Name Your Friends, but Only Five?

Report

Patrick Altmeyer

05 May, 2021

Brief summary

Surveys-based research of peer effects in networks usually set a cap on the number of neighbours ("friends") of individuals. For example, surveys would typically ask individuals to nominate at most five friends. In this context Griffith (2020) defines the property of *order irrelevance*.

Definition 1 (Order irrelevance)**.** The distribution of peer characteristics does not change across nomination ranks. In the context of censoring this implies that the characteristics of censored and hence ubobserved peers follow the same distribution as those of observed peers.

Griffith (2020) works with the same linear-in-means model we discussed in class, so I will refrain from restating this here. Let k_i^* denote the true number of friends of individual i. Then the interesting point is that with censoring ($k < k_i^*$) we have $\tilde{g_{ij}}^{(k)} = \frac{1}{k}\sum_{j=1}^{k} g_{ij}$, which is of course different from $\tilde{g_{ij}}^* = \frac{1}{k_i^*}\sum_{j=1}^{k_i^*} g_{ij}$. The same holds for the sample means of outcome variable y and exogenous characteristics x, as discussed in Griffith (2020). These deviations from the "true" model also introduce deviations in the reduced-form (RF) parameter estimates of peer effects. In particular, Griffith (2020) uses a number of simplifying assumption to derive two main analytical results. Firstly, assuming that individual characteristics are randomly assigned we have for the RF estimator that

$$\hat{\alpha}_{RF}^{(k)} \rightarrow \begin{pmatrix} \beta_2 \\ \frac{\mathbb{E}\frac{1}{d^*}}{\mathbb{E}\frac{1}{d^{(k)}}}\beta_3 \end{pmatrix} \tag{1}$$

where as in class β_3 measures peer effects. Clearly $d^* > d^{(k)}$ and hence $\hat{\beta}_3 < \beta_3$, the downward bias discussed above. The second result (not restated here) shows that the bias works in the opposite direction as the correlation between individual i's characteristics and those of her friends. In case of homophily then, β_3 is attenuated. The author further shows that in the same setting β_2 is inflated, so updward-biased.

Griffith (2020) then turns to real data sets and in two empirical experiments lets the censoring threshold k. In the interest of brevity we will not cover the findings in detail here, but simply summarize in the author's words:

> "Results are clear: the more links that are censored, the more attenuated are peer effects parameters; that is, peer effects parameters trend away from zero as more links are observed." — Griffith (2020)

In light of these findings the author proposes a way to bias-correct the downward-biased peer effect parameters estimated from a reduced-form model, once again under the assumption of order irrelevance. A complete description of the approach is beyond the scope of this summary, but at a high level it involves the following transformation

$$\hat{\alpha}_{RF}^{(k)} = \mathbf{B}_k \alpha$$
$$\alpha = \mathbf{B}_k^{-1} \hat{\alpha}_{RF}^{(k)}$$

where $\mathbf{B}_k$ can be estimated and - when working under the same assumptions as before - is of the following intuitive form:

$$\mathbf{B}_k = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k \end{pmatrix}$$

Thoughts

Griffith (2020) picks up on an important issue, demonstrates the issue empirically and then provides a remedy to it. All in all this paper is therefore well set-out and makes some interesting methodological contributions. A common problem associated with the computations involved in network analysis is the curse of dimensionality, which likely plays a role in authors choosing to resort to censoring in the first place. Using the bias-correction proposed here, one may circumvent the issue of high dimensionality and still mitigate some of the issues that arise and as discussed in Griffith (2020). One potential caveat is the reliance on relatively strong assumptions about distributional properties of individual characteristics as the author himself confesses. I am not sure, for example, if we should expect the assumption of order irrelevance (1) to hold: is there not reason to assume that individuals who are ranked higher on average (hubs) have very different individual characteristics on average than those that remain unobserved?

References

Griffith, Alan. 2020. "Name Your Friends, but Only Five? The Importance of Censoring in Peer Effects Estimates Using Social Network Data."