ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis: International Bachelor in Econometrics and Operations Research

Treatment and Spillover Effects Under Network Interference

July 5, 2019

Author:	Louis Perriens, 429200
Supervisor:	dr. Sun, Y.
Second assessor:	dr. Fok, D.

Abstract

In this paper, treatment and spillover effects are estimated when the SUTVA assumption is violated, using non-parametric and regression methods. We calculate standard errors robust to network interference as in Leung (2019). Their results are then extended to obtain quantile treatment effects.

Keywords: Treatment effects; Spillover; Network formation; Quantile effects.

Contents

1 Introduction

Network models differ from most models in causal inference, as it is usually assumed that individuals are independent, in the sense that treatment only affects the treated unit. However in practice, it is common that individuals interact among each other, forming a network. Network formation could occur in settings where units in a study are geographically close, or when they interact among each other in social settings such in schools or at workplaces. Such interactions lead to what is referred to as spillover effects and network interference, where the outcome of a single individual could indirectly intervene on outcomes of multiple observations in the same network, regardless of whether they have received treatment.

In this paper, we study treatment and spillover effects, under the assumption that spillovers are mediated by a network which is partially observable, where treatment is assigned randomly. In such cases, the assumption Rubin (1990) refers to as the Stable-Unit Treatment Value assumption (SUTVA), which is almost systematically required for most causal inference methods, does not hold. The SUTVA states that an individual's treatment response should only depend on the treatment received,i.e. should be invariant of the treatment allocation to other units. Yet, there are many scenarios in which this assumption will fail to hold, as discussed in Leung (2019). For example, in vaccine studies the outcome of an individual would not only depend on an individual's treatment, but also on whether units around this individual received vaccination (Halloran and Struchiner, 1995). In most studies about causal effect under network interference, researchers collect data on plausibly independent networks, or partition networks into presumably independent sub-networks according to some specified rules in order to calculate clustered standard errors for their estimates. Such methods rely on the assumption of ”partial interference” (Sobel, 2006), allowing interference within clusters, but not across clusters. However, such methods ignore possible links across networks, thus could lead to wrong results.

As stated above, most studies in causal inference rely on the assumption that observations are independent (Rosenbaum and Rubin (1983), Rubin (2006), Van der Laan et al. (2003)). According to van der Laan (2014), there is lack of methodological framework when the SUTVA does not hold, and for this reason many practical studies assume interference away. However, ignoring interference could lead to biased estimates and wrong conclusions: when interference is present, the difference in means of the treated and control groups does not measure the average treatment effect, but the difference in means of two sub-populations (Sobel, 2006). Studying causal inference under general interference is particularly relevant, as implementing models with a single network rather than partitioning the network into plausibly independent sub-networks will severely reduce costs for researchers. This eliminates the need to get geographically isolated samples, and removes concerns of links between clusters under the partial interference assumption.

Manski (2013) studied the identification of treatment effect estimates in the presence of interference, making the response to treatment a function of the vector containing treatment assignments for the entire sample.

He began with identification of treatment response under the SUTVA, then extended his results to cases where social networks were present. Athey et al. (2018) propose a number of new methods for testing hypothesizes in single-network interference settings. These include testing the significance of treatment effects under network interference, testing the presence of high-order spillovers effects, and testing for peer effect heterogeneity. Liu and Hudgens (2014) study causal inference under the partial interference assumption, and derive asymptotic properties of the estimators for large samples. Aronow et al. (2017) also study the estimation of causal effects under general interference for a single network. In contrast to Leung (2019), who imposes restrictions on the structure of the network to induce weak-dependence between units, Aronow et al. (2017) impose conditions on the potential outcomes.

In a first part of this paper, we shall replicate the research done by Leung (2019), who proposes a super-population model, where a single large network is considered, thus bypassing the downfalls of clustered standard errors. The paper provides estimators for treatment and spillover effects under network interference. The latter are consistent and asymptotically normal under a set of conditions discussed in Leung (2019). Furthermore, an estimator for the variance robust to heteroskedasticity and network dependence is introduced. We shall first generate the network and outcomes for individuals in MATLAB, following the procedure in Leung (2019). Then, the treatment and spillover effects will be estimated through linear regressions and non-parametrically. Finally, we shall compare the heteroskedasticity-consistent standard errors with the new standard errors proposed in Leung (2019) as a performance check.

Next, we propose some extensions to Leung (2019). We extend the average structural function by implementing its quantile analogue, to obtain quantile treatment and spillover effect. The linear regression is also extended to a quantile-based linear regression, where observations above specified quantiles are assumed to be unobserved when estimating the model. Quantile methods can often be more informative than sample mean measures: Melly and Wüthrich (2016) argue that a policy which increases the lower tail of some income distribution would be perceived as better than one shifting the median income, while the average treatment effects for both policies would be the same.

2 Model

2.1 Network Formation

The network is represented by a $n \times n$ matrix A, where $A_{ij} = 1$ if i and j are connected in the network, and $A_{ij} = 0$ otherwise. To obtain A, we initially randomly generate geographical locations $\rho_i \in [0,1]^2$,$i = (1,..,n)$, where n is the sample size, and calculate the distance between each pair i and j as $r^{-1}||\rho_i - \rho_j||$. The distance function $d(i,j)$ takes value 0 if the distance $r^{-1}||\rho_i - \rho_j||$ is less or equal than 1, and infinity otherwise. The scaling factor r, which equals $(\kappa/n)^{1/d}, \kappa > 0$, ensures the sparsity of the formed network, as real-world networks are significantly smaller than the total population. Next for each i, we generate $\alpha_i \in \{0,1\}$ an individual-specific characteristic, such as being above a certain age or not. The network links are generated as:

$$A_{ij} = \mathbb{1}\{\theta_1 + \theta_2(\alpha_i + \alpha_j) + \theta_3 d(\rho_i, \rho_j) + \theta_4 S_{ij} + \zeta_{ij} > 0\}, \tag{1}$$

where S_{ij} represents a strategic interaction between units i and j, and $\zeta_{ij} \sim N(0,1)$ is a random-utility shock. We include the strategic interaction $S_{ij} = \max_k B_{ik}B_{kj}$, where B is the initial network of geographical neighbours, computed as $\mathbb{1}\{r^{-1}||\rho_i - \rho_j|| \leq 1\}$. Then, S_{ij} takes value 1 if i and j share a common neighbour k, 0 otherwise. Next, we impose that there are no self-links, implying $A_{ii} = 0$ for all individuals, and make A symmetric, as Leung (2019) assume that A is an undirected network.

We assign treatment $D_i = 1$ if unit i is treated, 0 otherwise. We then define $\gamma_i = \sum_{j=1}^{n} A_{ij}$, the number of neighbours for observation i, and $T_i = \sum_{j=1}^{n} D_i A_{ij}$, the number of treated neighbours for the same observation, to define $W_i = (Y_i, D_i, T_i, \gamma_i)$, the data observable for individual i. We further assume that the 1-neighbourhood the i's is observable, i.e. the set of neighbors j directly to i in A, as it is crucial for the construction of standard errors. We generate the outcomes $Y_i \in \mathbb{R}$ according to the model used to describe the treatment and spillover effects, as shall be shown in 2.2.

2.2 Outcome Generation

For the non-parametric estimate, we generate the outcomes Y_i according to the random-coefficient outcomes model:

$$Y_i = \theta_{i1} + \theta_{i2}D_i + \theta_{i3}T_i + \theta_{i4}T_i^2 + \theta_{i5}T_i\gamma_i, \tag{2}$$

where the parameters are independent, identically distributed and θ_{i1} captures the unobservable peer effects. whereas for the linear regression model, we shall generate the Y_i's as:

$$Y_i = \beta_1 + \beta_2 T_i + \beta_3 T_i/\gamma_i + \epsilon_i, \tag{3}$$

where $\epsilon_i = \gamma_i^{-1} \sum_j A_{ij}\nu_j$ for $\nu_i \sim N(0,1)$ captures the unobserved heterogeneous peer effects.

2.3 The Models

As it is common in practice, we shall use the linear-in-means with no endogenous effect regression model:

$$Y_i = \beta_1 + \beta_2 D_i + \beta_3 T_i/\gamma_i + \epsilon_i, \tag{4}$$

The treatment effect is given by β_2 and the spillover effect by β_3. However, even robust and clustered standard errors are invalid, as the data will not be identically and independently distributed across the population sample: the number of treated neighbours T_i will be correlated across the sample and ϵ_i's could also be correlated. Thus, we prefer to use a more general non-parametric model given by

$$Y_i = r(D_i, T_i, \gamma_i, \epsilon_i),$$

where $r(\cdot)$ is a real-valued function which does not assume particular distributions of its parameters.

3 Estimation and Inference

3.1 Non-Parametric Estimation

We define the conditional Average Structural Function (ASF) as follows:

$$\mu(d, t, \gamma) = r(d, t, \gamma, \epsilon_1(\gamma)). \tag{5}$$

The ASF gives the mean outcome value for a group with a given number of treated neighbours t and treatment assignment d. Through the conditional ASF, one can capture the average treatment and spillover effect, defined as

$$\mu(d, t, \gamma) - \mu(d', t', \gamma),$$

where $d, d' \in \mathbb{B}$, $t, t' \in \mathbb{N}$, and $t, t' \leq \gamma$.

Note that the estimates implicitly depends on the number of units for which $\gamma = \gamma_1$. The degree of sampled units is fixed at $\gamma = \gamma_1$ in order to identify the average treatment and spillover effects. Within this subpopulation, we compare outcomes of units with treatment assignment d and treated neighbours t versus those with treatment assignment d' and treated neighbours t', with $d, d' \in \{0, 1\}$, and $t, t' \in \mathbb{N}$ and $t, t' \leq \gamma$. For a fixed t, variation in d, d' identifies the treatment effect, and for a fixed d, variation in t, t' captures the spillover effect. In this paper, we will use the sample analog to the conditional ASF:

$$\hat{\mu}(d, t, \gamma) = \frac{\sum_{i=1}^{\tilde{n}} Y_i \mathbb{1}_i(d, t, \gamma)}{\sum_{i=1}^{\tilde{n}} \mathbb{1}_i(d, t, \gamma)}, \tag{6}$$

where $\mathbb{1}_i(d, t, \gamma) = \mathbb{1}\{D_i = d, T_i = t, \gamma_i = \gamma\}$.

Then, the average treatment/spillover effect is given by the difference

$$\hat{\mu}(d,t,\gamma) - \hat{\mu}(d',t',\gamma). \tag{7}$$

To define the variance estimator, we first need to introduce the terms:

$$a_i = \frac{\mathbb{1}_i(d,t,\gamma)}{\hat{\rho}(d,t,\gamma)} - \frac{\mathbb{1}_i(d',t',\gamma)}{\hat{\rho}(d',t',\gamma)},$$

$$b_i = \hat{\mu}(d,t,\gamma)\frac{\mathbb{1}_i(d,t,\gamma)}{\hat{\rho}(d,t,\gamma)} - \hat{\mu}(d',t',\gamma)\frac{\mathbb{1}_i(d',t',\gamma)}{\hat{\rho}(d',t',\gamma)},$$

$$\hat{\rho}(d,t,\gamma) = \frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}} \mathbb{1}(d,t,\gamma).$$

Furthermore, for each pair (i,j), we create a link in the dependence matrix G, where $G_{ij} = 1$ if observations i and j are connected, share a common neighbor k, or if they are the same unit:

$$G_{ij} = \mathbb{1}\{A_{ij} + \max_k A_{ik}A_{kj} + \mathbb{1}\{i = j\} > 0\}, \tag{8}$$

where we impose symmetry on G.

Then, we obtain standard errors for the treatment and spillover effects using the variance estimator $\hat{\sigma}^2_{TS}$ defined as:

$$\hat{\sigma}^2_{TS} = \frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}}\sum_{i=1}^{\tilde{n}} G_{ij}((Y_i a_i - b_i)(Y_j a_j - b_j)). \tag{9}$$

3.2 Linear Regression

For the linear regression, we estimate $\beta = (\beta_1, \beta_2, \beta_3)$ and define the error terms for the regression as $\epsilon_i = \gamma_i^{-1}\sum_j A_{ij}\nu_j$ and $\nu_j \sim N(1,1)$. The variance-covariance matrix is defined as:

$$\hat{\Sigma} = (X'X)^{-1}M'GM(X'X)^{-1}, \tag{10}$$

where $X_i = (1, D_i, T_i/\gamma_i)$, and $M = (X_1\hat{\epsilon_1}, .., X_n\hat{\epsilon}_{\tilde{n}})$, and G is defined as in (8). By letting G equal the identity matrix, we obtain standard Eicker-Huber-White standard errors (HC), corresponding to the case of independent observations when the SUTVA holds. Using these variance estimator, we shall calculate standard errors as $SE = \frac{\sigma}{\sqrt{\tilde{n}}}$, and shall compare the new standard errors estimators robust to network dependence to the HC standard errors. Intuitively, one would except the standard errors obtained using (10) to be larger than conventional heteroskedasticity-consistent (HC) standard errors, as we would expect $G - I$ to be positive definite.

4 Quantile-Based Estimation

4.1 Non-Parametric Estimation

In this section, we extend the results of Leung (2019) by replacing the conditional ASF by its quantile analogue, the conditional quantile structural function (QSF). First, we generate the network and the outcomes as in (1) and (3) respectively. For $\tau \in (0,1)$, we set the τth QSF equal to the largest $\mu(d,t,\gamma)$ in the τth quantile.

In this paper, we use sample analogues for $\mu(d,t,\gamma)$ as defined in (6). Thus, to obtain quantile estimates, it suffices to order the Y_i's, and solely analyze units i whose outcomes Y_i is lower than the τth quantile, while keeping unobserved heterogeneity ϵ_i unchanged for the whole sample. With the cumulative distribution function $F_Y(y) = P(Y \leq y)$, we can define the τth quantile of Y as $Q_Y(\tau) = inf\{y : F_Y(y) \geq \tau\}$. Next, let us define the set $\omega_\tau = \{i : Y_i < Q_Y(\tau)\}$, containing the units below the τth quantile. The conditional QSF can then be defined as:

$$\hat{\mu}(d,t,\gamma,\tau) = \frac{\sum\limits_{i \in \omega_\tau} Y_i \mathbb{1}_i(d,t,\gamma)}{\sum\limits_{i \in \omega_\tau} \mathbb{1}_i(d,t,\gamma)}, \tag{11}$$

and the quantile treatment/spillover effects are defined as:

$$\hat{\mu}(d,t,\gamma,\tau) - \hat{\mu}(d',t',\gamma,\tau). \tag{12}$$

Next, we define the new dependence graph G_τ as:

$$G^\tau_{ij} = \mathbb{1}\{A_{ij} + \max_k A_{ik}A_{kj} + \mathbb{1}\{i=j\} > 0\}, (i,j) \in \omega^2_\tau \tag{13}$$

Then, we can obtain the variance estimator for quantile treatment and spillover effects as in (9), using G_τ, and the variables a_i, b_i, and $\hat{\rho}(d,t,\gamma)$ defined as above, for units satisfying $i \in \omega_\tau$.

4.2 Quantile-Based Truncated Regression

Lastly, as more of an exploratory exercise, we extend the linear regression to a quantile-based truncated regression. In this exercise, we generate A as in (1), and for each unit we obtain $W_i = (Y_i, D_i, T_i, \gamma_i)$, obtained Y_i's as in (3). Next, for a chosen quantile τ, we run the regression

$$Y_i = X_i\beta + \epsilon_i, i \in \omega_\tau \tag{14}$$

where $X_i = (1, D_i, T_i/\gamma_i)$, to provide us with $\beta_\tau = (\beta_{\tau 1}, \beta_{\tau 2}, \beta_{\tau 3})$. That is, we exclude all observations for which $Y_i > Q_Y(\tau)$ in the estimation of β_τ. We calculate the corresponding network variance estimator as in (10), with G_τ defined as in (13), using $M = (X_1\hat{\epsilon_1}, .., X_\tau\hat{\epsilon}_\tau)$ for $i \in \omega_\tau$. Once again, we shall compare the network standard errors to their robust counterpart obtained by letting G equal the identity matrix.

5 Monte Carlo

As the data used in this paper is randomly generated, we replication the simulation 2000 times, and summarize the results as the mean output of the aggregated runs.

5.1 Network and Outcome Generation

We form the network A according to (1), using parameters $(\theta_1, \theta_2, \theta_3, \theta_4) = (-0.25, 0.5, -1, 0.25)$, $\alpha_i \sim Ber(0.5)$ and $\kappa = 3.28$ as this choice of parameters ensure that the weak-dependence requirement for the network is met as in (Leung, 2019). We assign treatment to units randomly, with a probability of being treated of 0.3. The outcomes for the random-coefficient model as generated as in (2), with $\theta_{i2}, \theta_{i3} \sim N(0,1)$, $\theta_{i4}, \theta_{i5} \sim N(1,1)$, and $\theta_{i1} = \gamma_i^{-1} \sum_j A_{ij}\nu_j$ where $\nu_j \sim N(1,1)$. For the linear regression, we generate the Y_i's using the linear-in-means model (3), once with $\beta_0 = (-0.25, 0.5, 0.25)$ then for $\beta_0 = (1, -0.5, 0.75)$.

5.2 Treatment and Spillover Effects

5.2.1 Non-Parametric Estimation

As mentioned in 3.1, estimates implicitly depend on the number of neighbours units have. Thus, throughout the paper, we fix the degree of sampled units at $\gamma = 3$ in order to identify the average treatment and spillover effects. For fixed t, variation in d identifies the treatment effect, and likewise for the spillover effect for a fixed d and variation in t. Thus, the average treatment effect (ATE) is estimated as:

$$\hat{\mu}(1,1,3) - \hat{\mu}(0,1,3),$$

and the average spillover effect (ASE) as:

$$\hat{\mu}(0,1,3) - \hat{\mu}(0,0,3).$$

Results for the non-parametric estimation of the average treatment and spillover effects can be found in Table 1. Rows 'ATE estimate' and 'ASE estimate' display the estimates, and their corresponding standard errors calculated using (9) are displayed below in rows 'SE'. The rows 'Effective n(d,t,γ)' calculated as $\sum_{i=1}^{n} \mathbb{1}_i(d_i, t_i, \gamma_i)$, give the average effective sample sizes for the aggregated simulations, rounded to the nearest integer. The estimates for the average spillover effect are consistently close to the true value of -3 (by construction in this design), and this result holds across population sizes.

Table 1: Non-Parametric Estimation

n	1000	2000	n	1000	2000
Effective n(1,1,3)	9	18	Effective n (0,1,3)	22	41
Effective n(0,1,3)	22	41	Effective n (0,0,3)	17	32
ATE estimate	0.96	1.01	ASE estimate	-2.98	-2.99
SE	1.31	0.97	SE	1.44	1.05

5.2.2 Linear Regression

Tables 2 and 3 contain results of the linear regression, where the outcomes Y_i are generated according to 3. The outcomes are regressed on $X_i = (1, D_i, T_i/\gamma_i)$. Thus we estimate $\beta = (\beta_1, \beta_2, \beta_3)$, where β_1 is the intercept, β_2 estimates the treatment effect, and β_3 captures the spillover effect. The estimates are close to their target values β_0 for both sample sizes, and across choices of the β_0 parameters. The network standard errors perform well, but are systematically bigger than the heteroskedasticity-consistent (HC) standard errors as expected. We can then conclude that their in a trade-off between efficiency and robustness to network interference.

Table 2: Linear regression for β_0=(-0.25,0.50,0.25)

		β_1	β_2	β_3
n =500	Estimate	-0.249	0.506	0.251
	Network SE	0.00332	0.00191	0.00780
	HC SE	0.00193	0.00194	0.00534
n=1000	Estimate	-0.245	0.0500	0.245
	Network SE	0.00167	0.00095	0.00394
	HC SE	0.000952	0.000954	0.00266

Table 3: Linear regression for β_0=(1,-0.50,0.75)

		β_1	β_2	β_3
n=500	Estimate	0.999	-0.497	0.753
	Network SE	0.00322	0.00191	0.00782
	HC SE	0.00193	0.00194	0.00535
n=1000	Estimate	0.995	-0.500	0.761
	Network SE	0.00167	0.000950	0.00393
	HC SE	.000950	0.000955	0.00266

5.3 Quantile-Based Treatment and Spillover Effects

5.3.1 Non-parametric estimation

Tables 4 and 5 display results from the quantile treatment and spillover effects, calculated using the conditional QSF (11). In this section, we consider a population consisting of $n = 1500$ units, and estimate treatment and spillover effects for quantiles $\tau = \{0.25, 0.5, 0.75, 1\}$. As the quantiles increase to gradually include the whole sample, the average treatment and spillovers effects also converge to their respective values of 1 and -3, as in Table 1. The associated standard errors also increase with the quantiles, due to the increasing range of the outcomes Y_i and observations $\tilde{n}$.

Table 4: Quantile Treatment Effects

τ	0.25	0.5	0.75	1
Effective n(1,1,3)	3	8	10	14
Effective n(0,1,3)	9	23	26	32
ATE estimate	0.508	0.391	0.481	1.036
SE	0.783	0.896	0.928	1.100

Table 5: Quantile Spillover Effects

τ	0.25	0.5	0.75	1
Effective n(0,1,3)	9	23	26	32
Effective n(0,0,3)	0	2	18	25
ASE estimate	-5.971	-3.560	-3.696	-3.020
SE	1.023	1.005	1.042	1.205

5.3.2 Quantile-Based Truncated Regression

A second extension to Leung (2019) is the quantile-based truncated regression. In this exercise, we estimate $\beta_\tau = (\beta_{\tau 1}, \beta_{\tau 2}, \beta_{\tau 3})$ for $\tau = \{0.02, 0.04.., 1\}$, with a sample size of $n = 750$. We generate the outcomes as in (3) with $\beta_0 = (-0.25, 0.5, 0.25)$. As seen in Figure 1, the estimates converge to their true values.

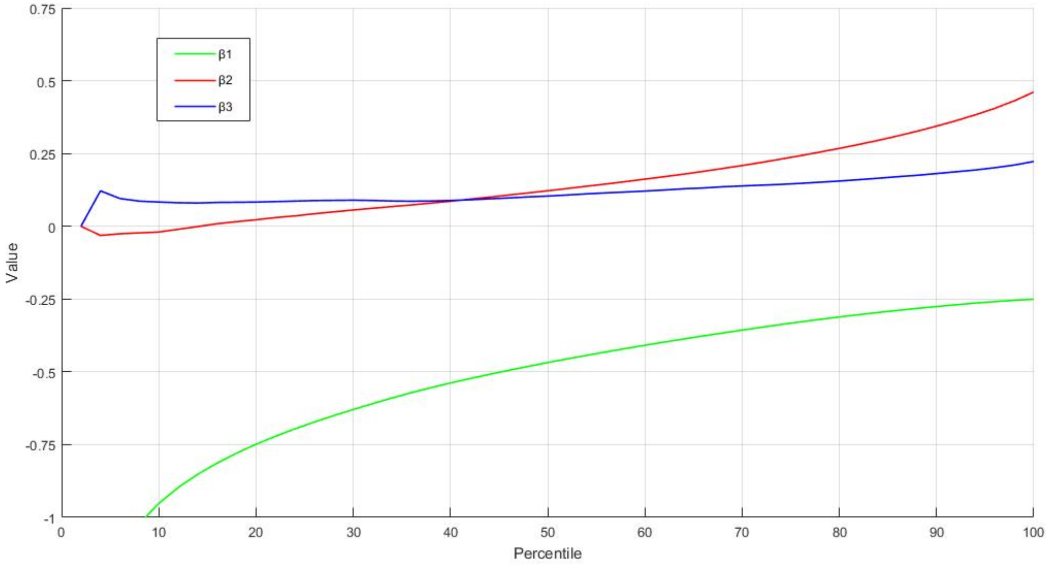

Figure 1: Treatment and spillover effects, per fraction of the population

From the graph, we observe that for units belonging to the $40th$ percentile, the spillover effect captured by β_2 is stronger than the treatment effect. This effect is also seen in Tables 4 and 5, where we notice that the average spillover effect is much larger (in absolute terms) than the average treatment effect, with respective values of -5.97 and 0.508 for the quantile $\tau = 0.25$.

Figure 2 shows the corresponding standard errors for the $\beta\tau$ estimates are also plotted. The network standard errors proposed in Leung (2019) are plotted with lines, whereas the heteroskedasticity robust ones are plotted dots. Most dots fall on the plotted line, thus from the graph we can not conclude that the HC standard errors outperform the network standard errors.

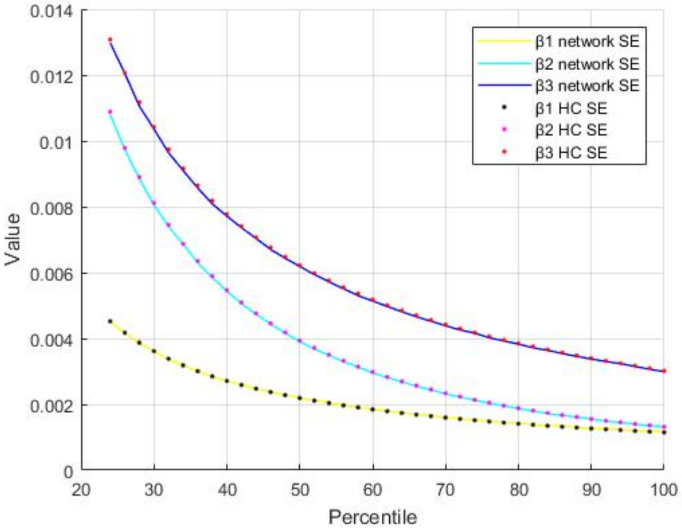

Figure 2: Network standard errors versus HC standard errors, per fraction of the population

6 Conclusion

In this paper, treatment and spillover effects mediated by network interference were estimated using non-parametric and linear regression methods. Following in the steps of Leung (2019), we used a super-population model with a single large network on which was imposed a number of restrictions to induce weak-dependence in the outcomes of the sampled units. We also construct standard errors robust to network interference. For the linear regression we compare these to conventional HC standard errors. The latter systematically perform better than the network standard errors in terms of efficiency, thus we conclude there is a small trade-off between efficiency and robustness to network interference. Under assumptions presented in Leung (2019), the estimates are identified and asymptotically normal. We extend the results of Leung (2019) by introducing quantile-based effects under network interference, to learn more about the outcomes' distribution. However, one must remain cautious in empirical applications, as estimates proposed in Leung (2019) using the super-population model are only valid under a number of restrictive assumptions on the network.

7 Appendix

Three programs were used to obtain all of the results in this paper. These were coded by myself, using MATLAB.

Thesis_Script(method,n,numberReplications,quantile)
Inputs: 'method' can take values (1) for the non-parametric estimation, (2) for the linear regression or (3) for the quantile non-parametric estimation. 'n' is the sample size, and 'numberReplication' the number of simulation runs. 'τ' is the quantile for quantile non-parametric estimation, only relevant for method (3).

This is my main script, which generates the network and data, then calculates the estimates and returns the output via a print function.

[Betas,Network,HC]=QuantileReg(n,numberReplications)
Inputs: 'n' is the sample size, and 'numberReplication' the number of simulation runs.
This script is contains the quantile-based truncated linear regression. It saves three 50 × 3 matrices 'Betas','Network' and 'HC' containing estimates for β_τ and corresponding network and HC standard errors, for each second quantile τ. These are then plotted with the program 'Plot'.

Plot
Plots the outputs of the program 'QuantileReg', giving Figures 1 and 2.

References

Aronow, P. M., Samii, C., et al. (2017). Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947.

Athey, S., Eckles, D., and Imbens, G. W. (2018). Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240.

Halloran, M. E. and Struchiner, C. J. (1995). Causal inference in infectious diseases. *Epidemiology (Cambridge, Mass.)*, 6(2):142–151.

Leung, M. P. (2019). Treatment and spillover effects under network interference. *Review of Economics and Statistics*, pages 1–42.

Liu, L. and Hudgens, M. G. (2014). Large sample randomization inference of causal effects in the presence of interference. *Journal of the american statistical association*, 109(505):288–301.

Manski, C. F. (2013). Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23.

Melly, B. and Wüthrich, K. (2016). Local quantile treatment effects. Technical report, Discussion Papers, Universität Bern, Department of Economics.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rubin, D. B. (1990). Formal mode of statistical inference for causal effects. *Journal of statistical planning and inference*, 25(3):279–292.

Rubin, D. B. (2006). *Matched sampling for causal effects*. Cambridge University Press.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476):1398–1407.

van der Laan, M. J. (2014). Causal inference for a population of causally connected units. *Journal of Causal Inference J. Causal Infer.*, 2(1):13–74.

Van der Laan, M. J., Laan, M., and Robins, J. M. (2003). *Unified methods for censored longitudinal data and causality*. Springer Science & Business Media.