

# Network endogeneity and ways to deal with it

Final report

Patrick Altmeyer

16 June, 2021

## Motivation

In the existing literature network exogeneity is commonly assumed: network properties are assumed to be independent of individual characteristics. This assumption is restrictive and in fact often violated in real-world applications. The most commonly cited example is peer-effects in the context of academic performance, where the assumption of network exogeneity imposes that there are no unobserved factors that affect both the formation of friendships and scholarly outcomes (**hsieh2016social?**). Clearly there are unobservable factors, for example concerning students' shared interests or family background, that may affect both the friendships they choose and how well they perform in school (**johnsson2021estimation?**). Other examples of endogenous networks come to mind: a central bank may be interested in studying the effect of interconnectivity in interbank money markets on the interest rates that individual market participants pay and receive. While there are observable variables such as size or geographical location of banks that may affect both their interconnectivity and the interest rates they face, it is easy to think of unobserved factors that lead to network endogeneity such as the banks' business models.

Unobserved individual heterogeneity is not unique to the study of peer effects. The literature on causal inference in social sciences has come up with a number of tools to deal with endogeneity, perhaps most notably the use instrumental variables (see for example Morgan and Winship 2015). Instrumental variables that affect the outcome only through the endogenous regressor are often hard to find and particularly so in the context of endogenous networks: linking this back to the example of students, it is difficult to think of an effective instrument without even knowing the mechanism of network formation (**hsieh2016social?**). Another popular way to deal with endogenous regressors is commonly referred to as the control function approach originally introduced by (**heckman1985alternative?**). Instead of attempting to resolve the issue of endogeneity by instrumenting the endogenous regressors, the control function approach tackles endogeneity head-on by modelling the endogeneity in the disturbance term of the model. This report reviews a number of recent papers that use the control function approach to resolve the issue of network endogeneity. (**johnsson2021estimation?**) is the most recent study that has emerged from that body of literature and will serve as the baseline for this review.

Morgan, Stephen L, and Christopher Winship. 2015. *Counterfactuals and Causal Inference*. Cambridge University Press.