Imperial College London

SOCIAL ACCESS STATISTIC: LINKING SOCIAL CONNECTIVITY TO HEALTH OUTCOMES

Sahil Loomba, Till Hoffmann and Nick Jones

Department of Mathematics, Faculty of Natural Sciences, Imperial College London

CONNECTIVITY KERNELS FROM EGOCENTRIC SURVEYS

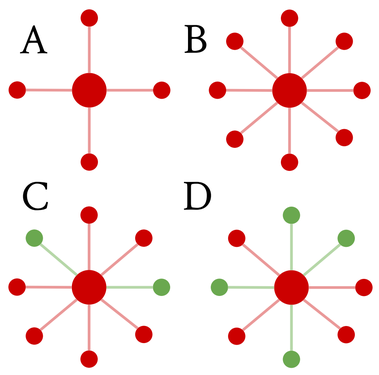
Social networks play a crucial role in determining health. Behaviours such as smoking [1] and outcomes like mental well-being [2] have network correlates of social support and influence. Social surveys are an inexpensive and efficient means of collecting data. Using aggregate relational data (ARD) type questions that ask **"what proportion of your friends are *x*?"**, one can make partial observations of how people make connections [3]. A stochastic block model (SBM) approach is proposed to infer a probabilistic connectivity kernel based purely on ARD-type questions.

Let $\pi \in [0,1]^m$ represent distribution of m observed communities in a society, across a single or multiple dimensions—sex, age, ethnicity and others assumed to interact independently in the model. A network realisation with n people can be sampled as: $z_q \sim \text{Multinomial}(\pi) \forall q \in \{0,1,\dots n\}$, $A \sim \text{Bernoulli}(Z\Psi Z^T/n)$, where $Z \in \{1\}^{n\times m}$ is an assignment matrix and $\Psi \in [0,\infty)^{m\times m}$ is an inter-community affinity matrix. Let ω_i be the mean degree and ρ_i the mean proportion of friends within the community for community i, both of which are simple aggregate statistics from the survey. Inference of the block matrix is straightforward: $\Psi_{ii} = \omega_i \frac{\rho_i}{\pi_i}$ $\Psi_{ij} = \omega_i \frac{1-\rho_i}{1-\pi_i}$

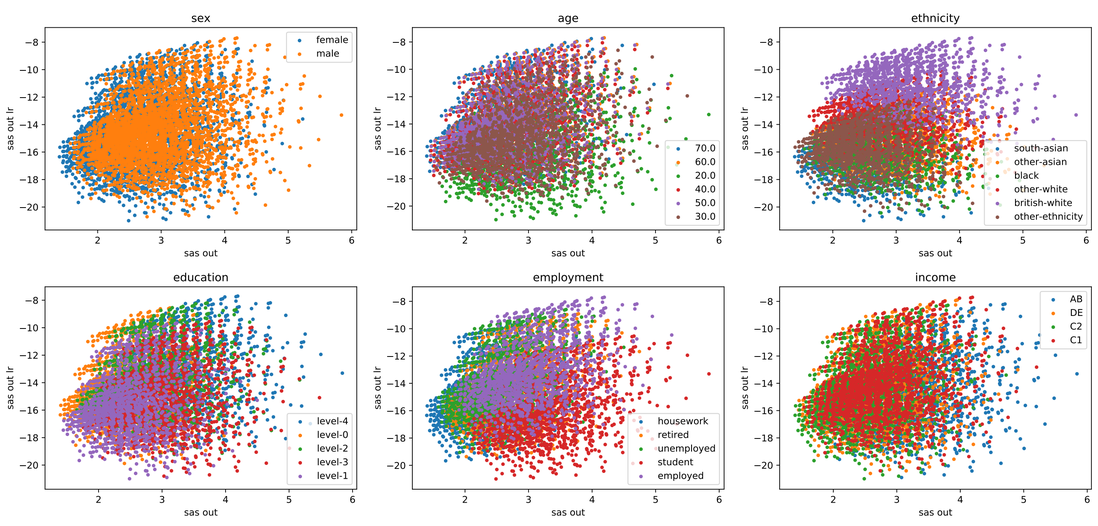
SOCIAL ACCESS STATISTICS

Treating the kernel as a compression of the social network, one can query it for insightful knowledge—a Social Access Statistic (SAS). **This is a network-structural measure of the amount of social capital a person's position in the social network offers them.**

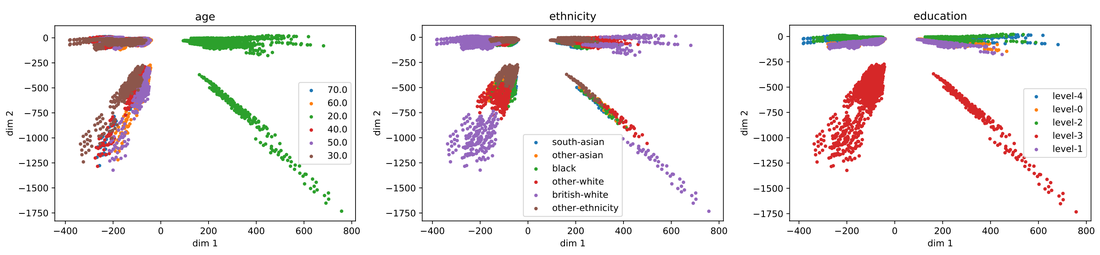
- **Absolute Social Access** is a measure of expected degree of an individual, that is, how many people they are friends with. For an individual of community i, this is given by $[\Psi\pi]_i$, and for an entire region define the "mean" access $\pi^T\Psi\pi$, and the "variance" in access $\sum_i \pi_i \left([\Psi\pi]_i - \pi^T\Psi\pi\right)^2$. Moreover, taking the transpose of Ψ gives an incoming notion of access.
- **Relative Social Access** is a measure of expected level of heterophily experienced by an individual, that is, how different their friends are from themselves. This is defined analogously, after a log-ratio of off-diagonal elements with the respective diagonal elements of Ψ.



Schematic of Social Access Statistics capturing forms of social capital. Absolute SAS indicates the expected degree of an individual (A, B) while Relative SAS (or SAS-lr) indicates the expected heterophily of an individual (C, D). These statistics are directed, and can indicate outgoing or incoming connections. They can be calculated for every individual, or for an entire region, depending on π of the regional population.



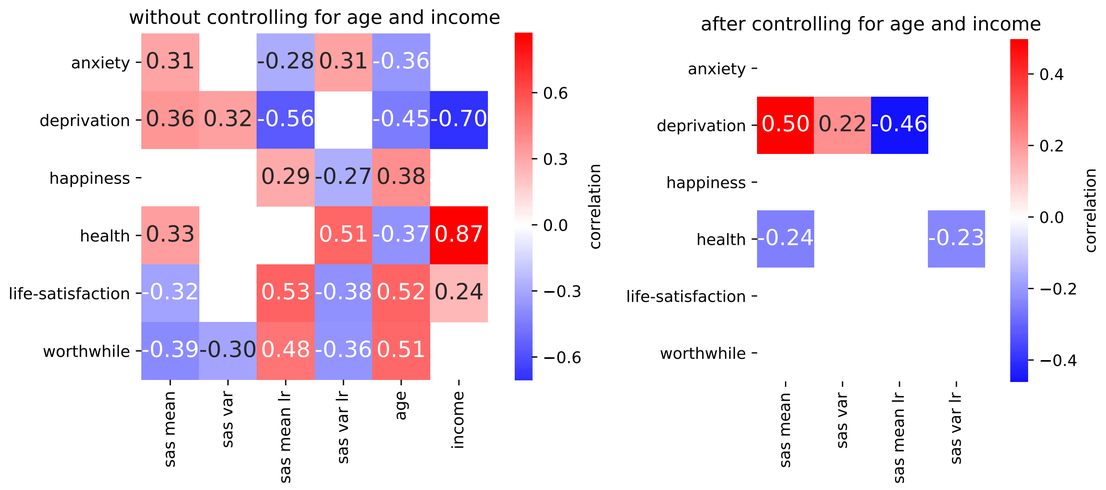
Absolute vs. Relative Social Access of all 7200 communities in the UK—along 6 observed dimensions of sex, age, ethnicity, education, employment and income—assuming π is the national distribution. Note that usually marginalized communities, such as ethnic minorities and the unemployed, experience lower absolute and relative social access.



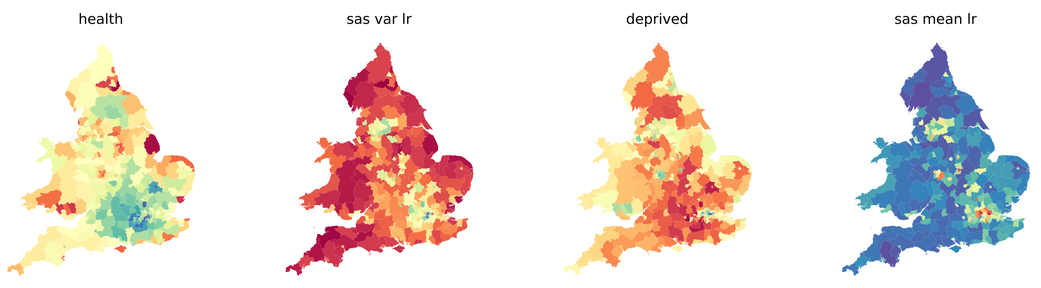
Top-2 Principal Components of the expected adjacency matrix generated from Ψ, assuming π is the national distribution. Dimension 1 seems to capture age, dimension 2 is indicative of education, while ethnicity appears to be arranged radially.

ANALYSIS: UNDERSTANDING SOCIETY AND UK CENSUS

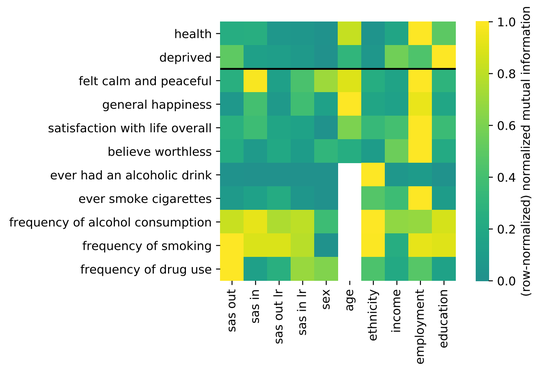
Block matrix Ψ is learnt from the 2011 wave of Understanding Society survey, which provides data along 6 observed sociodemographic dimensions for 43k respondents [4], alongside indicators of subjective well-being. Microdata of 2.8 million respondents in the 2011 UK Census is used to obtain distributions π, self-reported health and objective well-being (deprivation) [5]. Aggregate statistics on subjective well-being for 2011-13 at the grouped local authority level in England and Wales are obtained from the Census [6].



Social Access explains variation in self-reported health and objective well-being even after controlling for age and income. Analysis performed for 265 grouped-local-authorities (GLAs) across England and Wales, p-value<0.01 (Bonferroni corrected). While absolute SAS is positively correlated to deprivation, relative SAS is negatively correlated, suggesting that heterophily reduces deprivation while large degree increases it.



Health and Deprivation relate to Social Access as evident in maps of England & Wales.



Social Access is indicative of frequency but not adoption of substance abuse. Calculating the Normalized Mutual Information between Social Access and Well-being at the microdata level reveals that while adoption of behaviours of substance abuse are not informed by social connectivity, their frequency seems to be. Moreover, while alcohol consumption is informed by incoming social access, drug use is marked by outgoing social access.

SUMMARY AND FUTURE DIRECTIONS

We have developed a novel probabilistic framework to learn connectivity kernels from egocentric social survey data that can generate meaningful statistics capturing social access. **These statistics relate to self-reported health and material deprivation, even after controlling for age and income, indicating a role of social connectivity.** Future work includes:

- Relate social access to more diverse health outcomes and behaviours
- Derive other social access statistics, like a betweenness centrality analogue that measures "bridging" social capital
- Learn separate connectivity kernels for different regions of the UK
- Design and conduct our own survey in the UK and India

REFERENCES

Nicholas A Christakis and James H Fowler. "The collective dynamics of smoking in a large social network". In: *New England Journal of Medicine* 358.21 (2008), pp. 2249–2258.

Milton Greenblatt, Rosina M Becerra, and EA Serafetinides. "Social networks and mental health: An overview.". In: *The American Journal of Psychiatry* (1982).

Emily Breza et al. *Using Aggregated Relational Data to feasibly identify network structure without network data*. Tech. rep. National Bureau of Economic Research, 2017.

Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 11th Edition. http://doi.org/10.5255/UKDA-SN-6614-12. University of Essex, Institute for Social and Economic Research, NatCen Social Research, Kantar Public, UK Data Service, 2018.

UK Census Microdata. https://www.ons.gov.uk/census/2011census/2011censusdata/censusmicrodata/safeguardedmicrodata. Office of National Statistics.

UK Personal Well-being Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/datasets/headlineestimatesofpersonalwellbeing. Office of National Statistics.