CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Experimental Methods for Measuring Social Networks without Censoring

Eleonora Patacchini[1,2], Rebecca Morton[3], Paolo Pin[4], Jonathan Rogers[5] and
Tanya Rosenblat[6]

[1]Cornell University, Ithaca, USA, [2]Bocconi University, Milan, Italy, [3]NYU and NYU Abu Dhabi, New York,
USA, [4]Università di Siena, Siena, Italy, [5]UCLA School of Law, Los Angeles, USA and [6]University of
Michigan, Ann Arbor, USA
Corresponding author: Tanya Rosenblat; Email: trosenbl@umich.edu

Abstract
We elicit social networks among students in an Italian high school either by measuring the complete network in an incentive-compatible way or by using a truncated elicitation of at most five links. We find that truncation undercounts weak links by up to 90% but only moderately undercounts the time spent with strong friends. We use simulations to demonstrate that the measurement error induced by censoring might be particularly significant when studying phenomena such as social learning which are often thought to operate along weak ties. We then discuss how a modified network elicitation protocol might be able to reduce measurement error.

Keywords: social network measurement; censoring; weak links

Introduction

The study of social networks has emerged as an interdisciplinary field spanning sociology, political science, economics, and computer science. Social networks provide services such as coordination (Opp and Gern, 1993; Chwe, 2000), information (Duflo and Saez, 2003; Kremer and Miguel, 2007; Conley and Udry, 2010) and trust (Karlan et al., 2009; Ambrus et al., 2014).

A key insight from the theory literature on networks is that network structure matters. For example, Granovetter (1974)'s classic study on job search identified "the strength of weak ties" for social learning. It is therefore important to measure social networks correctly in order to evaluate the suitability of a particular instance of a network for services such as trust or social learning. While weak ties are important for

---

📊🎓 This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

learning, strong ties tend to form the backbone of trust networks (Karlan et al., 2009; Jackson et al., 2012; Gee et al., 2017; Patacchini and Rainone, 2017).

However, it is not easy to recover a network of informal acquaintances, even when asking directly to the subjects. A network may be too large for asking everyone,[1] people may have problems in recalling all their friends,[2] friendships are bilateral relations, and only one out of two persons in a dyad may consider it an actual friendship.[3] Empirical researchers have used a wide variety of methods for eliciting network data.

To solve for the recalling and for the nonreciprocating problems mentioned above, one basic design question revolves around *incentives*: reporting one's friends can be cumbersome and study participants might naturally tend to underreport friendships. This might be less of a concern with face-to-face surveys but becomes more salient when the elicitation is conducted over the phone or online. For this reason, incentivized network elicitation mechanisms have been developed starting with Leider et al. (2009).[4] Typically, the researchers ask participants to list a certain number of friends (typically anywhere between 3 and 10) as well as certain attributes of the friendship (such as time spent together per week, for example). If a link is reciprocated, both participants get a small stochastic payoff which increases the more the reported attributes of a friendship align between the two reports. However, most incentivized protocols only elicit a truncated friends' list with a uniform limit on the maximum number of elicited links.[5] Given the importance of network structure for network phenomena such as information diffusion and trust, this raises the question to what extent truncation matters.

A recent literature analyzes the importance of network mismeasurement on various estimation methods and contexts (Chandrasekhar and Lewis, 2016; Breza et al., 2020; Lee et al., 2021; Battaglini et al., 2022). The effects of censoring in peer effects estimates are studied in detail in Griffith and Peng (2023). Little is known however on structural differences between truncated and complete (in the sense of completely elicited) networks and when those differences are most likely to matter.

---

[1]In this case, a network is typically inferred asking the ego network to a representative sample of the overall population. Recent works, such as Almquist (2012), McPherson and Smith (2019), Griffith (2022) and Breza et al. (2020), have shown the statistical biases of such sampling procedures. On all these issues, see also the recent surveys by Stark (2018) and Souza (2021). In our experiment, we actually do not consider some of these issues, as we target small networks.

[2]While papers like Rice et al. (2014) focus exactly on this point, the problem is considered also in all the literature discussed in previous footnote. One method to solve for recalling problems is to ask for specific questions about the relationship, as suggested by Bukowski et al. (1994), and as we also do in our experiment.

[3]Ray et al. (2018) study this issue, providing a survey of the literature that is specifically focused on adolescent networks.

[4]Brañas-Garza et al. (2017) provide a survey on incentivized network elicitation. It would be interesting to compare incentivized and non-incentivized network elicitation similar to the way Brañas-Garza et al. (2023) compare incentivized and non-incentivized elicitation of time preferences.

[5]The question of how many links should be elicited in network surveys is rarely discussed. Exceptions are Moody (2001) (footnote 8 at page 690) and in Currarini et al. (2009) (footnote 12 at page 1009). Recently, the fact that people do have a limited number of contacts has also been used as a tool for the identification of network data (De Paula et al., 2018).

In this paper, we provide insights into this debate using an experiment with high school students. We measure both complete networks (asking about links to every student in a class) and truncated networks (asking about up to five links in a class) and complement this information with supplemental questions on the amount of time each subject spends with each of the other students in the school. Our analysis reveals that the network mismeasurement in our population is largest for weak links. While 90% of weak links with friends spending 0–5 h per week together are underreported with truncated network elicitation, only 29% of strong links with subjects spending 16–20 h a week together are omitted. We thus expect larger biases in estimating network effects in context where weak ties may be more important than strong ties. We test this hypothesis through simulations where we show that the truncation of weak links reduces the number of "shortcuts" in the graph compared to the complete network which reduces the rate of social learning. In contrast, a *randomly truncated* network with the same degree distribution as the truncated graph but where we remove strong and weak links with the same probability behaves more like the complete network for information diffusion.

This finding suggests that additional resources in eliciting fuller network data have the highest returns when studying network services such as social learning which tend to operate along weak links. In contrast, studies of network services such as trust might be less affected because they are often thought to operate along stronger friendship links.

This paper is organized as follows. In the "Model" section, we review conditions under which our incentivized network elicitation induces truth-telling for both complete and truncated networks. We describe our experimental design in the "Experimental design" section. In the "Results" section, we discuss the results and run a simulation study to help understanding the consequences of truncation in terms of network topology. The "Discussion" section discusses modifications to incentivized network elicitation protocols that can reduce censoring measurement errors.

Model

We consider an undirected social network among a group of N agents. Each agent i has a link to every other agent j in the network and we denote the *strength* of the relationship between i and j with $t_{ij} \geq 0$. In some studies, the intensity of relationship is simply measured by time spent together (such as Leider et al. (2009)). In our study, we directly asked subjects for both the quality of their relationship and time spent together. We code the absence of a relationship with $t_{ij} = 0$. We assume that both i and j know the strength of their relationship.

Network elicitation. A researcher elicits M social network links for every agent. We say that elicitation is *complete* if $M = N - 1$. Otherwise, we say elicitation is *censored*. The researcher asks every agent to list M links as well as the link intensity. We denote the agent's choice of links with R_i and the reported link intensity for every $j \in R_i$ with $\hat{t}_{ij}$: note that the report does not have to equal the true strength. For each agent, one of her reported links is chosen at random and her responses are compared to the responses of the other player. If two agents list each other *and* their reports

coincide, then both of them are eligible for a lottery prize L. If the reports do not coincide or the other person does not list them, then the payoff is 0.

To summarize, in this elicitation game an agent's action can be summarized by $\left(M_i, \left(\hat{t}_{ij}\right)\right)$. We say an equilibrium is *truthful* if all the reports are truthful.

Proposition 1. *Under complete network elicitation truth-telling is an equilibrium.*

Proof: Since network elicitation is complete, there is no strategic choice in whom to list. If other players tell the truth, then the best response is to be truthful, too.

Under censored network elicitation agents have a strategic choice in whom to list. In this case, we have to be careful about what agents know about their friends' friends.

Assumption 1. *Agents cannot observe the link intensity of network links that they are not a part of. They assume that each unobserved link is drawn i.i.d. from some distribution $F(t)$ over $[0, \infty]$.*

We call this the "limited knowledge" assumption.

With limited knowledge and censored elicitation, we can show that there exist truthful and *monotonic* equilibria where agents list their strongest links when asked.

Proposition 2. *Assume limited knowledge and censored network elicitation. Consider the strategy profile where all agents report their M strongest links truthfully. This is a Bayesian Nash equilibrium.*

Proof: Truth-telling is immediate: conditionally on being listed by the other agent it is a best response to tell the truth if the other agent does so. Now assume that the agent would swap out a stronger link j for a weaker ink j'. She knows that the probability that she is among the M strongest links of j is larger than the probability that she is among the M strongest links of j'. Hence, the probability of having a reciprocated link report is strictly larger if she reports j instead of j'. This shows that there is no profitable deviation by i.

Without limited knowledge, there are situations where an agent might not report her M strongest links. For example, if she knows that these friends are very popular and have much stronger links, then she might not want to list them because it is unlikely that they will reciprocate.

Monotonic equilibria are intuitive equilibria of this game because we expect that even completely non-strategic agents will behave similarly. Models of human memory from behavioral psychology, for example, predict that agents are more likely to retrieve links with high intensity from their memory. These models of "what comes to mind" are reviewed by Kahana (2012).

Experimental design

The experiment took place in April 2016 at a technical high school focusing on Science Technology, Engineering, and Mathematical Sciences. Students from across

the region attend the school and select a concentration from majors such as informatics, life sciences, or chemistry. They spend most of their school career with the same group of 15–20 students. About 13% of students have an immigrant background indicated by a non-Italian last name. Only 23.4% of the students are female.

We used the lottery method to elicit complete networks from 554 students in 30 classrooms. We presented subjects with a complete list of students in their group and asked them to classify each potential link along two dimensions. First of all, we asked them to qualitatively describe their link to another student in the group as a "strong friendship," "weak friendship," or an "acquaintance." We additionally asked a second question namely the amount of time the subject spent with that other student per week *outside the classroom* (with an ordinal scale of time spent with $1 = 0$–5 h, $2 = 6$–10 h, $3 = 11$–15 h, $4 = 16$–20 h, and $5 =$ more than 20 h). We considered the corresponding responses of two participants to be in agreement if they agree on both dimensions. In that case, both participants would receive a lottery ticket that would pay a monetary prize with some probability. This probability was fixed and did not depend on the number of total tickets assigned to students. A student could collect anywhere from 0 tickets (if their answers were always in disagreement) to a ticket count equal to the number of students in their classroom (if they always agreed). Participants only learned the final sum of realized payoffs but not the number of tickets that they earned. This was to ensure that participants would not be able to directly infer the number of reciprocated links and to avoid disappointment on the part of participants.

For 216 students in another random set of classrooms in the same school, we conducted a censored network survey in which we asked subjects to list five or fewer friends.[6] The incentives were otherwise the same as for the full network elicitation game. Due to the random assignment of classrooms to complete and truncated network elicitation, we can therefore compare means of network statistics to understand the causal impact of censoring.

Results

We start by comparing the types of links that are elicited using the complete and truncated elicitation protocol. We verify that truncation tends to omit weaker links where friends spend less time with each other.

We then show that truncation not only elicits a sparser network compared to complete network elicitation but also changes the structure of the social network *on top of this* by making the network more clustered. This *reduces* diffusion which we show through simulating a simple SI model.

Link-level comparisons of complete and truncated networks

For the complete network, we elicited $1,777$ (13.46%) strong friendships, $7,729$ (58.56%) weak friendships, and $3,692$ (27.97%) acquaintances. For the truncated

---

[6]A total of 280 students participated in the survey, but 64 listed more than 5 friends and are excluded from analysis.

Table 1. Breakdown of time spent, by friendship and network type

	Complete network			Truncated network		
	Strong	Weak	Acquaintances	Strong	Weak	Acquaintances
0–5 h	703	6,834	3,623	142	213	64
6–10 h	555	676	58	112	46	3
11–15 h	234	155	6	64	12	3
16–20 h	144	54	5	48	6	2
More than 20 h	141	10	0	0	0	0
Total	1,777	7,729	3,692	366	277	72

Table 2. Percentage of friendship type, by network type

	Strong ties	Weak tier	Acquaintances
Complete network	13.46	58.56	27.97
Truncated network	51.19	38.74	10.07

Table 3. Percentage of time spent with friends, by network type

	Complete network	Truncated network
0–5 h	84.56	58.60
6–10 h	9.77	22.52
11–15 h	2.99	11.05
16–20 h	1.54	7.83
More than 20 h	1.14	0.00

network, we elicited 366 (51.19%) strong friendships, 277 (38.74%) weak friendships, and 72 (10.07%) acquaintances. Table 1 presents the amount of time subjects reported spending with each of their elicited links.

Table 2 shows that in the truncated network, the proportion of close friends listed is nearly four times as great as it is in the complete network. The majority of links in the complete network (58.56%) are weak friends while the majority of links in the truncated network are strong friends (51.19%).

Table 3 illustrates the same insight when aggregating along time spent together (rather the quality of the friendship). For the truncated network, the distribution is shifted toward links with whom the respondent spends more time. Table 4 shows the mean value of the ordinal time variable, by link type, across the two treatments: conditional on reporting a strong link there is no significant difference in the amount of time spent together between the complete and truncated network. When reporting a weak link or an acquaintance, respondents in the truncated networks report links with whom they spend significantly more time together but the

Table 4. Average time with friend, by friendship and network type

	Complete network	Truncated network
Strong	2.136 (0.030)	2.049 (0.055)
Weak	1.153 (0.005)	1.318 (0.040)
Acquaintance	1.023 (0.003)	1.21 (0.076)
Total	1.249 (0.006)	1.681 (0.036)

Notes. Standard errors in parentheses.

Table 5. Extent of underreported links per agent

	Complete network	Truncated network	Underreporting (%)
0–5 h	20.14	1.94	−90.3
6–10 h	2.33	0.75	−68.0
11–15 h	0.71	0.37	−48.7
16–20 h	0.37	0.26	−29.2
More than 20 h	0.27	0	

difference is economically small (less than 20%). Truncation induces under-reporting of weak links *overall* rather than underreporting of "weakest-weak" links.

Table 5 summarizes the extent of the underreporting of links in the truncated protocol. In this table, we calculate the average number of links by time category and respondent. 90.3% of links in the 0–5 h per week category are missing compared to the complete network elicitation protocol while for links in the 16–20 h per week category only 29.2% of links are missing. Table 5 confirms that the largest measurement errors occur with weak links.

Simulating diffusion on truncated networks

In this section, we design a simulation experiment to provide further insights on the implications of truncation for diffusion processes, that area phenomena commonly studied on social networks (in particular for models of social learning). The truncated network is of course more sparse than the complete network; therefore, mechanically, diffusion will be faster on the complete network compared to the truncated network. In order to identify the marginal impact of truncated network elicitation on top of sparsity, we create "randomly truncated" networks where we randomly select, for each agent, five links from the completely elicited network. We draw 1,000 such randomly truncated networks to minimize simulation error.

Table 6 reports network topology measures for the complete, truncated, and randomly truncated networks *without* symmetrizing the networks while Table 7 shows statistics for symmetrized OR-networks (a link exists if either respondent names the other). Statistics such as diameter and clustering coefficient are commonly defined on the symmetrized networks only.

Table 6. Comparison of network statistics for complete, truncated, and randomly truncated networks (directed network)

	Nodes	Avg. degree	SI model
Complete network	23.67 (3.58)	17.56 (3.65)	0.89 (0.18)
Truncated network	20.73 (4.79)	4.39 (2.45)	0.66 (0.15)
Random truncation	23.67 (3.58)	3.92 (0.67)	0.75 (0.17)
Random addition	20.82 (4.68)	17.21 (2.26)	0.94 (0.08)

Note: Sample standard errors in parentheses. The complete network consists of 30 classrooms and the truncated network was elicited in 15 classrooms. The randomly truncated network was drawn 1,000 from complete elicitations (5 links per respondent). The random addition network does the reverse exercise and adds random friends to each truncated network to make it degree-comparable to the complete network. Degree is calculated on the directed network. The SI simulations pick a node at random as seed, infects outlinks with probability $\frac{1}{2}$, and continues for 3 steps. The percentage of infected nodes is reported.

Table 7. Comparison of network statistics for complete, truncated, and randomly truncated networks (symmetrized OR-network)

	Avg. degree	Assortativity	Diameter	Size GC	Clustering
Complete network	21.28 (3.65)	−0.24 (0.12)	1.90 (0.31)	1 (0)	0.88 (0.13)
Truncated network	6.00 (3.52)	−0.17 (0.13)	4.13 (1.60)	0.98 (0.06)	0.25 (0.12)
Random truncation	7.11 (1.08)	−0.06 (0.10)	3.27 (0.60)	0.99 (0.02)	0.07 (0.04)
Random addition	19.66 (3.58)	−0.09 (0.09)	1.67 (0.47)	1 (0)	0.81 (0.30)

Note: Sample standard errors in parentheses. The complete network consists of 30 classrooms and the truncated network was elicited in 15 classrooms. The randomly truncated network was drawn 1,000 from complete elicitations (5 links per respondent). The random addition network does the reverse exercise and adds random friends to each truncated network to make it degree-comparable to the complete network. Degree is calculated on the symmetrized network.

The most striking difference between the truncated and randomly truncated networks is that the truncated networks are far more clustered (0.25 versus 0.07): this is intuitive because truncation preserves strong links which tend to be more clustered than weak links. Therefore, the truncated network also has a larger diameter compared to the randomly truncated network (4.13 versus 3.23).

Because higher clustering and higher diameter reduce the speed of diffusion in truncated networks, we expect an important cost of truncation in terms of signal transmission over the network. To demonstrate this, we simulate an SI model which is one of the simplest models of infectious diseases and which can also be interpreted as a simple model of information diffusion. Time is discrete and a random node is picked at time $t = 0$. Outlinks are infected with probability $\frac{1}{2}$ and we simulate diffusion for 3 time periods and then record the share of infected nodes. Results are shown in the last column of Table 6. It appears that in the truncated network only 67% of agents are infected while in the randomly truncated network the share is 75% and closer to the diffusion speed in the complete network. This exercise therefore helps understanding how truncation can make us underestimate the speed of diffusion in real-world social networks by omitting weak links.

As an added exercise, we also "simulate" complete networks by starting with the truncated networks and randomly adding links to match the degree of the

complete networks. The results are reported in the last rows of Tables 6 and 7. The diameter of these randomly increased networks is smaller than the true complete network because the randomly drawn links create more shortcuts in the graph. However, clustering as well as speed of diffusion is comparable to the complete network and within the error bounds. These results support the intuition that the truncated weaker links are largely random links in the networks.

Discussion

We have shown that truncation induces serious measurement errors that affect our understanding of the structure of a social network. Intuitively, these errors should be most serious for popular respondents who have many friends and for whom the truncation cutoff is therefore more likely to bind. Based on this intuition, we are looking for protocols that *endogenize* the truncation cutoff: we would like popular respondents to report more friends compared to less popular respondents.

This can be achieved by making agent i's probability of receiving lottery prize L dependent on the number of her link submissions M_i. More precisely, we consider mechanisms where the probability of receiving a ticket is an **increasing and concave** function $p(M_i)$.

To state our next result, we introduce a notion of *popularity*: consider two agents i and j and their realized friendship intensities with all other $N-1$ neighbors. Assume that we order – for each agent – her links from the strongest to the weakest. We say that agent i is more popular than agent j if for any $1 \leq M \leq N-1$ agent i's Mth link has larger intensity than j's Mth link.

Proposition 3. *Assume that the agent has a fixed cost c for reporting on a link, that agents have limited knowledge and that agent i decides on the number of links M_i to submit. The agent is eligible for the lottery prize L with probability $P(M_i)$ where $P(0) = 0$ and $P(\cdot)$ is a strictly increasing and concave function. There is a truthful Bayesian Nash equilibrium where agent i reports her M_i strongest links truthfully. Moreover, in any such equilibrium more popular agents will report weakly more links.*

Proof: The marginal increase in the probability of receiving a link is equal to $p(M)/M$ which is decreasing with M. The probability that a link is reciprocated is also decreasing in the link intensity. An agent will only report link M if the marginal increase in the probability of getting lottery L times the probability of the link being reciprocated exceeds cost c. Therefore, if j reports her Mth link, then a more popular agent will certainly report it as well.

Note that our proposed mechanisms will generally not solve the measurement problem completely. However, it provides support for incentivized network elicitation with dynamic truncation.

Conclusion

We find in our population of high school students that when truncating the friendship list to five links we drastically undercount weak links but only moderately

undercount strong links. The truncated network has a larger diameter and is more clustered than a randomly truncated network which impedes information diffusion.

An implication of this finding is that truncation error might have more impact when studying network services such as social learning which tend to operate along weak links. In contrast, studies of network services such as trust might be less affected because they are often thought to operate along stronger friendship links.

References

Almquist, Zack W. 2012. "Random Errors in Egocentric Networks." *Social Networks* 34: 493–505.

Ambrus, Attila, Markus Mobius, and Adam Szeidl. 2014. "Consumption Risk-Sharing in Social Networks." *American Economic Review* 104: 149–82.

Battaglini, Marco, Eleonora Patacchini, and Edoardo Rainone. 2022. "Endogenous Social Interactions with Unobserved Networks." *The Review of Economic Studies* 89: 1694–747.

Brañas-Garza, Pablo, Natalia Jiménez, and Giovanni Ponti. 2017. "Eliciting Real-Life Social Networks: A Guided Tour." *Journal of Behavioral Economics for Policy* 1: 33–39.

Brañas-Garza, Pablo, Diego Andrés Jorrat, Antonio M. Esp'in, and Angel S'anchez. 2023. "Paid and Hypothetical Time Preferences are the Same: Lab, Field and Online Evidence." *Experimental Economics* 26: 412–34.

Breza, Emily, Arun G. Chandrasekhar, Tyler H. McCormick, and Mengjie Pan. 2020. "Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data." *American Economic Review* 110: 2454–84.

Bukowski, William M., Betsy Hoza, and Michel Boivin. 1994. "Measuring Friendship Quality during Pre- and Early Adolescence: The Development and Psychometric Properties of the Friendship Qualities Scale." *Journal of Social and Personal Relationships* 11: 471–84.

Chandrasekhar, Arun G., and Randall Lewis. 2016. Econometrics of Sampled Networks. Working Paper.

Chwe, Michael Suk-Young. 2000. "Communication and Coordination in Social Networks." *The Review of Economic Studies* 67: 1–16.

Conley, Timothy G., and Christopher R. Udry. "Learning About a New Technology: Pineapple in Ghana." *American Economic Review* 100: 1–28.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2009. "An Economic Model of Friendship: Homophily, Minorities, and Segregation." *Econometrica* 77: 1003–45.

Duflo, Esther, and Emanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118: 815–42.

Gee, Laura K., Jason J. Jones, Christopher J. Fariss, Moira Burke, and James H. Fowler. 2017. "The Paradox of Weak ties in 55 Countries." *Journal of Economic Behavior and Organization* 133: 362–72.

Granovetter, Mark. 1974. *Getting a Job*. Cambridge, MA: Harvard University Press.

Griffith, Alan. 2022. "Name Your Friends, but Only Five? The Importance of Censoring in Peer Effects Estimates using Social Network Data." *Forthcoming in Journal of Labor Economics* 40: 779–805.

Griffith, Alan, and Sida Peng. 2023. Identification of Network Structure in the Presence of Latent, Unobserved Factors: A New Result using Turán's Theorem. Working Paper.

Jackson, Matthew O., Tomas Rodriguez-Barraquer, and Xu Tan. 2012. "Social Capital and Social Quilts: Network Patterns of Favor Exchange." *The American Economic Review 102*: 1857–97.

Kahana, Michael. 2012. *Foundations of Human Memory*. Oxford, United Kingdom: Oxford University Press.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009. "Trust and Social Collateral." *Quarterly Journal of Economics 124*: 1307–61.

Kremer, Michael, and Ted Miguel. 2007. "The Illusion of Sustainability." *Quarterly Journal of Economics 122*: 1007–65.

Lee, Lung-Fei, Xiaodong Liu, Eleonora Patacchini, and Yves Zenou, 2021. "Who Is the Key Player? A Network Analysis of Juvenile Delinquency." *Journal of Business & Economic Statistics 39*: 849–857.

Leider, Stephen, Markus M. Mobius, Tanya Rosenblat, and Quoc-Anh Do. 2009. "Directed Altruism and Enforced Reciprocity in Social Networks." *Quarterly Journal of Economics 124*: 1815–51.

McPherson, Miller, and Jeffrey A. Smith. 2019. "Network Effects in Blau Space: Imputing Social Context from Survey Data." *Socius 5*: 2378023119868591.

Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology 107*: 679–716.

Morton, Rebecca, Eleonora Patacchini, Paolo Pin, Jonathan Rogers, and Tanya Rosenblat. 2023. "Replication Data for: Experimental Methods for Measuring Social Networks without Censoring." *Technical Report*. https://doi.org/10.7910/DVN/4XXIAG, Harvard Dataverse.

Opp, Karl-Dieter, and Christiane Gern. 1993. "Dissident Groups, Personal Networks, and Spontaneous Cooperation: The East German Revolution of 1989." *American Sociological Review 58*: 659–80.

Patacchini, Eleonora, and Edoardo Rainone. 2017. "Social Ties and the Demand for Financial Services." *Journal of Financial Services Research 52*: 35–88.

Paula, Áureo De, Seth Richards-Shubik, and Elie Tamer. 2018. "Identifying Preferences in Networks with Bounded Degree." *Econometrica 86*: 263–88.

Ray, Glen E., Robert Washington, Robert Cohen, Yeh Hsueh, and Zongkui Zhou. 2018. "The Relation of Reciprocated and Nonreciprocated Friendship Nominations to Peer Social Competence for Chinese Elementary School Children." *The Journal of Genetic Psychology 179*: 385–98.

Rice, Eric, Ian W. Holloway, Anamika Barman-Adhikari, Dahlia Fuentes, C. Hendricks Brown, and Lawrence A. Palinkas. 2014. "A Mixed Methods Approach to Network Data Collection." *Field Methods 26*: 252–68.

Souza, Pedro C.L. 2021. "Econometrics of Networks with Limited Access to Network Data: A Literature Survey." In *Handbook of Research Methods and Applications in Empirical Microeconomics*, ed. Nigar Hashimzade and Michael A. Thornton, pp. 416–437. Cheltenham, United Kingdom: Edward Elgar Publishing.

Stark, Tobias H. 2018. "Collecting Social Network Data." In *The Palgrave Handbook of Survey Research*, ed. David L. Vannette, Jon A. Krosnick, pp. 241–254. Cham and London, United Kingdom: Palgrave Macmillan.