Variables

* ind_i: industry code (3-digit, NACE Rev. 2.).
* ind_j: industry code (3-digit, NACE Rev. 2.).
* eid: edge ID, uniquely identifies and ind_i--ind_j combination.
* sr_norm: revealed skill relatedness value.

Notes

* ind_i == ind_j combinations excluded.
* Labour flows observed and aggregated from year-by-year changes in the main industry code of workers' workplaces, aggregated across 2013-2019.
* sr_norm is calculated based on the formulas of Neffke et al. (2017). See applications in Elekes et al. (2023, 2024).
* Edges are considered undirected. To do this, we take the average of sr_norm values for ind_i--ind_j and ind_j--ind_i. Hence, in the data, ind_i--ind_j has the same sr_norm value as ind_j--ind_i. Both ind_i--ind_j and ind_j--ind_i are present in the data.
* sr_norm has a theoretical range of -1 to 1 (higher the stronger relatedness). Data contains all edges, but sr_norm > 0 represents above-expected labour flows. In analysis, we often restrict our attention to this subset.

References

Elekes, Z., Baranowska-Rataj, A., & Eriksson, R. (2023): Regional diversification and labour market upgrading: local access to skill-related high-income jobs helps workers escaping low-wage employment. *Cambridge Journal of Regions, Economy and Society*, 16(3): 417-430.

Elekes, Z., Tóth, G., & Eriksson, R. (2024): Regional resilience and the network structure of inter-industry labour flows. Regional Studies, 58(12): 2307-2321.

Neffke, F. M., Otto, A., & Weyh, A. (2017): Inter-industry labor flows. *Journal of Economic Behavior & Organization*, 142: 275-292.