

General comments on the comments

- I've introduced a new numbering scheme for tables and equations in the appendix / supplementary material. We can revisit the exact format once we submit (e.g. Table S1 or Table B1, etc.).
- I have not yet added the Research Now data back in. My hunch is that, while those data confirm the overall picture, they are rather low quality because the response rate was lower than we hoped for. Having said that, adding another dataset now seems more feasible after moving the details to the appendix as you suggested. I'll dig in a bit more...
- I've had another go at the abstract and discussion in light of your comments—thank you!
- Regarding terminology, I've adopted the following conventions (which I hope clarifies things; happy to consider alternatives)
 - social separation refers to the pairwise segregation measure
 - social isolation refers to the individual segregation measure
 - social strain refers to the societal segregation measure
 - segregation measure is the collective term encompassing all of the above
 - distance is only used in the context where the social separation is a (semi) metric
- I've adopted the undirected approach throughout. Directed edges give a little more flexibility, but, as you pointed out, realistic reciprocity is likely more important than model flexibility.

See below for responses to comments. I hope any comments I don't address here are fully resolved.

Page 2

Quantifying homophily is not only important for understanding why social ties form between some people yet not between others, but the manifestation of homophily as poorly-connected social networks can have a significant impact on dynamics unfolding on social networks~\cite{Golub2012}.

how is citation 2 relevant here?

Golub et al. (2012) look at consensus times in social networks and found that homophily slows down the process because the network is not as well connected compared with a random graph.

The measures capture segregation at different scales: pairs of individuals, single individuals, and society as a whole.

re order and put single individuals first

Done. Although I'm interested in the reason for your order preference. I was thinking that pairs, individuals, society is a coherent progression up the "scales" of society. But maybe that's too early in the manuscript?

Page 4

Similarly, members of the ethnic majority have more social ties in social networks in US high schools~\cite{Currarini2009}.

Just a comment – could be mentioned in discussion

A natural alternative model is to consider radiative models of link formation, but these, by construction can eliminate some effects of isolation in Blau-space: likely the true model is some merger of the two "If my expected degree according to the geometric model is \bar{d} , I pick a number 1 up to F uniformly and then pick an additional number of "stretch" friends according to k -NN". This is a "stretch if you have-to" or "semi-radiative" model that partly respects local geometry but warps it if the degree is too low.

The nice thing about models like this is that they still give you a semi-metric *and a characteristic spatial scale* when you fit them (because there is a geometric part to things). They give isolated people in blau space a low degree but not no degree and they allow density variations to be partially compensated for.

I think these weird fusions of radiative and geometric network models are quite interesting on their own terms.

Yes, agreed that these semi-geometric models could be interesting. My hunch is that it's too much to add other than in passing. I added a sentence to the discussion.

\text{where } \logit \rho \triangleq \log \left(\frac{\rho}{1 - \rho} \right)

Slightly odd formatting.

I used the "where" in the equation to avoid having a lone "where" on its own line (because it's followed by another display-style equation). Left unchanged for now, but happy to change if you have a strong preference.

Page 5

If the connectivity kernel is homophilous, symmetric, and homogeneous, the pairwise segregation measure is a \emph{semi-metric}~\cite{Wilson1931}: it satisfies the properties of a metric, including non-negativity, the identity of indiscernibles, and symmetry with respect to exchange of arguments\hruleexcept the triangle inequality.

do you need to be a homogeneous kernel – could it be enough to merely be symmetric?

Consider the pairwise segregation measure under exchange of arguments:

$$\varphi(x, y) - \varphi(y, x) = \text{logit}\rho(y, y) - \text{logit}\rho(x, y) - \text{logit}\rho(x, x) + \text{logit}\rho(y, x).$$

If $\rho(x, y) = \rho(y, x)$, the second and fourth terms cancel, and we have

$$\varphi(x, y) - \varphi(y, x) = \text{logit}\rho(y, y) - \text{logit}\rho(x, x).$$

So, as you suggested, we don't actually need the kernel to be homogeneous (translationally invariant), we just need $\rho(x, x)$ to be constant for all x . Text is updated.

In the context of the semi-metric proposition.

I think, given the formulation of the proposition, you also need to give an example of how it does not always satisfy the triangle inequality.

Or you could say that the connectivity kernel is at least a semi-metric since it necessarily satisfies 3 conditions.

I added a pointer to the section where we discuss kernels that give rise to a proper metric.

where x and y denote the blocks and δ_{xy} is the Kronecker delta

you've some unclarified degeneracy between membership of a block and the block itself. Maybe you should unpick this (it's a bit cumbersome – but on the other hand this is the first example that we give)

I've changed to explicit block membership (as an attribute) in an attempt to clarify, but would be good to get your eyes on that section again.

Page 6

\text{where } \gamma = \sum_{x=1}^K P(x) \left(1 - P(x)\right)

Just check you're sure about this. I wasn't immediately sure – (worrying about whether you were doing the counting right for replacement). Maybe define K specifically.

For the stochastic block model, we have

$$\varphi(x, y) = (1 - \delta_{xy})\xi,$$

where $\xi = \text{logit}\rho_{\text{same}} - \text{logit}\rho_{\text{different}}$. Averaging with respect to the attribute of the alter y yields

$$\phi(x) = \xi \sum_{y=1}^K (1 - \delta_{xy})P(y) = (1 - P(x))\xi$$

because there are contributions for all $y \neq x$. Averaging again, we get the result in the manuscript

$$\Phi = \xi \sum_{x=1}^K P(x)(1 - P(x)).$$

We could add the steps to the manuscript, but my hunch is that it's too much. I've added a few extra equation references.

Page 7

Didn't you do some sanity check synthetic inference? Could these be put in a supplement? If possible it will strengthen the ms.

Yes, I'll add them to the appendix, but thought it would be good to go through another iteration in parallel.

Page 17

Regarding the visual map of Blau space figure.

If a figure like this could be produced for USoc and also displayed that would be cool rhetorically. Helps make the case for Universal Metrics.

Yes, possible in principle, but it might be tricky in practice: USoc has more features than GSS, and it'll be tricky to try and visualise them because I've already used colour and shape for each point in the scatter plot. Left it for now, but let me know if it would contribute significantly.

Page 19

Regarding discussion of standardised regression coefficients vs social strain statistic.

I get what you're saying here – but it adds nuance (murkiness :)) at exactly the time when we need to be clear. The kernel (though not θ) shouldn't depend on $P(x)$ – in the sense that if I added lots more people of age t that shouldn't affect the inferred kernel (all else equal). But that will massively affect ϕ . That message about the separation of the kernel and $P(x)$ is central for our argument.

Could you put it another way? That generically there is no necessary relationship between the properties of the kernel and the properties of ϕ – because the latter depends on the combination of both the kernel *and* the blau-space co-ordinates of the society? [There might be a relationship between the parameters θ and ϕ – but that depends on how the individual f is scaled]

I've reordered the argument to highlight that Φ captures the effect of both the kernel and the attribute distribution. And then discuss the shortcomings of using standardised regression coefficients as a measure of importance instead.

Page 20

Our observations, together with a study by \textcite{Mossong 2008} finding that “mixing patterns [\ldots] were remarkably similar across different European countries”, suggest that a universal connectivity kernel for friendships may exist. To test this hypothesis, further surveys should be conducted in a unified fashion to minimise the effects of question wording and how the survey is administered.

common default

I couldn't quite figure out which bit the “common default” comment referred to.