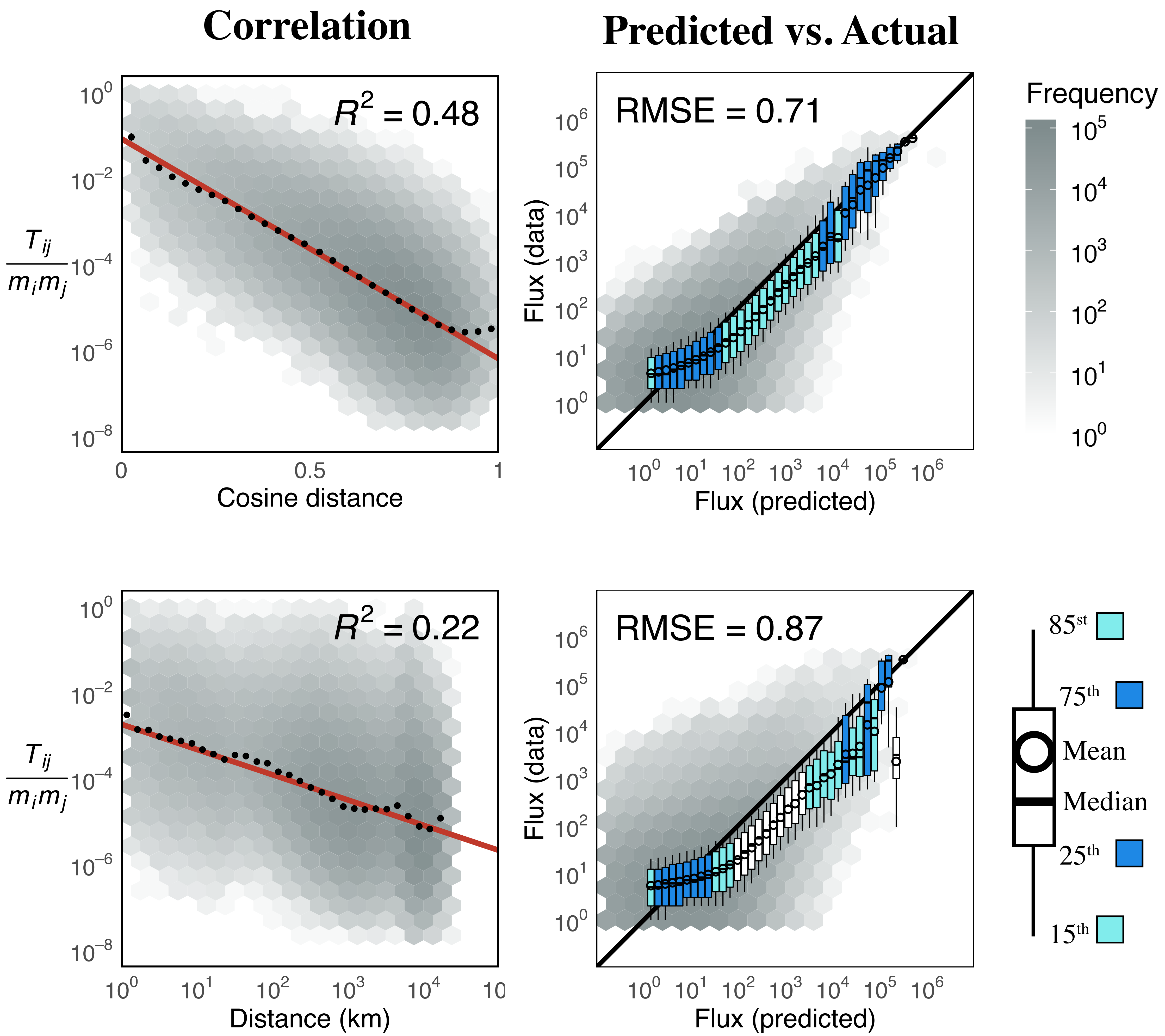


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Embedding Distance

Geographic Distance

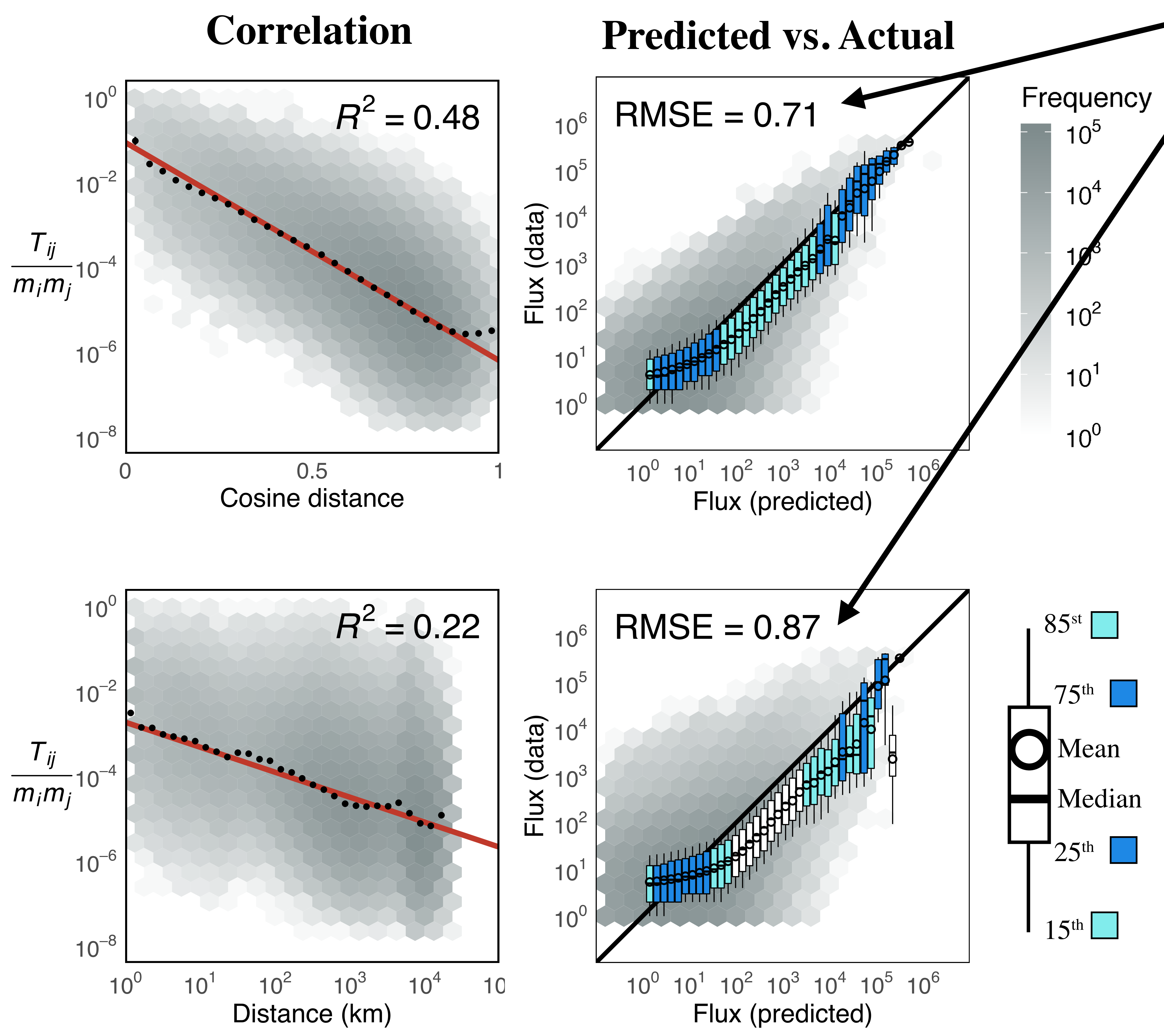


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Embedding Distance

Geographic Distance



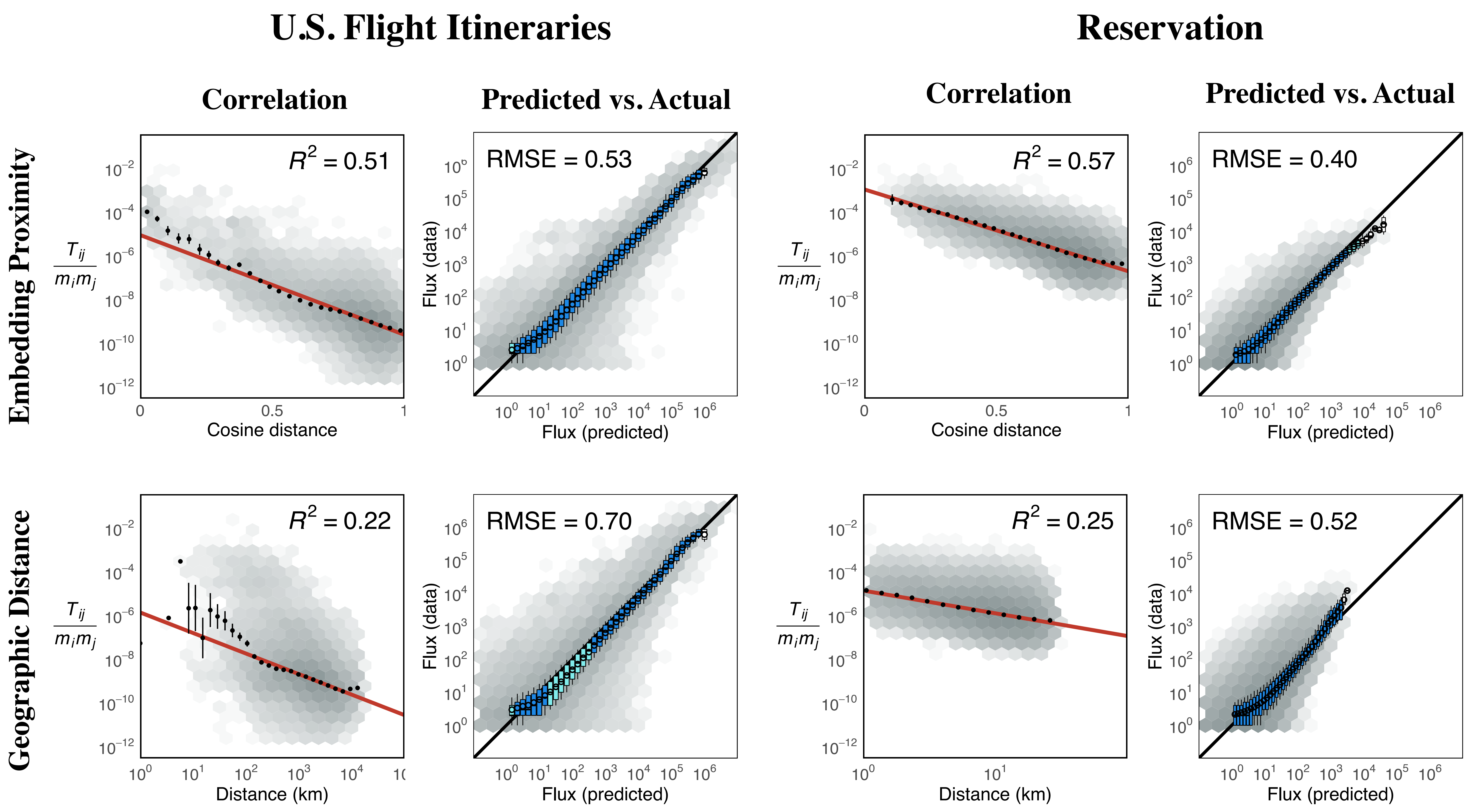
Root mean squared error of prediction

Embedding distance leads to better predictions

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The embedding performs well in other domains!

U.S. domestic flight itineraries and South Korean hotel reservation trajectories

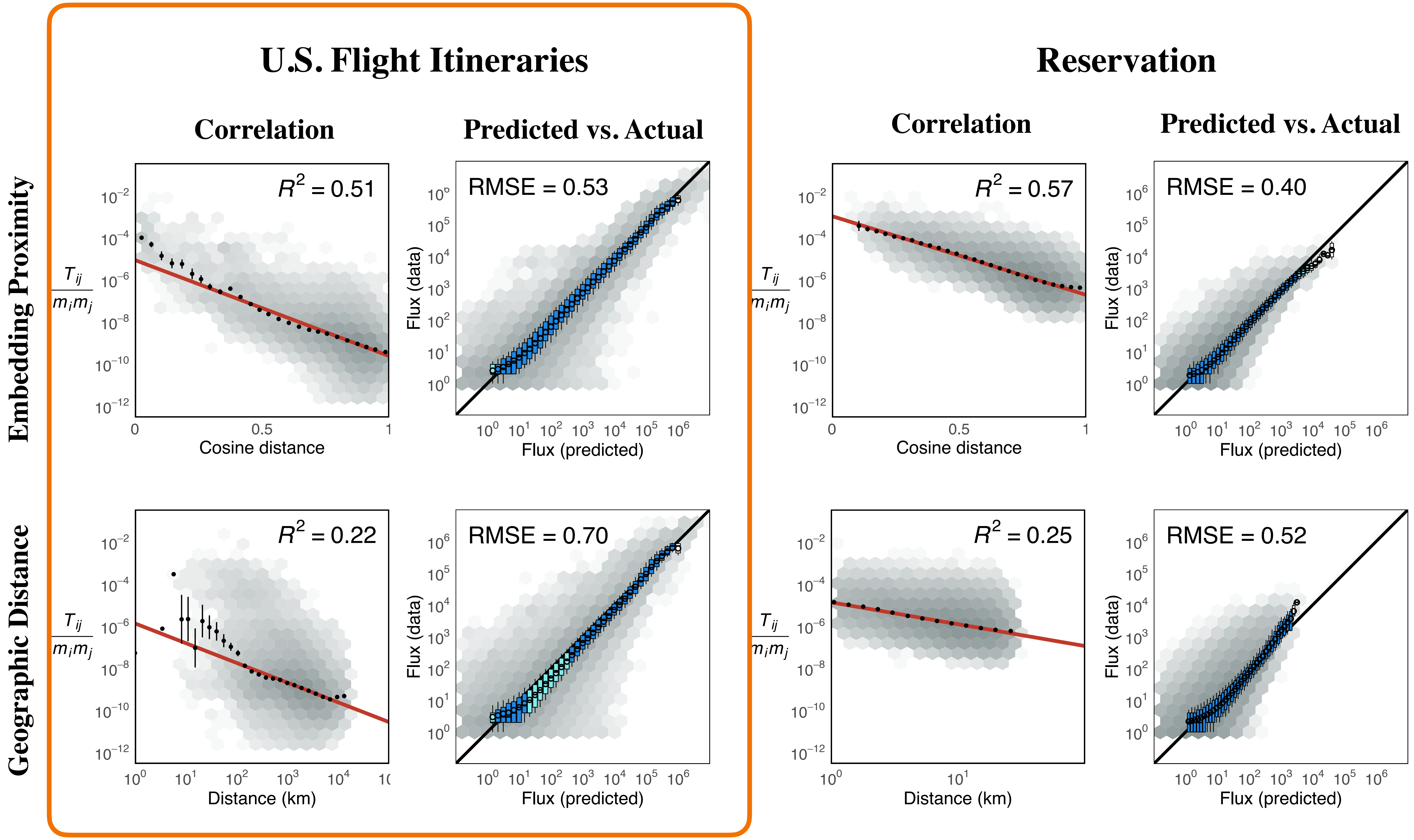


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The embedding performs well in other domains!

U.S. domestic flight itineraries and South Korean hotel reservation trajectories

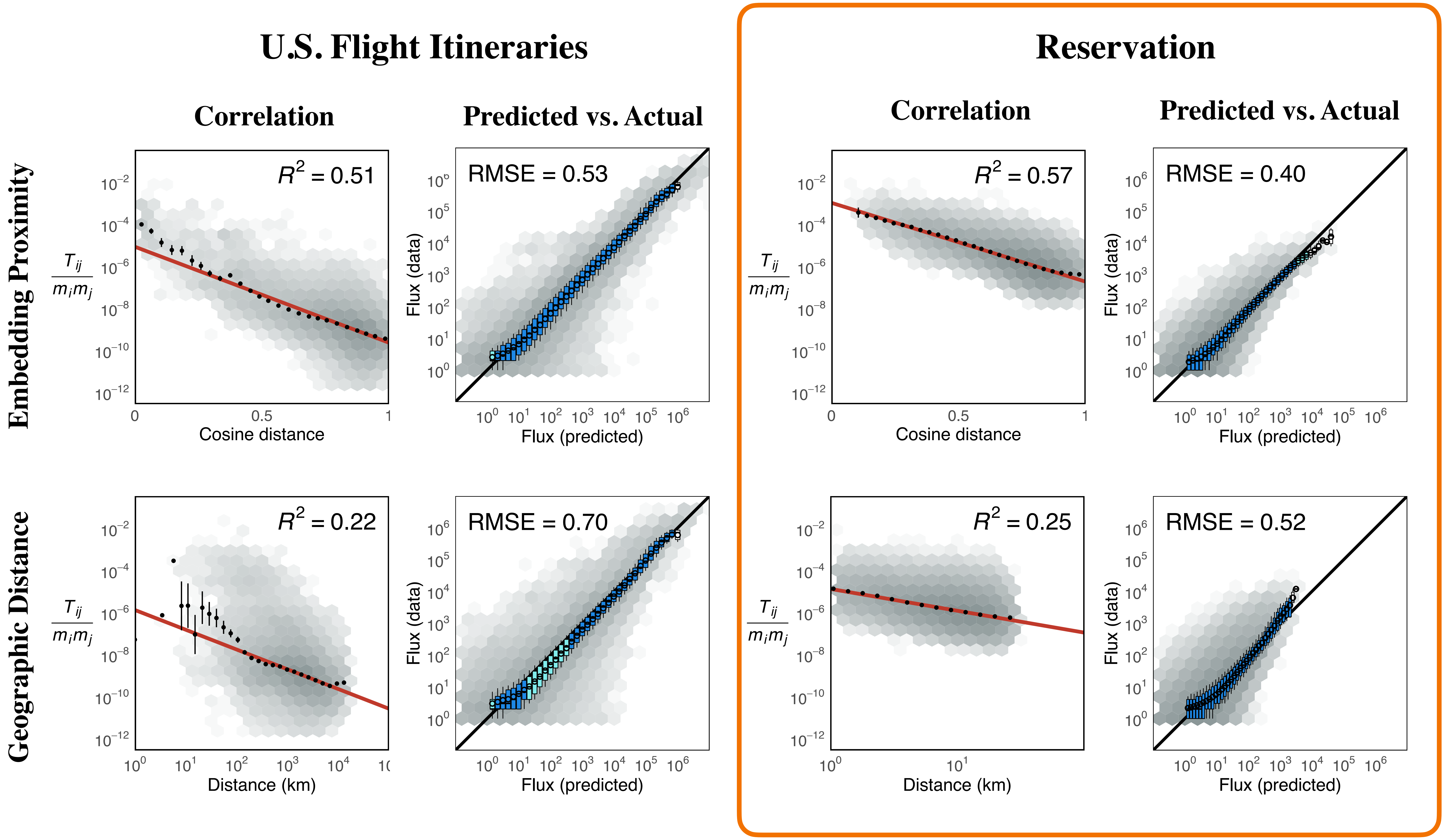
Better performance using embedding distance on trajectories derived from U.S. domestic flights



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The embedding performs well in other domains!

U.S. domestic flight itineraries and South Korean hotel reservation trajectories



And again for South Korean hotel reservation trajectories

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Why does it work so well?

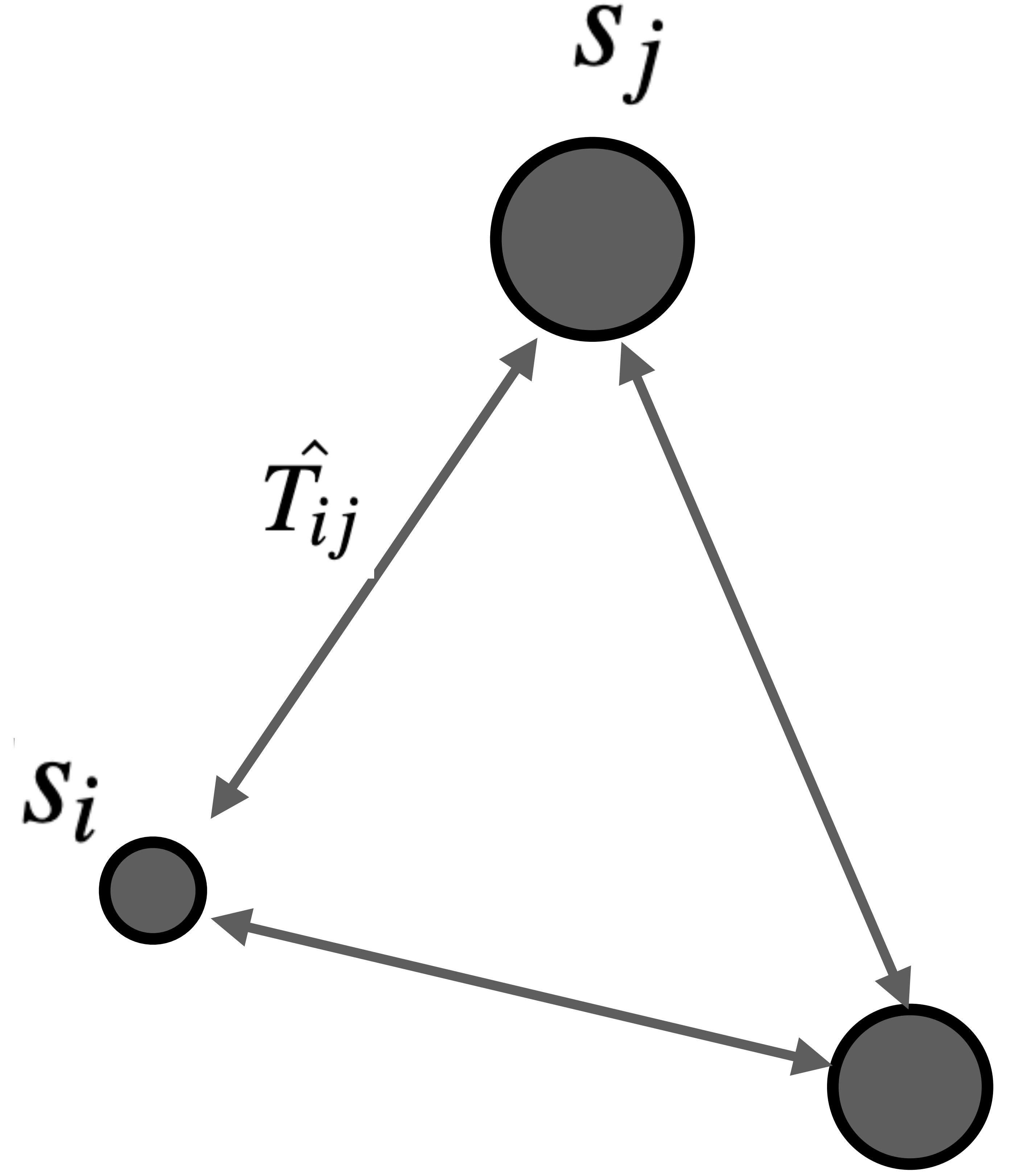
Word2vec is equivalent of a gravity model

$$\hat{T}_{ij} = C s_i s_j \exp(\boldsymbol{v}_j \cdot \boldsymbol{v}_i)$$

$\hat{T}_{ij}$: Flow between two locations i and j

s_i : Frequency of location i in all trajectories

$\boldsymbol{v}_j$: Embedding vector for location i



Jisung Yoon & Sadamori Kojaku

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

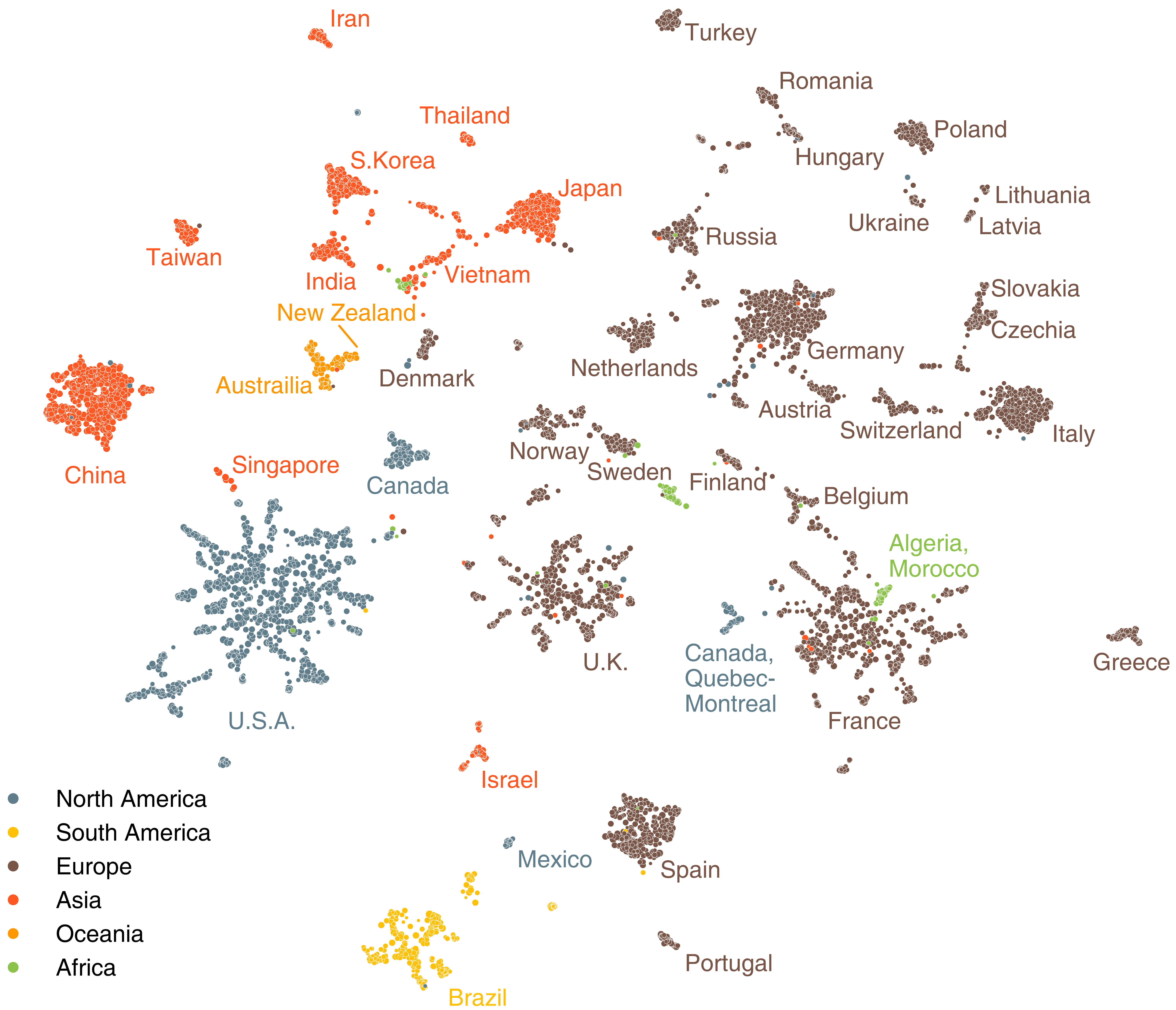
word2vec captures the latent structure of mobility

word2vec captures the latent structure of mobility

What is that structure?

Visualizing the embedding space

UMAP projection of organizations

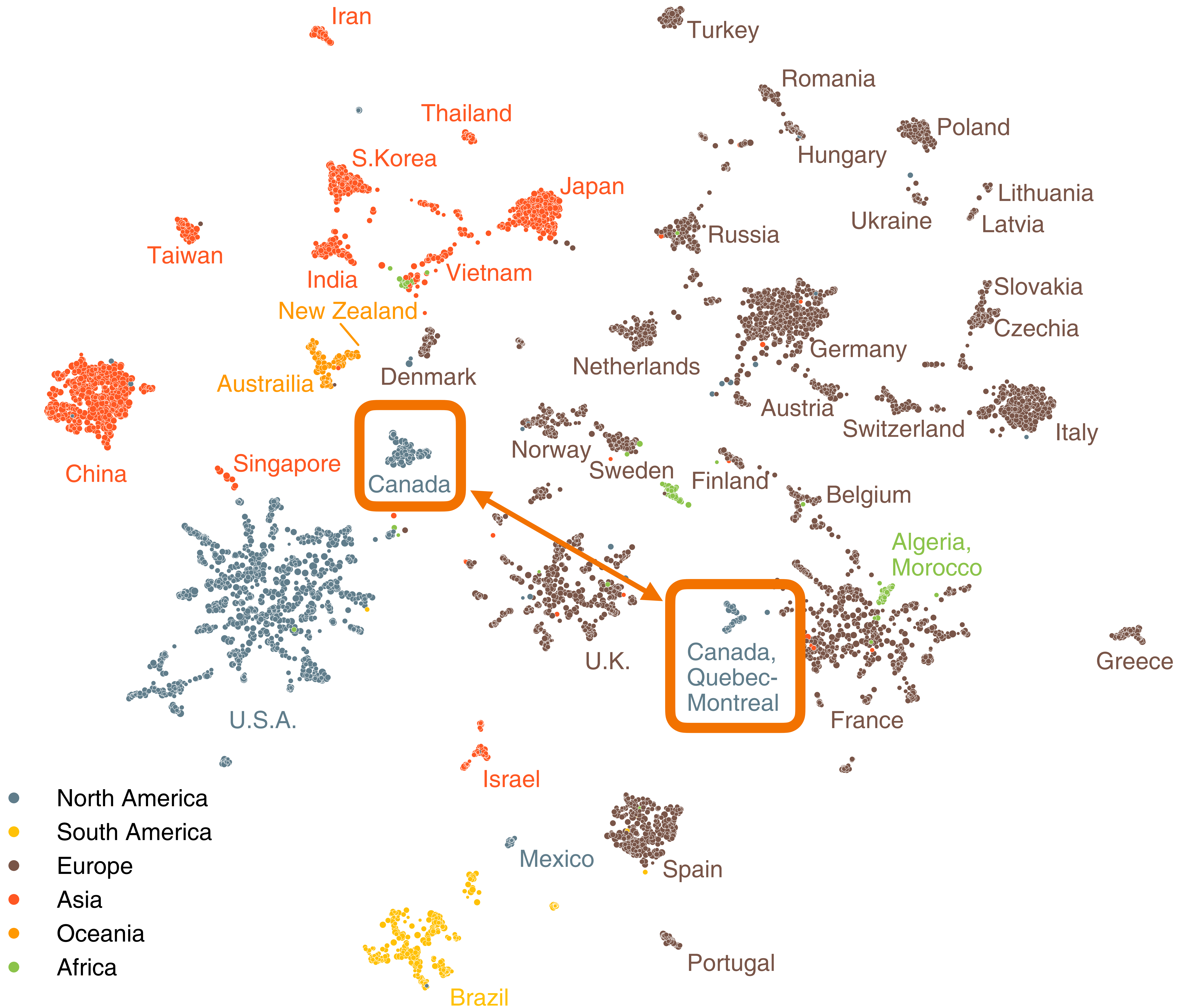


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

dakotamurray.me/talk/2021-copenhagen/

Visualizing the embedding space

Canada, Quebec, & French

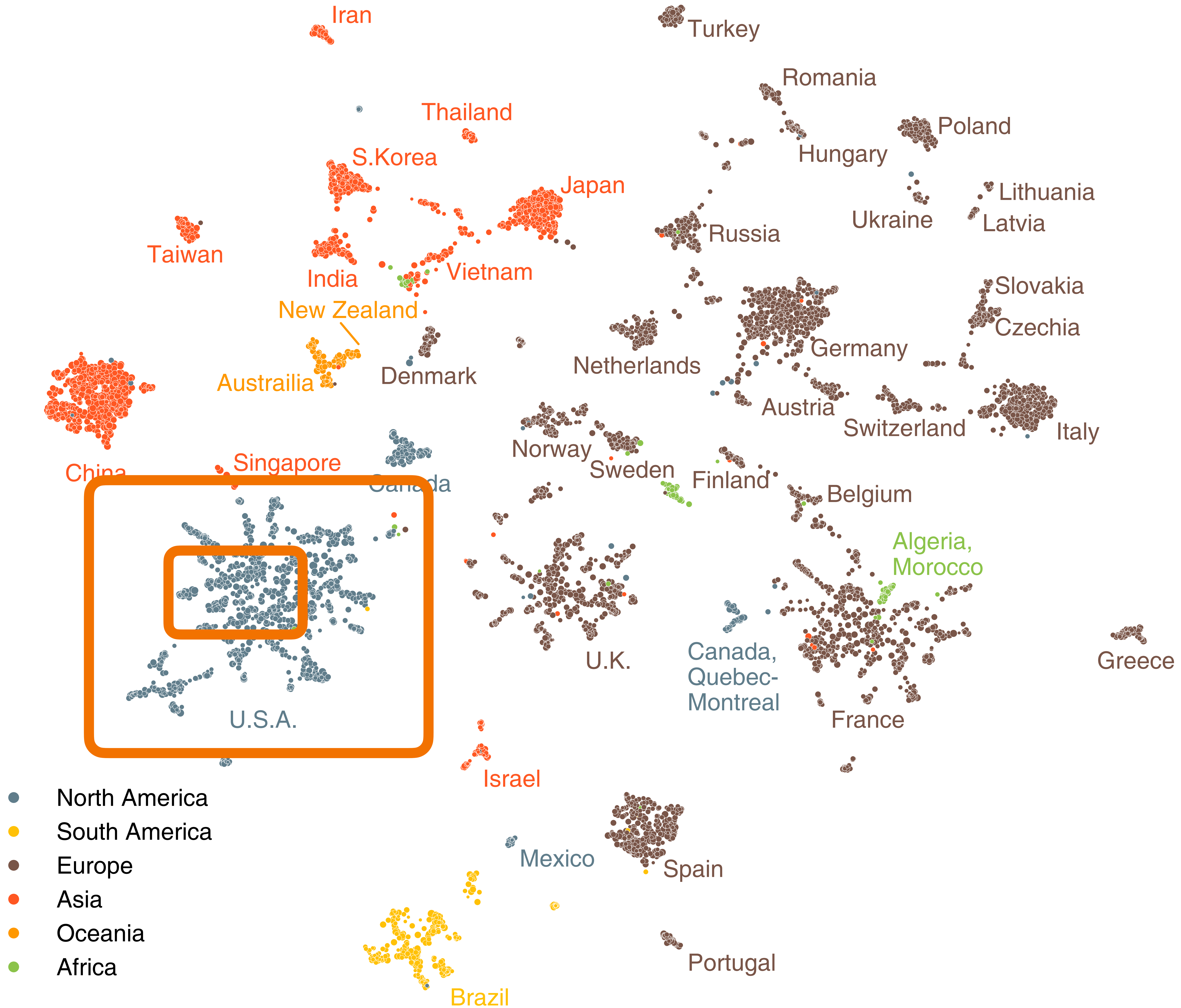


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Visualizing the embedding space

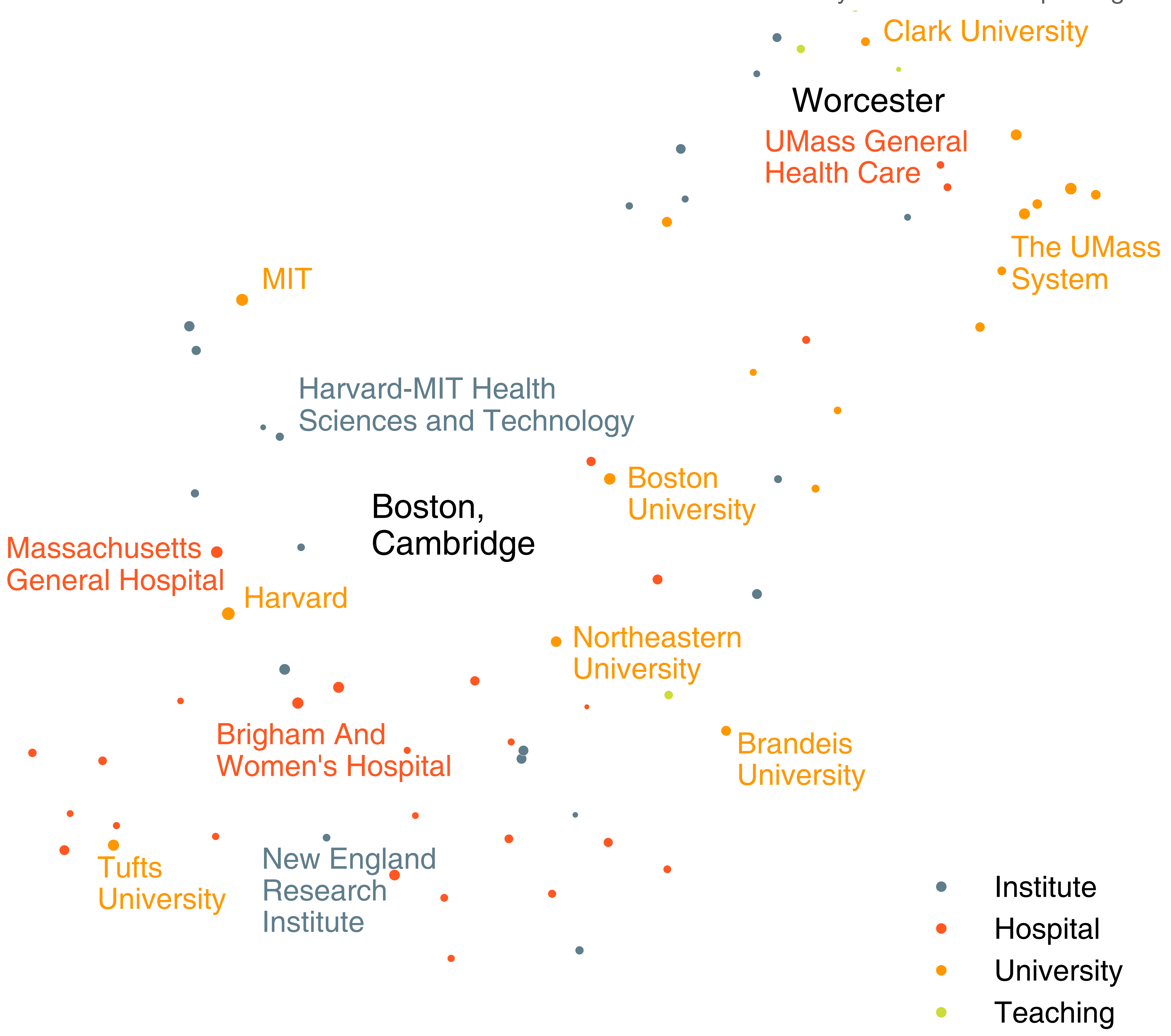
We can “zoom in”

US - Massachusetts



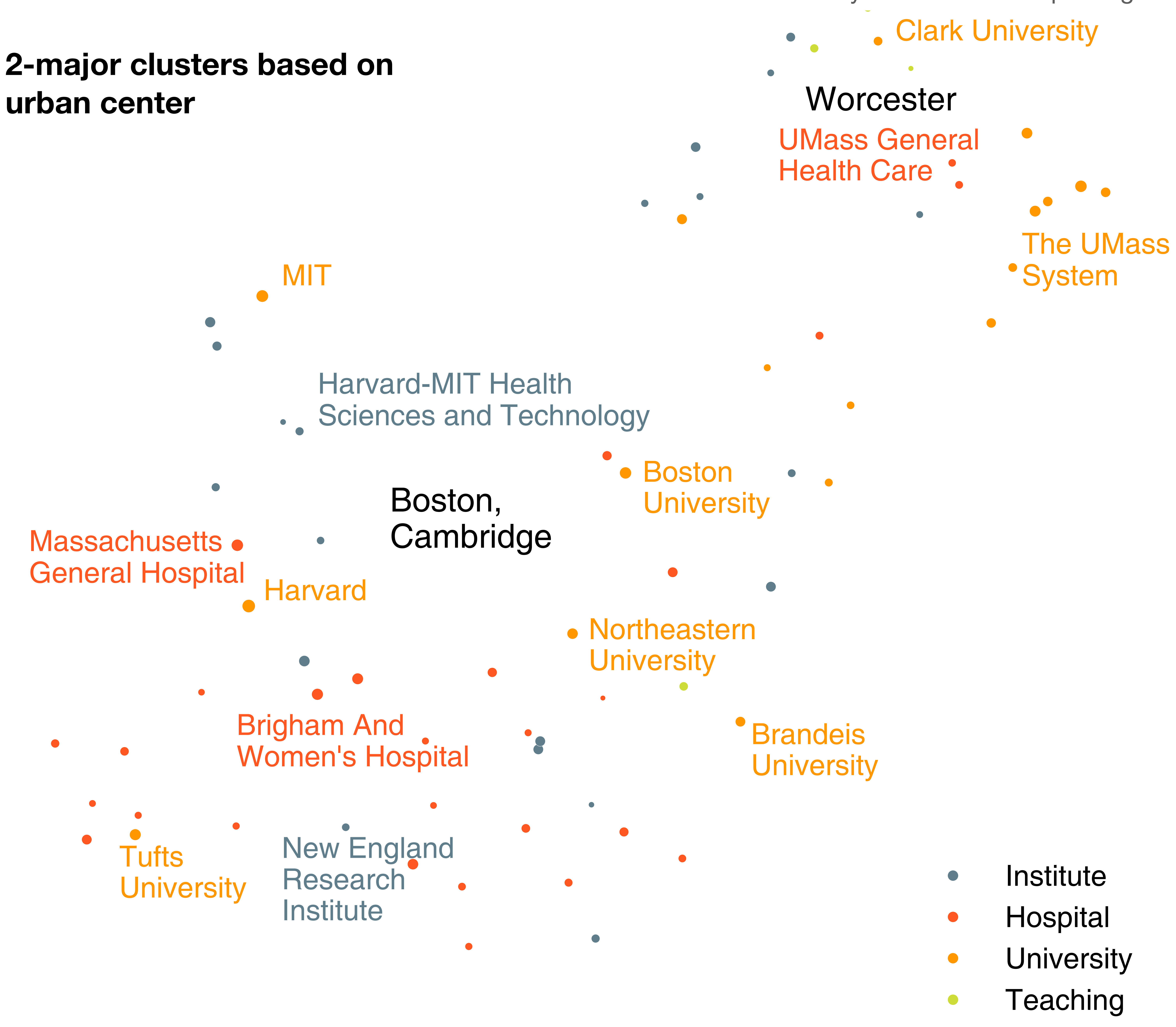
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Massachusetts



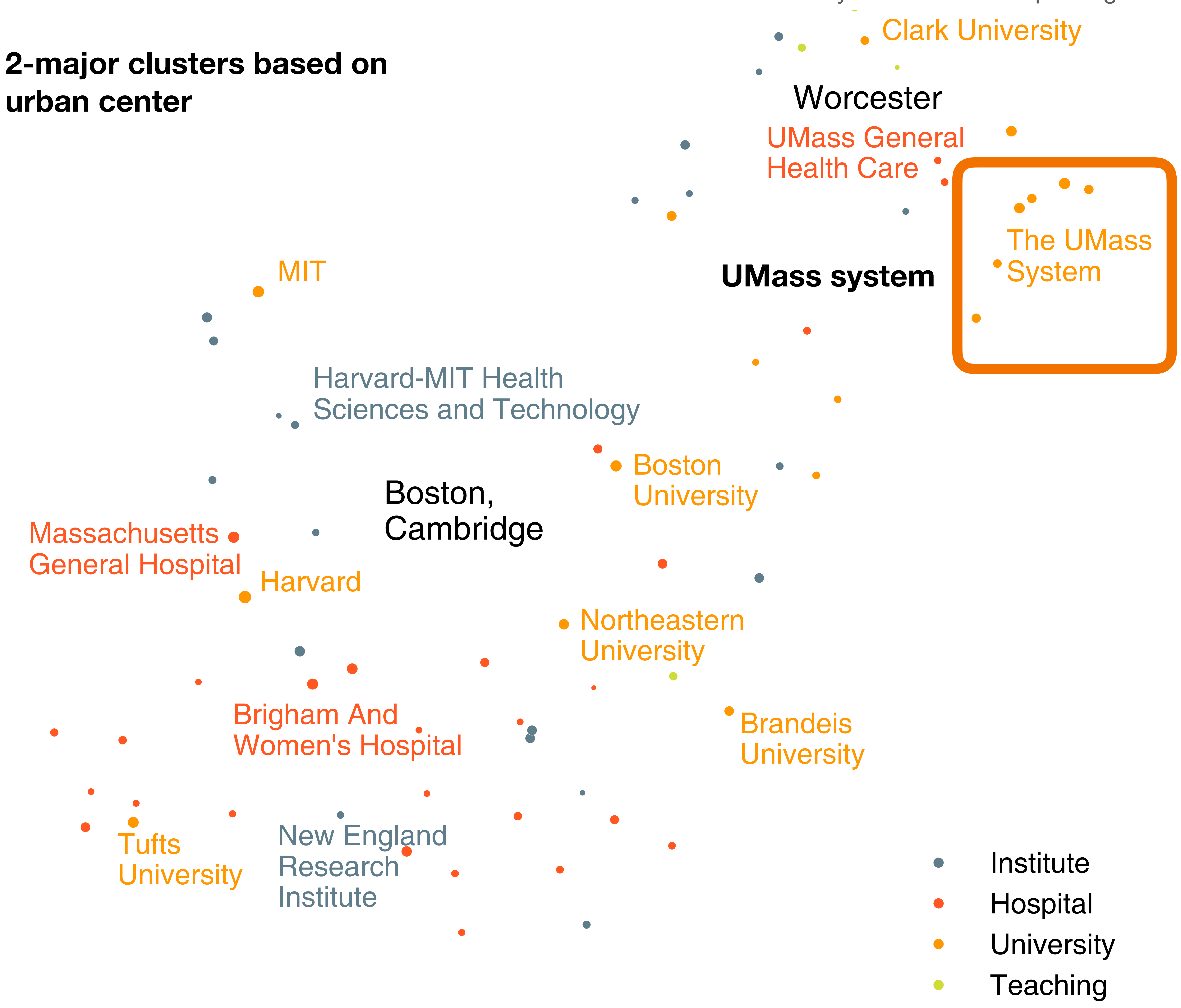
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Massachusetts



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

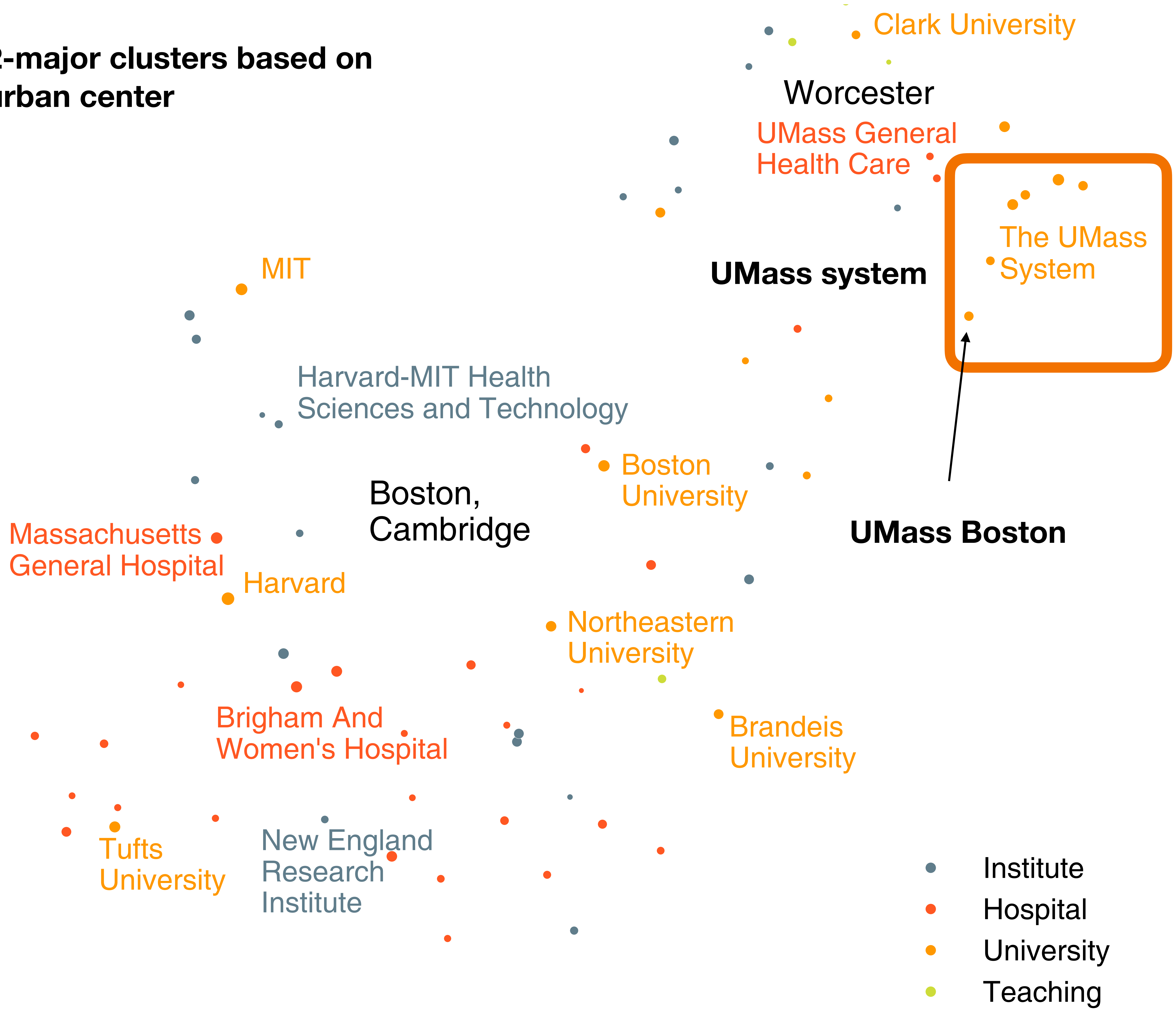
Massachusetts



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

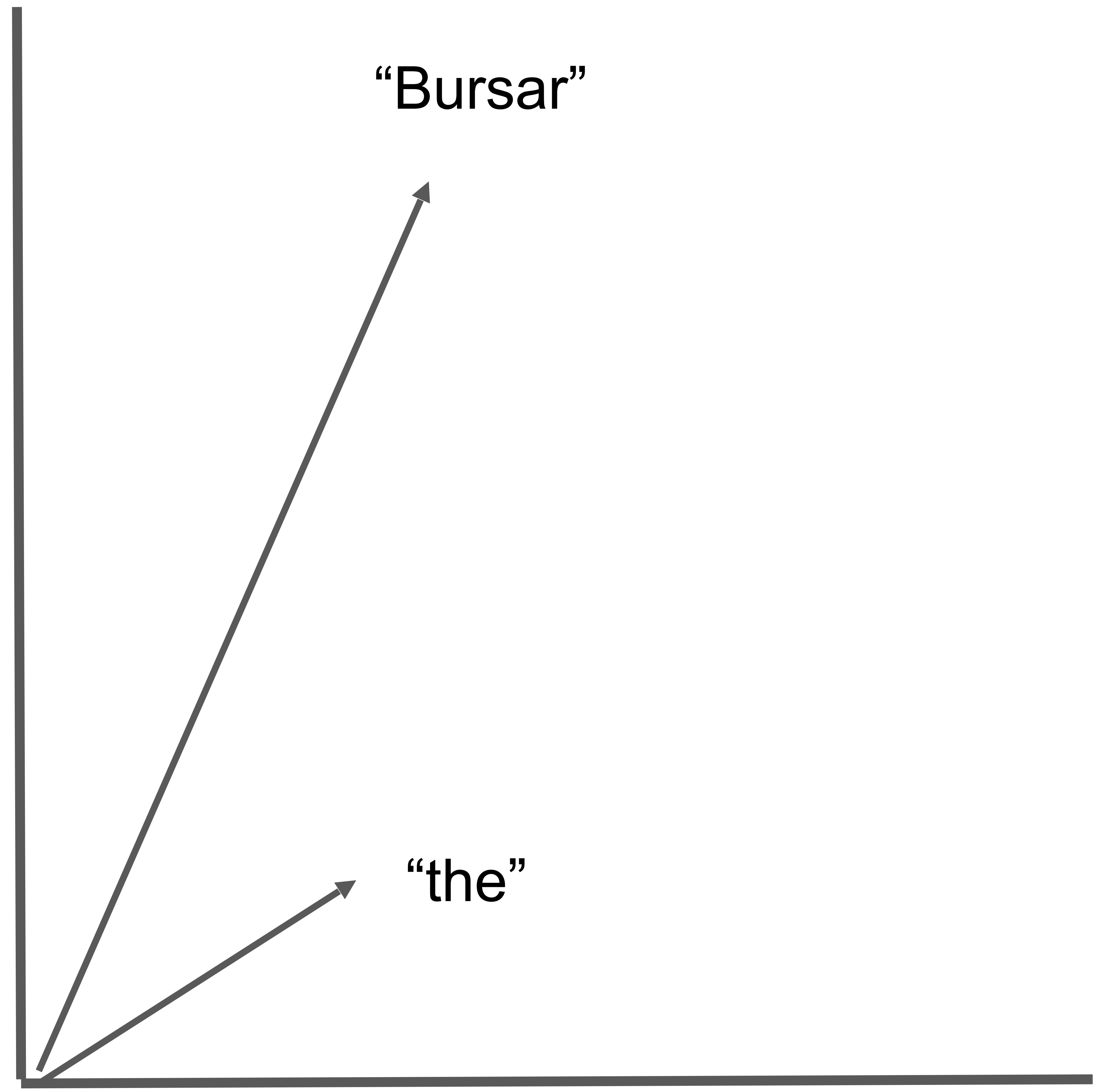
2-major clusters based on urban center

Massachusetts



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

What else is encoded? Vector length

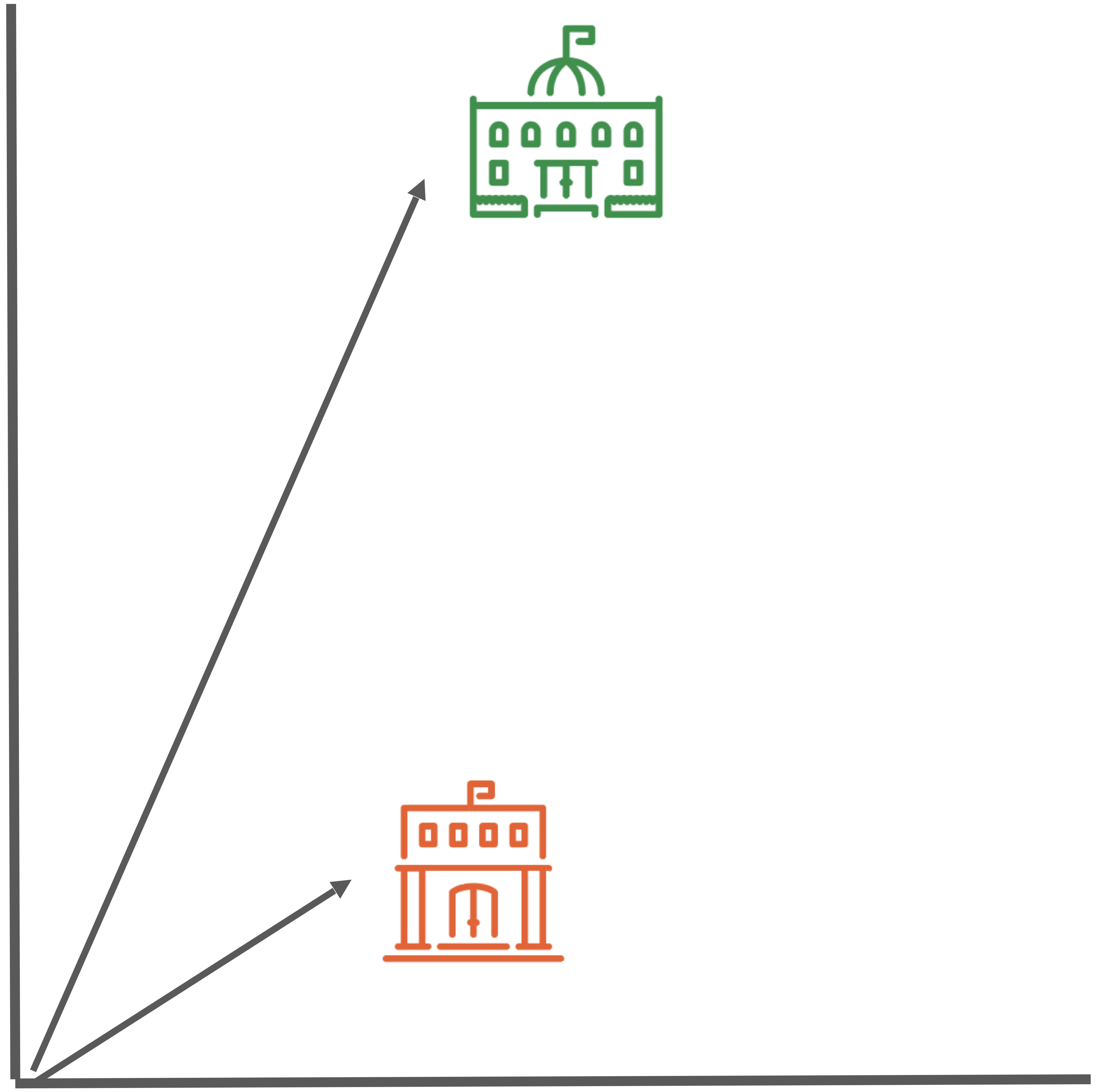


- In word embeddings, larger vectors (by magnitude) tend to appear in a single context

- Shorter vectors tend to appear in more and more different contexts—they are more universal. More *central*

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

What else is encoded? Vector length



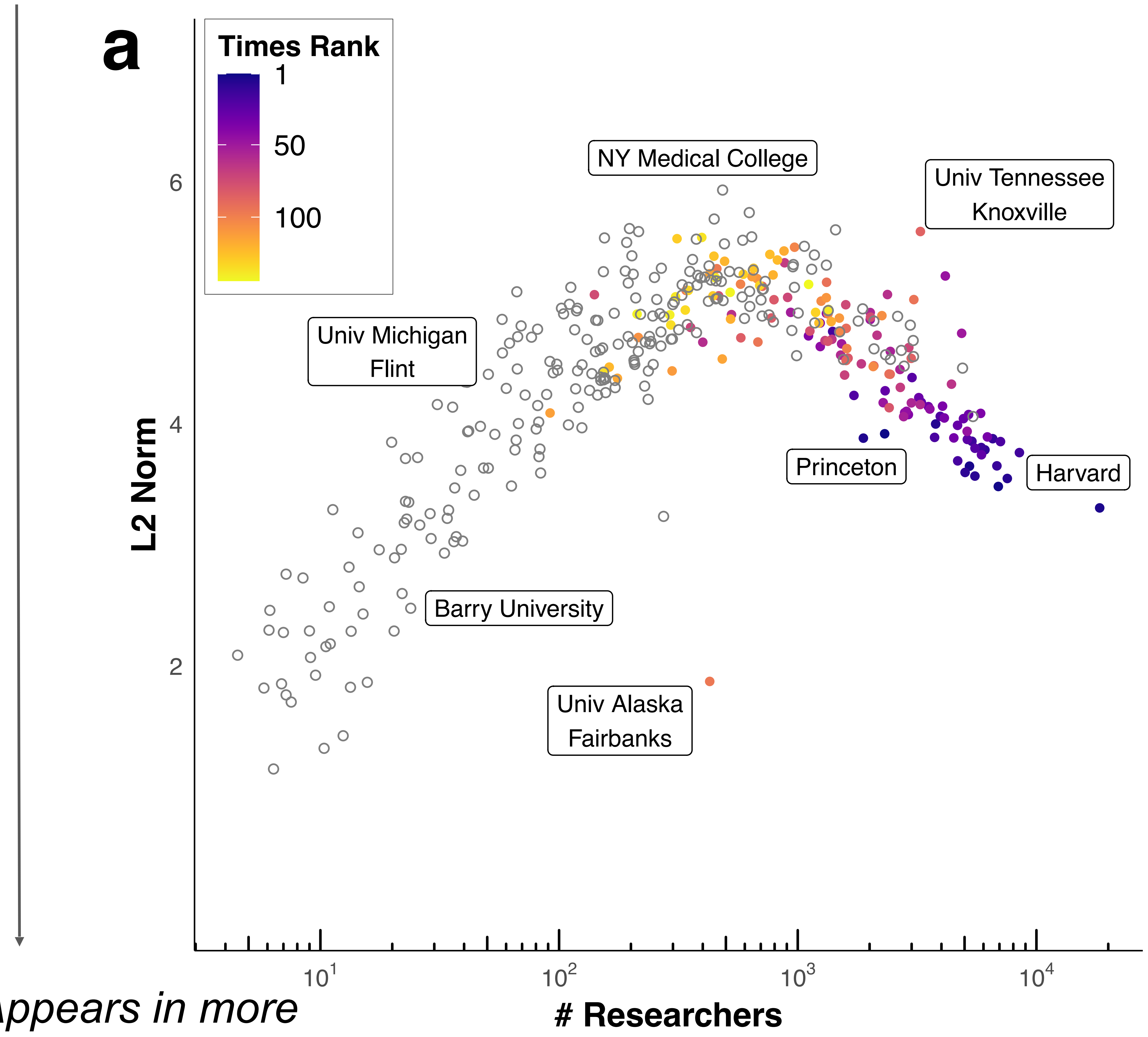
- In word embeddings, larger vectors (by magnitude) tend to appear in a single context

- Shorter vectors tend to appear in more and more different contexts—they are more universal. More *central*

- Also works for organizations

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

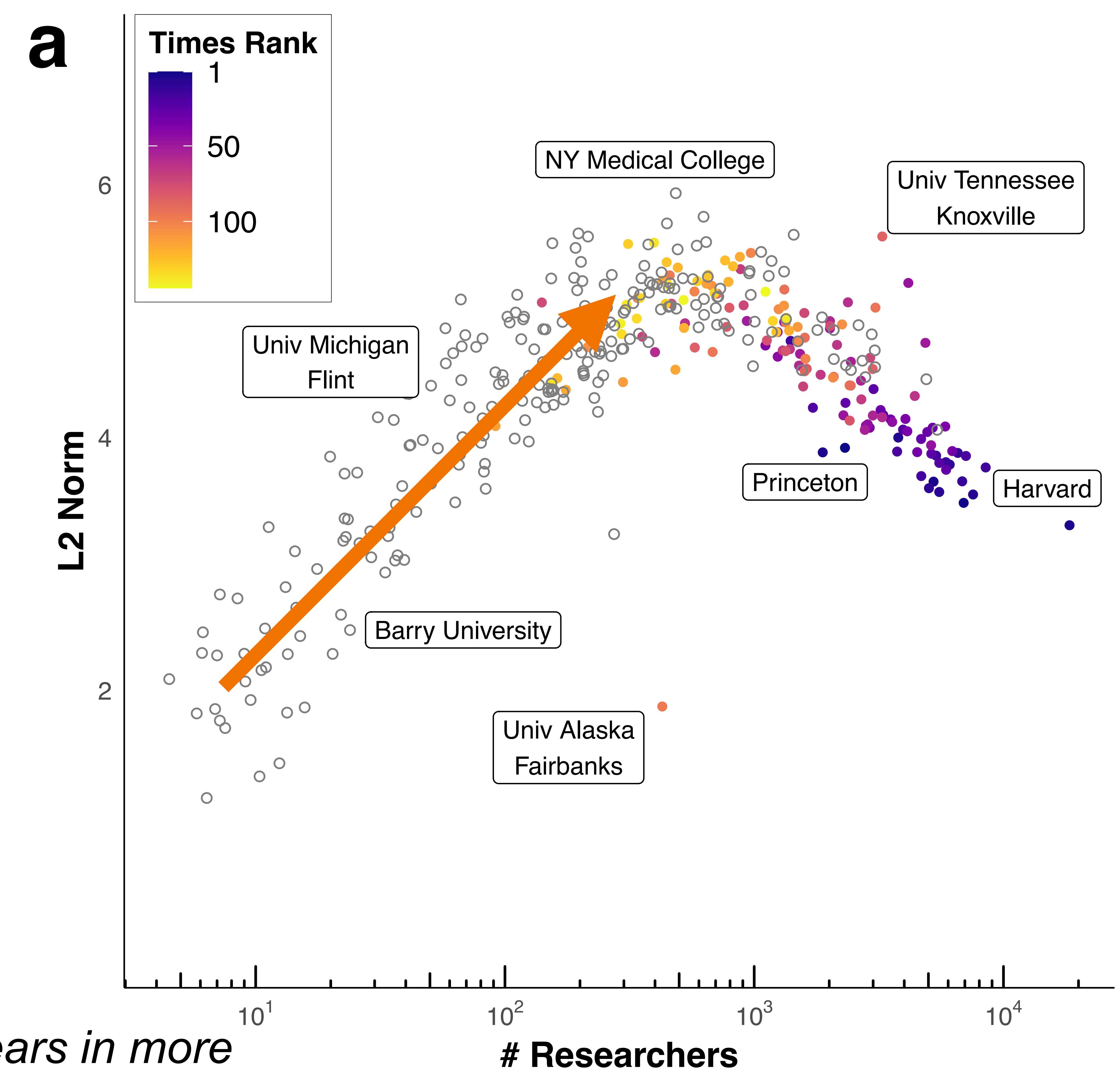
Prestigious U.S. universities appear in more diverse contexts



Appears in more diverse contexts

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Prestigious U.S. universities appear in more diverse contexts

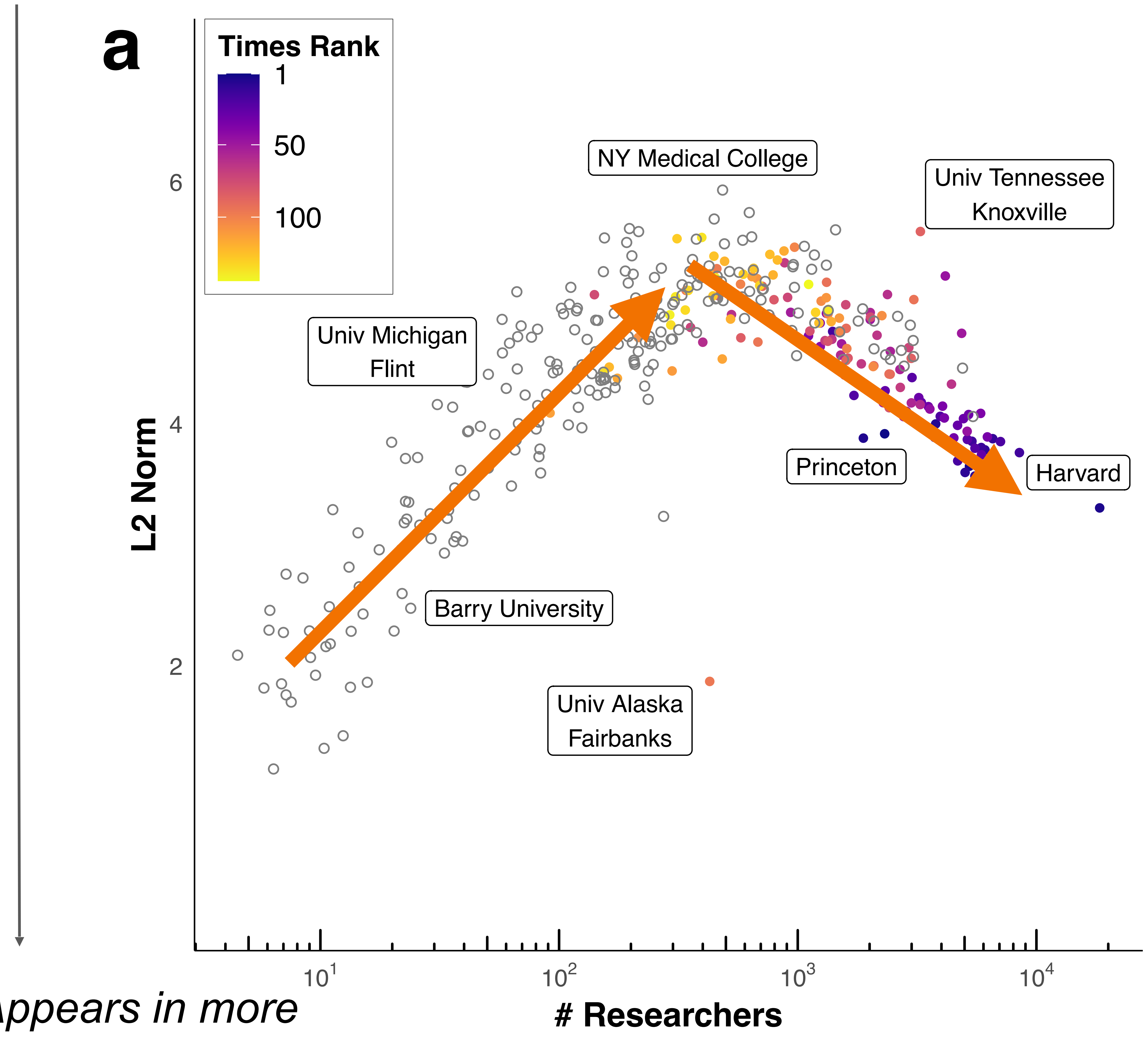


Appears in more diverse contexts

Bigger organizations are more isolated….unless they are prestigious

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Prestigious U.S. universities appear in more diverse contexts



Appears in more diverse contexts

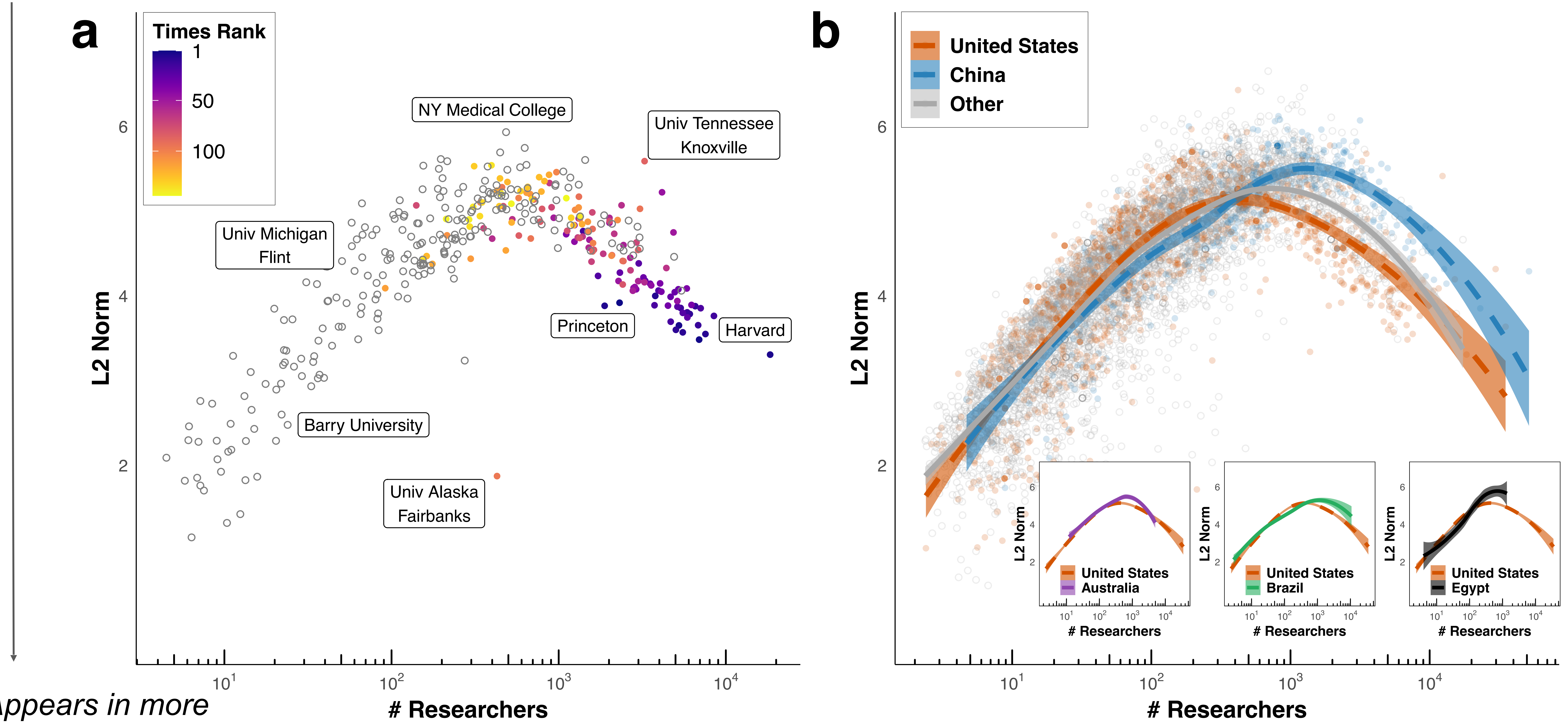
Bigger organizations are more isolated….unless they are prestigious

Those from prestigious universities are more central

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Prestigious U.S. universities appear in more diverse contexts

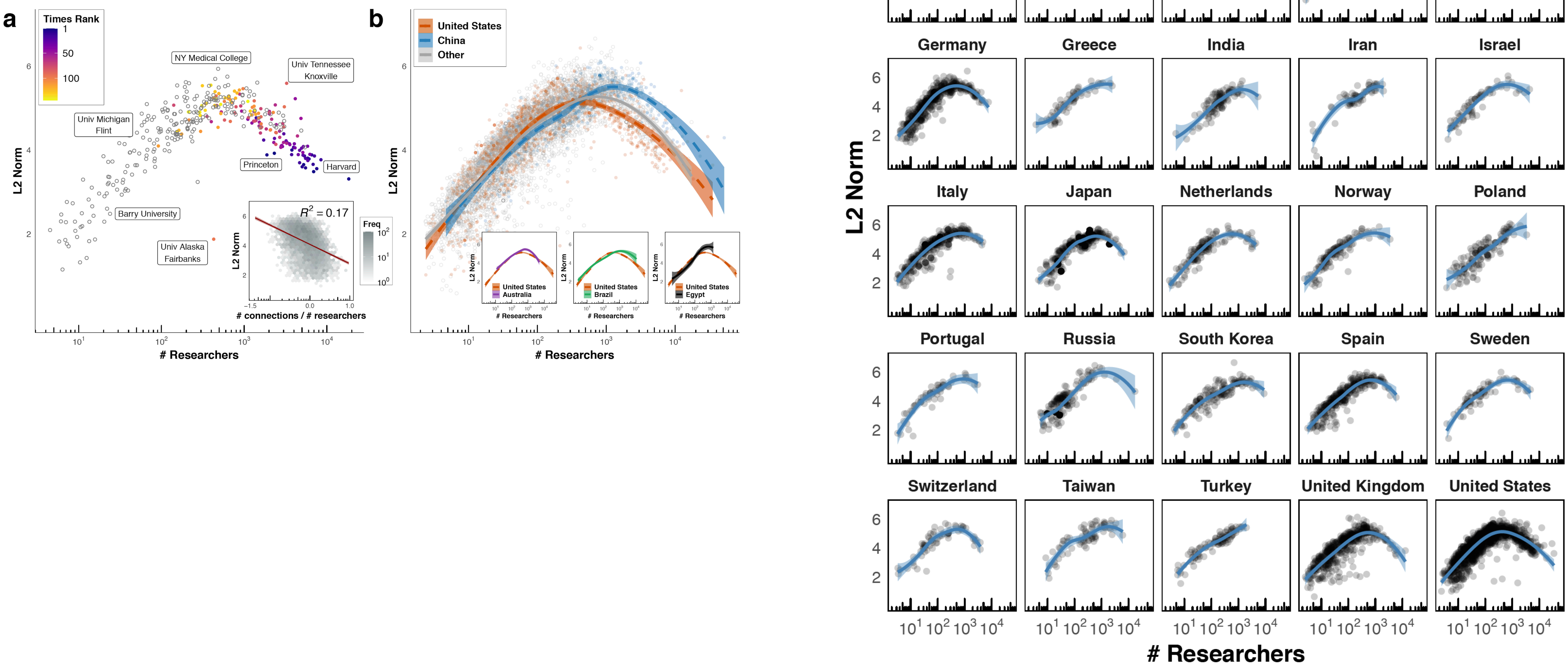
Repeats across many countries



Appears in more diverse contexts

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

The universal boomerang



Conclusion

- Embeddings capture the latent structure of mobility
- Underlying connection between the gravity model and *word2vec*
- Geography, language, culture, and prestige structure global scientific mobility
- Embeddings offer an opportunity for deeper understandings of mobility in many domains

Thank you!

Dakota Murray
@dakotasmurray
dakota.s.murray@gmail.com

Slides at: dakotamurray.me/talk/2021-copenhagen/

Appendix

Mobility is central to science

Institutionalized in evaluation

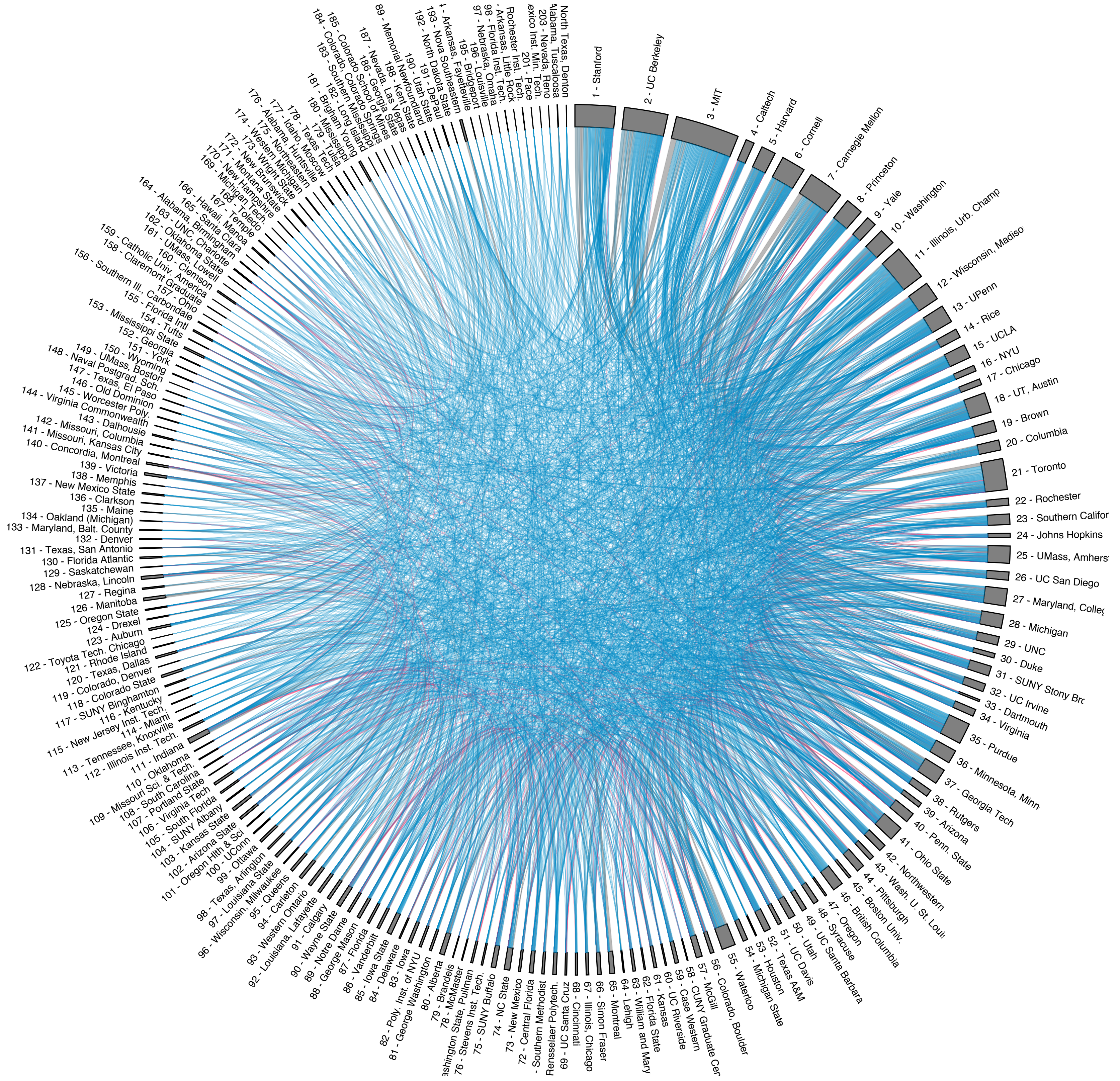
Article 19: Requirement for International Visits
When applying for promotion to full professor or equivalent rank, the applicants who were born after January 1, 1970, must complete at least a 6-month international visit.

- Hangzhou Dianzi University

Thanks to Fei She

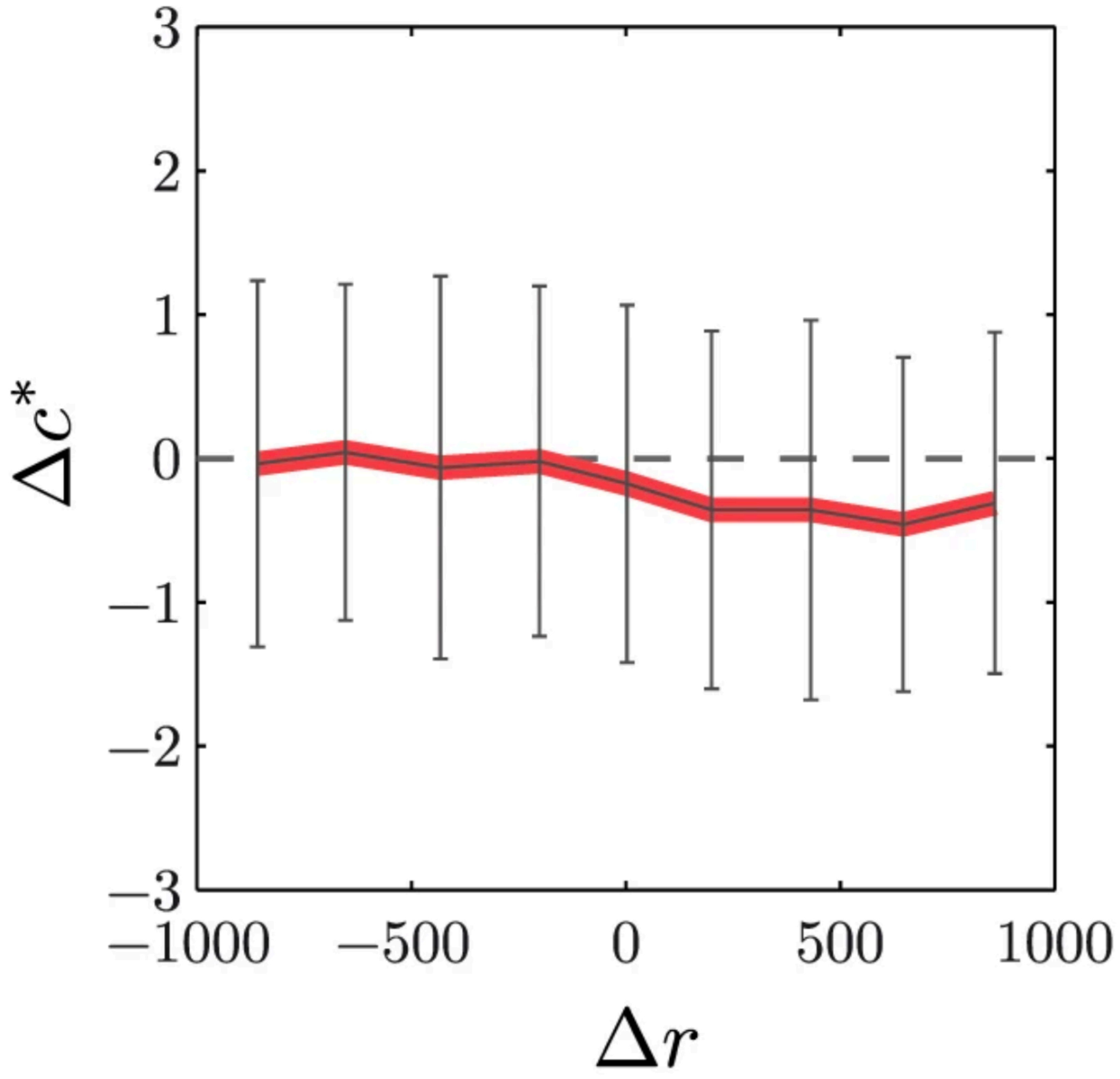
Mobility is intimately tied to evaluation and performance

Mobility an output of evaluation in faculty hiring



Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*(1), e1400005.

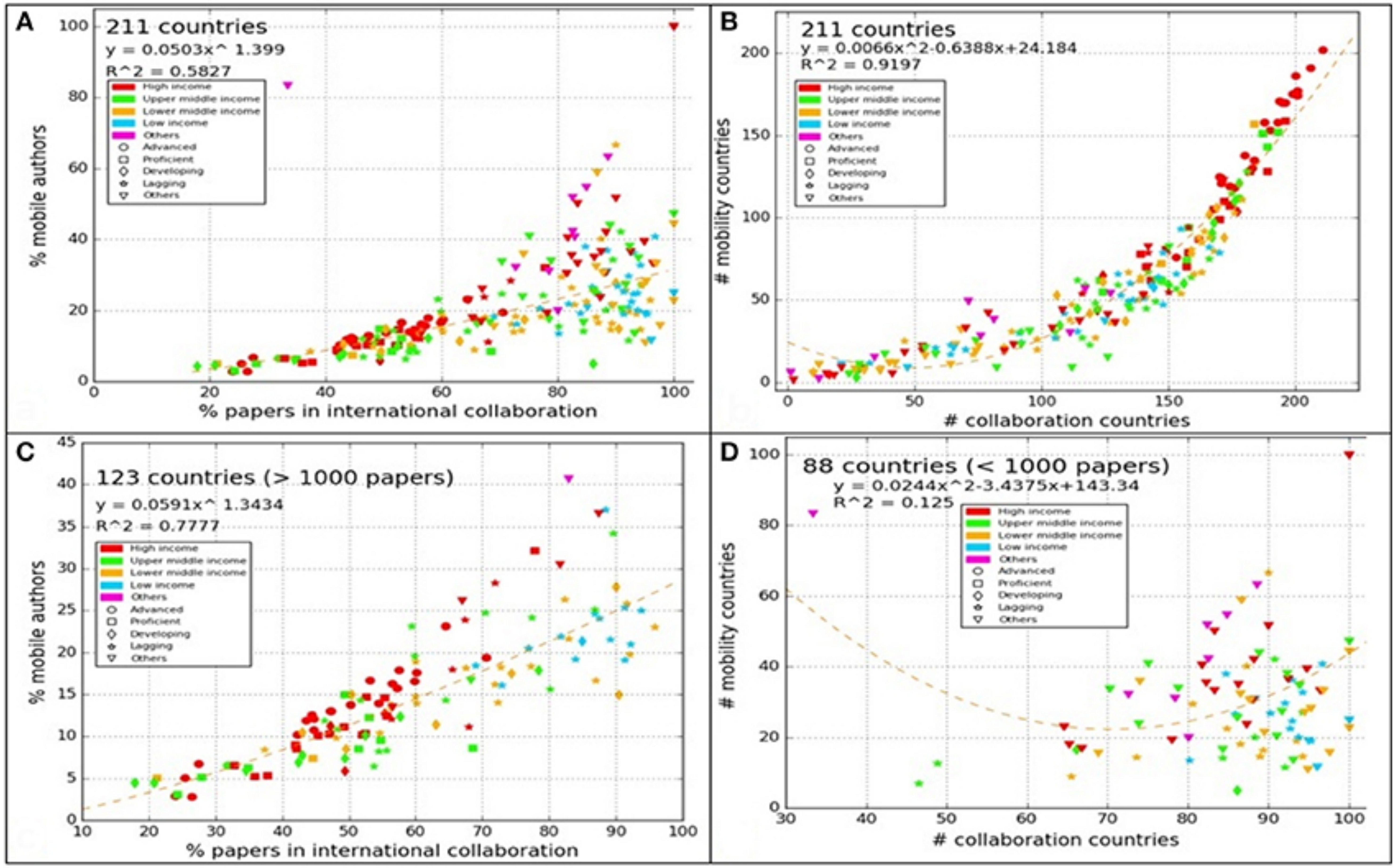
Moving to a low-rank institution lowers impact



Deville, P., Wang, D., Sinatra, R., Song, C., Blondel, V. D., & Barabási, A.-L. (2014). Career on the Move: Geography, Stratification and Scientific Impact. *Scientific Reports*, *4*(1), 1–7.

Mobility develops international relations

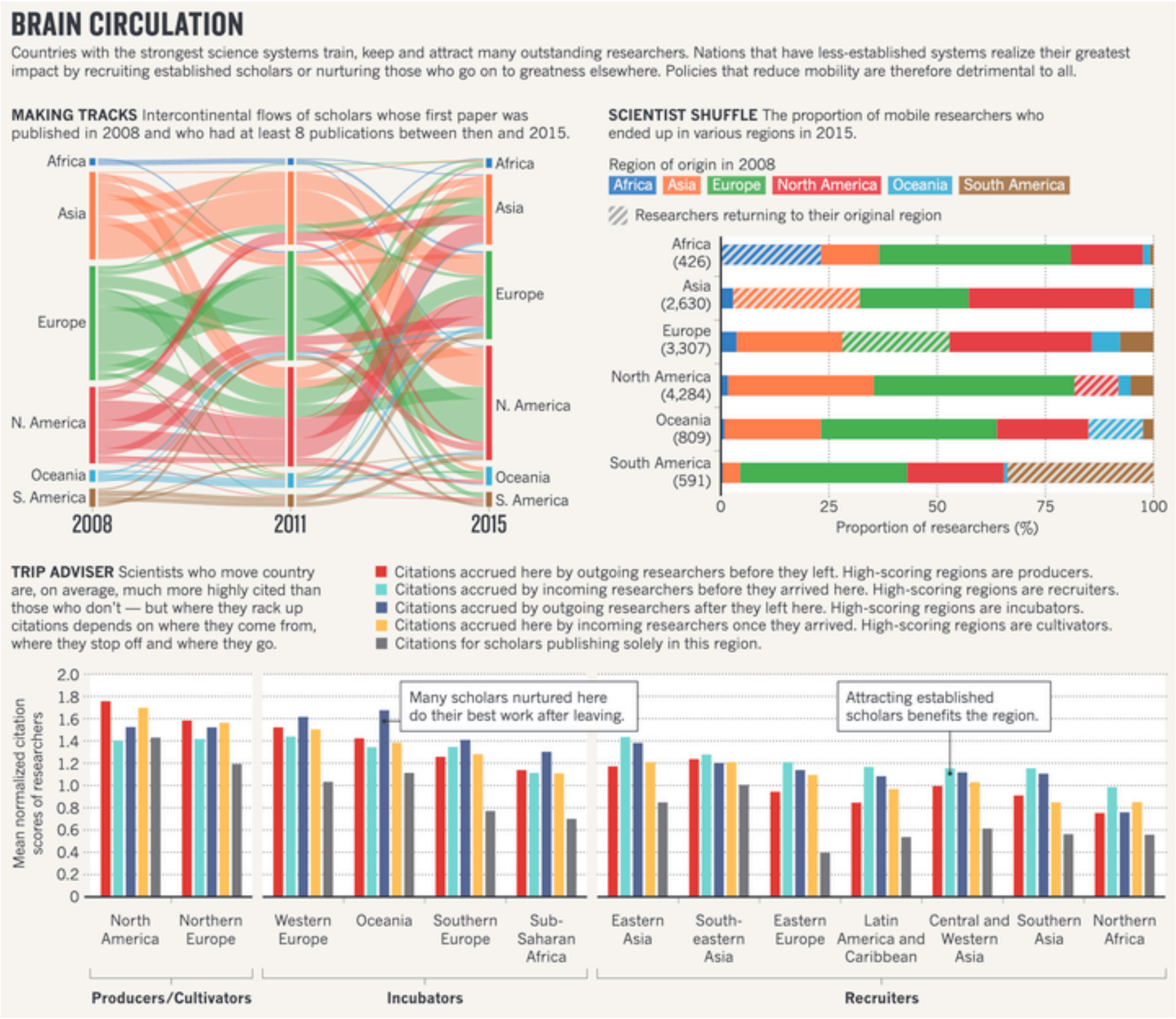
Countries with more mobility between them, also have more collaboration!



Chinchilla-Rodríguez, Z., Miao, L., **Murray, D.**, Robinson-García, N., Costas, R., & Sugimoto, C. R. (2018). A Global Comparison of Scientific Mobility and Collaboration According to National Scientific Capacities. *Frontiers in Research Metrics and Analytics*, *3*.

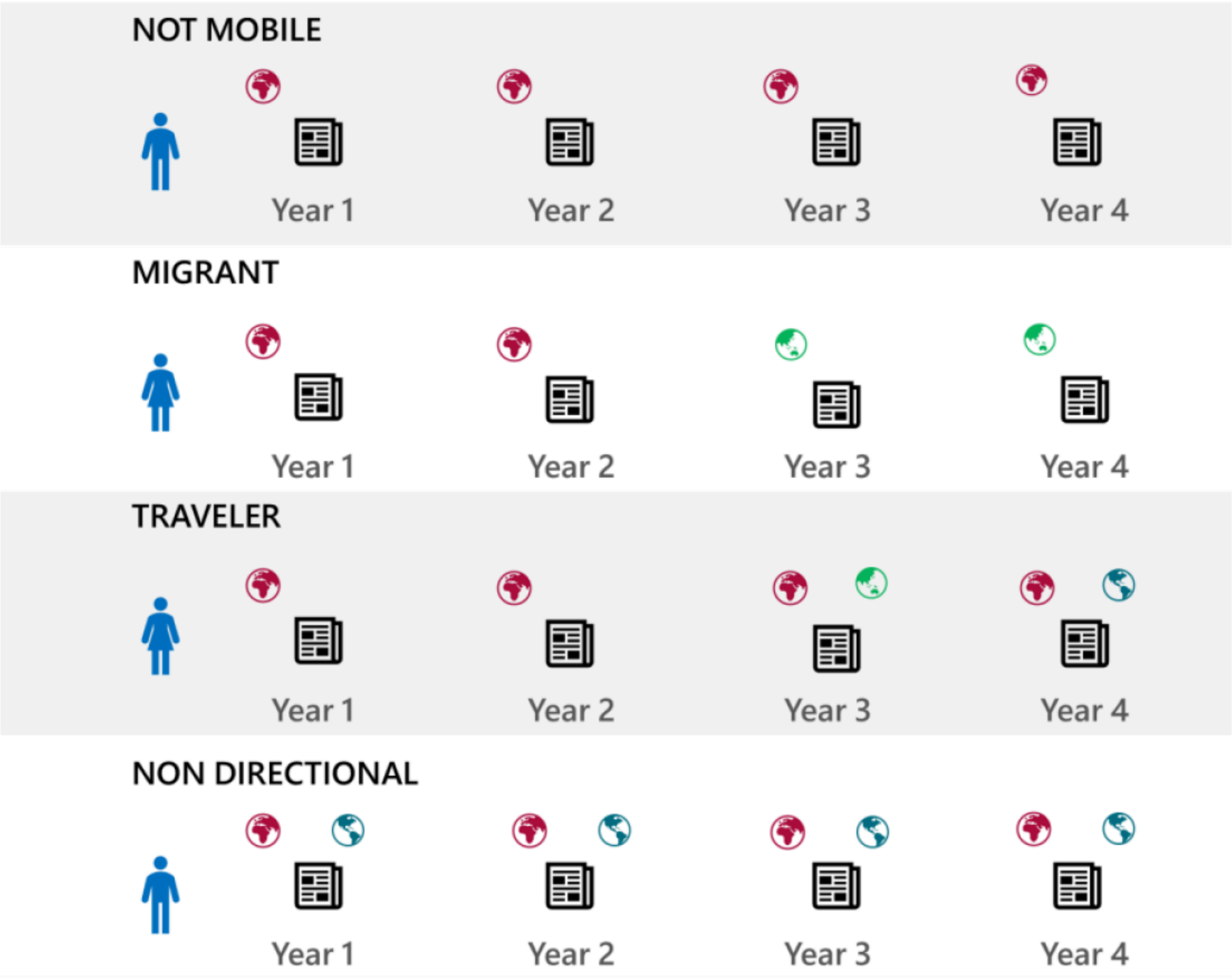
Mobility fosters improved performance

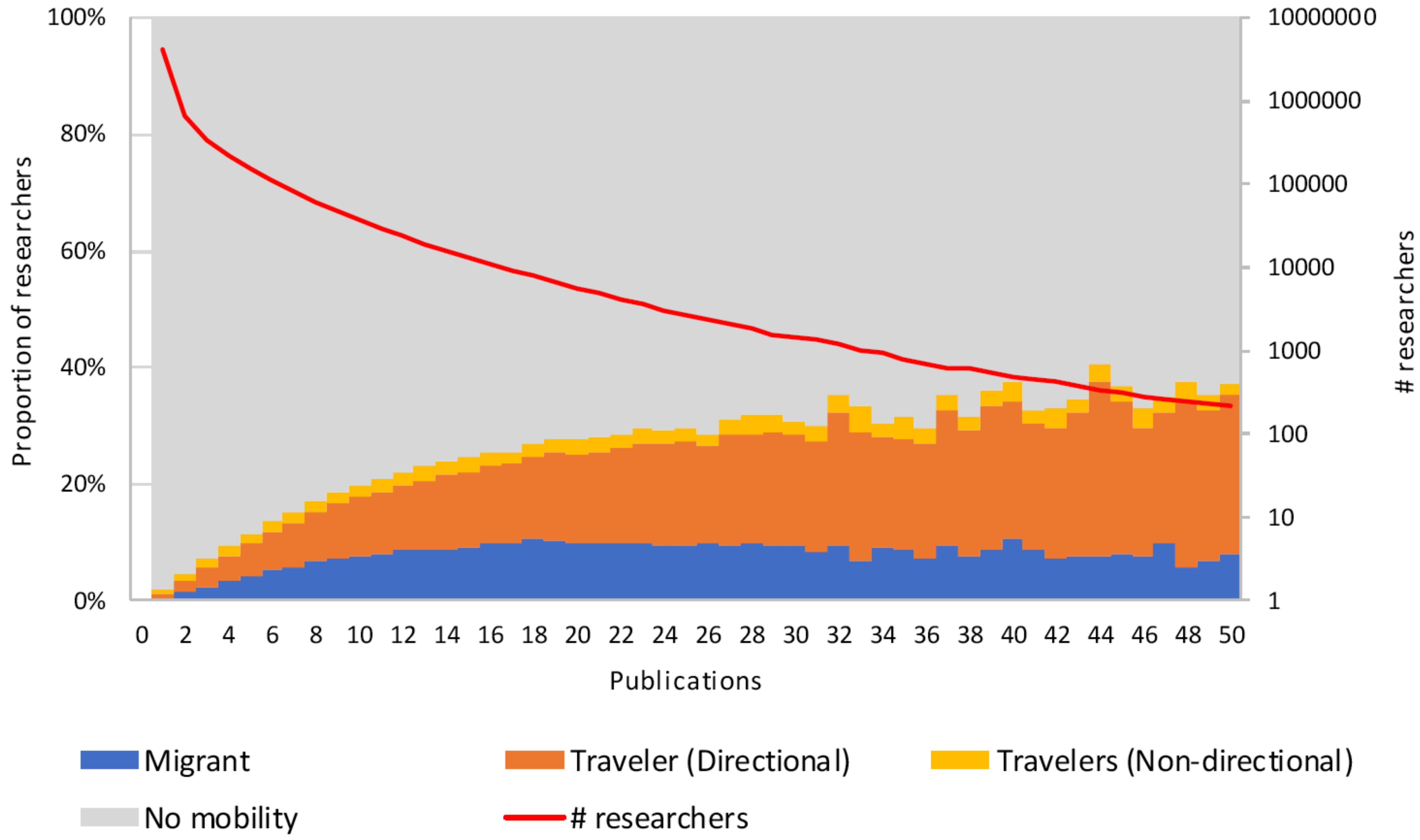
Scholars who are mobile
have higher citation impact!



Sugimoto, C. R., Robinson-Garcia, N., **Murray, D. S.**, Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, *550*(7674), 29–31.

Mobility is complicated





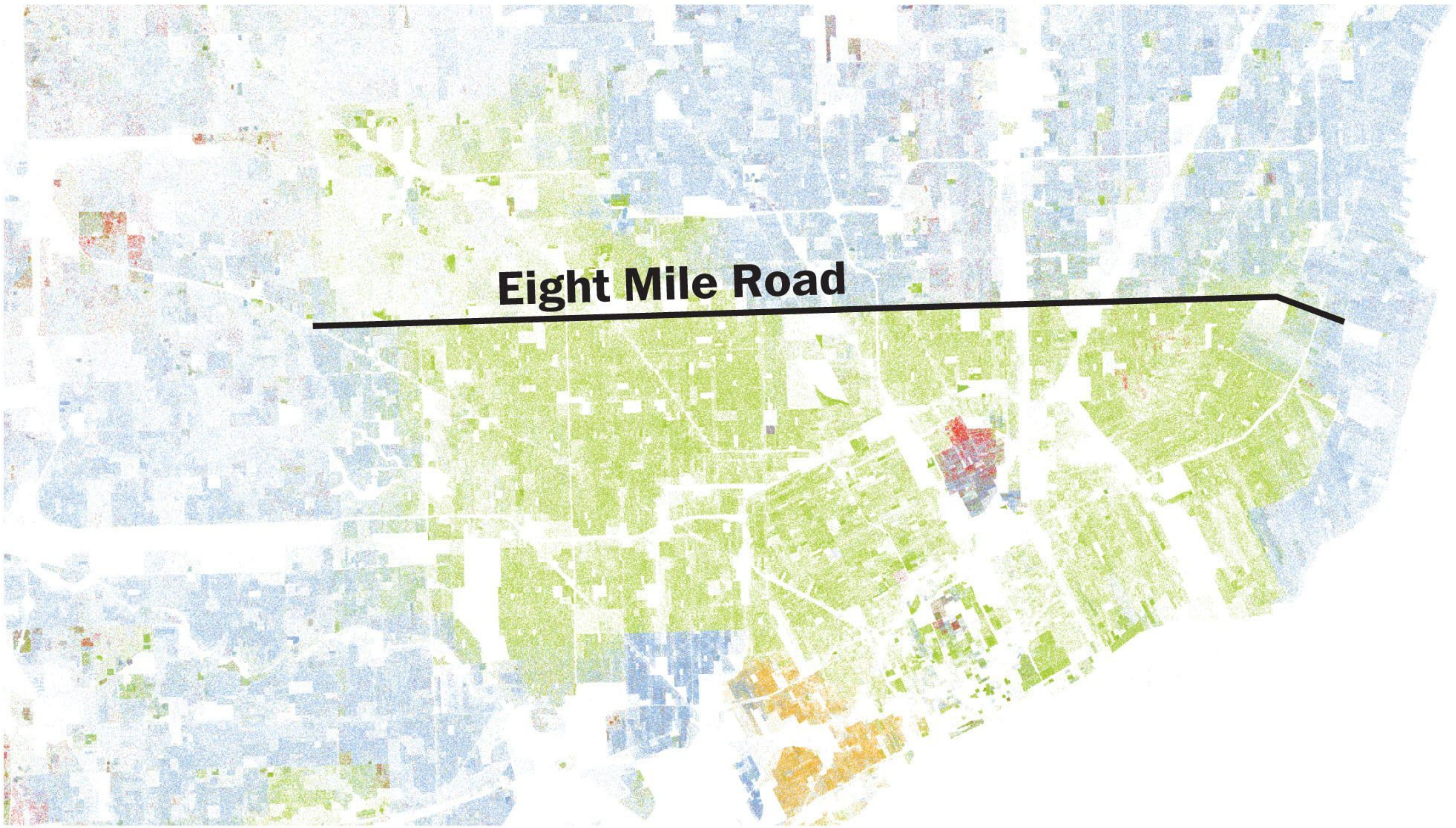
Robinson-Garcia, N., Sugimoto, C. R., **Murray, D.**, Yegros-Yegros, A., Larivière, V., & Costas, R. (2018). Scientific mobility indicators in practice: International mobility profiles at the country level. *El Profesional de La Información*, *27*(3), 511.

Robinson-Garcia, N., Sugimoto, C. R., **Murray, D.**, Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, *13*(1), 50–63.

Lines of segregation in Detroit

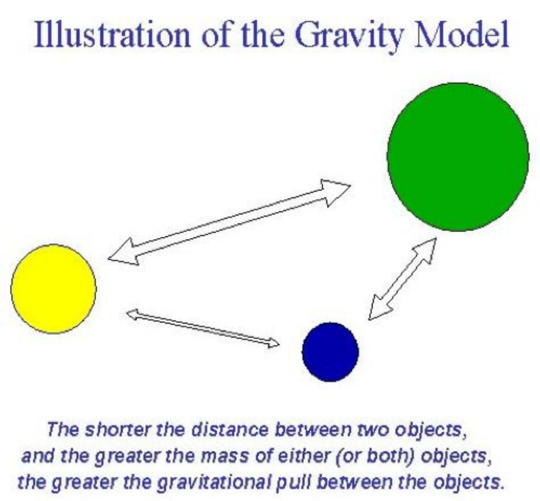
Economic opportunity and demographic makeup can wildly diverse in just a few city blocks

1 dot = 1 person White Black Asian Hispanic Other



Source: U-Va. Cooper Center analysis of 2010 Census data

THE WASHINGTO



Gravity Model

- Popular model of mobility
- ***Flows*** between places (co-affiliations) a function of their size and distance $T_{ij} = C\frac{m_i m_j}{f(r_{ij})}$
- We use two kinds of distance measure.
    1. Geographical distance
    2. Embedding distance $d_{ij} = 1 - \frac{v_i \cdot v_j}{|v_i||v_j|}$

A basic principle: a good *representation* allows *prediction*.

$$P(w_t|w_c) \gg P(w_{\mathrm{random}}|w_c)$$

Target Context

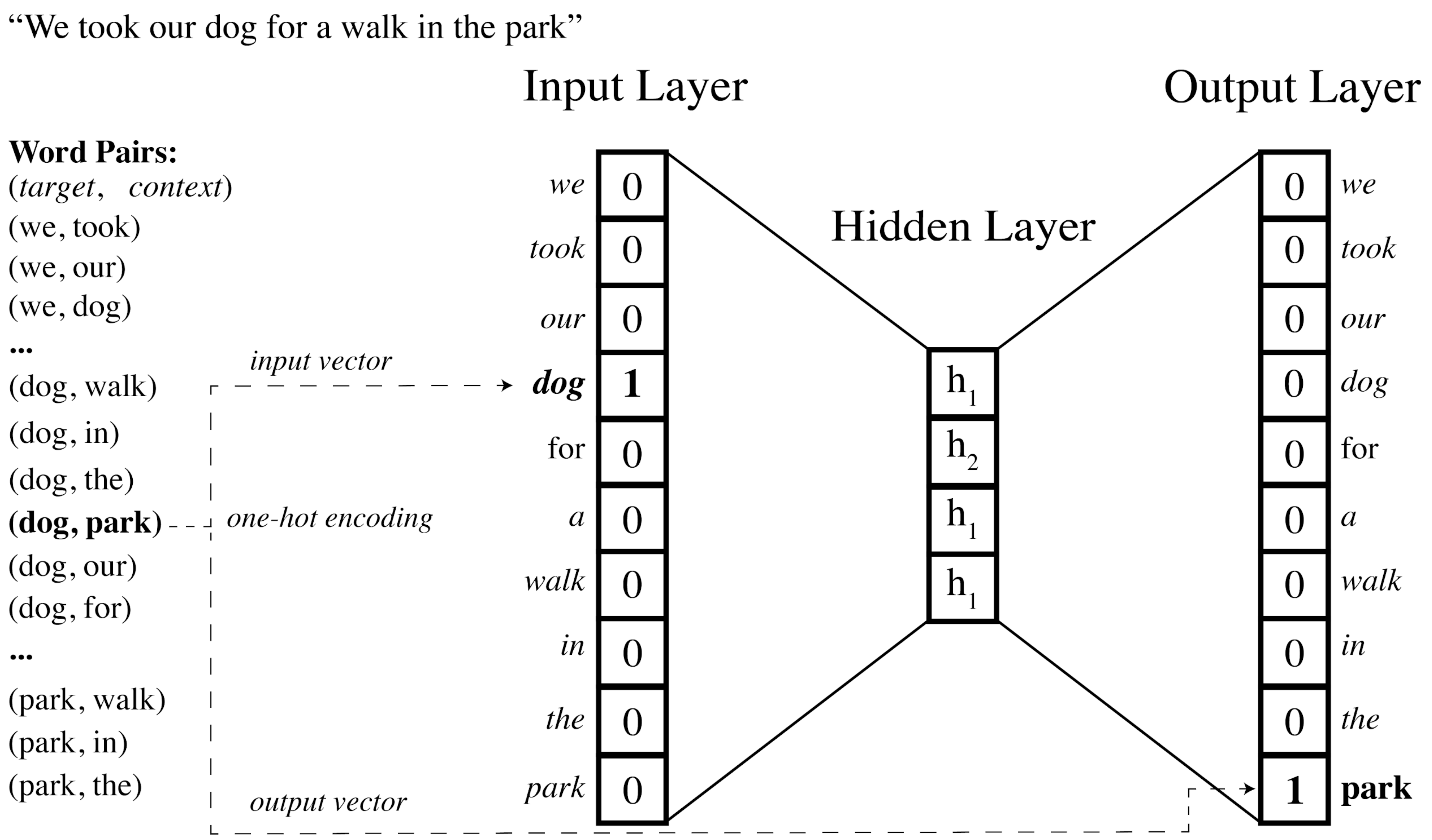
Let's assume that we can calculate the conditional probability with a function (e.g. dot product) of word **vectors** and **learn those vectors with a neural network**.

Learning word embeddings

Train a neural network to predict context words given a target

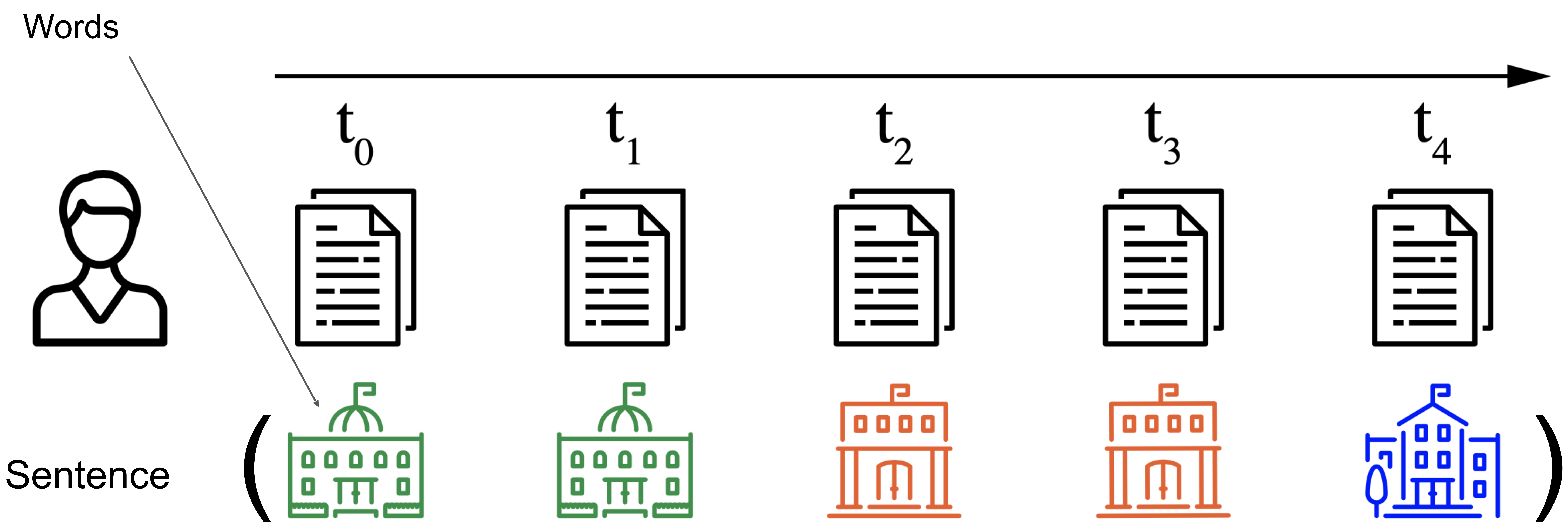
The hidden layer maps targets to concepts!

Words with similar contexts will have a similar “mapping” vector in the hidden layer



They don't have to be real "words" or "sentences"

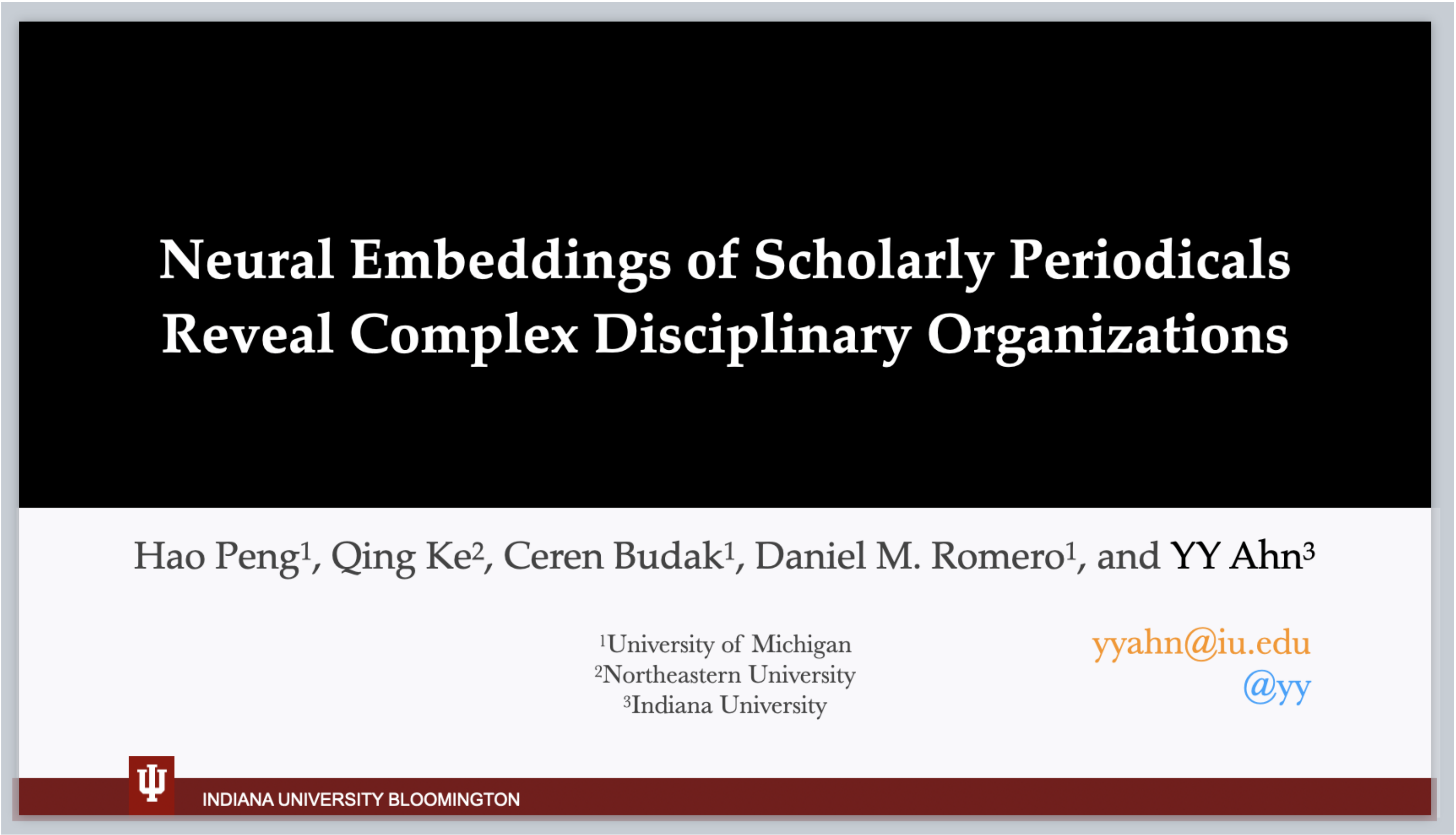
Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use **trajectories of scientists** as **sentences** and **organizations** as **words**.



They don't have to be real "words" or "sentences"

Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use the **trajectories of scientists** as **sentences** and **institutions** as **words**.

(

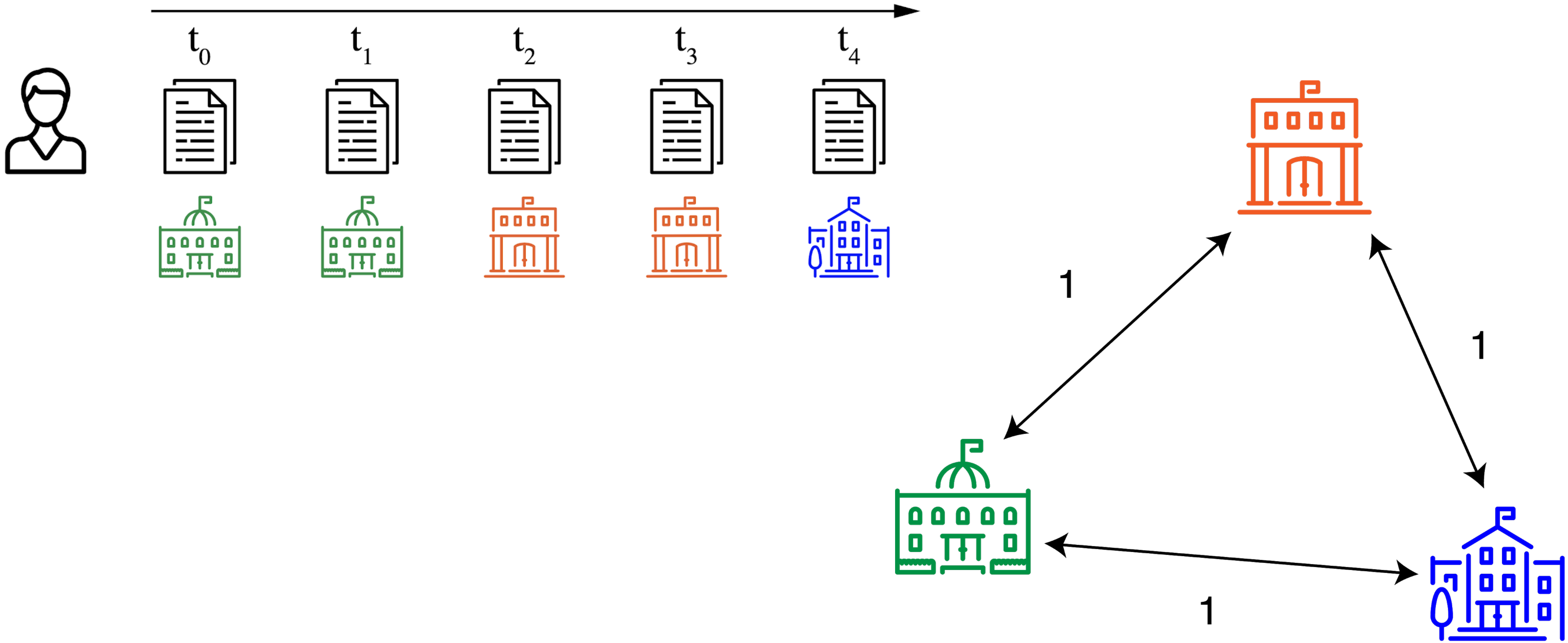


IC2S2 "embedding" session

)

Each a trajectory of organizations

Derive "flux" from scientists career trajectories



Aside: why does it work?

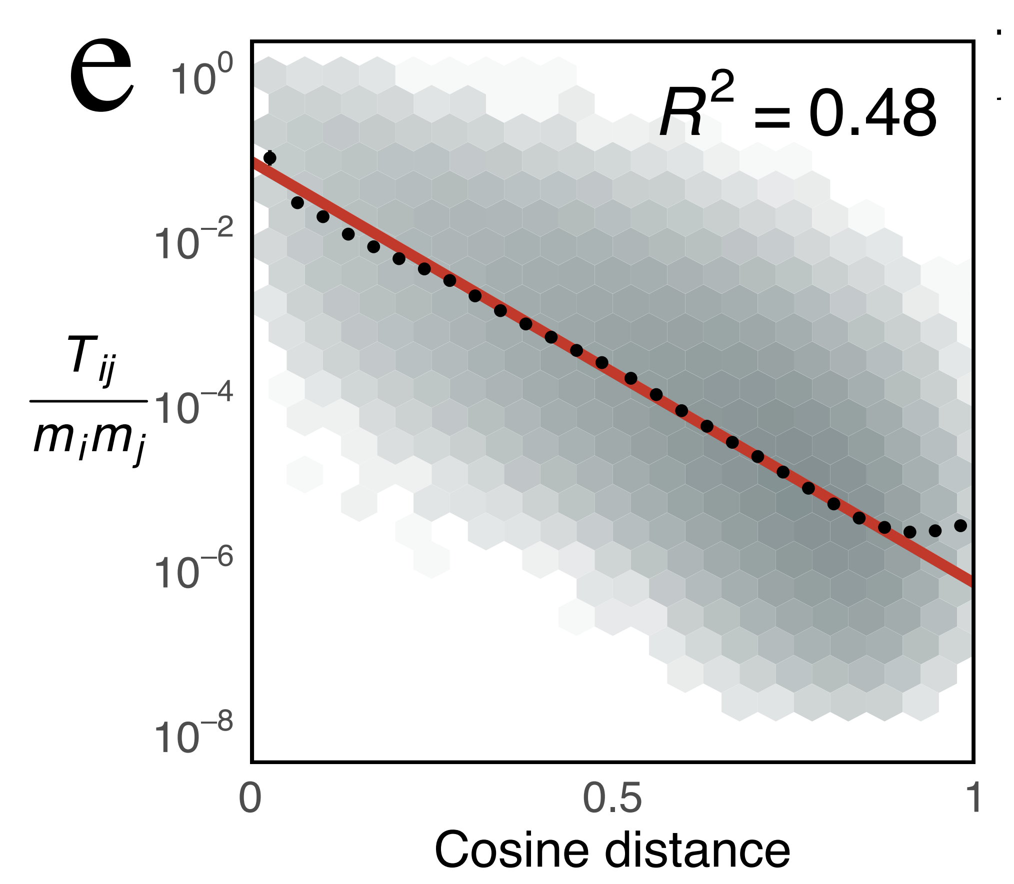
Relationship between word2vec and the Gravity Law

- Distance in the embedding space maps well to the gravity model
- Both correlation and prediction
- Word2vec finds a representation to predict adjacent words, but the gravity law emerges
- Deep connection between them?
- Preliminary work, but ideas welcome!

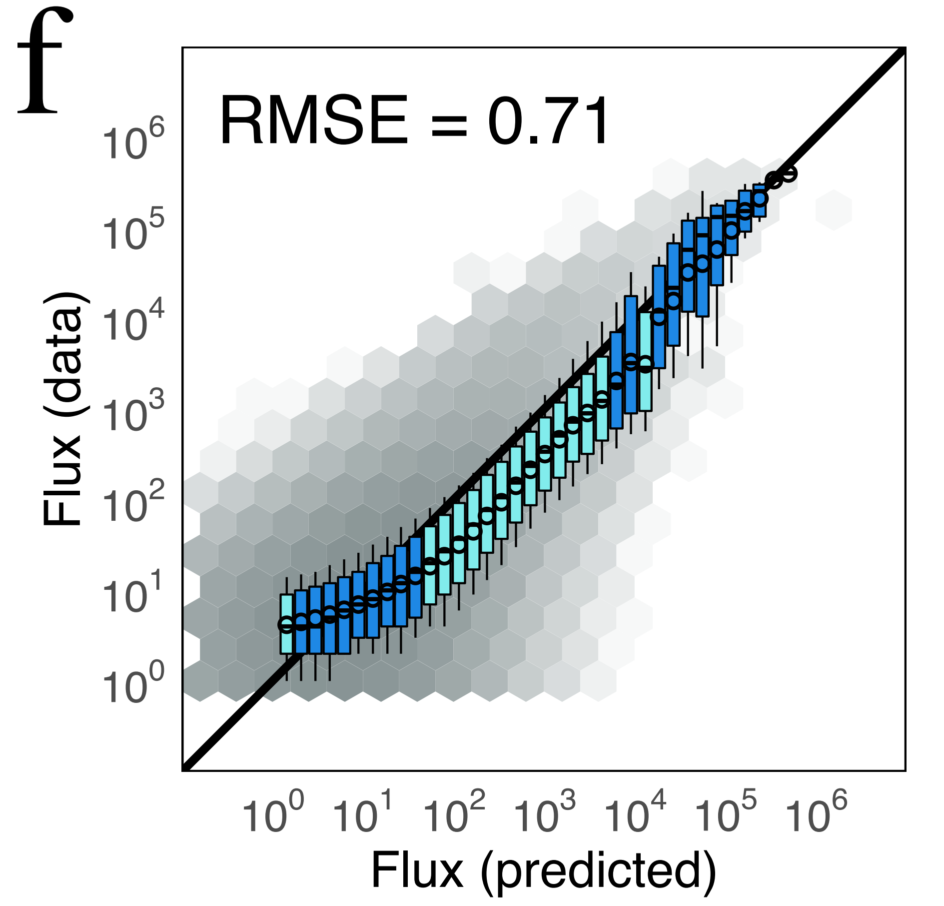
$$T_{ij} = C m_i m_j f(r_{ij})$$

$$\frac{T_{ij}}{m_i m_j} = f(r_{ij})$$

Correlation



Predicted vs. Actual



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

Gravity model

$$\hat{T}_{ij} = C m_i m_j f(r_{ij}) \quad (1)$$

word2vec

$$p(j \mid i) = \frac{\exp(\boldsymbol{u}_j \cdot \boldsymbol{v}_i)}{Z_i} \quad (2)$$

Window size = 1

$$\hat{T}_{ij} = s_i P(j \mid i) = \frac{s_i \exp(\boldsymbol{u}_j \cdot \boldsymbol{v}_i)}{Z_i} \quad (3)$$

Taking into account the effect from the negative sampling

$$P(j \mid i) := \frac{p_0(j) \exp(\boldsymbol{u}_j \cdot \boldsymbol{v}_i)}{Z'_i} \quad (4)$$

$$p_0(j) = s_j / \sum_{\ell'} s_{\ell'} \quad (5)$$

$$\boldsymbol{u}_i = \boldsymbol{v}_j \quad (6)$$

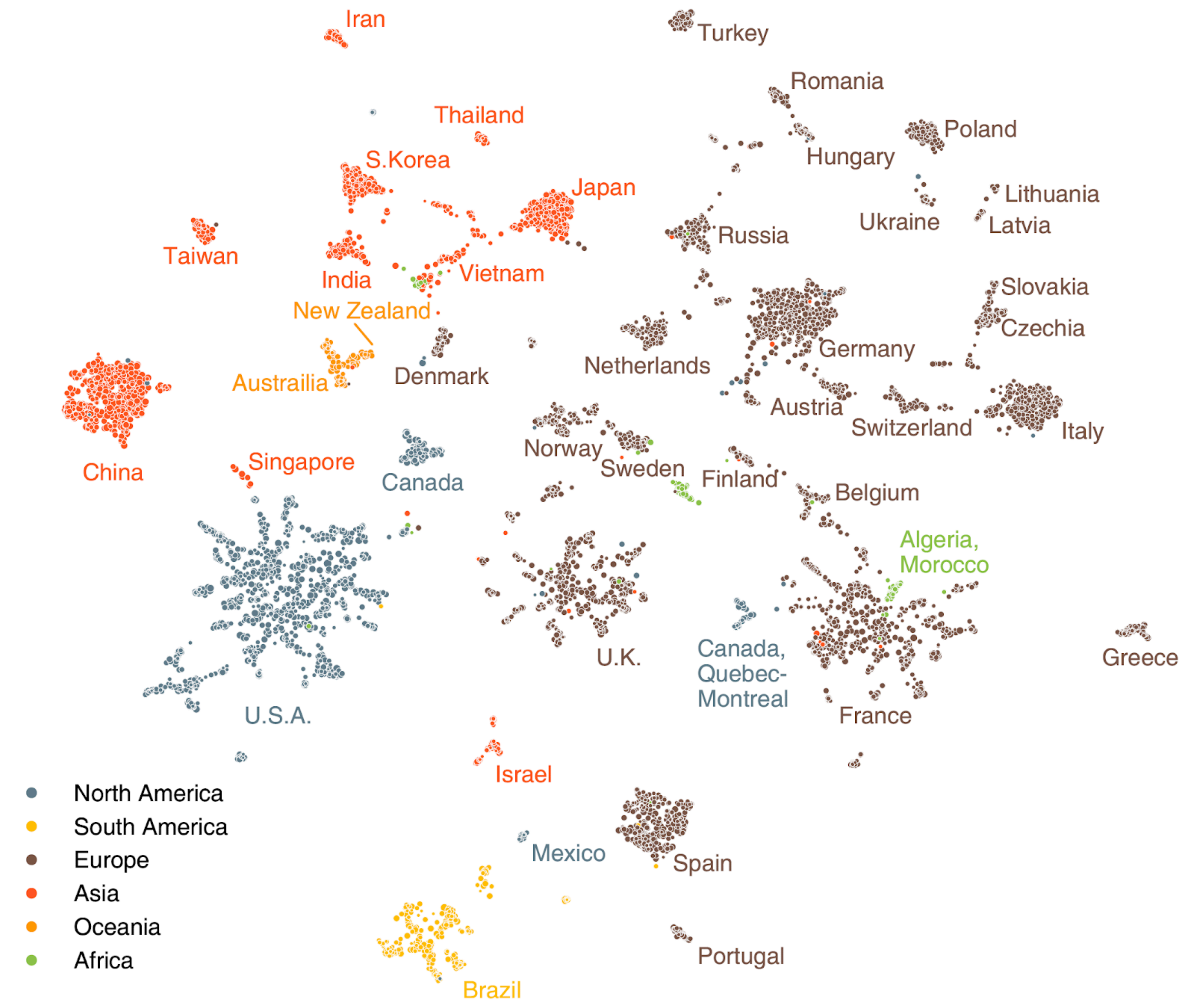
$$\hat{T}_{ij} = s_i P(j \mid i) = \frac{s_i s_j \exp(\boldsymbol{v}_j \cdot \boldsymbol{v}_i)}{Z'_i \sum_{\ell'} s_{\ell'}} \quad (7)$$

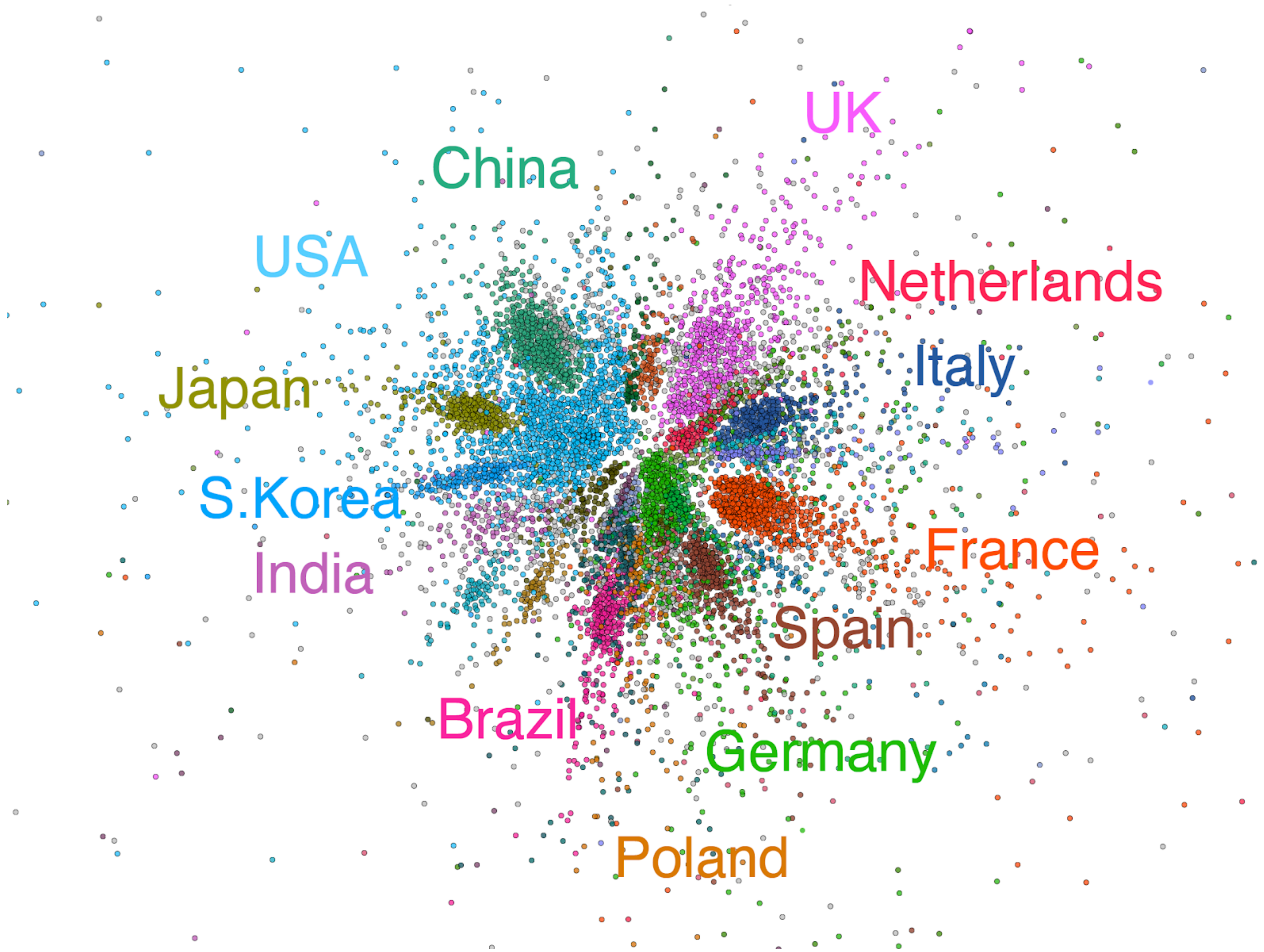
$$T_{ij} = T_{ji} \Rightarrow Z'_i = Z'_j, \quad (8)$$

$$\boxed{\hat{T}_{ij} = C s_i s_j \exp(\boldsymbol{v}_j \cdot \boldsymbol{v}_i)} \quad (9)$$

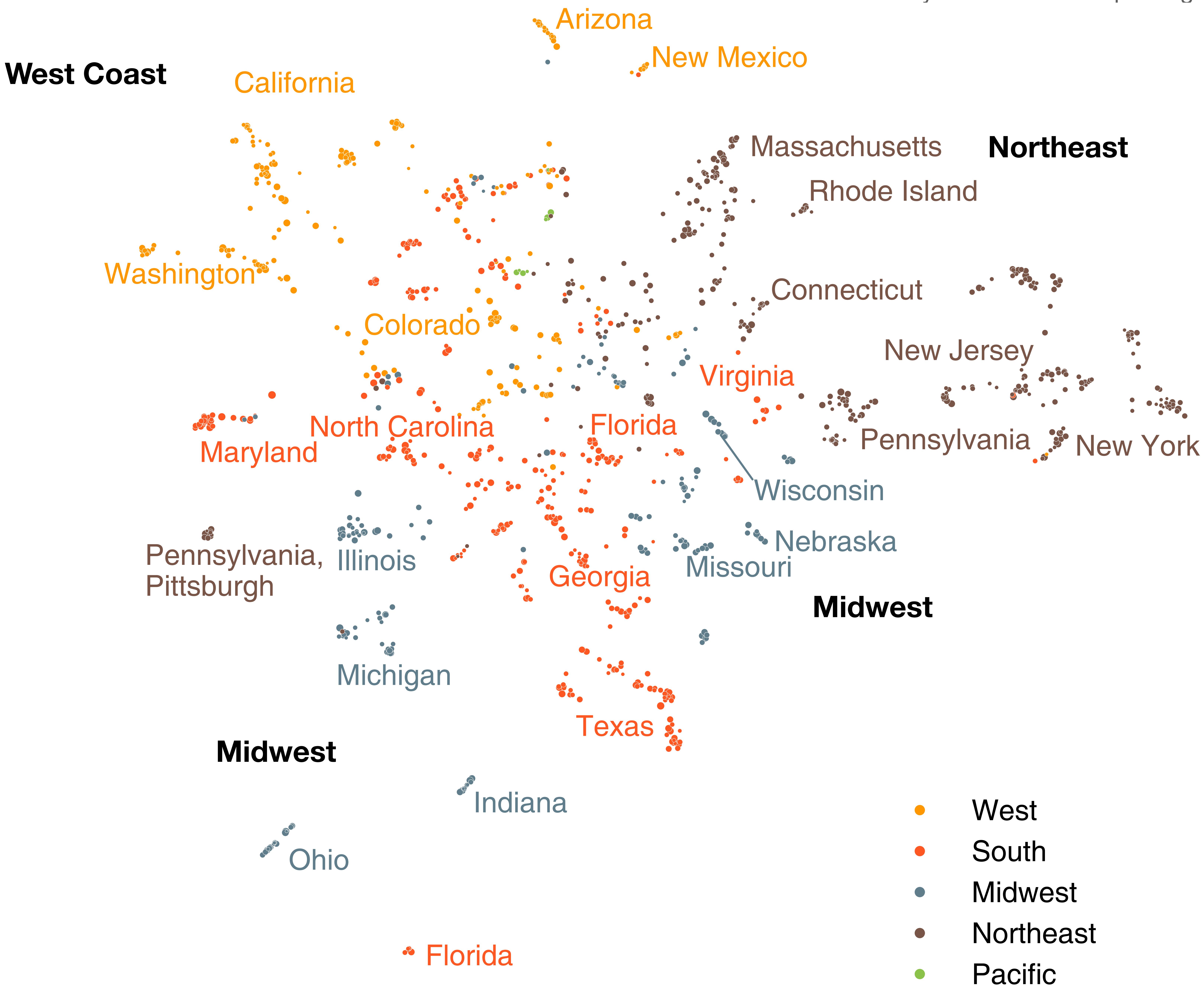
Why not networks?

- Structure obfuscated in visualizations
- Poor performance (reported in the paper)
- Many techniques, tuning required
- Missing edges
- Embeddings provide access to many interesting techniques





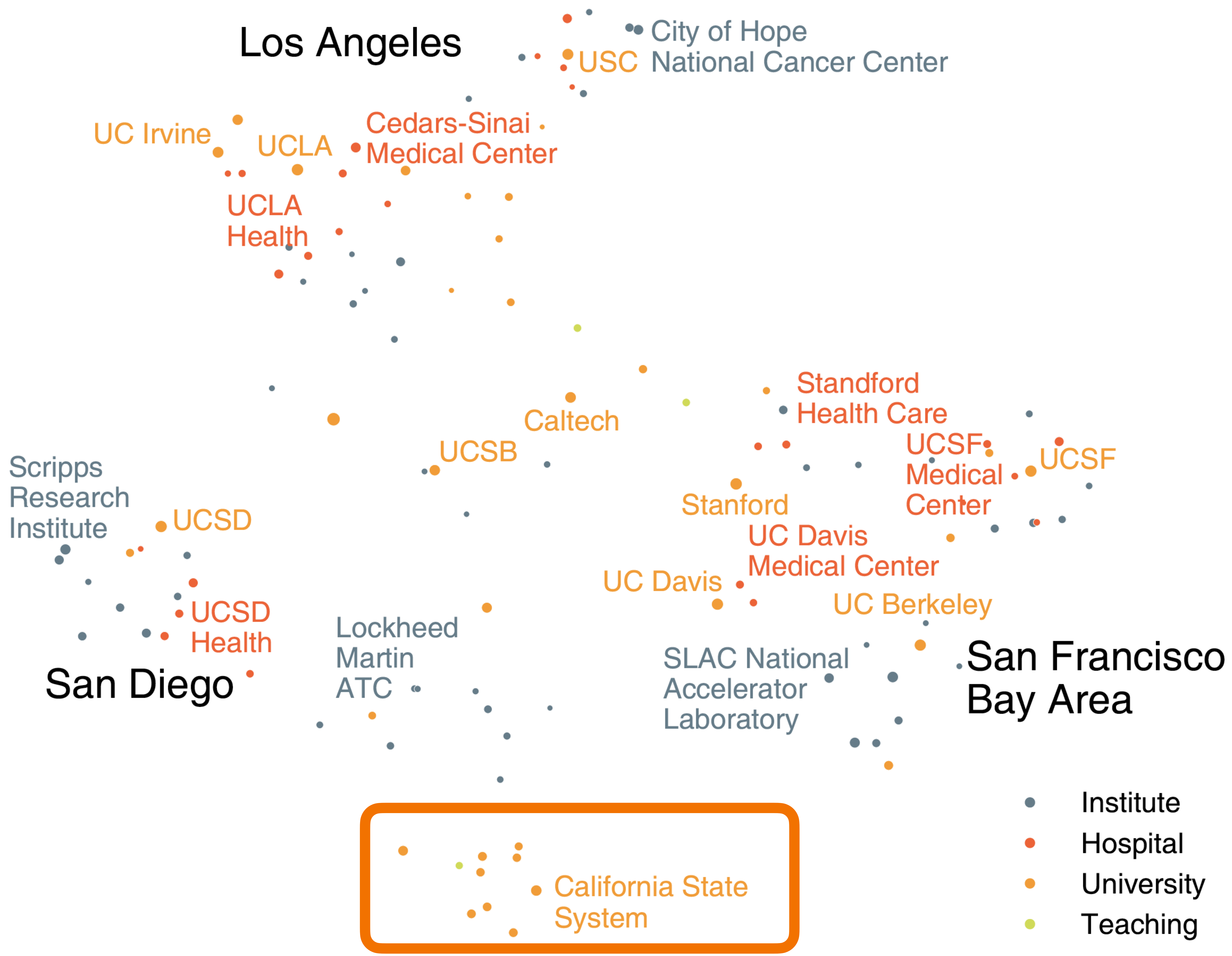
United States



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

3-major clusters based on urban center

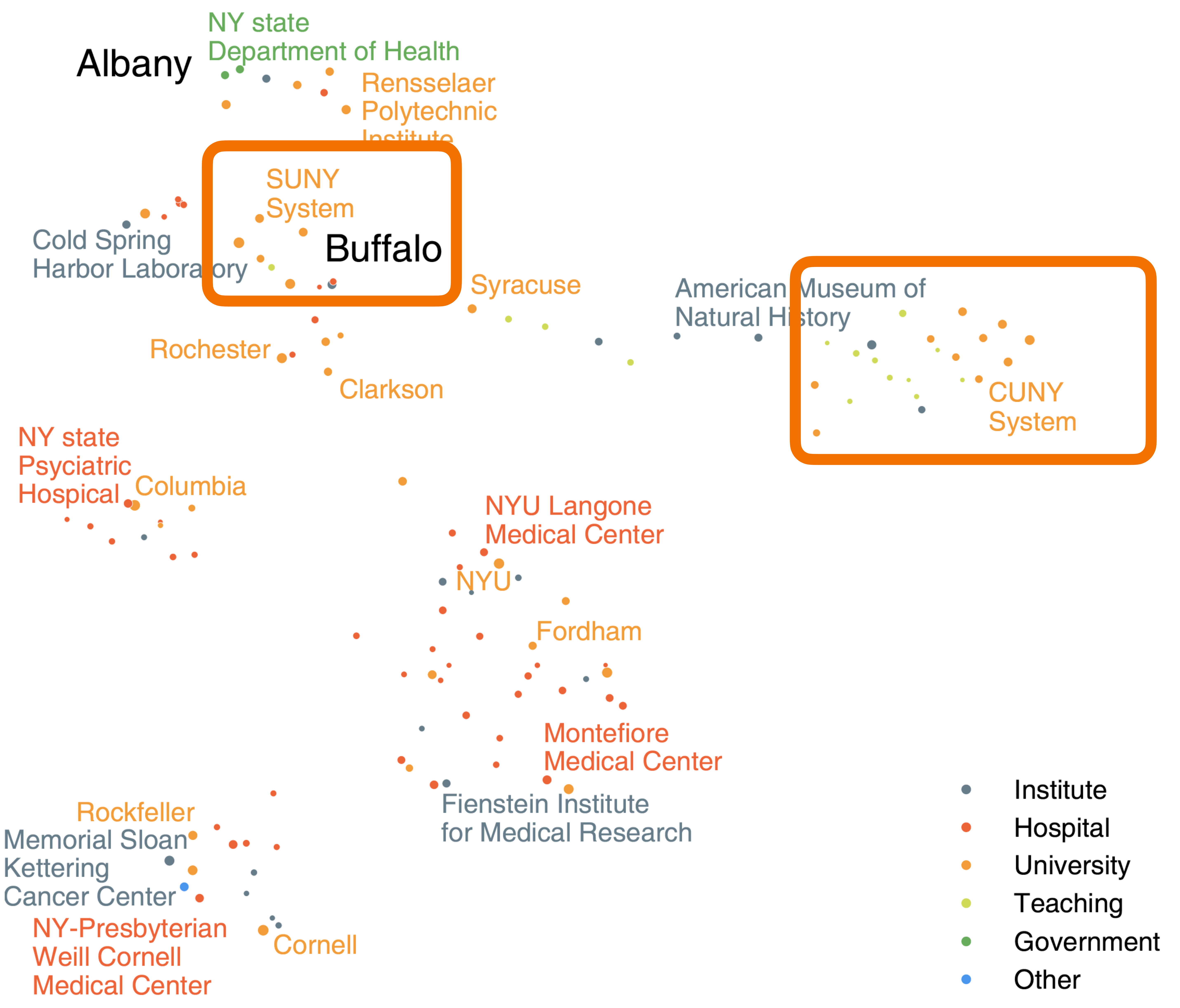
California



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

3-major clusters based on urban center

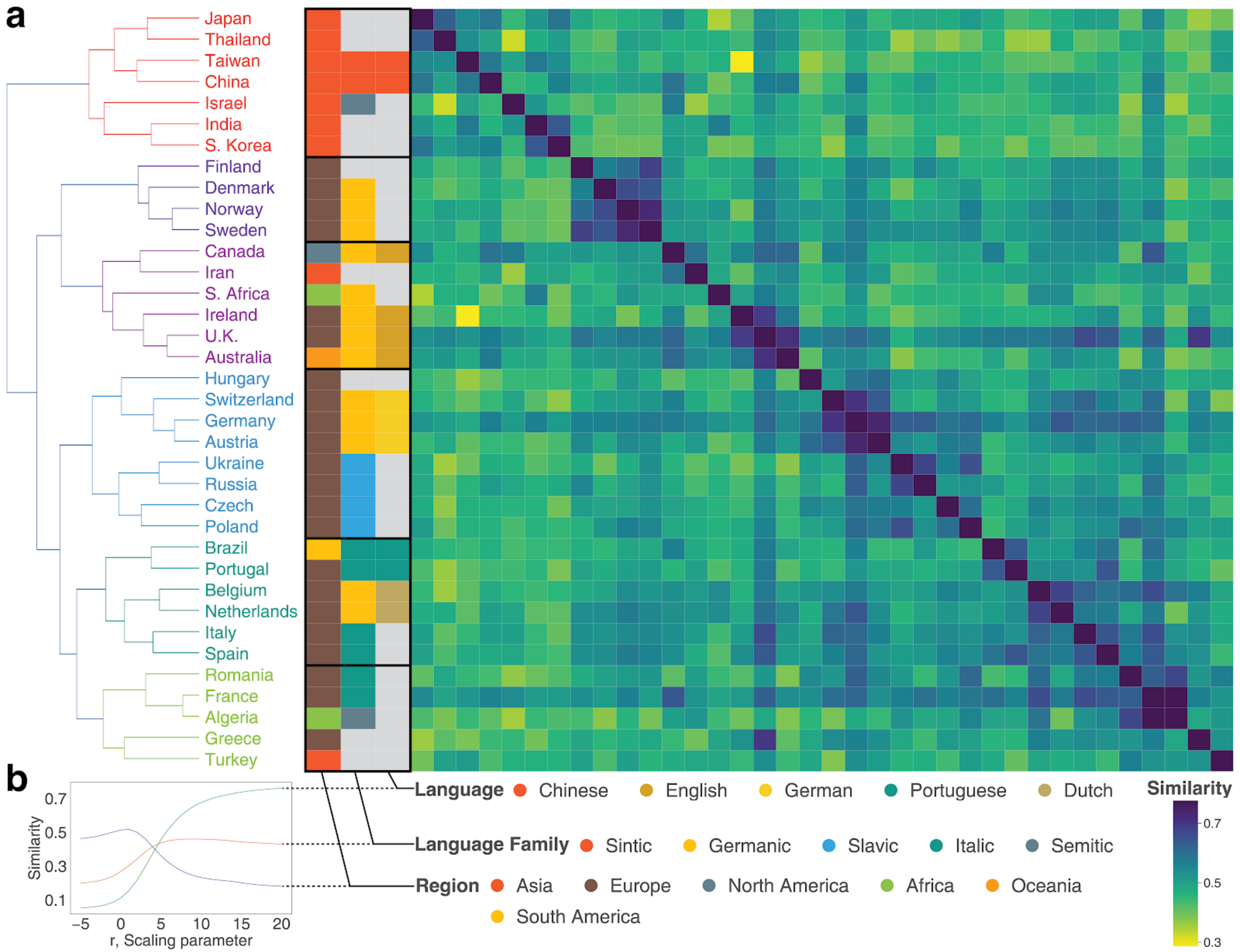
New York



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

dakotamurray.me/talk/2021-copenhagen/

Geography, then language, structure the vector space

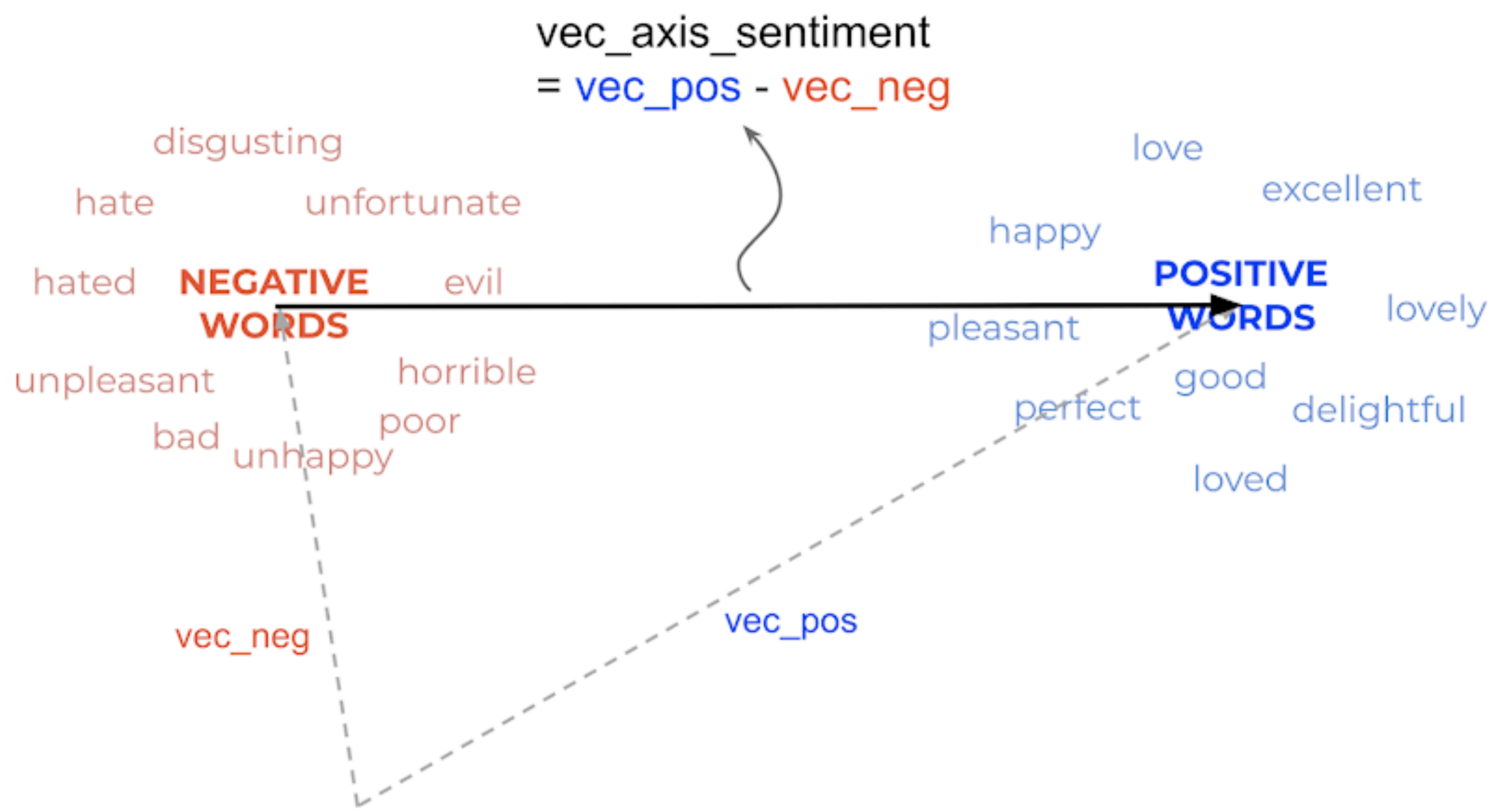


But also prestige…

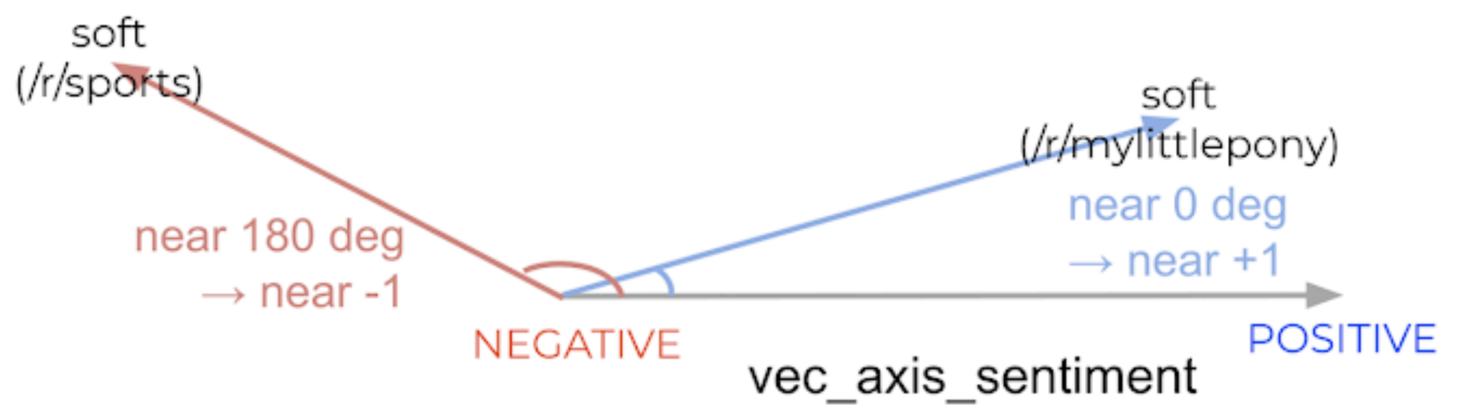
dakotamurray.me/talk/2021-copenhagen/

Leverage the semantic properties of the embedding

Step 1: Define a semantic axis

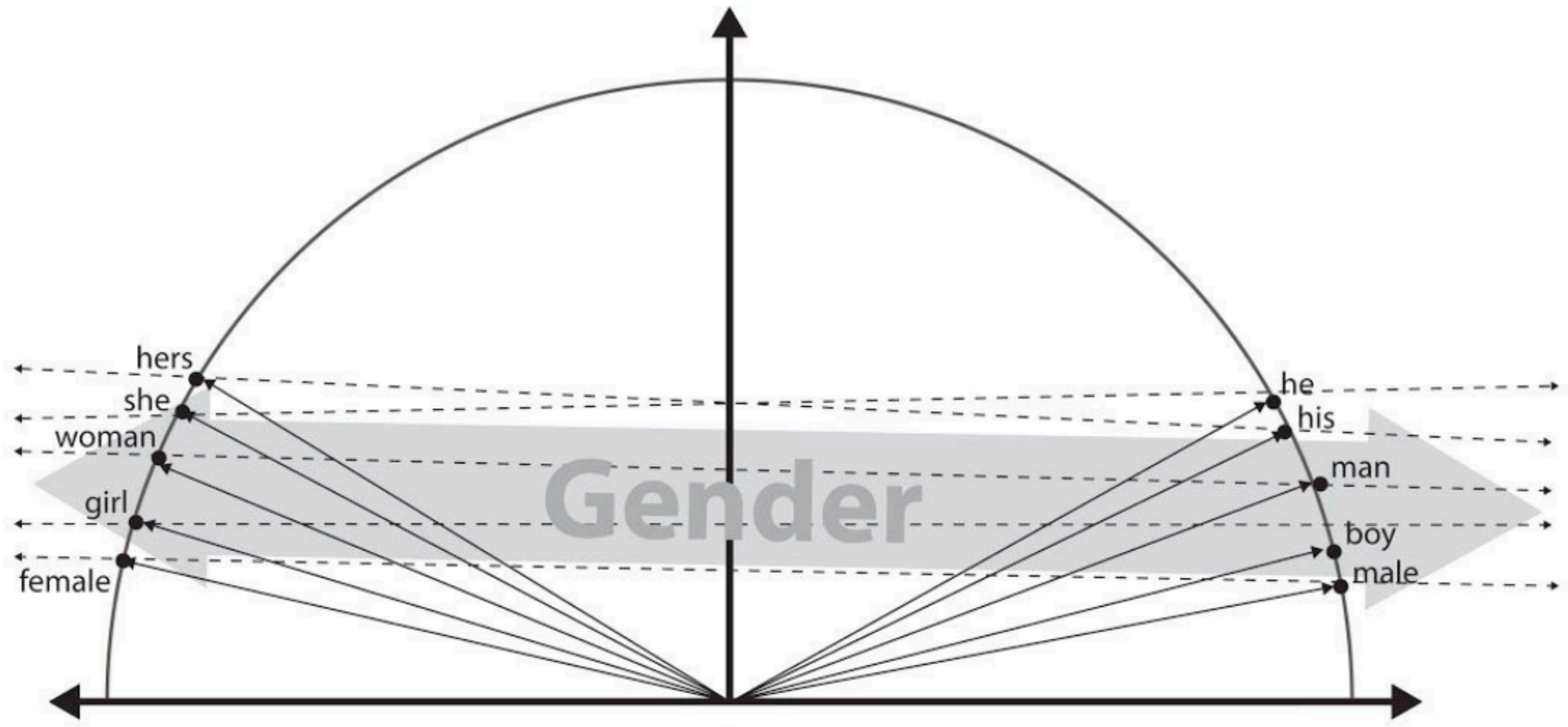


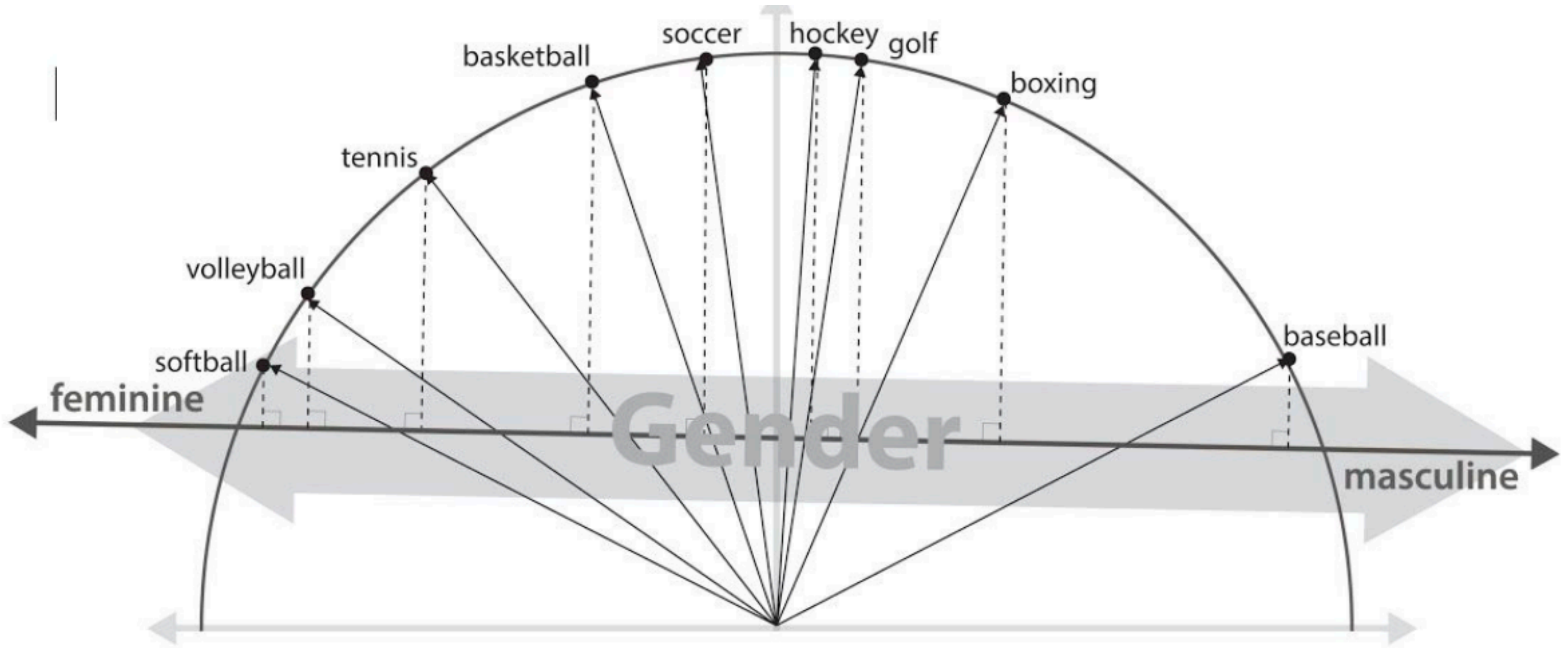
Step 2. Locate words on the semantic axis



Compute a cosine similarity between a word vector and axis vector

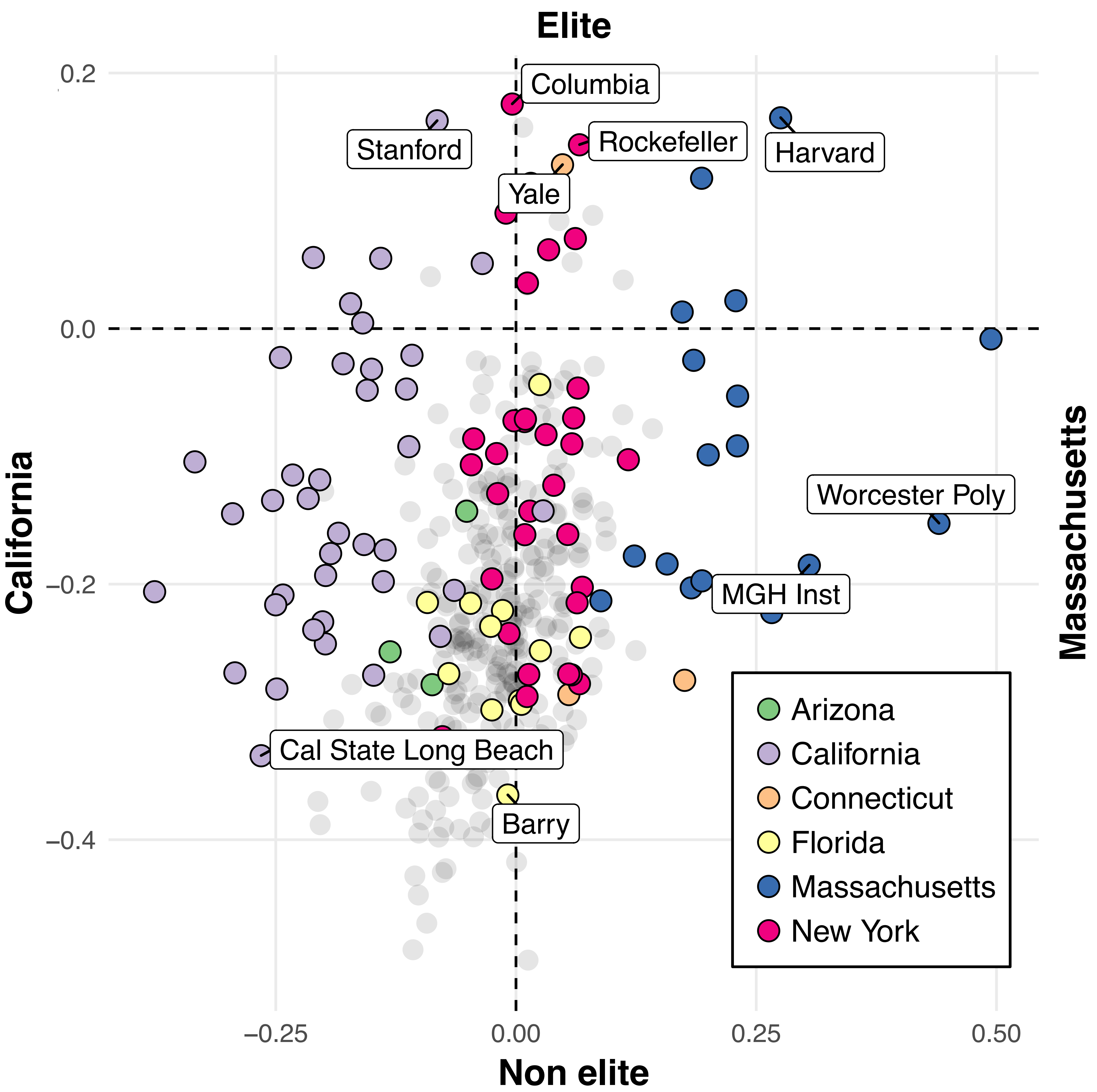
An, J., Kwak, H., & Ahn, Y.-Y. (2018). SemAxis: A Lightweight Framework to Characterize Domain-Specific Word Semantics Beyond Sentiment. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2450–2461.





Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*, *84*(5), 905–949.

SemAxis using Geography and Prestige



SemAxis reconstructs university prestige