Network structure of global remittances

Alex Stivala

Università della Svizzera italiana, Lugano, Switzerland
Swinburne University of Technology, Melbourne, Australia

ASNAC 2018, November 27–28 2018, Canberra

Background

- Starting from these two publications
  - Lillo, F., García, L., & Santander, V. (2017). Dynamics of global remittances: A graph-based analysis. *Mathematical Social Sciences*, 87, 64-71.
  - Lillo, F., & Molina Garay, J. A. (2018). The global remittance network: an inflow and outflow analysis. *The Journal of Mathematical Sociology* DOI: 10.1080/0022250X.2018.1496917
- which:
  - Analyzed publicly-available bilateral remittance data from the World Bank for the four years 2010–2013.
  - Constructed global remittances network by creating arcs for remittance flows $\geq$ threshold value δ (e.g. $\delta = 100$ million USD).
  - Described power-law degree distributions and two-vertex cycles.
  - Described log-normal inflow and outflow distributions, and “quasi transshipment countries”, where the total inflow is approximately equal to the total outflow.

What we will do

- Using the most recent World Bank bilateral remittances data (2017)
- and other data from the World Bank (GDP, population, etc.)
- A similar look at the distributions of remittance flows,
- but with a new normalized measure as well,
- And some more advanced network analysis, including community detection with spatial null models and stochastic block-modeling.

Big trouble with little data (1)

- “Credible national data on bilateral remittances are not available” (Ratha, D., & Shaw, W. (2007). South-South migration and remittances. *World Bank Working Paper 102*.)
- Therefore the World Bank estimates them from its bilateral migration matrices (data from census bureaus and other sources) and remittance inflows data (collected from IMF Balance of Payments Statistics including employee compensation and personal transfers).
- The bilateral remittance matrix is then estimated from this data, with a model weighting by per capita income in source and destination countries (Ratha & Shaw 2007).
- Lillo et al. (2017,2018) analyze this model probabilistically and explore the structure of graphs constructed from it.

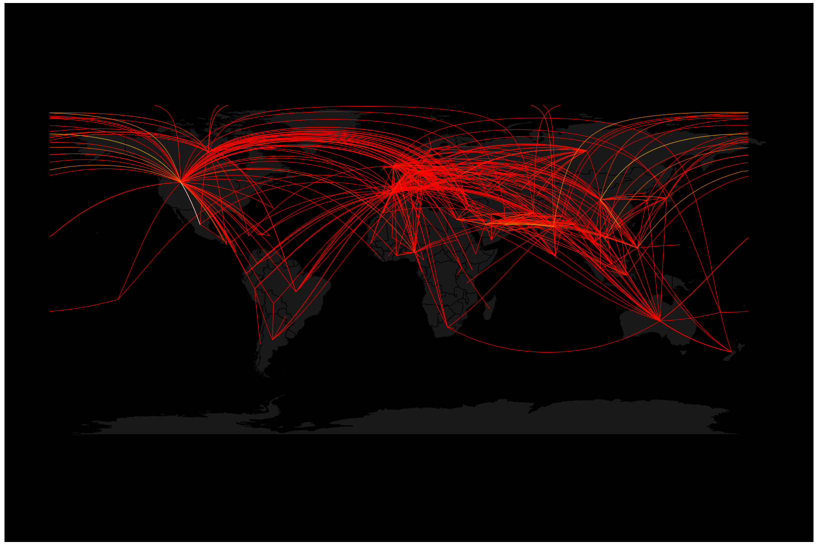
Big trouble with little data (2)

- So we have to be careful to remember we are investigating structure arising from a model based on data (migration flows and total national remittance inflows) — not real data directly.
- So e.g. making inferences from models such as ERGM may be problematic.
- Such models are sensible for the migration data directly: Windzio, M. (2018). The network of global migration 1990–2013: Using ERGMs to test theories of migration between countries. *Social Networks*, 53, 20–29.

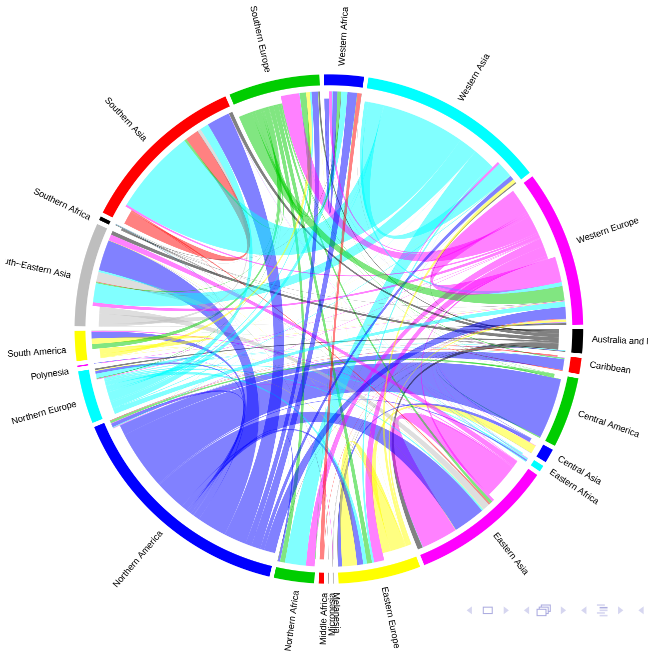
Largest remittance in and out flows as % of GDP

Top outflow/GDP countries	Outflow (% of GDP)	Top inflow/GDP countries	Inflow (% of GDP)
Gambia, The	28	Tonga	34
New Caledonia	21	Kyrgyz Republic	33
French Polynesia	19	Tajikistan	31
Nepal	13	Haiti	29
Belize	11	Nepal	28
Kuwait	10	Liberia	27
Andorra	10	Bermuda	27
Togo	9	New Caledonia	24
American Samoa	9	Comoros	21
Benin	9	Gambia, The	21
United Arab Emirates	9	El Salvador	20
Cameroon	8	Moldova	20
Bhutan	8	Honduras	19
Liberia	8	Yemen, Rep.	18
Bahrain	8	French Polynesia	17
Jordan	7	Jamaica	17
Solomon Islands	7	Samoa	17
Gabon	7	Lesotho	16
Northern Mariana Islands	7	Lebanon	15
Saudi Arabia	7	Marshall Islands	15

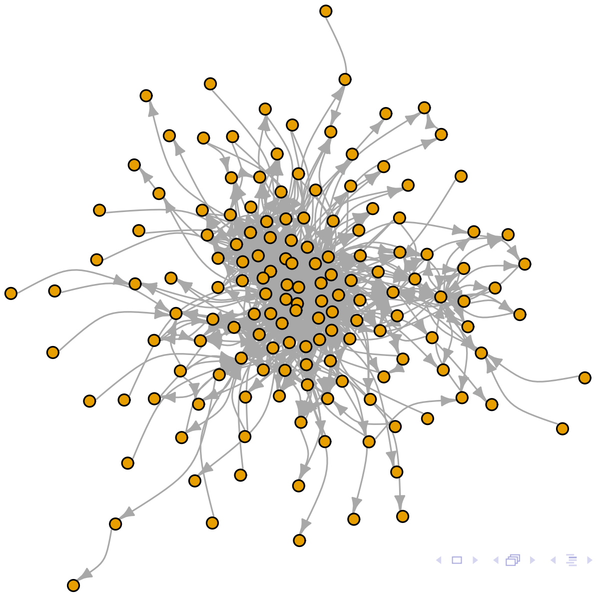
The global remittances network, $\delta = 100$



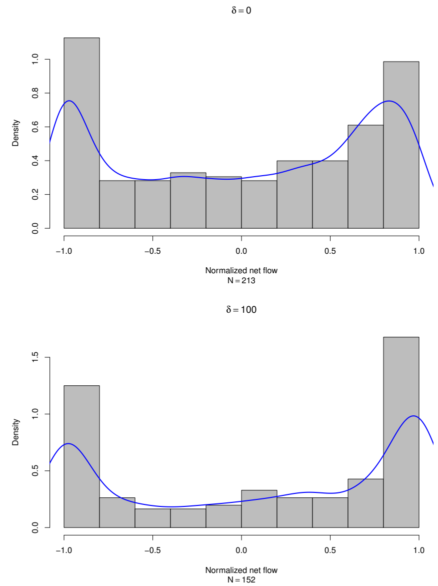
Circular plot of global remittances network $\delta = 100$



The global remittances network, $\delta = 100$ giant component ($N = 152$) "graphopt" graph layout



Distribution of normalized net flow (net flow / total flow)



Countries with smallest magnitude normalized net flow ($\delta = 0$)

Smallest normalized net flow countries	NNF
Finland	-0.070
Togo	-0.051
French Polynesia	-0.040
Malawi	-0.011
Korea, Rep.	0.028
Burkina Faso	0.038
New Caledonia	0.062
Cambodia	0.073
Mauritius	0.076
France	0.077

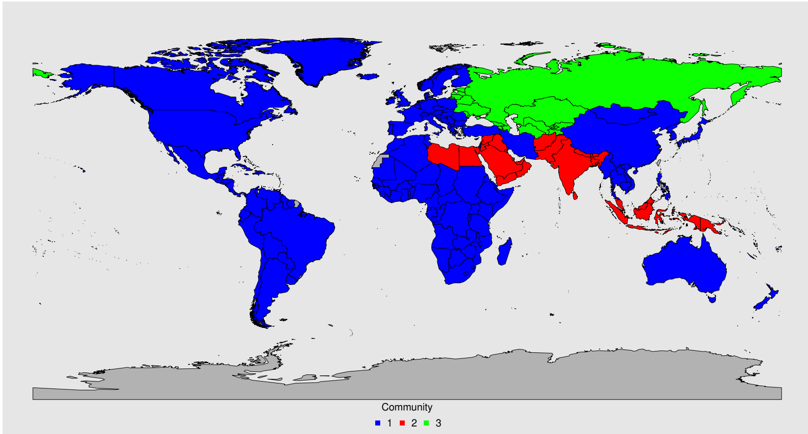
Top ten betweenness centrality countries ($\delta = 0$)

Country	Betweenness centrality
United States	0.06
France	0.04
Australia	0.03
China	0.03
Canada	0.03
Russian Federation	0.03
United Kingdom	0.02
Egypt, Arab Rep.	0.02
Italy	0.02
Turkey	0.02

Finding space-independent communities

- Expert et al. 2011 "Uncovering space-independent communities in spatial networks" PNAS 108(19):7663–7668
- Instead of using the Newman-Girvan null model (preserve node degrees on average), use instead a null model that preserves weighted average for an edge to exist at a given distance.
- The null model is similar to a "gravity" model: edge probability between two nodes is proportional to the product of the node "masses" (or importances) over function of the distance between them.
- For "importance" we try node degree (similar to Newman-Girvan null model) as well as GDP and population.
- For the clustering algorithm we use a generalized Louvain method (Jeub et al. 2011) and implement the Expert et al. (2011) null model in the modularity matrix.

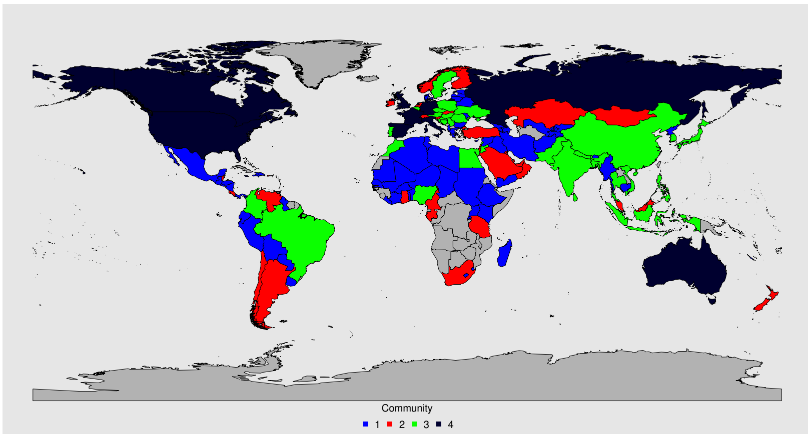
Louvain algorithm spatial null model with degree mass



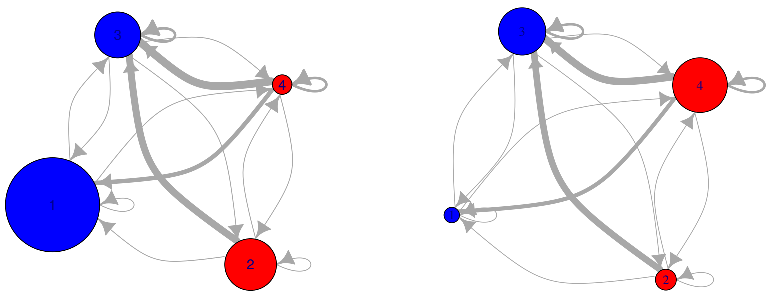
Stochastic Block Model (SBM)

- A model of networks with unobserved classes (blocks) where the probability of a tie between nodes depends only on the classes to which they belong (Nowicki & Snijders, 2001).
- Much more general than community detection, which (by definition) can only find assortative (i.e. community) structure. SBM can also find disassortative, core-periphery, and other structures.
- A large literature on this and the computationally difficult problem of finding the blocks.
- We will use a variational Bayesian method to find the blocks in weighted directed networks: Aicher, C., Jacobs, A. Z., & Clauset, A. (2014). Learning latent block structure in weighted networks. *Journal of Complex Networks*, 3(2), 221-248.

Blocks found by WSBM $\delta = 100$ network with log-normal weight and Bernoulli edge distribution

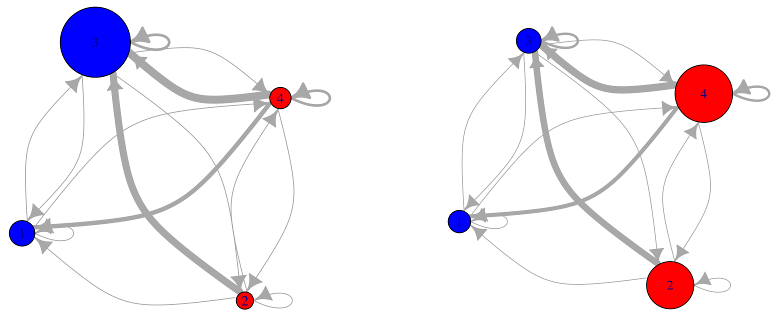


WSBM edge blockmodel of $\delta = 100$ network (1)



Blue nodes have a net inflow of remittances and red nodes a net outflow. Node size proportional to number of countries in block (left) and total GDP of countries in block (right).

WSBM edge blockmodel of $\delta = 100$ network (2)



Blue nodes have a net inflow of remittances and red nodes a net outflow. Node size proportional to total population of countries in block (left) and mean per capita GDP of countries in block (right).

Acknowledgments

- We thank Christopher Aicher, Abigail Jacobs and Aaron Clauset for making their WSBM MATLAB code available, and Lucas Jeub, Marya Bazzi, Inderjit Jutla, and Peter Mucha for making their generalized Louvain method MATLAB code available.

Social Network Analysis 5-Day Workshop: Theory, Method and Application

- Monday 18 February – Friday 22 February, 2019
- Swinburne University of Technology, Hawthorn VIC
- Cost: $3,000 (Full-time PhD students $1,500)
- Enquiries: Dr Peng Wang, Centre for Transformative Innovation: pengwang@swin.edu.au
- `https://www.eventbrite.com.au/e/social-network-analysis-5-day-workshop-theory-method-and-application-tickets-52032527691`

Hidden bonus slides

What are the gray countries on the maps?

- Areas coloured gray on the maps are regions for which there is no data, because they are not recognized by the World Bank as "countries".
- The large one in the northwest of Africa is Western Sahara, "a disputed territory in the Maghreb region of North Africa, partially controlled by the self-proclaimed Sahrawi Arab Democratic Republic and partially Moroccan-occupied, bordered by Morocco proper to the north, Algeria to the northeast, Mauritania to the east and south, and the Atlantic Ocean to the west" (Wikipedia)
- The one in the north of South America is French Guiana, an overseas department and region of France.
- To the north of Norway is Svalbad, a Norwegian archipelago.
- Between China and the Philippines is Taiwan, not considered separately in the World Development Indicators.

More on remittances data

- Many problems with country reporting of remittances (World Bank Group. 2016. Migration and Remittances Factbook 2016, Third Edition. Washington, DC: World Bank):
  - missing data (not reported to IMF), arbitrary classifications, citizenship rather than residency,
  - central banks using data from commercial banks but not e.g. money transfer operators, post offices, mobile transfers
  - Not accounting for flows through informal channels at all. New surveys required for this, household surveys only indicative.
- See particulary for the importance of transaction costs in motivating a high proportion of remittances via informal channels:
  - Freund, C., & Spatafora, N. (2008). Remittances, transaction costs, and informality. *Journal of Development Economics*, 86(2), 356-366.

Remittances household survey data

- There is some household survey microdata for some African countries and African diaspora in Belgium available from the World Bank at `http://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data`
- This data has been used in some publications in regression models, without any network aspects, e.g.:
  - Musumba, M., Mjelde, J. W., & Adusumilli, N. C. (2015). Remittance receipts and allocation: a study of three African countries. *Applied Economics*, 47(59), 6375-6389.
  - Bang, J. T., Mitra, A., & Wunnava, P. V. (2016). Do remittances improve income inequality? An instrumental variable quantile analysis of the Kenyan case. *Economic Modelling*, 58, 394-402.
  - Bredtmann, J., Martnez Flores, F., & Otten, S. (2018). Remittances and the brain drain: Evidence from microdata for Sub-Saharan Africa. *Journal of Development Studies*. DOI: 10.1080/00220388.2018.1443208

Research idea using some of the World Bank microdata

- The household survey data of households in Belgium with people from D.R. Congo, Nigeria, and Senegal (2005) could be used to build a personal remittances network.

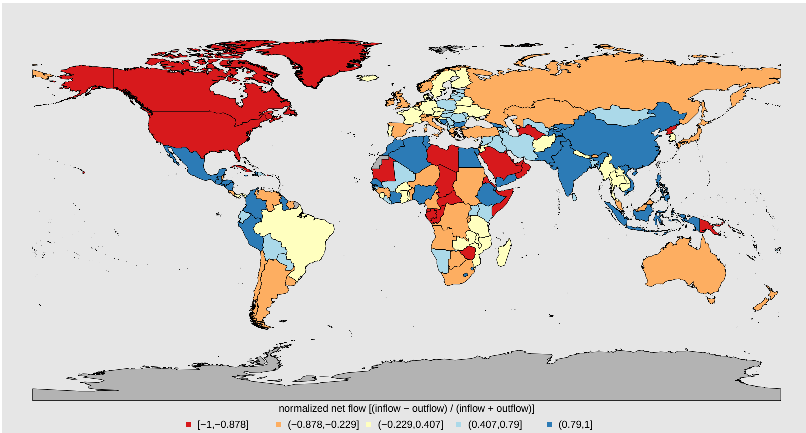
Ambitious social networks research idea (for someone else)

- As noted by the World Bank, data on remittance flows relies on bank reporting and excludes very important informal channels.
- Informal value transfer systems such as *hawala* are often used for remittances (often due to lower cost than formal channels, or the absence of functioning formal banking).
- They do not transfer cash or other financial instruments such as promissory notes, but are based entirely on honour (or trust — indeed in Arabic *hawala* can mean "transfer" or "trust" (source: Wikipedia)).
- Hence an ideal study (ethnography / sociology) on social networks and their place in systems of trust and relation to migration and remittances.
- Published research on them seems quite limited (despite increased interest due to supposed use in money laundering or terrorist financing — although this is overblown according to some researchers [see citations on next slide]).

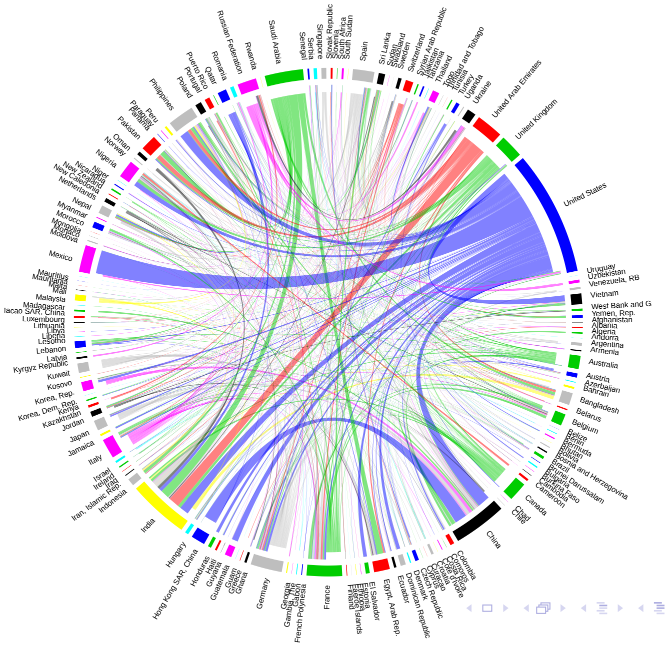
Some papers on informal remittance channels

- Maimbo, S. M. (2003). The money exchange dealers of Kabul: A study of the Hawala system in Afghanistan. *World Bank Working Paper*, 13.
- Passas, N. (2006). Demystifying Hawala: A look into its social organization and mechanics. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 7(S1), 46-62.
- Passas, N. (2006). Fighting terror with error: the counter-productive regulation of informal value transfers. *Crime, Law and Social Change*, 45(4-5), 315-336.
- McCusker, R. (2005). Underground banking: legitimate remittance network or money laundering system? *Trends & Issues in Crime & Criminal Justice*, 300.
- Razavy, M., & Haggerty, K. D. (2009). Hawala under scrutiny: Documentation, surveillance and trust. *International Political Sociology*, 3(2), 139-155.
- Siegel, D., & van de Bunt, H. (2014). Underground Banking in the Netherlands. In *Organized Crime, Corruption and Crime Prevention* (pp. 251-261). Springer, Cham.

Remittance normalized net flows



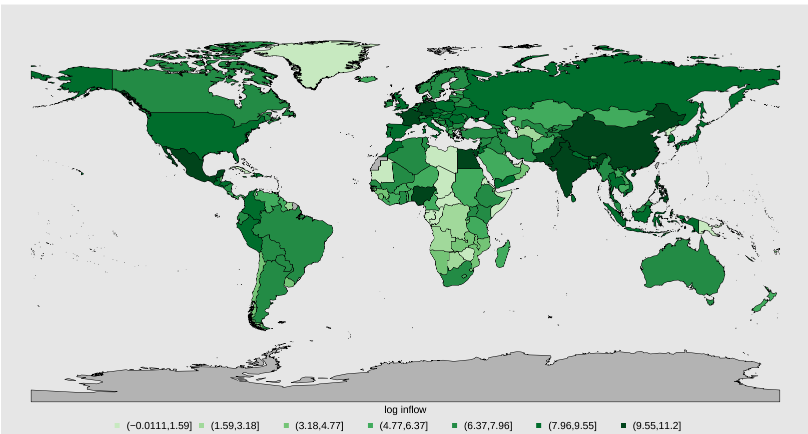
Circular plot of global remittances network $\delta = 100$



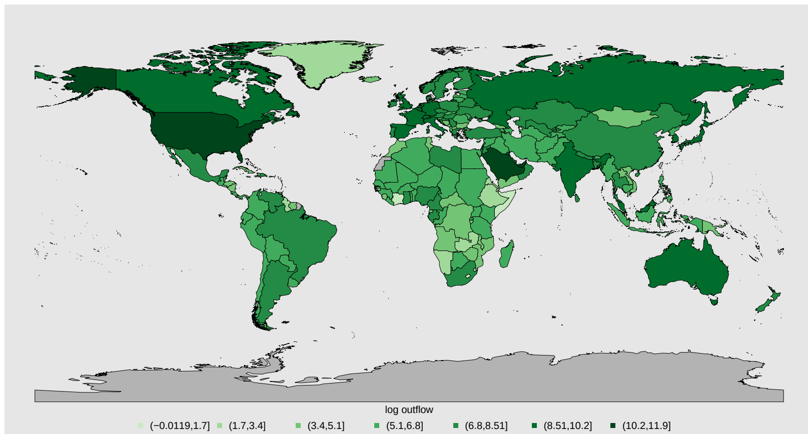
Summary statistics of the network

δ	N	Components	Mean degree	Density	Clustering coefficient	Assortativity coefficient	Average path length directed	Average path length undirected
0	214	2	111.25	0.26116	0.64017	-0.29174	1.74	1.61
100	214	63	6.61	0.01551	0.26134	-0.24028	3.09	2.60

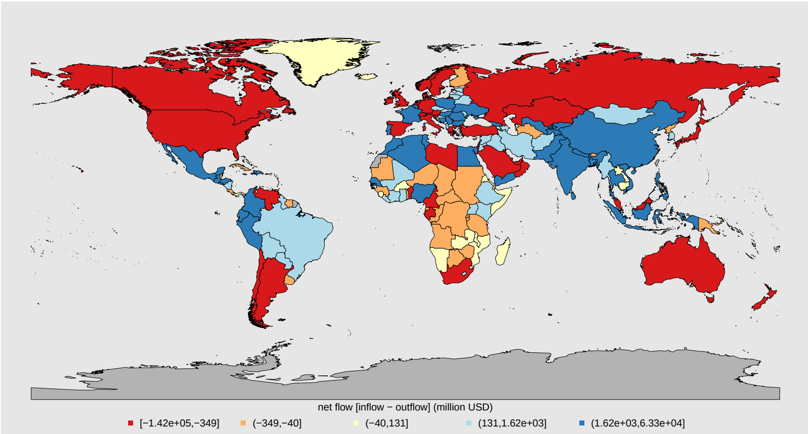
Remittance inflows



Remittance outflows



Remittance net flows



Largest inflows and outflows

Top net outflow countries	Net flow (million USD)
United States	-141868
Saudi Arabia	-46438
United Arab Emirates	-32978
Canada	-23219
United Kingdom	-22428
Hong Kong SAR, China	-16691
Australia	-14947
Kuwait	-11729
Qatar	-10009
Russian Federation	-8477

Top net inflow countries	Net flow (million USD)
India	63258
China	61032
Philippines	32271
Mexico	27851
Nigeria	20824
Egypt, Arab Rep.	19582
Pakistan	19298
Vietnam	13676
Bangladesh	11356
Guatemala	8359

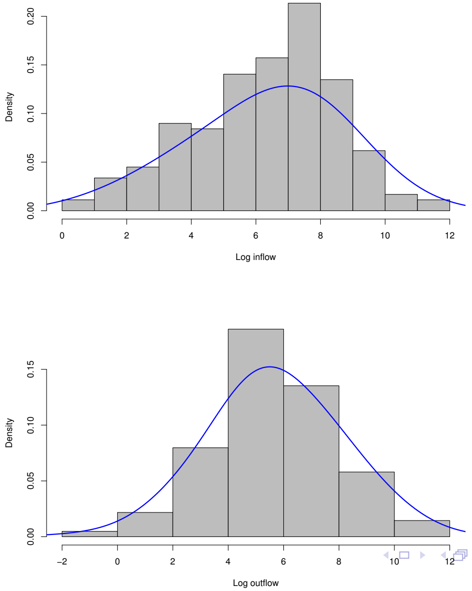
Smallest magnitude net flows

Smallest net flow countries	Net flow (million USD)
San Marino	-15.081
Eritrea	-12.490
Mozambique	-9.927
Antigua and Barbuda	-5.374
Grenada	-1.974
Malawi	-0.871
Somalia	-0.869
St. Martin (French part)	0.000
St. Kitts and Nevis	2.076
Tuvalu	4.127

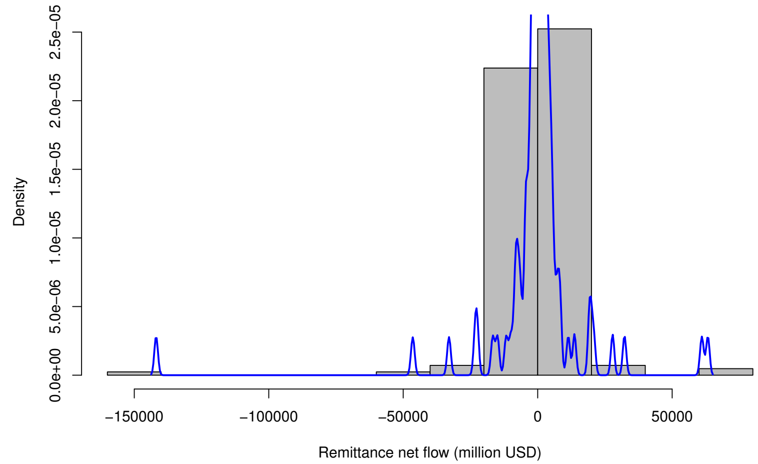
Countries with smallest magnitude normalized net flow ($\delta = 100$)

Smallest normalized net flow countries	NNF
Austria	-0.157
Sweden	-0.059
Belarus	-0.043
Korea, Rep.	0.005
New Caledonia	0.010
Burkina Faso	0.016
Finland	0.042
France	0.084
Ghana	0.147
Luxembourg	0.159

Inflow and outflow distributions appear to be log-normal



Not so clear for net flows however

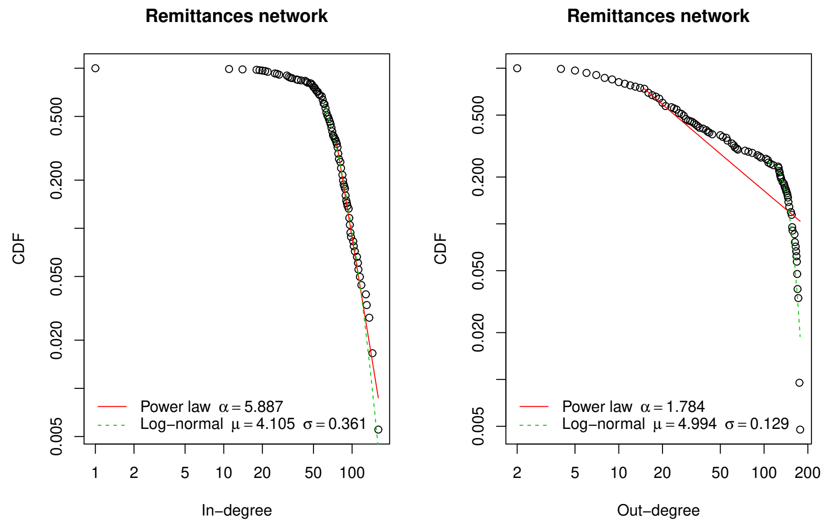


Small-world proprties of the network

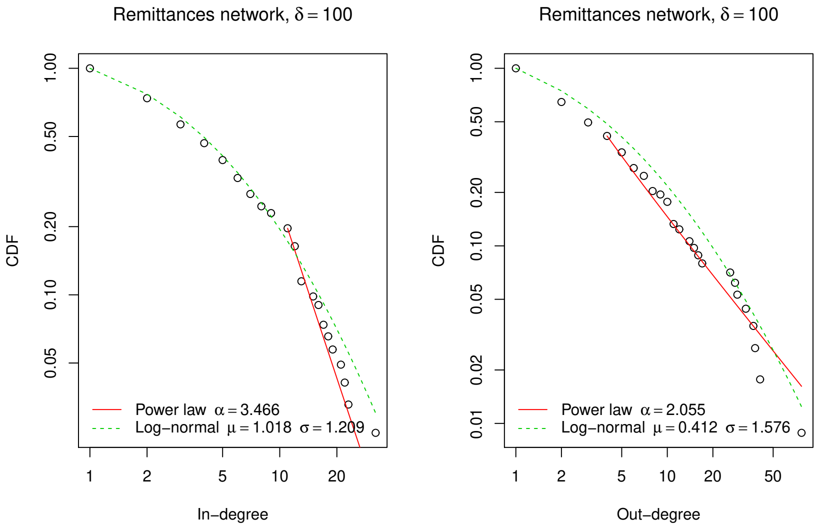
δ	N	L_{g}	L_{r}	L_{l}	C^*_{g}	C^*_{r}	C^*_{l}	SWI
0	213	1.61	1.14	0.953	0.560	0.525	0.743	0.575
100	152	2.60	2.25	8.17	0.434	0.061	0.660	0.585

- The networks are small-world according to the S^{Δ} significance test of Humphries & Gurney (2008).
- Small World Index (SWI) (Neal 2017) ranges from 0 to 1.
- L_{g} is the average shortest path length of the network
- C^*_{g} is its clustering coefficient.
- L_{r} and C^*_{r} are, respectively, the average shortest path length and clustering coefficient for an Erdős-Renyi random graph with same size and mean degree.
- L_{l} and C^*_{l} are, respectively, the mean path length and clustering coefficient for a ring lattice graph with the same size and mean degree.

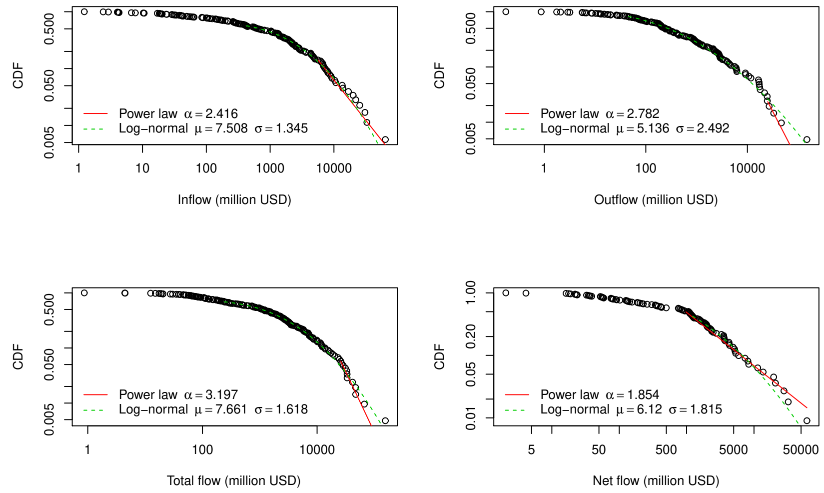
The in-degree distribution is consistent with both power law and log-normal disribution, but out-degree is not



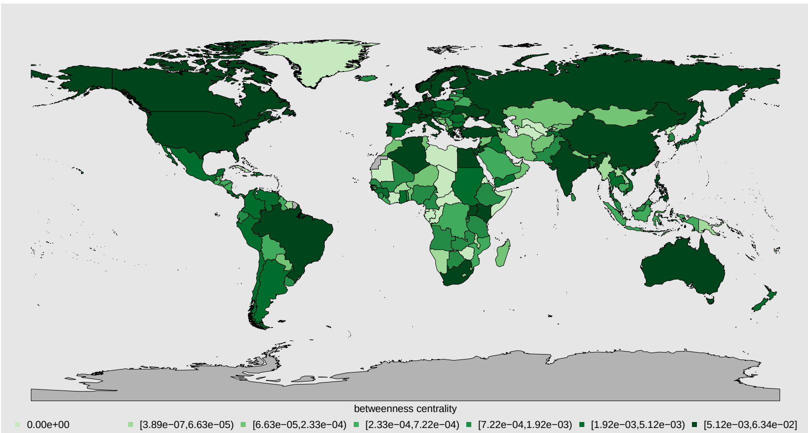
For threshold $\delta = 100$ network, log-normal is a better fit but also consistent with power law



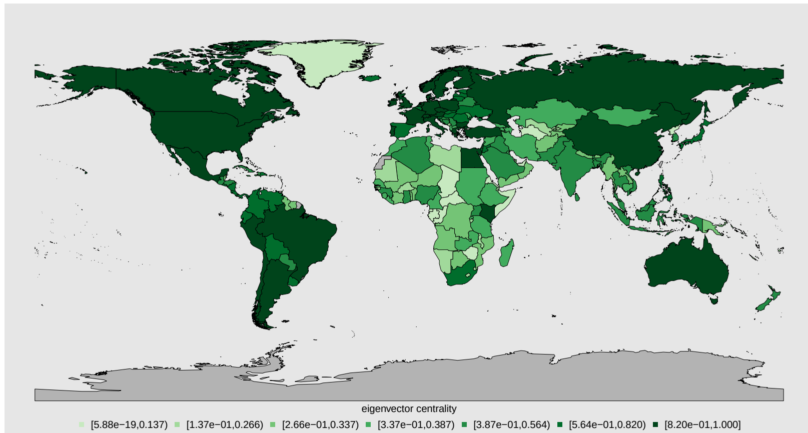
The log-normal distribution is a better fit for all except net flow for which power law is better



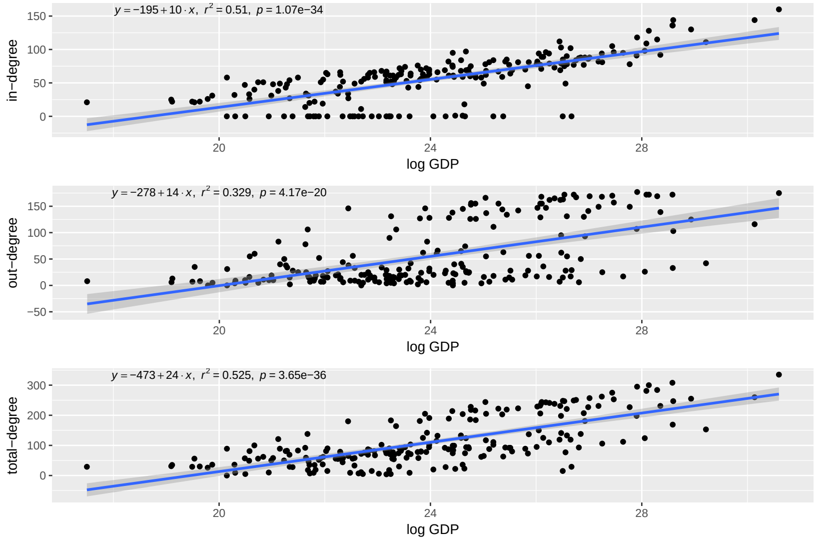
Betweenness centrality



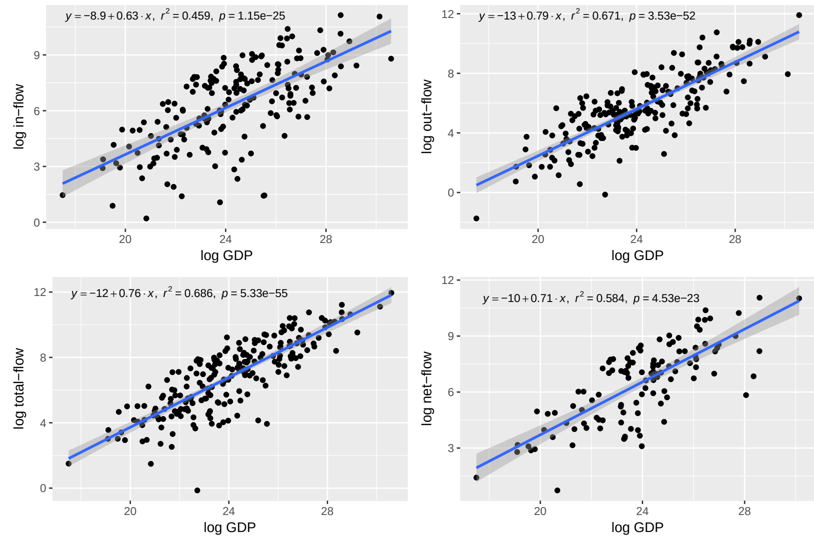
Eigenvector centrality



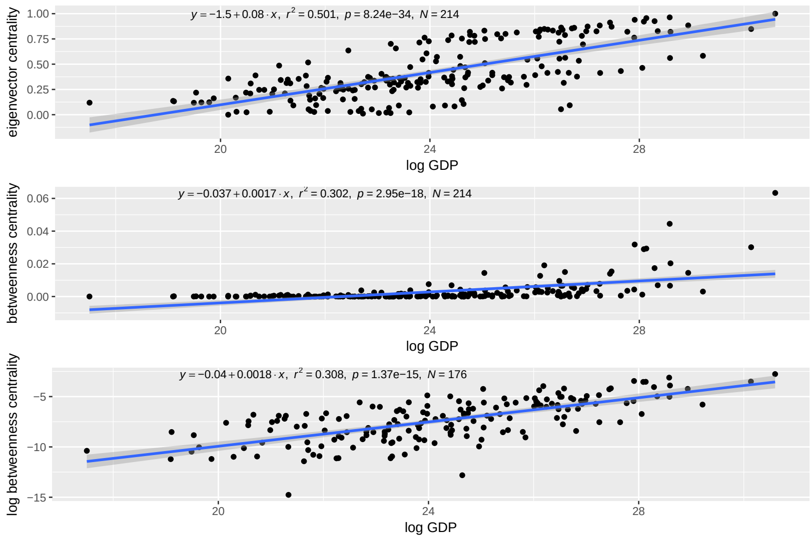
Degree in the remittances network is linearly correlated with log GDP



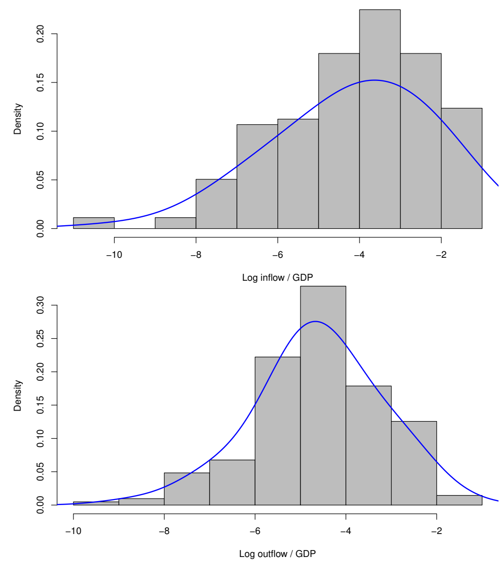
Logarithms of weighted degrees in the remittances network are linearly correlated with logarithm of GDP.



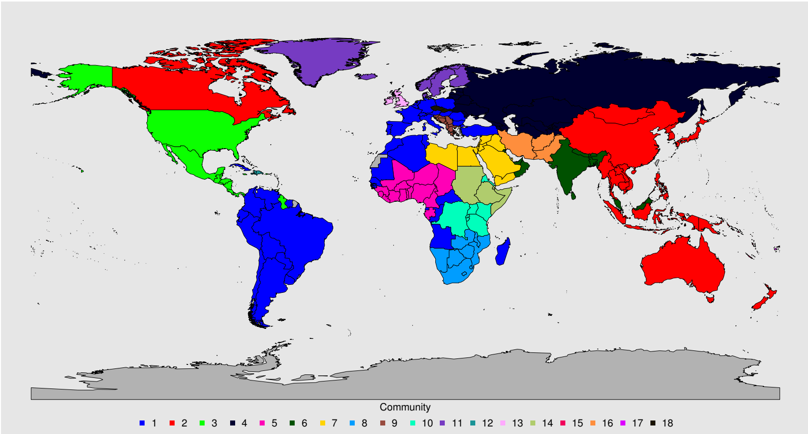
Eigenvector centrality is correlated with log GDP; betweenness centrality, not so much



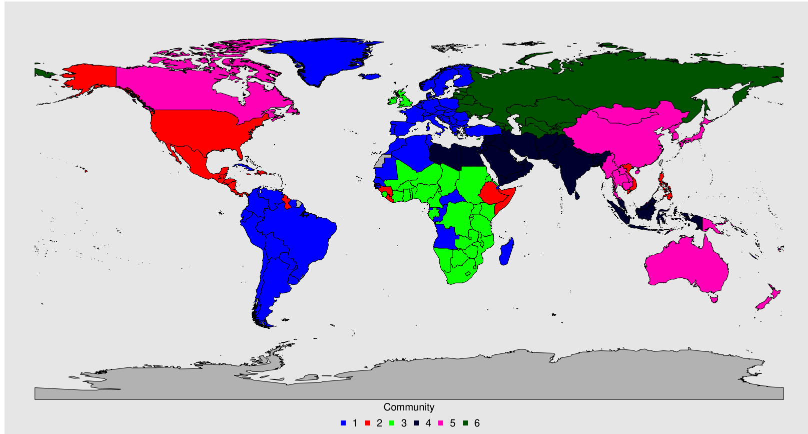
Distribution of inflows and outflows as fraction of GDP



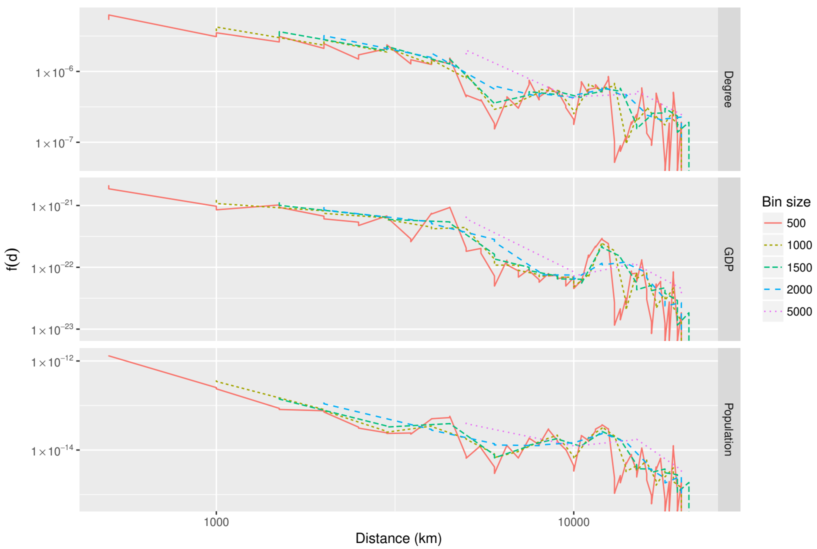
Communities detected with the Infomap algorithm



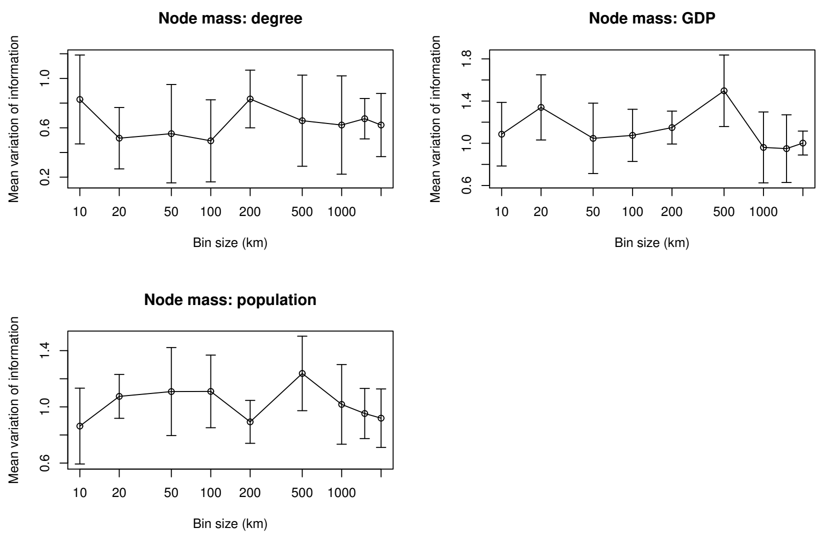
Communities detected with Louvain algorithm



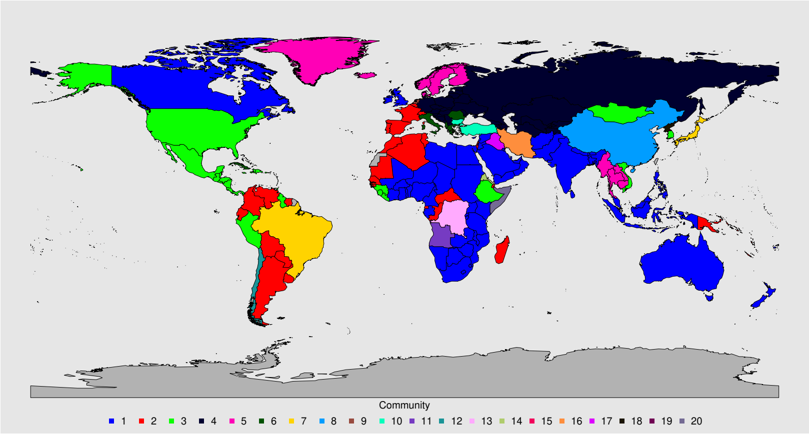
Deterrence function for different bin sizes



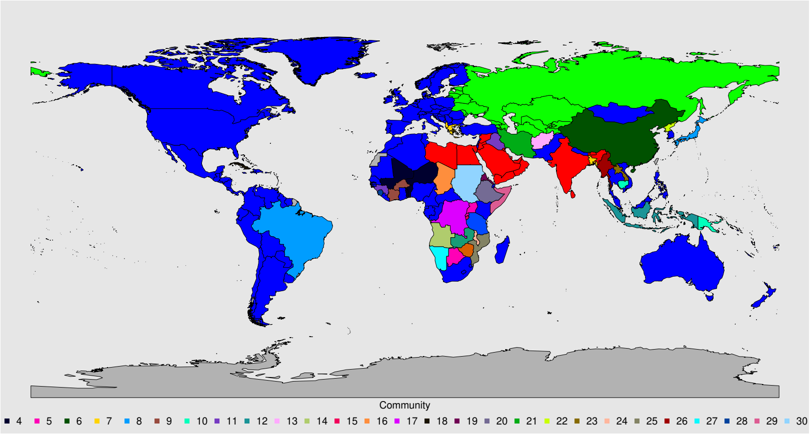
Mean variation of information for different bin sizes



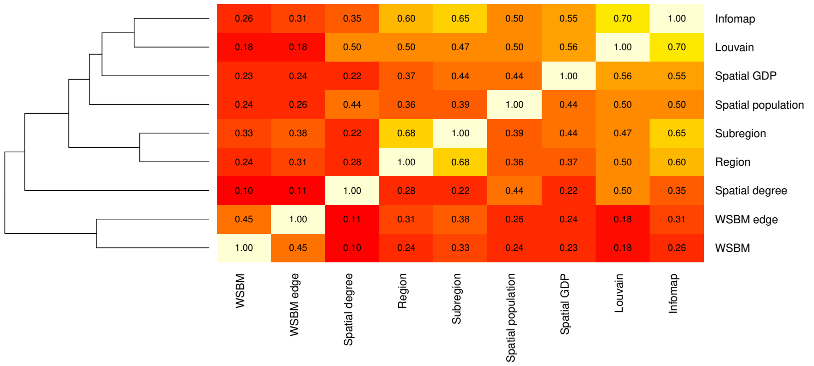
Louvain algorithm spatial null model with GDP mass



Louvain algorithm spatial null model with population mass



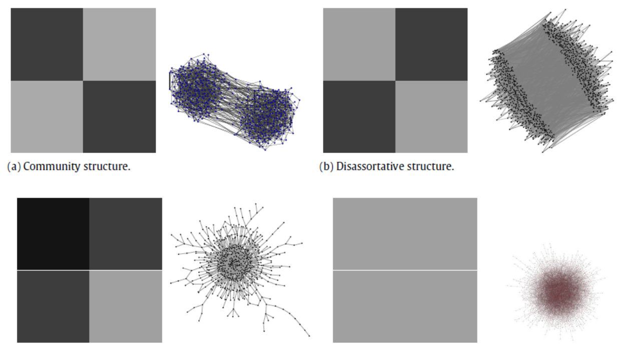
Comparing the different communities found using NMI



	WSBM	WSBM edge	Spatial degree	Region	Subregion	Spatial population	Spatial GDP	Louvain	Infomap
Infomap	0.26	0.31	0.35	0.60	0.65	0.50	0.55	0.70	1.00
Louvain	0.18	0.18	0.50	0.50	0.47	0.50	0.56	1.00	0.70
Spatial GDP	0.23	0.24	0.22	0.37	0.44	0.44	1.00	0.56	0.55
Spatial population	0.24	0.26	0.44	0.36	0.39	1.00	0.44	0.50	0.50
Subregion	0.33	0.38	0.22	0.68	1.00	0.39	0.44	0.47	0.65
Region	0.24	0.31	0.28	1.00	0.68	0.36	0.37	0.50	0.60
Spatial degree	0.10	0.11	1.00	0.28	0.22	0.44	0.22	0.50	0.35
WSBM edge	0.45	1.00	0.11	0.31	0.38	0.26	0.24	0.18	0.31
WSBM	1.00	0.45	0.10	0.24	0.33	0.24	0.23	0.18	0.26

SBM structure examples

S. Fortunato, D. Hric / Physics Reports 659 (2016) 1–44 9

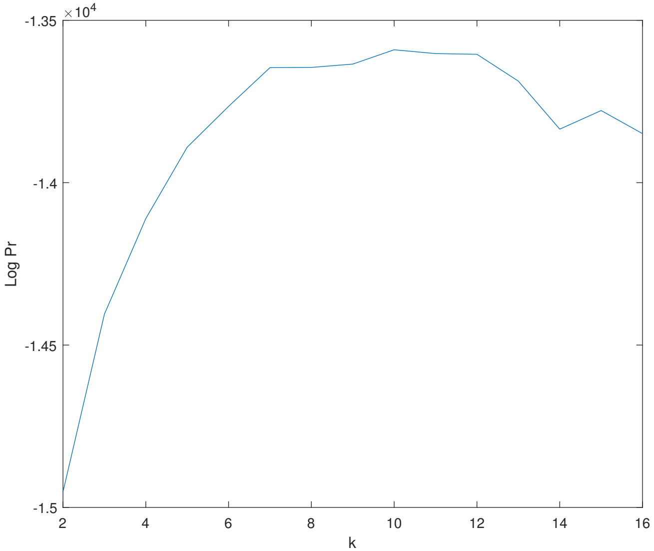


(a) Community structure. (b) Disassortative structure.

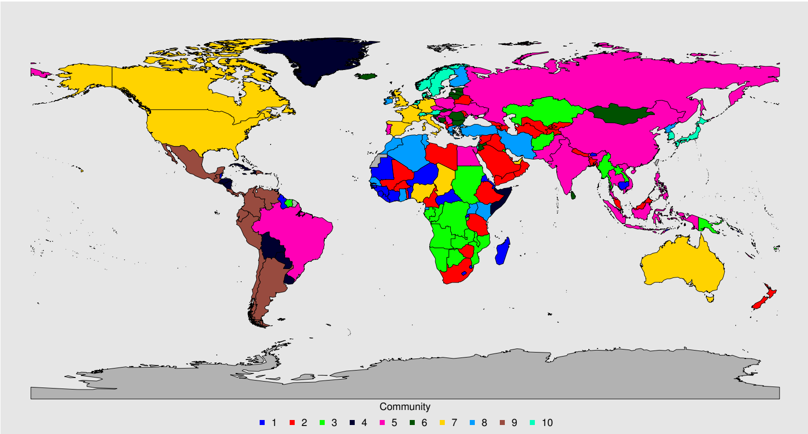
(c) Core–periphery structure. (d) Random graph.

Fig. 8. Stochastic block model. We show the schematic adjacency matrices of network realisations produced by the model for special choices of the edge probabilities, along with one representative realisation for each case. For simplicity we show the case of two blocks of equal size. Darker blocks indicate higher edge probabilities and consequently a larger density of edges inside the block. (a) Illustrates community (or assortative) structure: the probabilities (link densities) are much higher inside the diagonal blocks than elsewhere. (b) Shows the opposite situation (disassortative structure). (c) Illustrates a core–periphery structure. (d) Shows a random graph à la Erdős and Rényi: all edge probabilities are identical, inside and between the blocks, so there are no actual groups.
Source: Adapted figure with permission from [19].

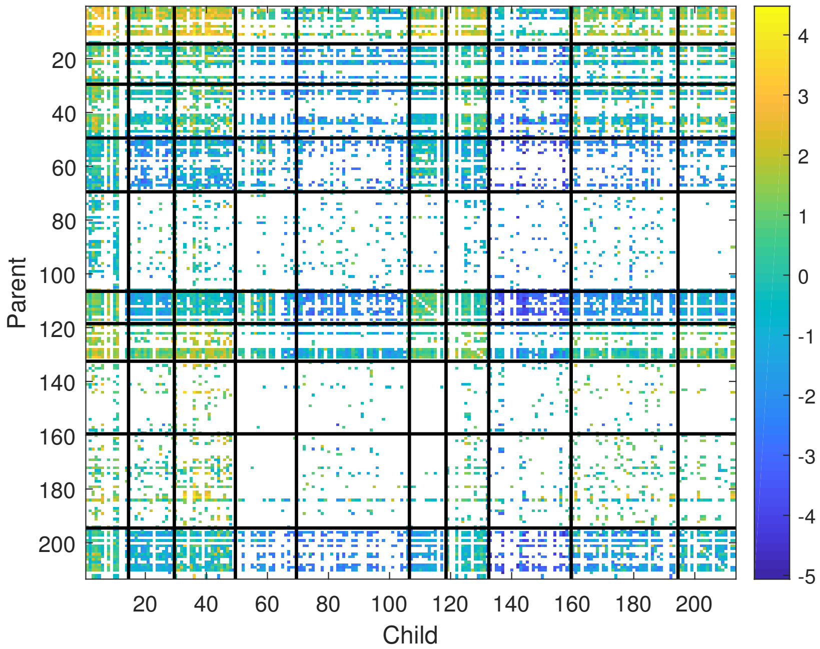
Finding k for pure WSBM with log-normal weight distribution



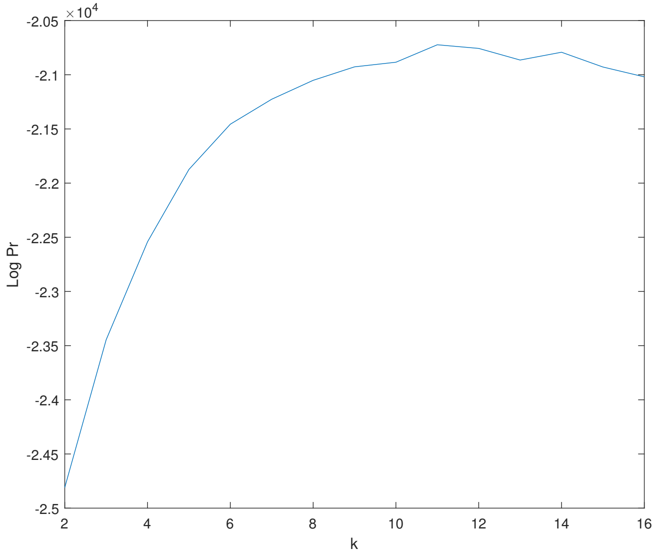
Blocks found by pure WSBM with log-normal weight distribution



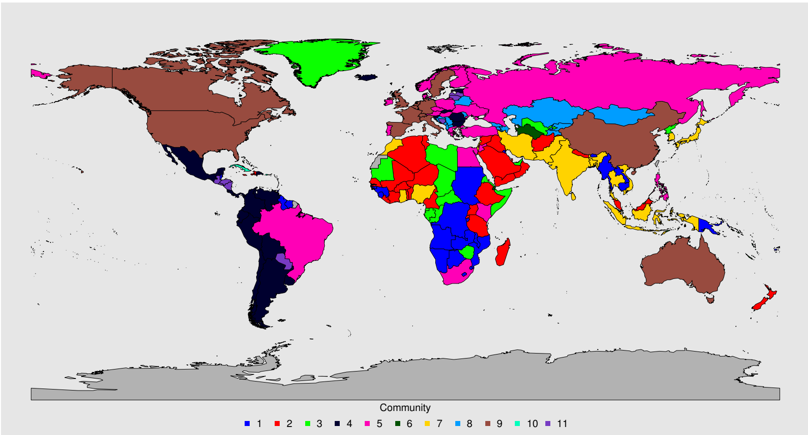
Adjacency matrix sorted by community from pure WSBM



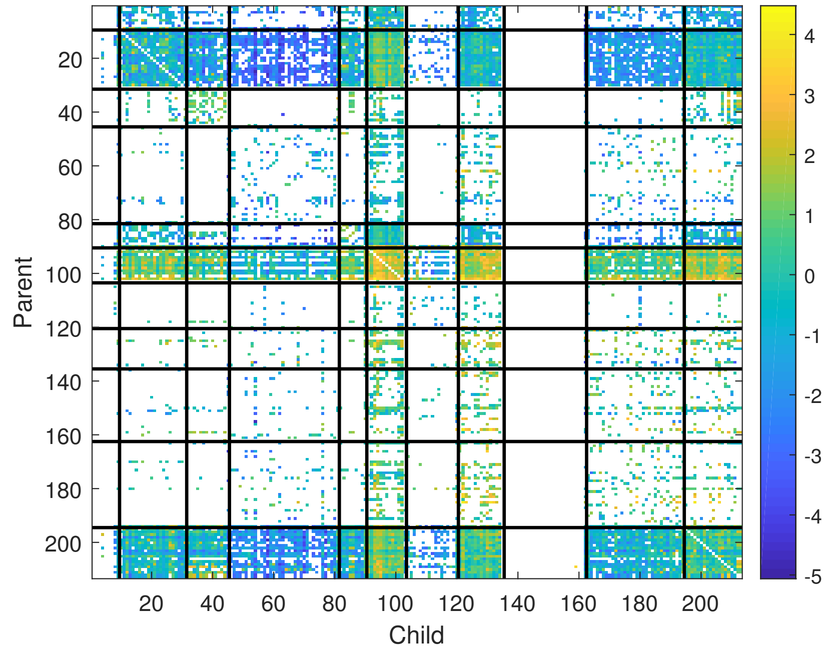
Finding k for WSBM with log-normal weight and Bernoulli edge distribution



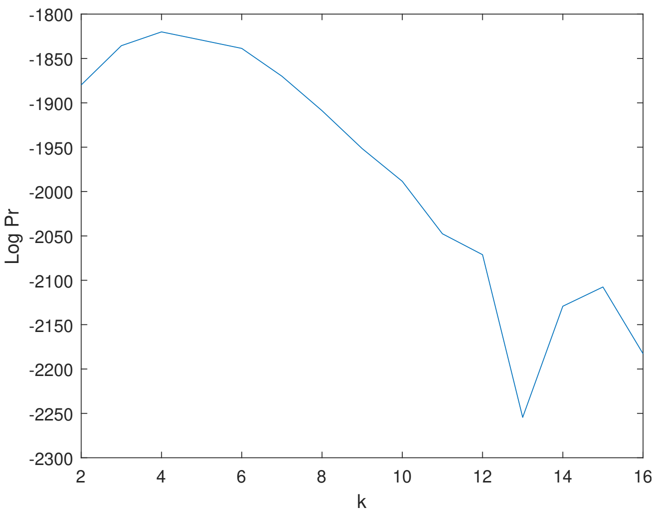
Blocks found by WSBM with log-normal weight and Bernoulli edge distribution



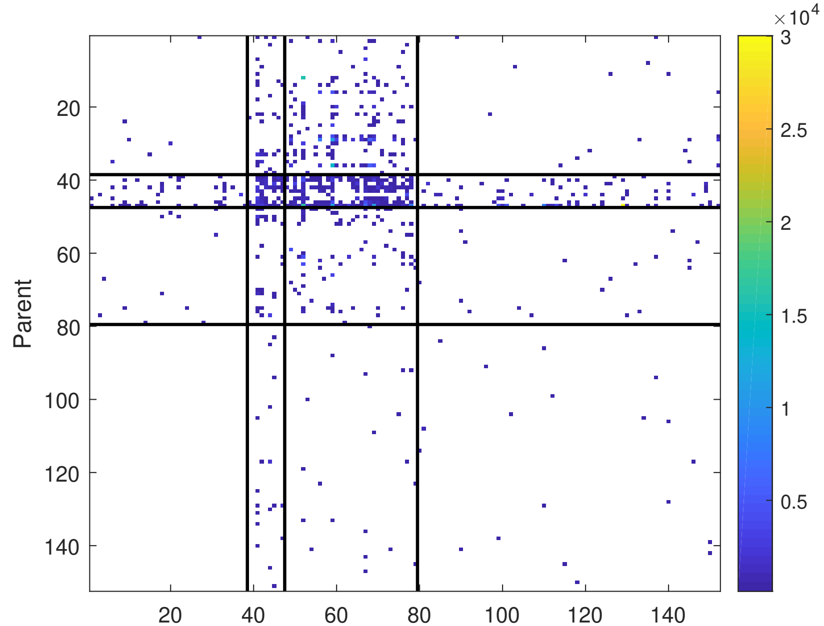
Adjacency matrix sorted by community from WSBM with edge model



Finding k for WSBM $\delta = 100$ network with log-normal weight and Bernoulli edge distribution



Adjacency matrix $\delta = 100$ network sorted by community from WSBM with edge model



Blocks found by WSBM $\delta = 100$ network with log-normal weight and Bernoulli edge distribution

Cluster	Countries
1	Afghanistan; Albania; Algeria; Armenia; Azerbaijan; Belarus; Benin; Bermuda; Bhutan; Bolivia; Bulgaria; Burkina Faso; Cambodia; Chad; Comoros; Cote d'Ivoire; Curacao; Cyprus; Denmark; Dominican Republic; Ecuador; El Salvador; Estonia; Ethiopia; Faeroe Islands; French Polynesia; Georgia; Greece; Guam; Guatemala; Guyana; Haiti; Honduras; Iran, Islamic Rep.; Iraq; Jamaica; Kenya; Korea, Dem. Rep.; Kosovo; Kyrgyz Republic; Latvia; Lesotho; Liberia; Libya; Lithuania; Madagascar; Mali; Malta; Mauritania; Mauritius; Mexico; Moldova; Monaco; Myanmar; New Caledonia; Nicaragua; Niger; Panama; Paraguay; Peru; Rwanda; Senegal; South Sudan; Sudan; Swaziland; Syrian Arab Republic; Tajikistan; Togo; Trinidad and Tobago; Tunisia; Uganda; Uruguay; Uzbekistan; Yemen, Rep.
2	Andorra; Argentina; Bahrain; Belize; Brunei Darussalam; Cameroon; Chile; Costa Rica; Finland; Gabon; Gambia, The; Ghana; Hong Kong SAR, China; Ireland; Israel; Kazakhstan; Kuwait; Luxembourg; Macao SAR, China; Malaysia; Mongolia; Netherlands; New Zealand; Norway; Oman; Puerto Rico; Qatar; Saudi Arabia; Singapore; Slovak Republic; Slovenia; South Africa; Switzerland; Tanzania; Turkey; United Arab Emirates; Venezuela, RB
3	Austria; Bangladesh; Belgium; Bosnia and Herzegovina; Brazil; China; Colombia; Croatia; Czech Republic; Egypt, Arab Rep.; Hungary; India; Indonesia; Japan; Jordan; Korea, Rep.; Lebanon; Morocco; Nepal; Nigeria; Pakistan; Philippines; Poland; Portugal; Romania; Serbia; Sri Lanka; Sweden; Thailand; Ukraine; Vietnam; West Bank and Gaza
4	Australia; Canada; France; Germany; Italy; Russian Federation; Spain; United Kingdom; United States