Inferential clustering reveals administrative boundaries in Austrian migration networks

NetSci 2025, Maastricht, 5th June 2025

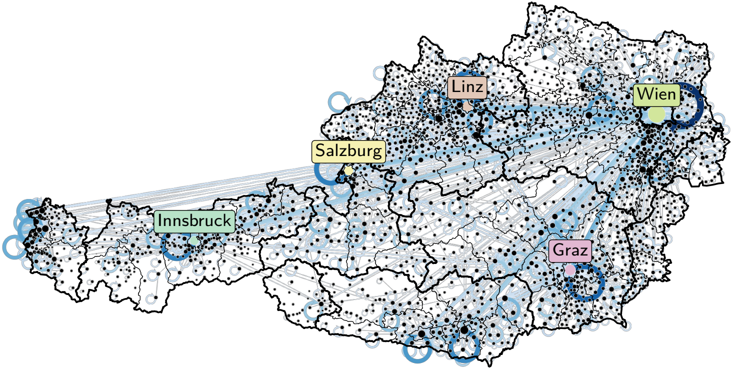
Martina Contisciani

Central European University, Vienna, AT

with Thomas Robiglio, Márton Karsai, Tiago P. Peixoto

Austrian Internal Migration Network[1]

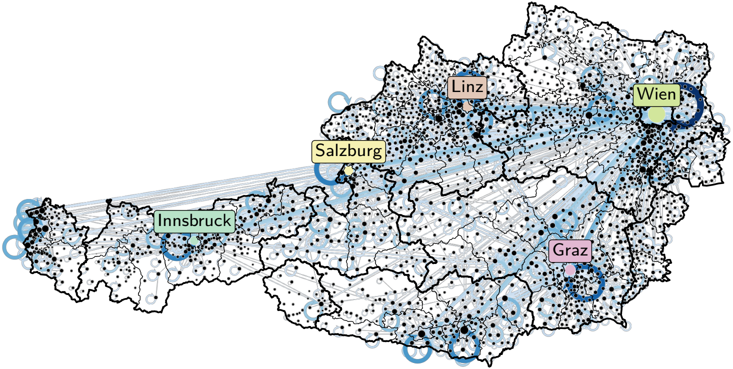
- Node i: municipality ($N = 2093$)
- *Directed and weighted* edge x_{ij}: relocations ($E \sim 70K$)
- Years 2002-2021, aggregated annually



[1] https://data.statistik.gv.at/

Austrian Internal Migration Network[1]

- Node i: municipality ($N = 2093$)
- *Directed and weighted* edge x_{ij}: relocations ($E \sim 70K$)
- Years 2002-2021, aggregated annually



We analyse twenty distinct networks that capture migration flows for each year. The results in this presentation refer to the year 2013.

[1] https://data.statistik.gv.at/

Gravity Model

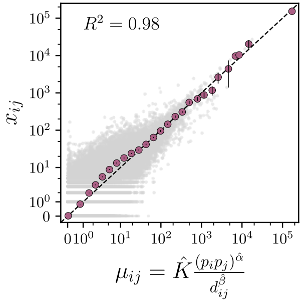
The rate of movement (I_{ij}) between two locations tends to increase with the product of their population densities (p_i, p_j), and to decay with their distance (d_{ij}):

$$\mathbb{E}[I_{ij}] := \mu_{ij} = K \frac{\left(p_i p_j\right)^{\alpha}}{d_{ij}^{\beta}}$$

Gravity Model

The rate of movement (I_{ij}) between two locations tends to increase with the product of their population densities (p_i, p_j), and to decay with their distance (d_{ij}):

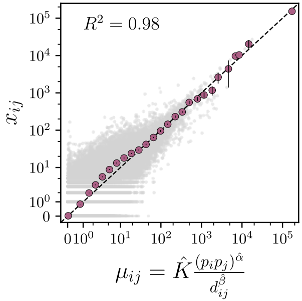
$$\mathbb{E}[I_{ij}] := \mu_{ij} = K \frac{(p_i p_j)^{\alpha}}{d_{ij}^{\beta}}$$



Gravity Model

The rate of movement (I_{ij}) between two locations tends to increase with the product of their population densities (p_i, p_j), and to decay with their distance (d_{ij}):

$$\mathbb{E}[I_{ij}] := \mu_{ij} = K\frac{\left(p_i p_j\right)^{\alpha}}{d_{ij}^{\beta}}$$



But, hidden discrepancies in relation to geographical and urban-rural information.

Weighted Stochastic Block Model[2]

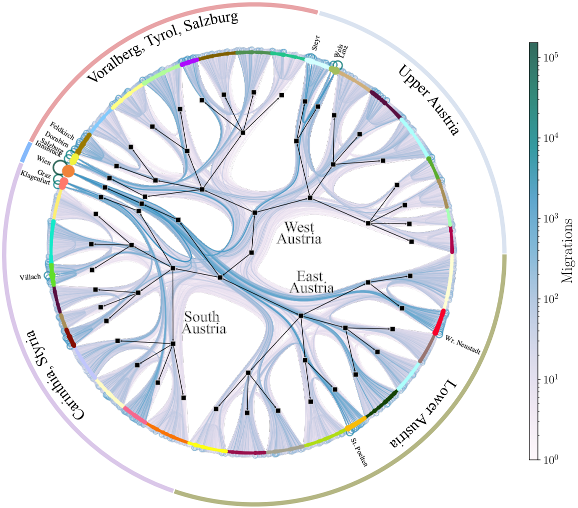
Given a partition $\mathbf{b}$ of the municipalities into B groups, the migrations between two locations are sampled only according to their group memberships:

$$P(\mathbf{x} \mid \theta, \mathbf{b}) = \prod_{ij} P(x_{ij} \mid \theta_{b_i, b_j})$$

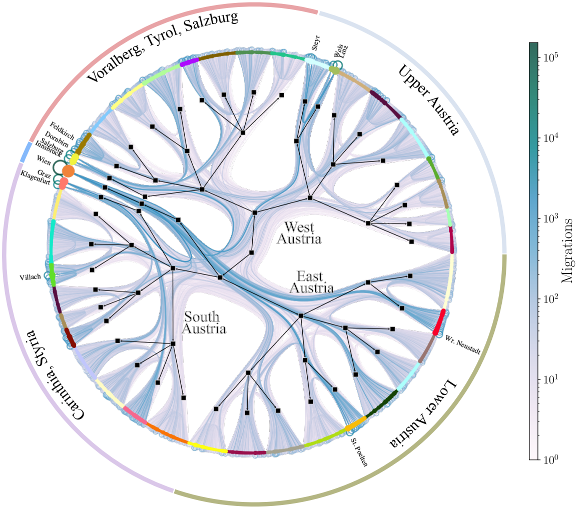
- $P(x_{ij} \mid \theta_{b_i, b_j})$ is a kernel distribution conditioned only on the groups
- Number of groups B inferred from data
- Hierarchical partition

[2] T. P. Peixoto, Physical Review E 97, 012306 (2018)

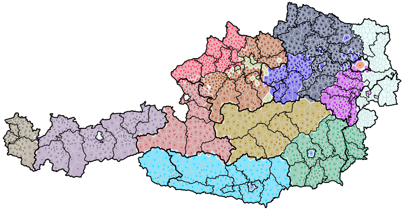
Inferred Hierarchical Partition



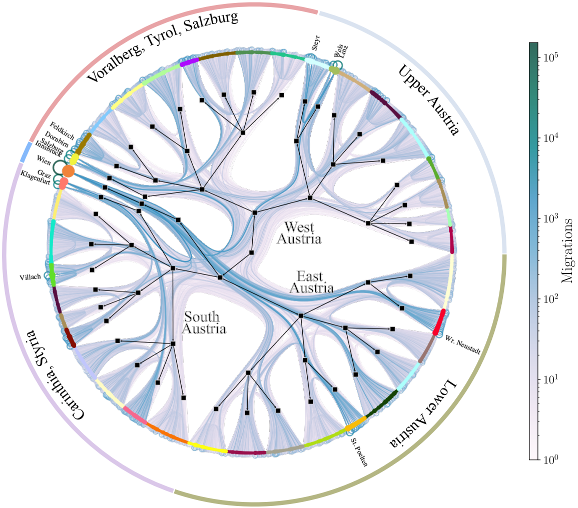
Inferred Hierarchical Partition

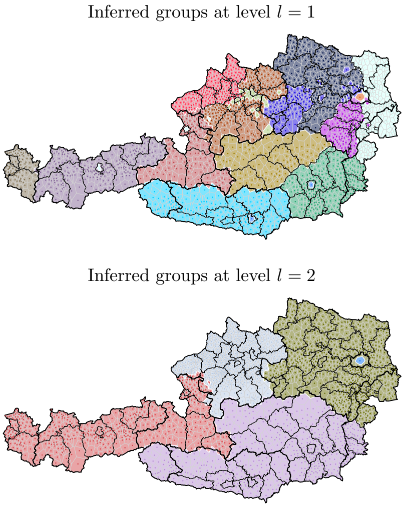


Inferred groups at level $l = 1$



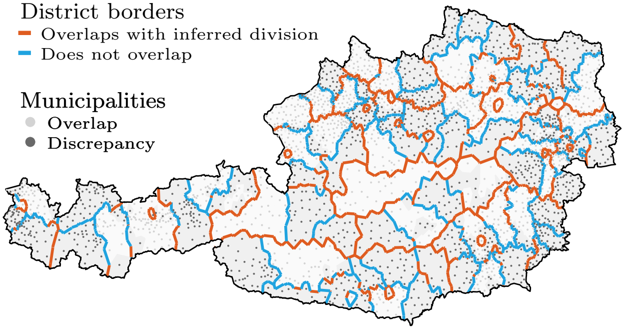
Inferred Hierarchical Partition





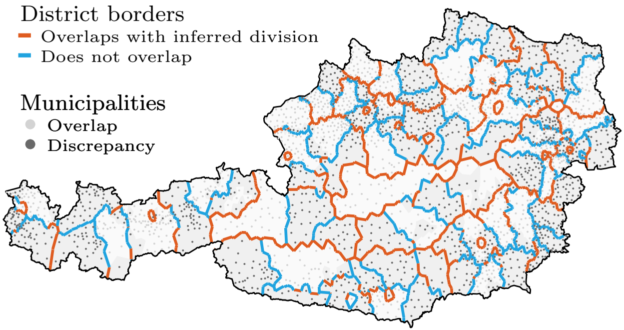
Administrative Boundaries

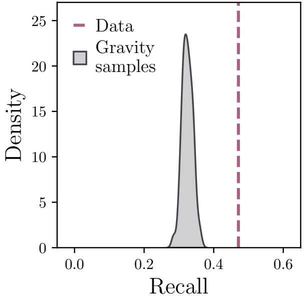
Around 47% of the district borders coincide exactly with the boundaries between the inferred groups, and the same holds for $\sim 72\%$ of the federal state boundaries.



Administrative Boundaries

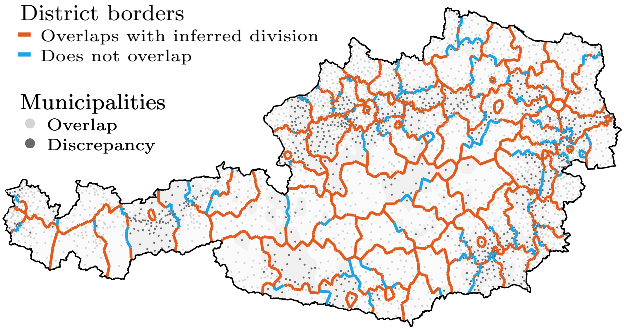
Around 47% of the district borders coincide exactly with the boundaries between the inferred groups, and the same holds for $\sim 72\%$ of the federal state boundaries.

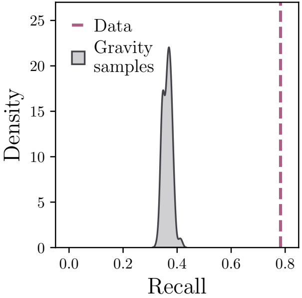




Administrative Boundaries in Binary Network

District-level effects become more visible when the magnitudes are excluded, and the match between district borders and inferred boundaries reaches 78%.





Main Takeaways

- Migration flows in Austria are driven by **more than gravity**
- Inferential clustering reveals effects of:
  - ◇ **administrative boundaries**
  - ◇ **urban-rural divide**
- Patterns consistent over twenty years

Main Takeaways

- Migration flows in Austria are driven by **more than gravity**
- Inferential clustering reveals effects of:
  - **administrative boundaries**
  - **urban-rural divide**
- Patterns consistent over twenty years

Next step of the MOMA project: provide explanations of the observed patterns

Thank you!

Stay tuned... Soon on arXiv!

contiscianim@ceu.edu
mcontisc.github.io

Thomas Robiglio

Márton Karsai

Tiago P. Peixoto

Gravity Model

The migration flows between two locations are modelled as Poisson-distributed random variables

$$I_{ij} \sim \mathrm{Pois}(\mu_{ij})$$

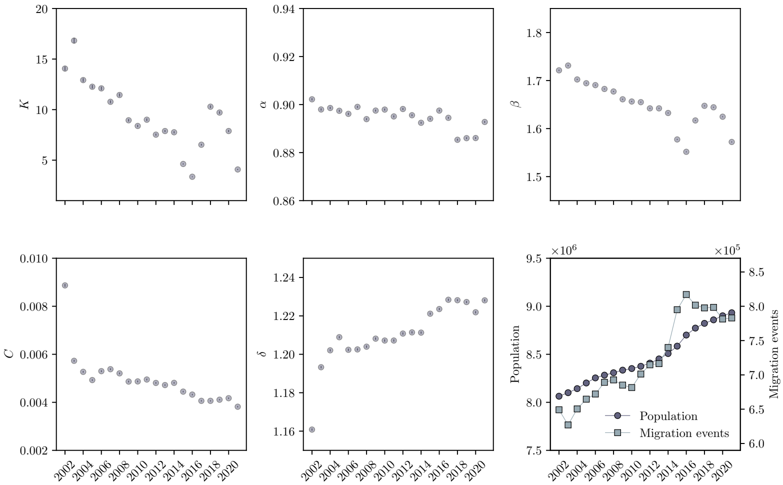
with

$$I_{ij} = \begin{cases} x_{ij} + x_{ji} & \text{if } i \neq j \\ x_{ii} & \text{if } i = j \end{cases} \quad \text{and} \quad \mu_{ij} = \begin{cases} K \frac{(p_i p_j)^\alpha}{d_{ij}^\beta} & \text{if } i \neq j \\ C p_i^\delta & \text{if } i = j \end{cases}.$$

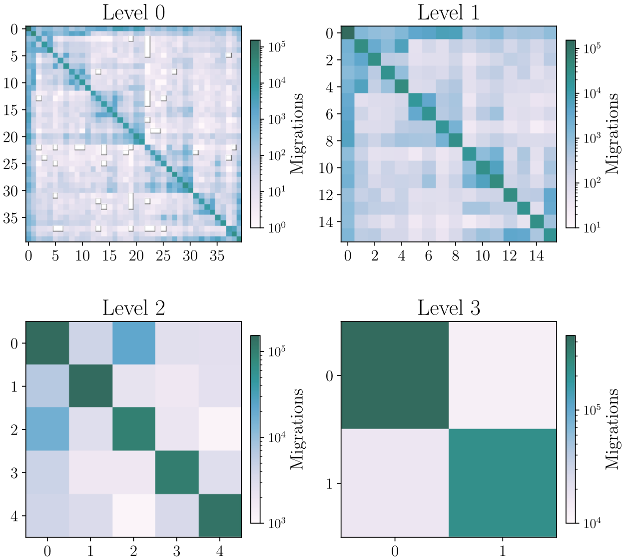
To generate directed synthetic networks, we sample the edge weights $\hat{x}_{ij}$ from the estimated Poisson gravity model rates as

$$\hat{x}_{ij} \sim \begin{cases} \mathrm{Pois}(\mu_{ij}/2) & \text{if } i \neq j \\ \mathrm{Pois}(\mu_{ii}) & \text{if } i = j \end{cases}.$$

Inferred Parameters Gravity Model

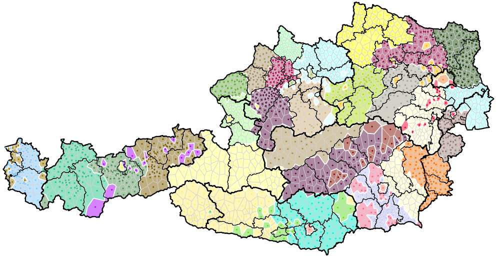


Inferred Affinity Matrices

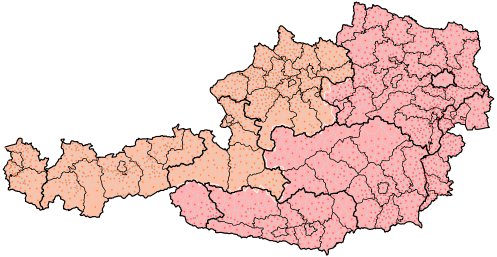


Inferred Partitions

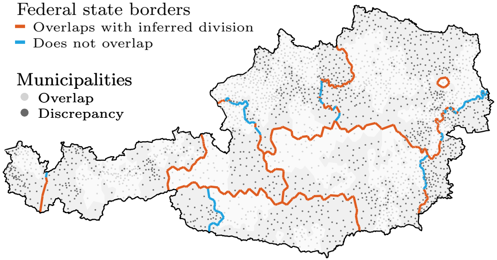
Inferred groups at level $l = 0$

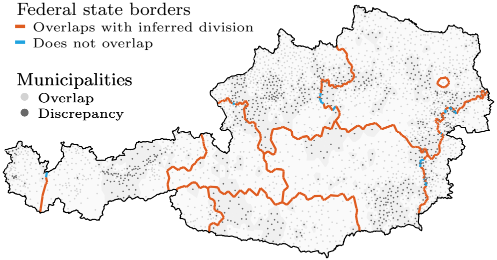


Inferred groups at level $l = 3$

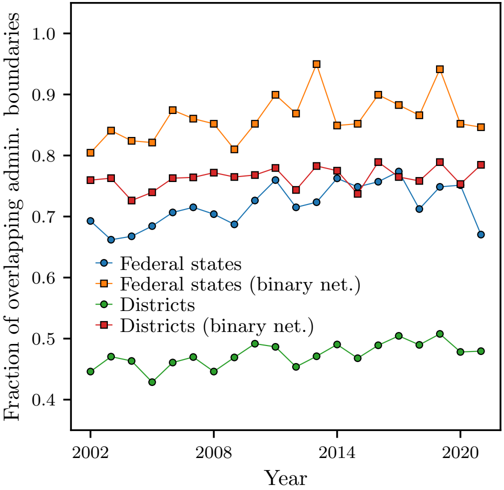


Federal State Boundaries

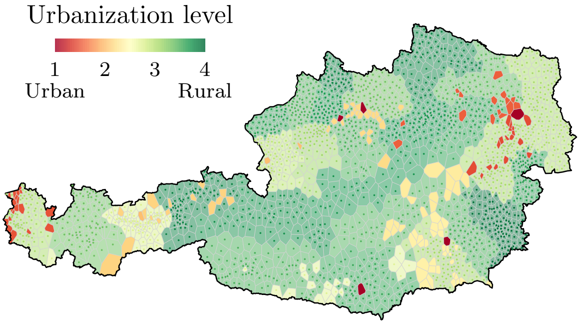


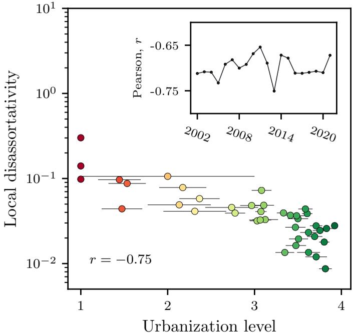


Administrative Boundaries Over Time



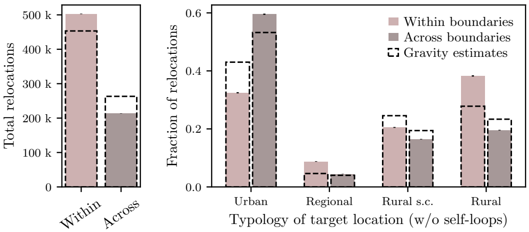
Urban-Rural Classification



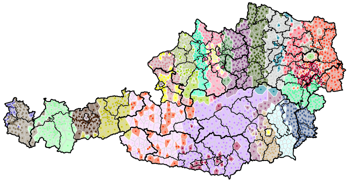


Additional Results

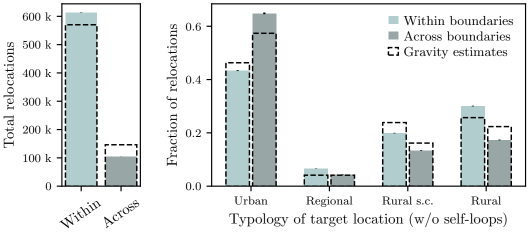
(a) Migration volumes in relation to districts



(b) Inferred groups from a gravity model sample



(c) Migration volumes in relation to federal states



(d) Comparison with gravity model samples