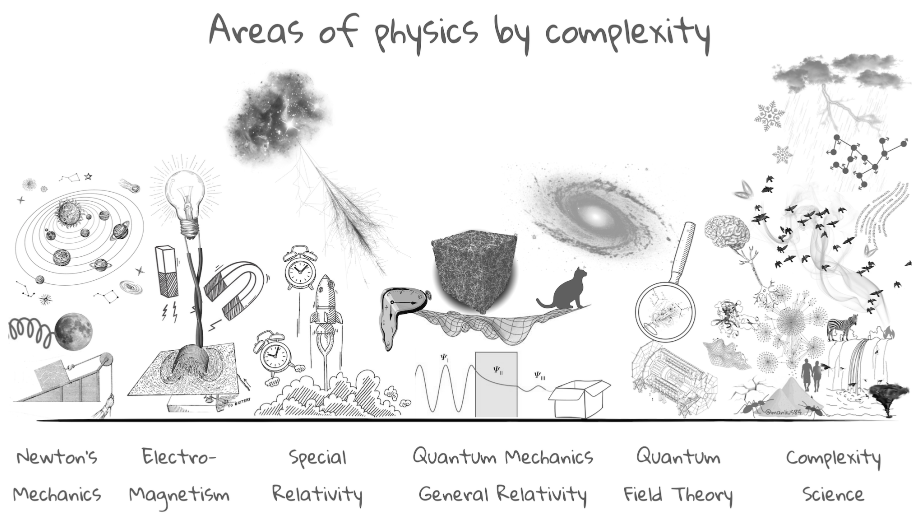
Università degli Studi di Padova

Final Report

Physics of Complex Networks: Structure and Dynamics



Projects Report: # 34, # 41

Quaglio Emanuele

Last update: February 13, 2026

Contents

1 | How heterogeneity affects generosity

A game-theory on networks perspective

Task leader(s): *Quaglio Emanuele*

1.1 | Ultimatum Game with co-evolving strategies on a graph

Simulating evolutionary game dynamics on networks [6] requires defining local payoffs as functions of the node strategies, and update rules for these strategies as functions of the payoffs. In this work I study how *generosity* and *cooperation* are affected by heterogeneity both in the network topology and in the population of update rules.

In particular, in the Ultimatum Game (UG) on a network, each agent i carries two continuous strategy parameters $(p_i, q_i) \in [0,1]^2$, respectively offer and acceptance threshold.

Playing in both directions, offers are accepted if $p_i \geq q_j$. Payoffs are then accumulated over neighbors: $\Pi_i = \sum_{j\in\Gamma_i} \Big[\mathbf{1}(p_i \geq q_j)(1-p_i) + \mathbf{1}(p_j \geq q_i)p_j\Big]$, where Γ_i is the neighbor set of node i, following [7].

On top of strategy evolution, mimicking the method applied by [4] on a different game, I study heterogeneous evolutionary update rules: nodes may update their strategy by different rules, and the update rule itself can spread/transform.

The rules from [4] are:

- **Replicator (REP):** pick a random neighbor j. If $\Pi_j > \Pi_i$, copy $(p_j, q_j, \text{rule}_j)$ with probability $\Pr(i \leftarrow j) = \frac{\Pi_j - \Pi_i}{2\max(k_i,k_j)}$.
- **Unconditional Imitation (UI):** if the neighbor with maximum payoff exceeds Π_i, copy that neighbor's (p, q, rule).
- **Moran-like (MOR):** copy (p, q, rule) from a neighbor sampled with probability proportional to its payoff.

Additionally, [7] applies **Social-Penalty (SP):** re-randomize the closed neighborhood of the node that has globally minimal payoff.

To study the effect of degree heterogeneity, a good choice is the parametric model by Gómez-Gardeñes–Moreno [5], used by [4], which interpolates between scale-free (BA-like) behavior and homogeneous (ER-like) behavior with a parameter $\alpha \in [0,1]$ proportional to homogeneity.

1.2 | Methods and Results

Synthetic network generation. I implement the model proposed by [5] in the following way:

1. Start from an empty graph on N nodes.
2. Create an initial clique of size m_0.
3. For each new node $i = m_0 + 1, \ldots, N$, add m edges and:
    - with probability α: attach to a uniformly random node,
    - with probability $1 - \alpha$: attach preferentially $\propto k$.
4. During (3), enforce a simple graph (no self-loops / multi-edges).

Ultimatum Game implementation. Since convergence of the evolutionary dynamics could require a large number of generations, maximum speed is obtained by employing `Rcpp` instead of plain `R`. The UG round-robins are implemented by looping over the edges, computing an outcome in both directions.

Co-evolution of update rules. In the explored settings, update rules are expected to change the state of a given node as a (non-linear, possibly stochastic) function of its neighborhood; thus, looping over nodes and quickly accessing their connections becomes necessary. For this reason I decide to implement the co-evolution step in `Rcpp` using CSR format for the adjacency matrix. Initialization is uniformly random from a set of rules.

Co-evolutionary Social-Penalty variant. I modify the SP implementation in [7] to suit the co-evolutionary settings of [4]. Indeed, I want the update of a node to only be influenced by the update rule of the node itself. In standard SP, all neighbors are affected[3]; instead, my variant:

1. Finds the node $i_{\min}$ with minimum payoff $\min_i \Pi_i$.
2. For nodes in $\Gamma[i_{\min}] = \{i_{\min}\} \cup \Gamma(i_{\min})$ **that currently have rule SP**: re-randomizes (p, q, rule) by uniform sampling.

Preliminary analysis. Networks with 100, 1000, 10000, and 100000 nodes are tested to verify that even smaller sizes manifest a qualitatively similar behaviour to their larger counterparts. This permits using reduced sizes to check simulations up to 10^8 generations for the SP update rule, characterized by slow convergence to the limiting distribution. This allows me to choose size and time-length reasonably low, though still sufficient to qualitatively reproduce the main findings of [7] and [4].

Combinatorial analysis. I explore the (α, rule) space cumulating statistics over 100 instances per configuration. The following quantities are stored:

- degrees k_i;
- p_i, q_i, rule_i at multiple generations over the time-length of the simulation;

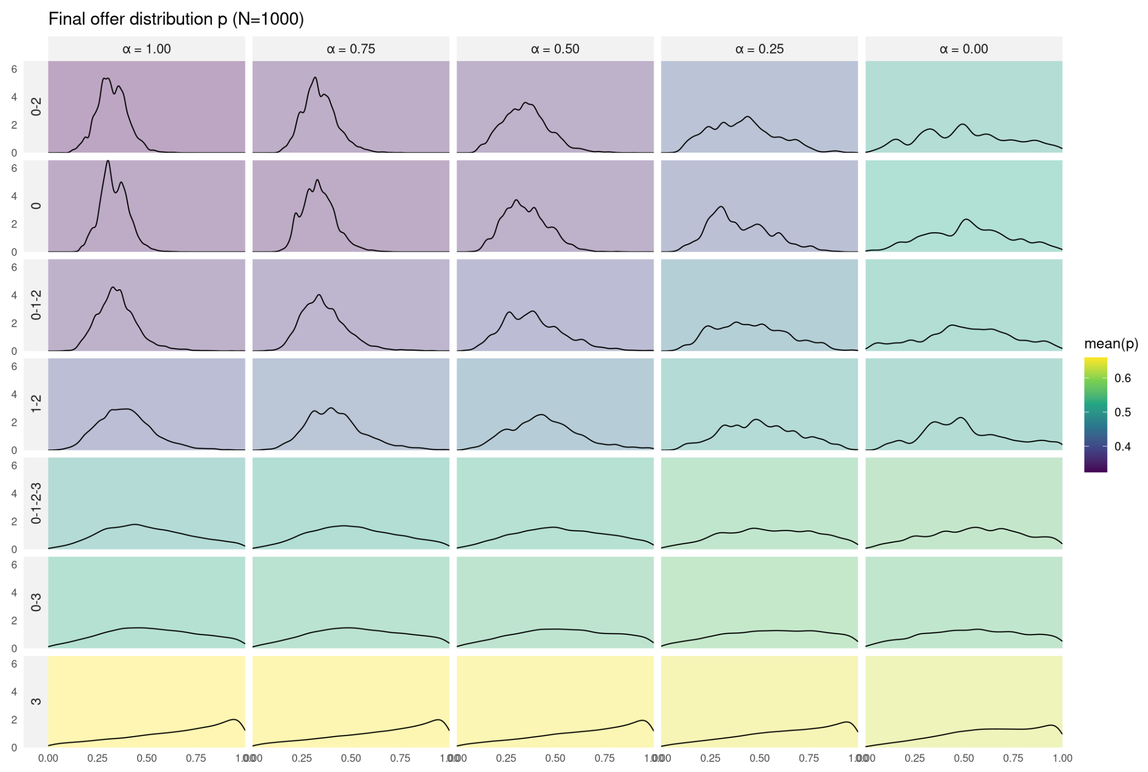


Figure 1.1: **Effect of heterogeneity on generosity:** distributions of offer p over 100 instances of networks of size $N = 1000$, under different degree heterogeneity $1-\alpha$, and different initial populations of update-rules (0: REP, 1: UI, 2: MOR, 3: SP).

- rule–rule edge endpoints at the final generation;

and the following summaries are computed:

- mean final offer per degree $\mathrm{E}[p|k]$;
- mean and standard deviation of p per rule over time $\mathrm{E}[p|\texttt{rule}]_t$, $\sqrt{\mathrm{Var}[p|\texttt{rule}]}_t$;
- rule abundance fractions over time $\mathrm{P}(rule)_t$;
- neighbor-rule joint distribution $\mathrm{P}(\mathrm{rule}_i, \mathrm{rule}_j)$ over edges.

Results. The results confirm and extend the findings in [7] and [4] that:

- degree heterogeneity positively affects mean offer p;
- the Social-Penalty rule has the greatest effect in increasing offer;
- compared to the Replicator rule alone, heterogeneous mixtures of update rules increase offer even without SP;
- for all α and all rule sets without SP, strategies with $q_i > p_i$ are strongly disfavoured, while SP dynamics preserve very greedy, low-offer/high-acceptance-threshold strategies.

2 | Public transport in large cities worldwide

Task leader(s): *Quaglio Emanuele*

2.1 | Dataset and framework

The dataset at [1] report the development over time of public transportation networks (different modes and potentially different lines per mode) across a variety of cities around the world. I build for each city a node list (nodeID, nodeLabel, latitude, longitude, mode, year) and and edges list (nodeID_from, nodeID_to, mode, line, year). The data is essentially tabular, with entities Stations and Sections roughly corresponding to nodes and edges, plus additional tables for transportation lines, systems, and modes, and tables storing relationships between these entities. The data is geometric in nature: each station features a position, and each section features a polygonal chain. Position and chain are represented respectively as Point and LineString according to GeoJSON standard [2], based on `(longitude, latitude)` coordinates following WGS 84 system. The schemas of the tables are the following:

- stations: id, name, geometry, buildstart, opening, closure, city_id;
- sections: id, geometry, buildstart, opening, closure, length, city_id;
- cities: id, name, coords, start_year, url_name, country, country_state, length;
- lines: id, city_id, name, url_name, color, system_id, transport_mode_id;
- systems: id, city_id, name, historic, length; - transport_modes: id, name;
- section_lines: id, section_id, line_id, created_at, updated_at, city_id, fromyear, toyear, deprecated_line_group;
- station_lines: id, station_id, line_id, city_id, created_at, updated_at, fromyear, toyear, deprecated_line_group.

2.2 | Pipeline and preliminary analysis

2.2.1 From raw JSON to specified node/edge files.

Inspection and conversion. From an inspection of the JSON data files, I confirm they have columnar structure, so I decide to convert them into a more convenient tabular representation. I choose Parquet over CSV for type-preservation (since they are present in the original JSON) and higher speed / memory efficiency.

Graph construction. After filtering table rows by city, I define the desired fields for the nodes and edges files, skipping cities with no stations.

For the nodes:

- nodeLabel: use `name` field or (if not available) fallback to original ID;
- nodeID: sort by original IDs and remap to dense indexes;
- year: use `Stations.opening` of the station or (if not available) fallback to its earliest `station_lines.fromyear`;
- mode: in order not to lose or distort mode information, I preserve all the (sorted, unique) modes associated by joining `station_lines->lines->transport_modes`, concatenated. The idea is that for mixing analysis it'll be possible to convert it into a custom `multiple-modes` mode feature, to properly count edges that connect stations "heterogeneously". But at the same time, it will be still possible to do finer analysis (checking presence of a specific mode in the set obtained from splitting by `"|"` separator).

For the edges, I decide to **admit multigraphs** (and simplify them when necessary during analysis). Indeed, I identify them by both keys `section` and `line`. The motivation behind this choice is that I think it would be useful to be able to identify a station as a "hub" when multiple lines depart from it, even when they share the same physical section (e.g. shared rails).

- line: use `Lines.url_name` or fallback to `Lines.name`;
- year: use `section_lines.fromyear` or fallback to `Sections.opening`;
- `nodeID_from`, `nodeID_to`: associate each section endpoint to a node based on the geometric procedure described below; furthermore, sort endpoint IDs and drop self-loops.

Edge-to-nodes snapping procedure. Each endpoint of a `section` is snapped to the `station` at minimal Haversine distance. To avoid spurious links due to noisy data, a section is kept only as long as said minimal distances are below a certain city-dependend threshold:

$$thr_{city} = \min(\max(200m,\ d_{q99}),\ 1500m) \tag{2.1}$$

where d_{q99} is the 99th percentile distance between endpoint and nearest station, distribution computed over all city section endpoints. The acceptance threshold is higher for cities where said minimal distances are generally higher (due to noisy data acquisition / more sparse transport system / etc...), but can never be below 200m or above 1500m.

2.2.2 Sanity checks, filtering, global analysis.

Preliminary global analysis. To check the quality of the produced networks I follow these steps:

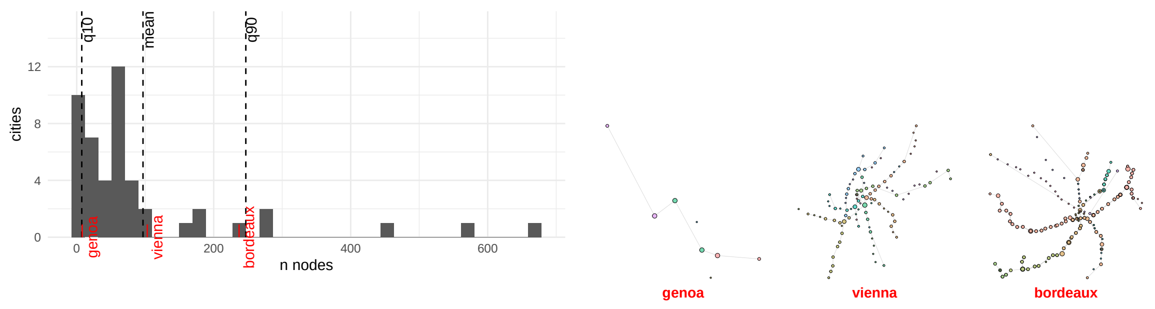


Figure 2.1: Left: size distribution of public transport networks across several cities in the world. The cities are filtered by mean degree ≥ 1. Right: the reconstructed networks of Genoa, Vienna and Bordeaux, respectively representing the 10th percentile, mean, and 90th percentile of the global size distribution. Node size encodes degree, node color encodes community according to Louvain method.

1. I build per-city networks with `igraph`: both muligraphs and *simplified* graphs.
2. I compute, per each city/network (mostly on the simplified graph), some relevant summary metrics, stored in `per_city_summary.csv`: number of nodes, number of edges of the simple and the multigraph, density of the simple graph; mean degree of the simple and the multigraph; number of connected components, giant component fraction, global clustering coefficient; mean, median, and 90th percentile distance, and diameter, of the LCC; number of communities and resulting modularity after applying Louvain and Infomap method; mean and max edge multiplicity of the multigraph.
3. I generate and visualize across-cities distributions of said metrics.
4. To better inspect each global metrics and the underlying individual networks, I select some representative cities: in particular, those approximately located at the mean value, at the 10th percentile and at the 90th percentile of the across-cities distribution. This is done in order to have a glimpse of both the most typical and some more atypical networks, still neglecting fully pathological reconstructions.
5. Finally, I plot the simple and multigraph and the degree distribution of these representative cities.

Pathology filtering. Upon inspection, aided by the previous procedure, I observe that a substantial proportion of cities present a patological network reconstruction: the most common phenomenon is presence of very long edges connecting non-subsequent stations, in some cases just the first and last of a line. I suspect this could be caused by incomplete data on the sections `LineString`, by sections that weren't updated upon later introduction of new stops and stations, or simply by noisy geometric data. To partially amend this problem and reconstruct cleaner (although possibly biased) global statistics, I attempt filtering out this family of pathological networks by discarding networks with a mean degree lower than a certain threshold (since typical transport lines are made of linear chains of stops connected by hub stations). By setting this threshold to 1, more than 20% of the cities are kept for further analysis, whose networks appear reasonably well reconstructed upon visual inspection.

A | Appendix task 34

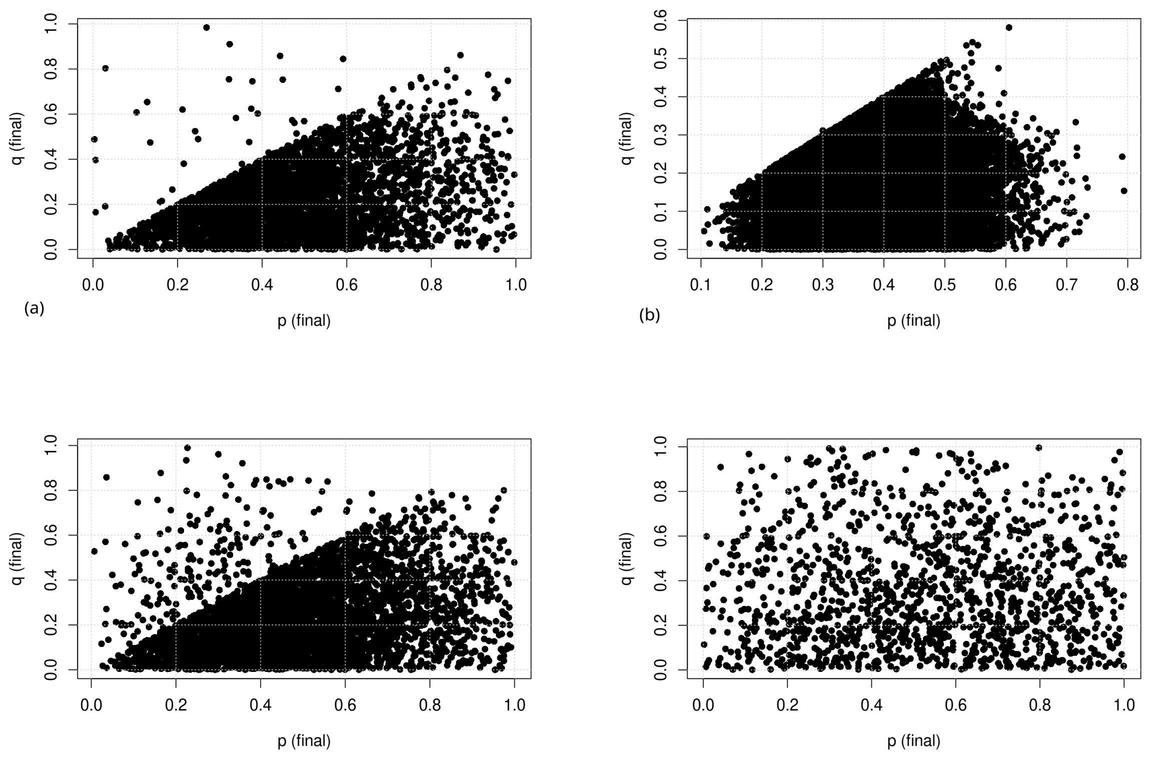


Figure A.1: **Relationship between offer and acceptance threshold:** scatterplot of (q_i, p_i) at final generation, showing strong selective pressure towards $q \leq p$ for all cases except when Social-Penalty update rule is used. (a) $\alpha = 0$, update-rule REP; (b) $\alpha = 0$, update-rule REP; (c) $\alpha = 0$, update-rules: REP, UI, MOR; (d) $\alpha = 0$, update-rules REP, UI, MOR, SP.

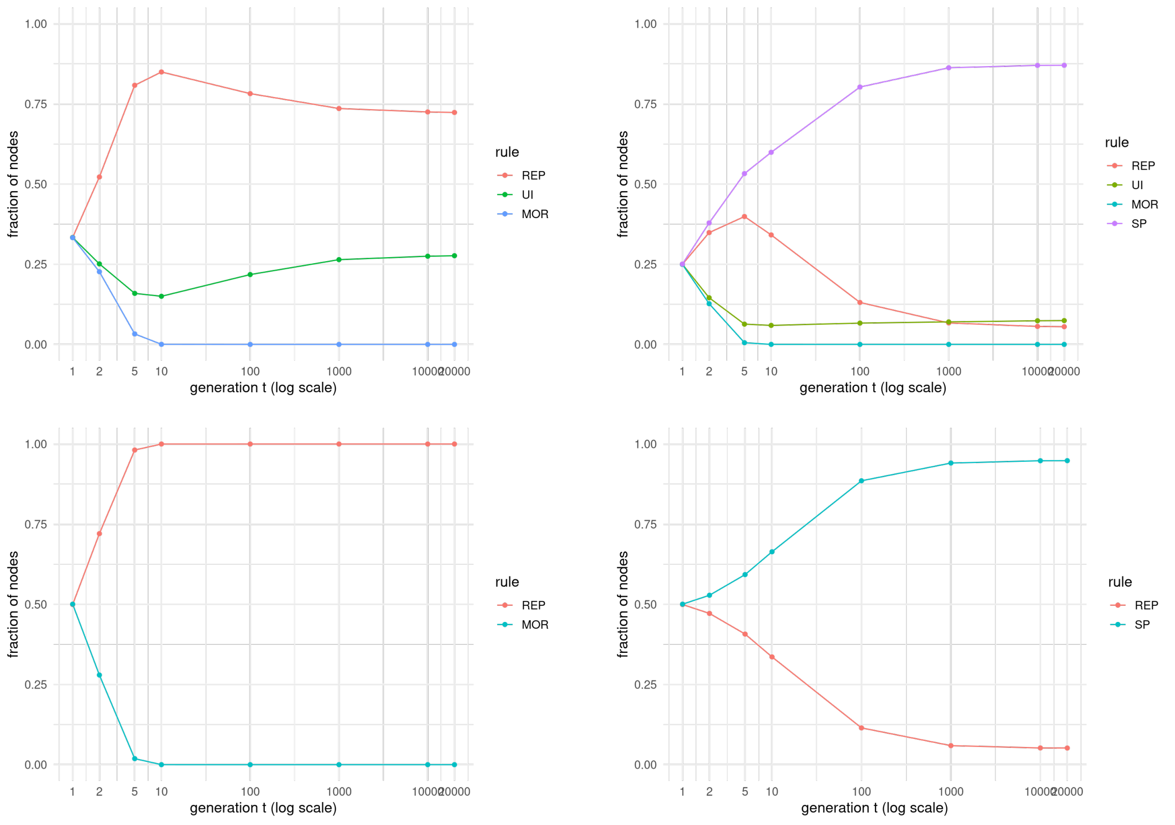


Figure A.2: **Rules relative abundances over time:** $N = 10000$, $\alpha = 0.25$

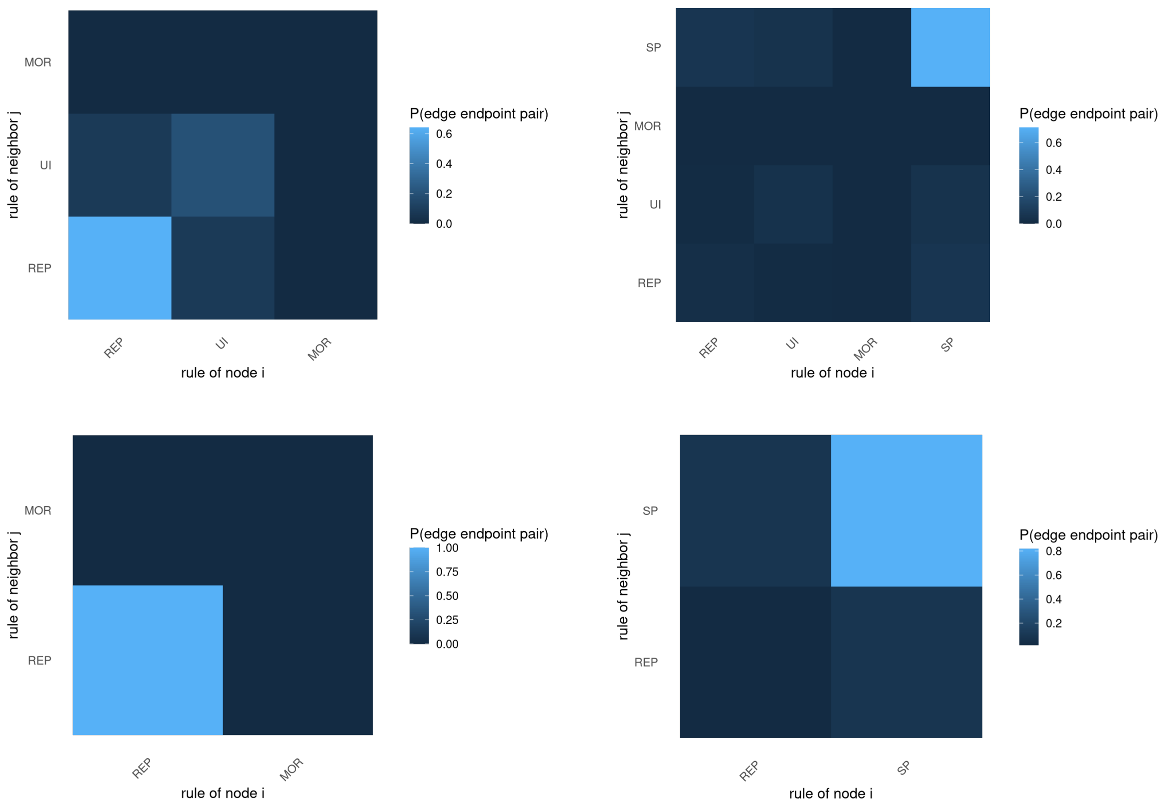


Figure A.3: **Neighbor-rule joint probabilities:** $N = 10000$, $\alpha = 0.25$

B | Appendix task 41

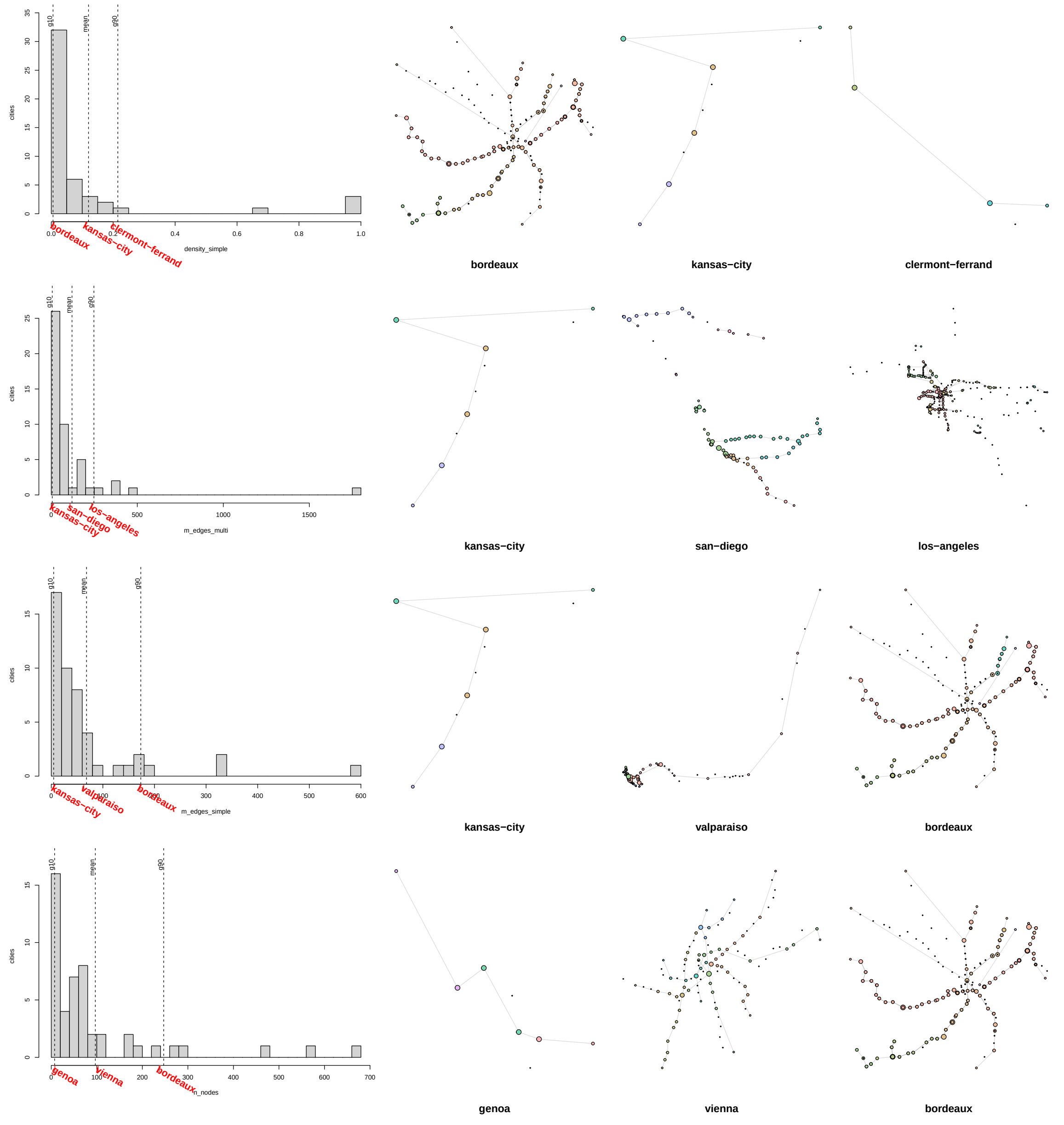


Figure B.1: **Size and density:** node size encodes degree, node color encodes Louvain community

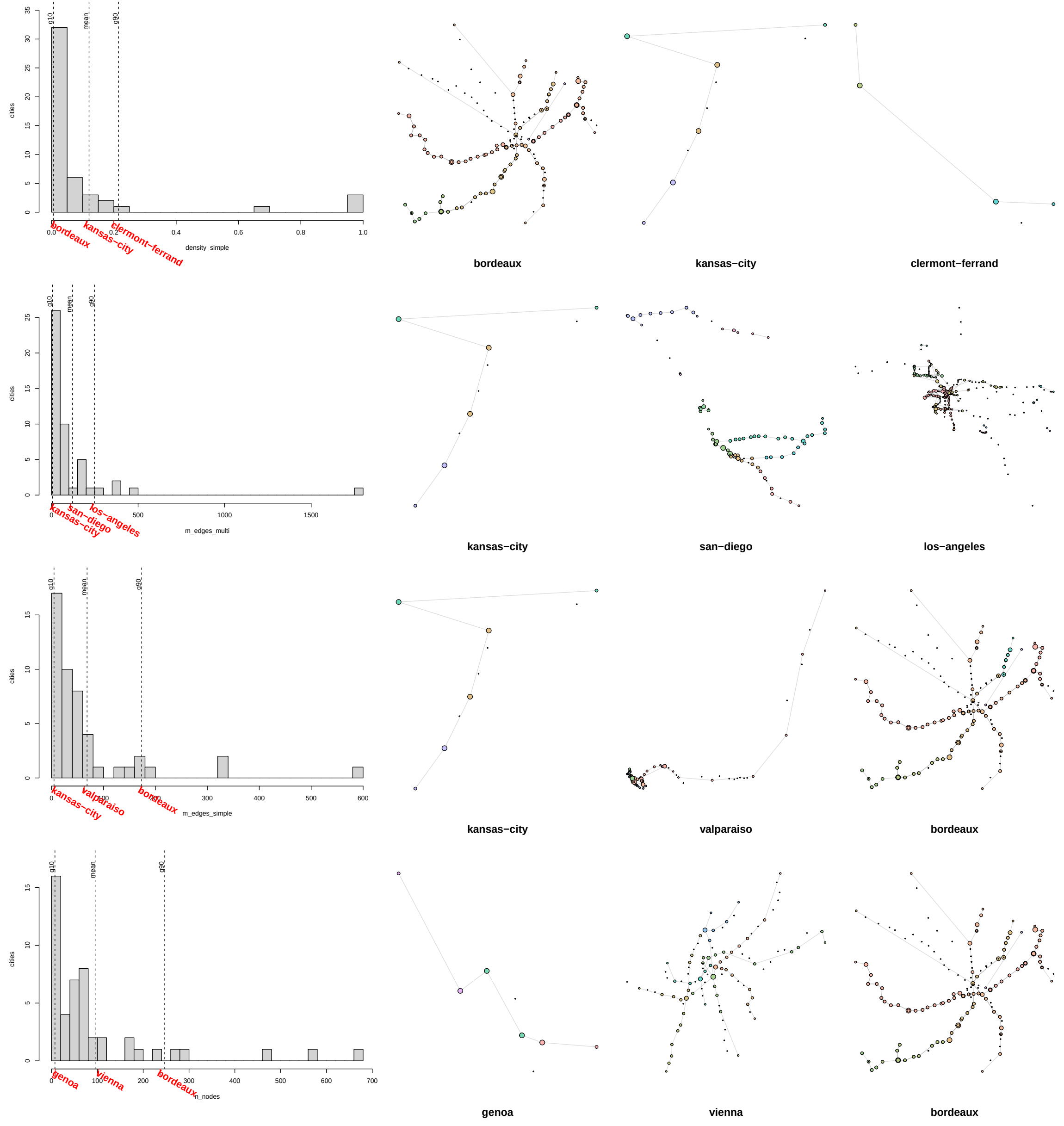


Figure B.2: **Degree and multiplicity:** node size encodes degree, node color encodes Louvain community

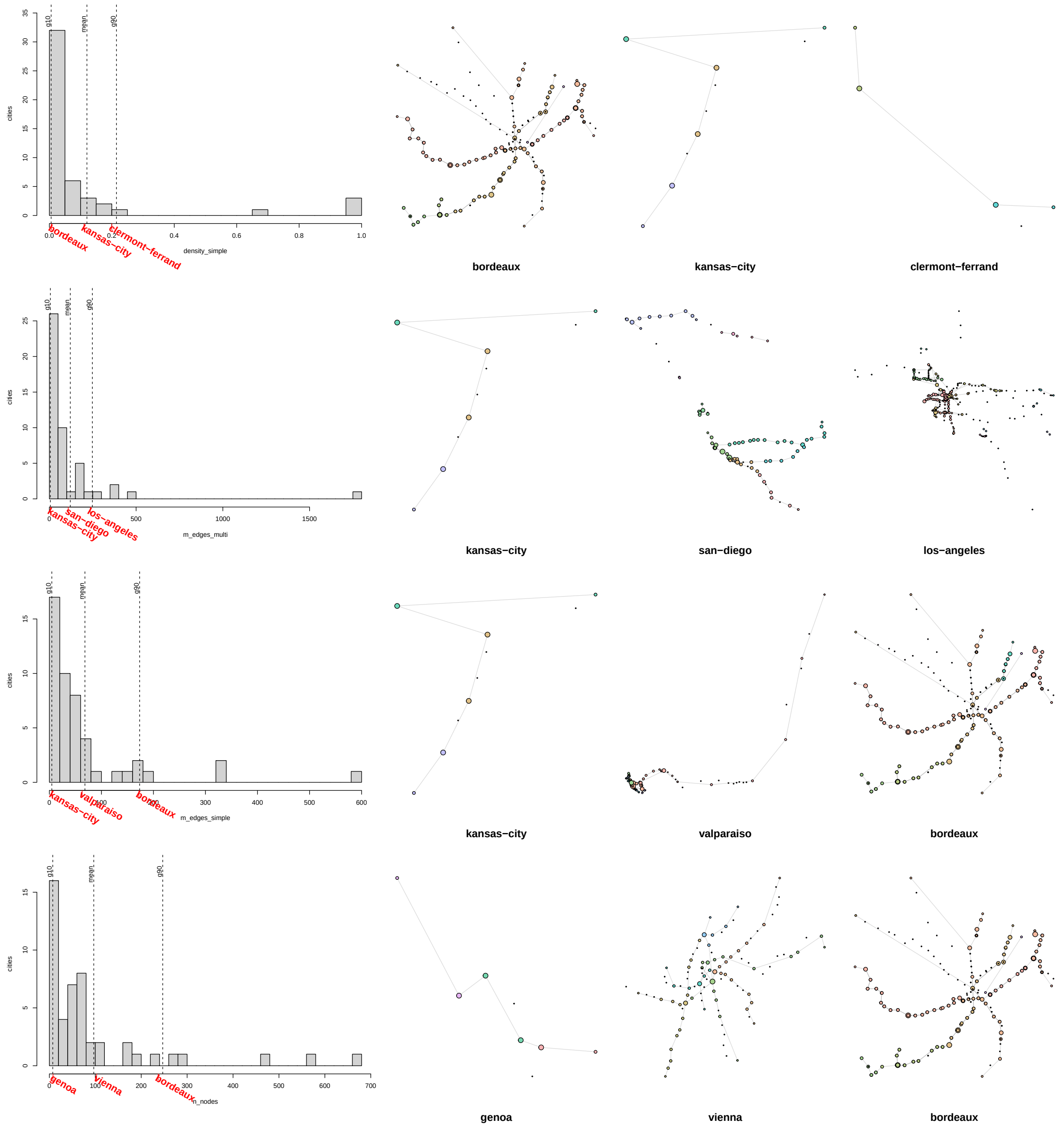


Figure B.3: **Connectivity and clustering:** node size encodes degree, node color encodes Louvain community

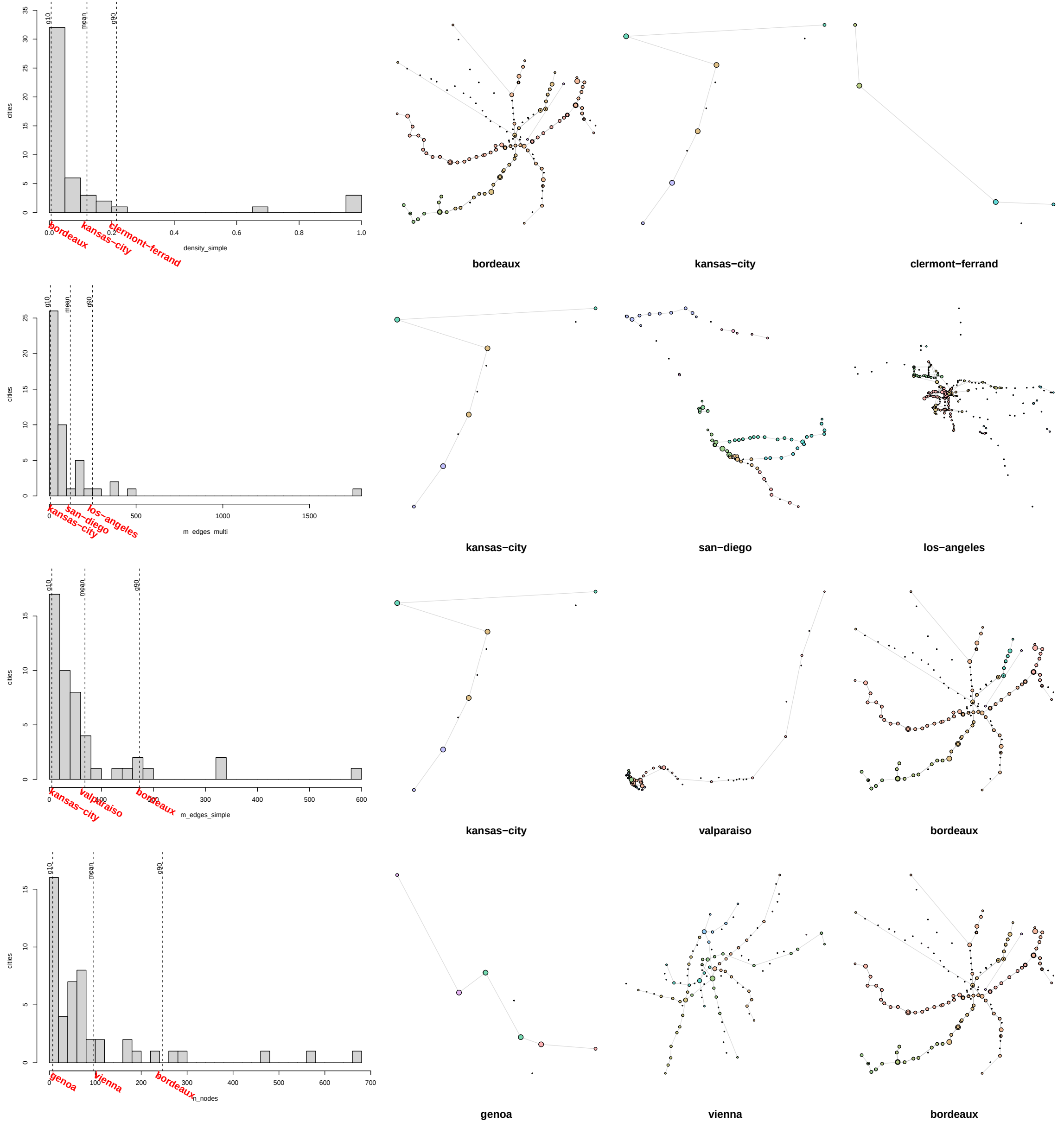


Figure B.4: **Path length statistics:** node size encodes degree, node color encodes Louvain community

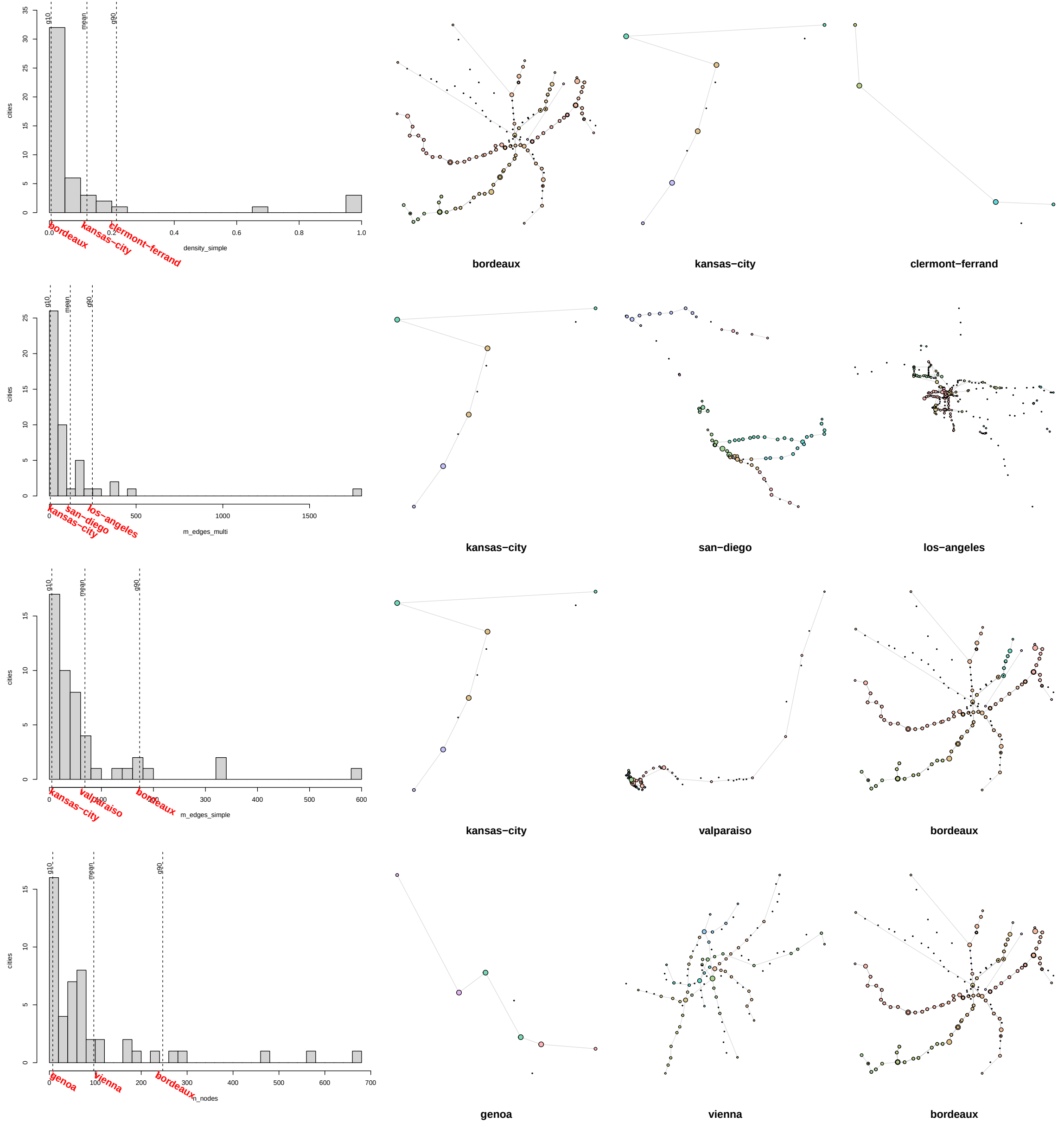


Figure B.5: **Community and mesoscale structure:** node size encodes degree, node color encodes Louvain community

C | Bibliography

[1] Citylines.co. *"We want to map all the transit systems of the world."*. https://www.citylines.co/data. [Accessed 3-Feb-2026].

[2] Geojson. https://geojson.org/. [Accessed 3-Feb-2026].

[3] Per Bak and Kim Sneppen. Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.*, 71:4083–4086, Dec 1993. doi: 10.1103/PhysRevLett.71.4083. URL https://link.aps.org/doi/10.1103/PhysRevLett.71.4083.

[4] Alessio Cardillo, Jesús Gómez-Gardeñes, Daniele Vilone, and Angel Sánchez. Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks. *New Journal of Physics*, 12(10):103034, oct 2010. doi: 10.1088/1367-2630/12/10/103034. URL https://doi.org/10.1088/1367-2630/12/10/103034.

[5] Jesús Gómez-Gardeñes and Yamir Moreno. From scale-free to erdos-rényi networks. *Phys. Rev. E*, 73:056124, May 2006. doi: 10.1103/PhysRevE.73.056124. URL https://link.aps.org/doi/10.1103/PhysRevE.73.056124.

[6] J. Hofbauer and K. Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003. doi: 10.1090/S0273-0979-03-00988-1. URL https://pure.iiasa.ac.at/id/eprint/6820/.

[7] R Sinatra, J Iranzo, J Gómez-Gardeñes, L M Floría, V Latora, and Y Moreno. The ultimatum game in complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(09):P09012, sep 2009. doi: 10.1088/1742-5468/2009/09/P09012. URL https://doi.org/10.1088/1742-5468/2009/09/P09012.