

AN UPDATE ON THE LINK BETWEEN NONRESPONSE RATE AND NONRESPONSE BIAS: A NETWORK ANALYSIS

SHANNON S. DICKSON^{1*}

¹DEPARTMENT OF METHODOLOGY AND STATISTICS, UTRECHT UNIVERSITY

Abstract

A paradox confronting survey researchers is that the persistent decline in response rates witnessed over the decades is not accompanied by an increase in nonresponse bias. Yet it is a common conviction that maximising the response rate will mitigate nonresponse bias in a survey and considerable expenditure in time and budget has been dedicated to do so. Empirical evidence of a link between response rates and nonresponse bias is limited or even contradictory. Our study implements a network modelling approach to clarify what mechanisms – if any – lead to nonresponse bias. Our data arise from an original and systematic review of the literature published between 1978 and 2020, featuring a diverse set of survey design characteristics across 187 unique studies making it one of the most extensive meta-analyses on the topic. Furthermore, we present the first application of network modelling to the issue of nonresponse rates and nonresponse bias, discussing the merits and limitations of the method. Our findings are mostly in agreement with prior literature: nonresponse bias is generally independent of the response rate and survey design. We note one exception: a conditional dependence between nonresponse rate and nonresponse bias in surveys administered by mail/paper. Even so, these relationships are complex and nuanced, meriting careful reflection from experts in the field.

Keywords: response rate, participation rate, survey bias, selection bias, mixed graphical modelling, causal inference

*Correspondence

Shannon S. Dickson, Utrecht University

Email: s.s.dickson@students.uu.nl

Journal of Survey Statistics and Methodology

1 INTRODUCTION

Survey response rates have suffered a consistent decline over the years Lipps (2009). One implication of poor response rates is the increased risk for nonresponse bias. Smaller sample sizes also reduce statistical power, and the overall precision of estimates suffers (Lohr 2021). Theoretically, traditional probability samples require a 100% response rate to achieve completely unbiased estimates. Consequently, practitioners are willing to spend a considerable portion of their survey budget on enhancing fieldwork efforts, advertising campaigns, and offering financial incentives to boost response rates Curtin et al. (2000). A key question is if, given the empirical evidence, directing significant effort to maximising response rates is really worth it when reducing nonresponse bias is our goal.

The European Social Survey (ESS) is a cross-national and longitudinal survey measuring attitudes, beliefs, and behaviours across countries in Europe that puts considerable investment in limiting nonresponse. The ESS has an extensive fieldwork period of six weeks over the course of five months to achieve their target response rate of 70%. As such, fieldwork efforts have increased wave-to-wave with more non-contact attempts made, more attempts to convert refusals, increasing the use of prenotification and offering incentives to name a few (Beullens et al. 2018). The overall cost of fieldwork comes to €4.2m (ESS 2016). Yet, ESS response rates are still declining at each wave. Fortunately for the ESS, the differences between respondents and nonrespondents are minimal and nonresponse bias appears stable (Beullens et al. 2018). There are even some cases when increasing the response rate has led to an *increase* in nonresponse bias (Beullens and Loosveldt 2012). This can arise if the converted nonrespondent is very unwilling such that their responses actually increase the measurement error, in turn increasing nonresponse bias (Tourangeau et al. 2010). Having a very long fieldwork period can similarly lead to increased nonresponse bias if certain types of individuals are more likely to respond early or late. A study by Struminskaya and Gummer (2022) found that a fieldwork period of 2 weeks is optimal in web surveys (although note that this is in the context of a panel survey which is very different to cross-sectional surveys). This also indicates that the length of fieldwork should be determined by the point at which nonresponse bias is no longer significant, rather than when response rates reach equilibrium. The paradoxical trend of declining response rates that coincide with low nonresponse bias add further confusion. In a cross-national online survey conducted in Great Britain, Estonia, and Slovenia, the most representative sample came from the country with the lowest response rates (Great Britain) (Maslovskaya and Lugtig 2022). Other auxiliary information was more indicative of nonresponse bias, such as age (in Great Britain) and level of education (in Estonia and Slovenia). These studies and more indicate that response rates alone are a poor predictor of nonresponse bias (Groves et al. 2006; Groves and Peytcheva 2008; Koch et al. 2014; Lee et al. 2009).

The most extensive work on this issue comes from two studies (Groves et al. 2006; Groves and Peytcheva 2008). In one, the authors attempted to produce nonresponse bias by experimentally manipulating survey design and response propensity. Participants were randomly assigned to survey topics they had varying degrees of interest in and that either offered incentives or not but failed to find any consistent pattern. In part, this was due to the different effects of survey mode on participation; in this case mail surveys yielded lower response rates as there was less opportunity to make the topic salient to respondents.

In a second experiment, the authors compared the differences in estimates of the birding population by surveying those with and without a special interest in birding. The special interest group and the general group yielded very different estimates, but this practically vanished when incentives were offered to the disinterested group. Groves et al. (2006) concluded that the relationship between response rates and nonresponse bias is complex and that survey topic, mode, saliency, and incentives have the power to change this relationship.

In a later meta-analysis of 59 different studies on nonresponse, the authors sought to extend these findings by incorporating a wider range of survey design features and sample attributes. Only mode (self- or interviewer-administered), sponsorship, population type (specific or general), and the type of question (behavioural, attitudinal, demographic) showed significant differences between respondents and nonrespondents. Specifically, bias was greater for interviewer-administered surveys, surveys involving the general population, and questions about attitudes (Groves and Peytcheva 2008). Unfortunately, many of the existing studies have important methodological drawbacks that compromise their claims. For instance, Groves and Peytcheva (2008) meta-analysis suffered a small sample size that overrepresented self-administered surveys, health surveys, and limited them to simplistic bivariate analyses that are unlikely to capture the multivariate nature of nonresponse bias. Others focus on particular surveys and modes (Bethlehem 2010; Kohler et al. 2022; McGonagle 2020; Rybak 2023) or are based on theoretical simulations rather than empirical data (Beullens and Loosveldt 2012; Hua 2023). Another shortcoming of the literature is the limited use of causal inference in investigating the link between nonresponse and bias especially given the multivariate nature of it (but see Kohler et al. (2022) and Mercer et al. (2017)). These studies demonstrate that nonresponse bias is a variable-specific phenomenon. A number of theoretical frameworks have been proposed over the years that explain why this is the case. Leverage-saliency theory (Groves et al. 2000), response-propensity theory (Brick and Tourangeau 2017), and benefit-cost theory (Singer 2011) all present a similar case: survey participation decisions balance the participation-benefits and participation-costs, certain aspects of a survey design have leverages on a person's participation and the saliency of these leverages can be manipulated to increase their response propensity, and when response propensities are correlated with the survey outcome nonresponse bias occurs. Consensus on these theories has urged a shift in how we think about nonresponse bias.

In this paper, we first explain how nonresponse and nonresponse bias arise using different causal models of survey response. Next, we describe our own meta-analysis of surveys on nonresponse and analyse the statistical relations among a detailed set of survey design characteristics using graphical network modelling. Our aim is to encourage more causal thinking when it comes to survey inference and to explore the utility of an innovative method for analysing causal structures in survey data.

2 BACKGROUND

2.1 NONRESPONSE BIAS

Nonresponse bias occurs when the sampled units (realised respondents) differ systematically from the non-sampled units (nonrespondents). The Total Survey Error (TSE) framework identifies nonresponse (along with noncoverage) as one of the major contributors to bias (Biemer 2010). Indeed, the American Association of Public Opinion Research (AAPOR) insist practitioners should be “maximising the number of people who respond to your survey” by increasing incentives, contact attempts, and converting refusals (AAPOR 2022). The persistence of response rates as an indicator of survey quality may, in part, be due to this stance. An additional attraction is the theoretical simplicity with which nonresponse rate can relate to nonresponse bias, highlighted by the formula from Bethlehem (1988):

$$NR_{bias}(\bar{y}_r) = \frac{\sigma_{p_i, \mu_i}}{\bar{P}} \quad (1)$$

where,

\bar{y}_r = the value of the survey outcome for respondent r ,

σ_{p_i, μ_i} = the covariance of survey outcome and response propensity for individual i ,

\bar{P} = the mean response propensity

Equation 1 implies that when \bar{P} is high the nonresponse bias is low, *ceteris paribus*. However, if σ_{p_i, μ_i} increases then nonresponse bias will too. Increasing the response rate is only worth it if the decrease in bias that it causes is greater than the increase caused by σ_{p_i, μ_i} . In response, researchers make the maximisation of response rates a primary goal, even moving away from traditional probability-based sampling procedures in favour of lower cost nonprobability surveys that can be fielded faster and achieve higher response rates (Cornesse et al. 2020). It is also standard for cross-sectional and longitudinal surveys always report response rates (Draugalis et al. 2008; Fincham 2008) despite underwhelming or even contradicting evidence of its impact on bias (Groves et al. 2006; Groves and Peytcheva 2008; Gummer 2019; Hedlin 2020; Keeter et al. 2000; Lee et al. 2009). Some have even called for new indicators of survey representativeness, the so-called R-indicators (Bethlehem and Cobben 2008). Before we conclude an absence of a relation between response rate and nonresponse bias, we should consider the situations where one does give rise to the other and how we might control for this.

2.2 CAUSAL MECHANISMS OF NONRESPONSE BIAS

Nonresponse is typically categorised as ignorable or non-ignorable (Fay 1986; Hedlin 2020). If nonresponse has no connection to the survey outcome, we can safely assume that responses are missing completely at random (MCAR) (Little and Rubin 2002)

and therefore is ignorable (i.e., $\sigma_{p_i, \mu_i} = 0$ and nonresponse bias is not a concern). In surveys, this is why probability-based sampling is the gold standard. In probability surveys, all sampled units have an equal probability of responding and any associated missingness would be random. In the causal inference literature, this is called strong ignorability which has two requirements: (conditional) exchangeability and positivity.

Conditional exchangeability is met when all possible confounding variables are known and measured for all of the sampled units such the causal effect is estimable when these variables are controlled for (Greenland and Robins 1986). In this case data are missing at random (MAR) (Little and Rubin 2002). Exchangeability is unconditional when data are MCAR.

Positivity simply ascertains that all sampled units have an equal probability to participate in the survey (Hernán and Robins 2006; Petersen et al. 2012). Probability sampling ensures this. If positivity and (conditional) exchangeability are met, we can be confident that the composition of the sample is sufficiently alike the population to estimate causal effects. In reality, probability-sampling rarely achieves the 100% response rate required to guarantee the conditions for causality. Nonprobability sampling is also becoming more popular (Cornesse et al. 2020) under which we can not assume anything about σ_{p_i, μ_i} . Numerous statistical methods have been developed to relate the sample back to the population when nonresponse is present and nonresponse bias is a concern (Gary 2007; Lee 2011; Lundström and Särndal 1999; Toepoel and Schonlau 2017). Still, conditional exchangeability (thus causal inference) relies on correctly specifying these models.

The assumptions in the MAR case are only met if we condition on the *right* variables. The challenge is in deciding what these variables are. If we believe that the design of a survey influences the response rate and in turn the amount of nonresponse bias, we should condition on these. For example, certain survey modes such as face-to-face achieve higher response rates and one might expect nonresponse bias to be low. Conversely, mail surveys have lower response rates and so should have greater nonresponse bias. Within a survey mode, however, other choices such as the survey topic or incentivising participants will also bear consequences for the response rate and corresponding bias. It becomes obvious that understanding the complex mechanisms that give rise to nonresponse bias is pertinent to accurate estimation of causality in surveys. In the section below, we outline some potential mechanisms that give rise to nonresponse bias and threaten causal conclusions in surveys.

2.3 GRAPHICAL MODELS OF NONRESPONSE BIAS

Graphical causal models for survey inference are not a new idea (Fay 1986; Groves 2006; Kohler et al. 2022; Mercer et al. 2017), but their use is limited and mostly informal. Figures 1 and 2 depict three causal models that have been adapted from Groves (2006). In each of the causal graphs:

- P is the propensity to respond
- Y is the estimated survey outcome
- Z and X are independent survey variables

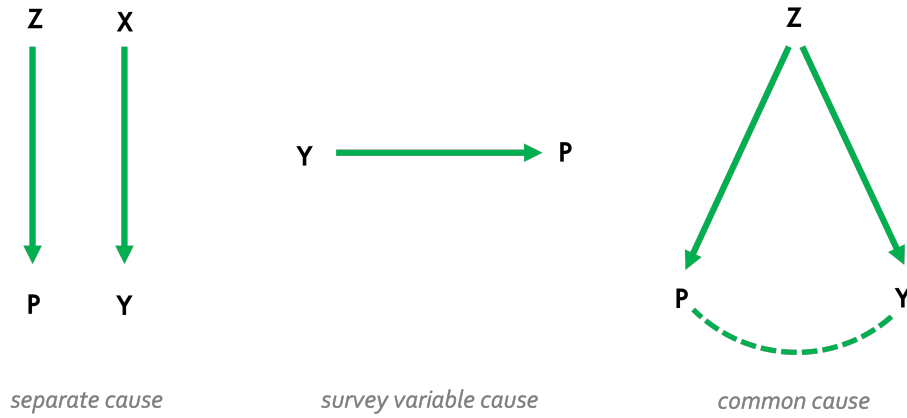


Figure 1: Three causal graphs, adapted from Groves (2006). P is the latent response propensity and Y is the survey outcome. Z and X indicate separate potential causes for P and Y . Nonresponse bias occurs when P and Y are correlated (centre and right).

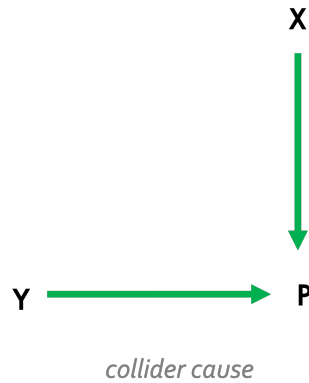


Figure 2: Extension of the survey variable cause graph.

The *separate cause* scenario is the most ideal. Response propensity, P , and the survey outcome, Y , have independent and separate causes, Z and X . Thus, it is impossible for a covariance to exist between the survey outcome and response propensities, and we do not expect any nonresponse bias. The *survey variable cause* scenario illustrates a less desirable scenario where the survey outcome of interest at least partly causes response propensities to change. For example, a survey that intends to estimate the effectiveness of a new medication for depression may overestimate the benefits because people who responded well to the medication are more likely to respond. The *collider cause* is an extension of this scenario where an additional independent variable, X , also affects the response propensity but is independent of Y . Because we are forced to rely on those who have a high response propensity, we may suffer from collider bias where a spurious connection between X and Y exists in-sample but not in the population. Continuing with our example, in addition to receiving more responses that are positive we may also receive more responses from women (it is often observed that women response more often to surveys than men (Becker 2017; Elliott et al. 2005)). As a result, a spurious correlation is induced making it seem as if women respond more positively to the new medication than men. Lastly, in the *common cause* scenario the response propensity, P , and the survey outcome, Y , share a common cause, Z . In our motivating example, Z could be adherence to a treatment plan.

Adherence affects the treatment outcome because better taking the medication consistently is of more benefit. Adherence also affects the propensity to respond because those who adhered are also more invested in their health and more likely to participate in health-related surveys. In the same way, those who did not adhere to the treatment plan are less likely to respond to avoid judgement. Again, we see a positive bias in our estimate of the treatment effect because response propensity and the survey outcome are related through their shared cause of treatment adherence. In the *separate cause* scenario, the survey outcome and response propensity are d-separated (independent) (see Pearl (2009) for a thorough discussion of d-separation). In the other scenarios nonresponse covaries with the survey outcome, resulting in non-ignorable nonresponse bias. However, d-separation is possible by conditioning on the correct covariate. If we control for both gender and treatment adherence, with some statistical tricks, we can estimate a more accurate causal effect by blocking the spurious paths.

Although research has indicated that nonresponse bias is likely to be outcome-specific within surveys, there is a lack of consensus on which survey-related characteristics may induce a relationship between nonresponse and nonresponse bias at the between-survey level. Furthermore, some between-survey characteristics may influence those within-survey aspects that relate to nonresponse and bias (Tourangeau 2017). For instance, a health survey is intrinsically salient to those who are invested in their health, but offering an incentive can increase the saliency to those feeling less invested.

Our study aims to understand these complex mechanisms by estimating a structural network of conditional (in-)dependencies among a broader set of survey design characteristics based on our own meta-analysis of surveys about nonresponse rate and nonresponse bias that greatly expands the sample size of previous meta-analyses. We also discuss the utility of network modelling as method for potential causal inference in surveys. In the next section we describe our data collection and analysis methods.

3 METHODS

In our research, we conducted a systematic review of the literature on nonresponse rates, nonresponse bias, and survey characteristics (section 3.1; Table 1). Using a network modelling approach, we estimated the conditional (in-)dependencies between several survey design characteristics, estimating centrality indices (i.e., node strength) for each network, and assessing the robustness of the network edge-weights and centrality indices (section 3.3). All code used in preprocessing, statistical analysis, and processing of the results is publicly available (section 3.4). Ethical approval for the research conducted was granted by the FETC under file number 22-2063.

3.1 LITERATURE SEARCH AND DATA

Our data arises from a meta-analysis of estimates of nonresponse rate and nonresponse bias from 187 studies. Studies were identified from a systematic search of the literature that was based on the strategy implemented by Groves and Peytcheva (2008). Articles were identified from five electronic databases: Google Scholar, Web of Science, Scopus, Web Survey Methodology and PsychInfo. Additional sources included conference proceedings (e.g., the American Statistical Association Survey Re-

search Methods section), articles cited in book chapters (e.g., Wiley Series in Survey Methodology), and internet searches of the key terms to obtain a wide range of articles and grey literature. The primary search criteria were that the study contained a survey and recorded nonresponse rates and nonresponse bias for that survey. Example of the search terms are *response bias*, *nonresponse bias*, *selection bias*, *survey bias*, *participation rates*. Further inclusion criteria were:

1. Article contains an empirical study about nonresponse rates.
2. A probability-sample is drawn from a population.
3. A survey is conducted.
4. Sample frame information is available for respondents and nonrespondents.

Screening of the initial search results was aided by the automatic screening tool ASReview (v0.19.1) ([Schoot et al. 2021](#); [Van de Schoot et al. 2022](#)). Since all articles considered in the meta-analysis by Groves and Peytcheva (2008) were granted automatic inclusion, we used them as the training set for ASReview. ASReview returns the articles predicted as relevant or irrelevant according to the training set. Five independent researchers quality checked the output from ASReview by manually coding their relevance, stopping when 25 consecutive articles were coded as irrelevant. Articles that were incorrectly predicted by ASReview were amended. Articles about panel surveys were excluded as the pattern of nonresponse is more complicated than for cross-sectional surveys. The process resulted in 187 studies as described in Figure 2

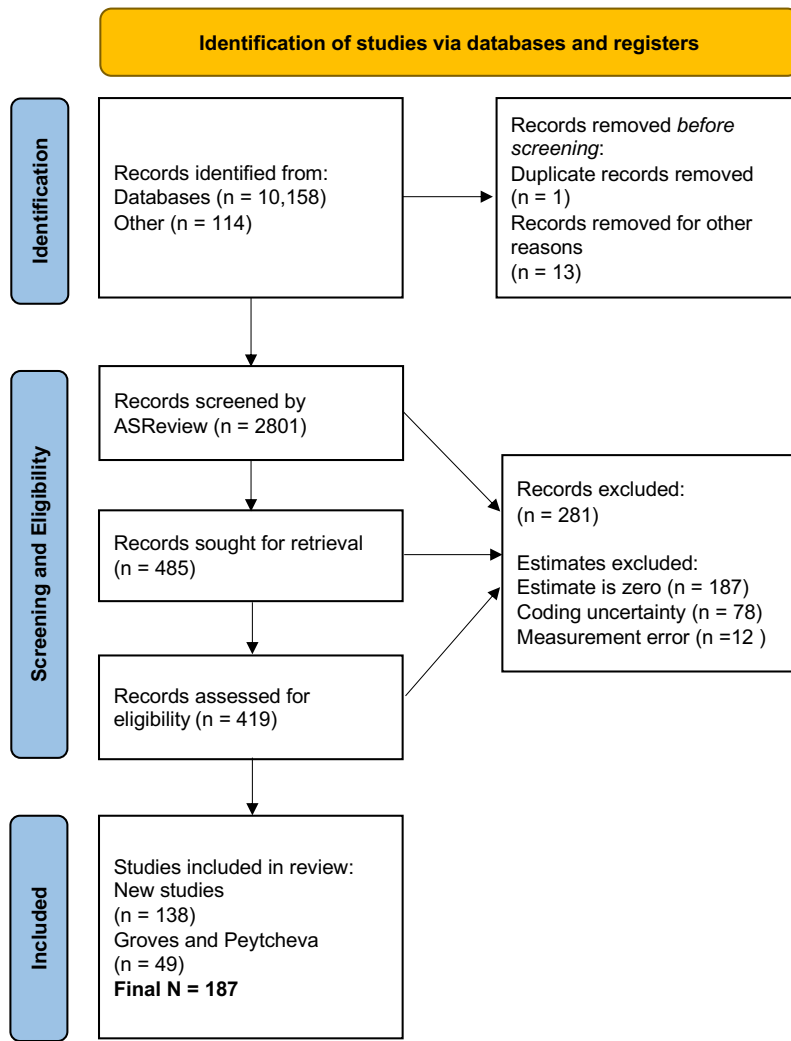


Figure 3: PRISMA (2020) Flow Diagram

The same five independent researchers undertook the task of coding each survey's design characteristics. Most of the characteristics are related to the design of the study (e.g., nonresponse rate, nonresponse bias, survey mode, survey topic, survey incentivisation, survey sponsorship, type of population, use of reminders). Some characteristics are related to the type of outcome statistics (e.g., mean, median, total, or percentage) and the type of question (e.g., attitudinal, demographic, behavioural). An operationalisation of the design characteristics selected for our analysis is presented in Table 1. We note that due to the selected sample, the variation in survey design characteristics is somewhat limited: some categories of the dichotomous variables have very few observations within a particular survey mode. To mitigate this, we selected the characteristics with the most complete cases and that we believe are important for nonresponse rate and nonresponse bias.

Table 1: Descriptive statistics showing the frequencies of survey design characteristics in different survey modes.

Survey Characteristic	Primary Survey Mode				
	Paper drop, N = 178	F2F, N = 523	Mail, N = 1,759	Telephone, N = 374	Web, N = 333
Topic					
Health	164 (92%)	257 (49%)	1,189 (68%)	197 (53%)	77 (23%)
Other	14 (7.9%)	266 (51%)	558 (32%)	177 (47%)	256 (77%)
Unknown	-	0 (0%)	12 (0.7%)	0 (0%)	0 (0%)
Question					
Demographic	95 (53%)	271 (52%)	613 (35%)	195 (52%)	214 (64%)
Unobservable	69 (39%)	87 (17%)	800 (45%)	96 (26%)	114 (34%)
Observable	14 (7.9%)	159 (30%)	346 (20%)	83 (22%)	5 (1.5%)
Unknown	-	6 (1.1%)	0 (0%)	0 (0%)	0 (0%)
Population					
Specific	167 (94%)	203 (39%)	1,178 (67%)	276 (74%)	262 (79%)
General	11 (6.2%)	320 (61%)	581 (33%)	98 (26%)	71 (21%)
Salient topic					
Yes	164 (92%)	190 (36%)	1,298 (74%)	311 (83%)	113 (34%)
No	6 (3.4%)	140 (27%)	240 (14%)	56 (15%)	220 (66%)
Unknown	8 (4.5%)	193 (37%)	221 (13%)	7 (1.9%)	0 (0%)
Incentives					
No	178 (100%)	396 (76%)	1,397 (79%)	349 (93%)	290 (87%)
Yes	-	127 (24%)	352 (20%)	25 (6.7%)	42 (13%)
Unknown	-	0 (0%)	10 (0.6%)	0 (0%)	1 (0.3%)
Reminders					
No	146 (82%)	427 (82%)	627 (36%)	316 (84%)	82 (25%)
Yes	32 (18%)	96 (18%)	1,132 (64%)	58 (16%)	251 (75%)
Unknown	-	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Sponsorship					
Yes	131 (74%)	359 (69%)	1,482 (84%)	309 (83%)	223 (67%)
No	47 (26%)	154 (29%)	267 (15%)	33 (8.8%)	110 (33%)
Unknown	-	10 (1.9%)	10 (0.6%)	32 (8.6%)	0 (0%)

Note: 134 observations were missing a coded survey mode.

3.2 NONRESPONSE RATE AND NONRESPONSE BIAS

Nonresponse rates and nonresponse bias were calculated for each survey following the same procedure as Groves and Peytcheva (2008). Calculation was possible because all studies had information on the respondents, nonrespondents, and the population or “truth”. Nonresponse rate is defined as a percentage of all the potentially eligible units that do not respond to all items in the survey (i.e., unit nonresponse). Response rate, RR , is calculated first (Equation 2) and the nonresponse rate is $1 - RR$.

$$RR = 100 * \frac{n_R}{n_E} \quad (2)$$

where,

n_R = the number of eligible units in the responding sample,

n_E = is the total number of eligible units invited to respond.

Nonresponse bias consists of two components: the nonresponse rate and the difference between the respondents and the nonrespondents on the estimate or survey outcome. Equation 3 is the *absolute relative bias* following Groves and Peytcheva (2008):

$$Bias(\bar{y}_r) = \left| \frac{100 * (\hat{y}_r - \hat{y}_n)}{\bar{p}_n} \right| \quad (3)$$

where,

\hat{y}_r = the survey estimate for respondents,

\hat{y}_n = the survey estimate for nonrespondents,

p_n = the nonresponse rate.

3.3 STATISTICAL ANALYSIS

Network modelling estimates the (conditional) statistical dependencies among nonresponse rates, nonresponse bias, and the survey design characteristics described in Table 1. Network models are families of probability distributions that satisfy several statements of conditional (in)-dependence that are represented in graphical form (Haslbeck and Waldorp 2015). Networks consist of nodes (variables) and edges (statistical relations between variables). In our case nodes represent the survey design characteristics and are from different probability distributions (i.e., categorical and continuous variables). The Mixed Graphical

Model (Chen et al. 2015; MGM: Haslbeck and Waldorp 2015; Yang et al. 2014) caters to this by modelling the joint probability density of the appropriate distribution for a given node. A pairwise interaction between two continuous nodes, denoted s and r , is represented by a single parameter, $\theta_{s,r}$, when the interaction is non-zero. A pairwise interaction between two categorical variables, denoted m and u , is represented by $R = (m - 1) * (u - 1)$ parameters corresponding to indicator functions (dummy variables) for R states (categorical levels). A pairwise interaction between a continuous and a categorical variable is represented by $R = 1 * (m - 1)$ parameters corresponding to $(m - 1)$ corresponding indicator functions for all R states (Haslbeck and Waldorp 2020). Estimation results in an undirected graph, denoted $G = (V, E)$, that has nodes $V = \{1, 2, \dots, p\}$ and edges $E \subseteq V * V$. Our analysis is exploratory; we estimate a network including all survey modes and then estimate networks stratified by mode (paper and mail modes are combined for adequate sample size). All other parameters are standard across estimations.

3.3.1 MIXED GRAPHICAL MODELS

MGM networks were estimated using LASSO (Least Absolute Shrinkage and Selection Operator) regularization and EBIC (Extended Bayesian Information Criterion) model selection with $\gamma = 0.25$. The MGM is estimated in two steps: First, the conditional distribution of a node is estimated separately holding all other nodes constant, using the generalised linear regression framework. This results in two estimations per node going in opposite directions (i.e. $N1 - N2$ and $N2 - N1$). Second, the average of these two estimates is computed and it is this estimate that is represented in the network structure. An “OR” rule was specified meaning that an edge is represented in the network if either set of parameters for a node is non-zero. The regularization and model selection procedures lead to a conservative network featuring good specificity (low false-positives) at the cost of sensitivity. Estimation was conducted using *mgm* (Haslbeck and Waldorp 2020) and network visualisation with *qgraph* (Epskamp et al. 2012).

3.3.2 ROBUSTNESS CHECK

Non-parametric bootstrap (i.e., resampling rows with replacement) with $n = 1000$ was implemented to estimate the accuracy of edge-weights. Case-dropping bootstrap (i.e., dropping subsets of the sample) with $n = 1000$ was implemented to estimate the stability of centrality indices, indicated by the correlation of strength centrality between the original sample and bootstrap subsets. The Correlation Stability (CS) coefficient (R_{CS} ; Epskamp et al. 2018) derived from this represents the percentage of our sample that can be dropped while maintaining a correlation $R_{CS} \geq 0.70$ (95%CI). Accuracy and stability analysis is conducted using *bootnet* (Epskamp et al. 2018).

3.4 SOFTWARE

All analyses were conducted in R Core Team (2021) version 4.1.2 on macOS Ventura version 13.2.1 CPU Intel Core i5 1.4GHz Quad-Core. All code used in the study can be found on GitHub: https://github.com/shandickson/thesis_masters. Further details are in the Supplementary Materials.

4 RESULTS

We present the descriptive analysis (section 4.1), showing how nonresponse rate and nonresponse bias compare in our new studies compared to the original 49 from Groves and Peytcheva (2008) (Table 2). We also present correlational analyses on the association between nonresponse rate, nonresponse bias, and how this changes over the years for the original and new studies (Figures 4 and 5). The results of the network models, centrality indices (section 4.2), and network robustness (section 4.3) is also presented.

4.1 DESCRIPTIVE ANALYSIS

Our first step was to compare the nonresponse rate and nonresponse bias in our sample compared to the original meta-analysis from Groves and Peytcheva (2008), to understand if there are substantial changes. As seen in Table 2, nonresponse rate has increased from 36% to 44% (+8%) which is consistent with the literature. Nonresponse bias has barely changed (+2%), also as expected. There are some differences in the nonresponse rate and nonresponse bias for different survey modes (Table 2); face-to-face surveys have the lowest response rate (26.50%) and lowest nonresponse bias (8.44%) whereas paper surveys have the highest nonresponse rate (55.23%) and web surveys have the highest nonresponse bias (12.67%). It is clear that the relationship between nonresponse rates and nonresponse bias is not monotonic across survey (e.g., mail surveys have comparatively low bias to other modes with similar nonresponse rates), indicating that this relationship has other likely influences (i.e., survey design characteristics).

We find similar correlations between nonresponse rate and nonresponse bias to Groves and Peytcheva (2008) (Figure 4). Taking all estimates, the correlation in the present study is 0.20 (compared to 0.23; Groves and Peytcheva (2008)) and the correlation weighted by the study's sample size is 0.40 in the present study (compared to 0.47 in Groves and Peytcheva 2008). A substantial between-study correlation is similarly reported by (Brick and Tourangeau 2017), but our result arises from a much larger sample. The upward trend of nonresponse rates and the stability of nonresponse bias over the years is also evident (Figure 5).

Table 2: Nonresponse rate and nonresponse bias

	Nonresponse Rate	Absolute Relative Bias	Number of estimates
Studies			
New studies	44 %	10.25 %	2342
Groves	36.01 %	8.32 %	959
Mode			
Mail	41.54 %	8.98 %	1759
F2F	26.5 %	8.44 %	523
Telephone	42.57 %	11.14 %	374
Web	53.74 %	12.67 %	333
Paper drop	55.23 %	11.13 %	178

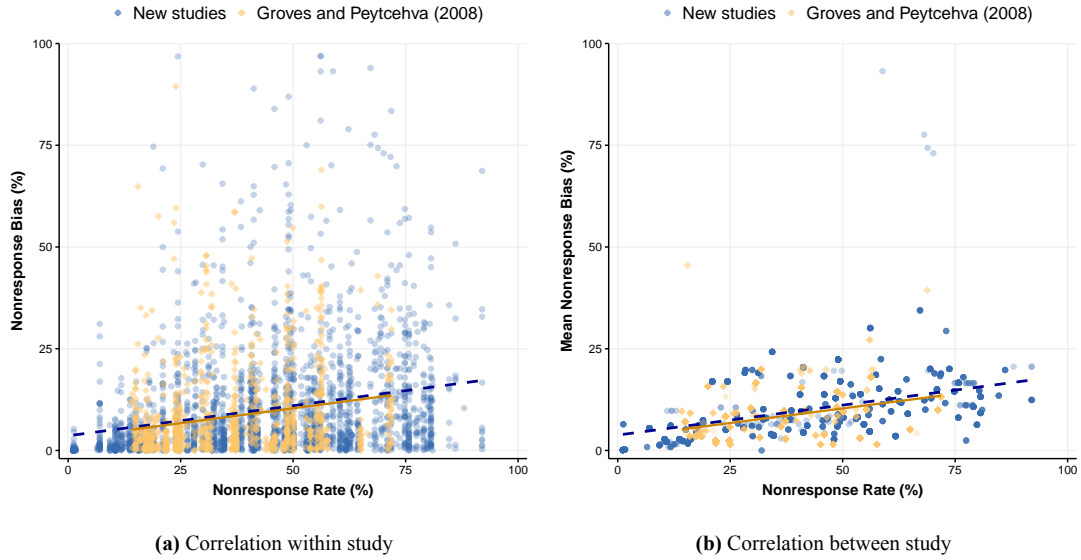


Figure 4: Similar to Groves and Peytcheva (2008): Nonresponse rates are weakly correlated with nonresponse bias.

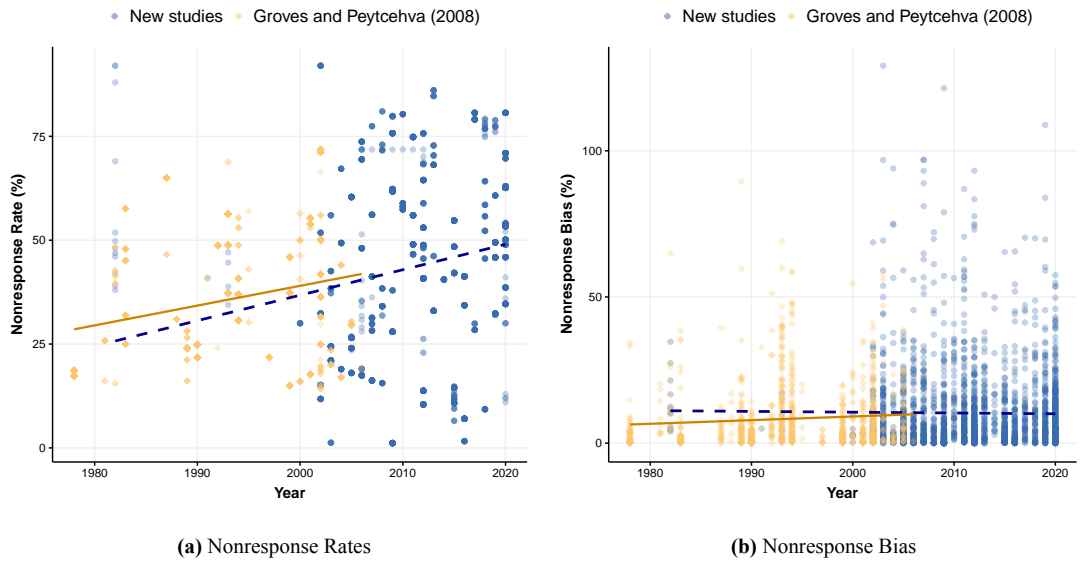


Figure 5: Similar to Groves and Peytcheva (2008): Nonresponse rate is increasing over time but nonresponse bias is relatively stable.

4.2 NETWORK MODELS

Each network model involves the survey design characteristics detailed in Table 1, but only the first network includes *all* survey modes: paper (reference category), face-to-face, mail, web, and telephone. The ensuing networks are stratified by mode (paper and mail are combined). Table 3 contains the aggregated nodewise edge-weights for all networks. For disaggregated edge-weights (i.e., per category level) see the Appendix Tables 4 to 8. The pie charts indicate the explained variance or classification accuracy for a given node, showing generally that little variance in nonresponse bias is explained by the other nodes in the network. Slightly more variance in the nonresponse rate is explained by other nodes, but is still low. Classification accuracies are sometimes unreliable (i.e., approaching 1), demonstrating the aforementioned class imbalances within some of the categorical variables. Simply, this shows how certain design characteristics are overrepresented in certain modes and is still a reasonable representation of survey design. Centrality indices are also presented alongside the relevant networks, showing the node with the strongest (absolute sum of) connections (Figures 6b and 8). In the next section, we pay attention to specific edges of interest and compare them across the different networks.

4.2.1 OVERALL NETWORK MODEL

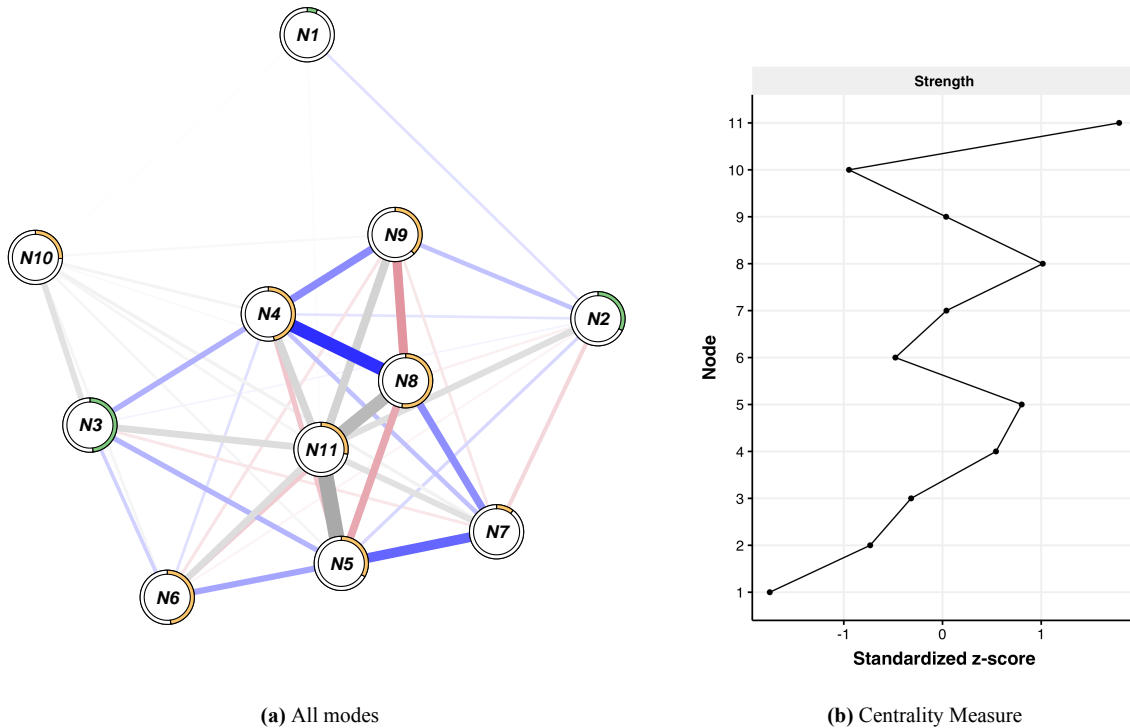


Figure 6: Overall network with all survey modes: paper, face to face, mail, web, telephone. Nodes are nonresponse bias (N1), nonresponse rate (N2), year (N3), topic saliency (N4), incentives (N5), reminder (N6), sponsored survey (N7), health/miscellaneous topic (N8), specific/general population (N9), question type (N10), and mode (N11). Edges are positive (blue), negative (red), or undefined for over 2 categories (grey). Pie charts represent explained variance (green) and classification accuracy (yellow). Centrality measure shows node 11 (survey mode) is the most central node in the overall network.

A network graph on the associations between survey design characteristics, including modes, has 11 nodes, a mean edge-weight of 0.22, and high overall density: 46/55 (~84%) of all possible edges (Figure 6a). There is a direct association between nonresponse rates and nonresponse bias, after all other nodes are accounted for and very weak edges are regularized out. Although unseen in the graph (due to the cut argument), there are direct associations between nonresponse bias to reminders and survey mode (see Table 3). The positive association to reminders means that nonresponse bias increases when reminders are used, which is surprising at first. However, reminders are typically distributed to nonrespondents *at the time* in a bid to boost response rates (and in the network are indeed negatively associated with nonresponse rates). This means that despite the use of reminders increasing response rates, they also increase nonresponse bias (directly and indirectly). This could arise if reminders boost response rates from a particular group within the target population, biasing estimates towards them.

Nonresponse rate is more densely connected: positively associated with (publication) year, salient survey topics, using incentives, and surveys in specific populations; and negatively associated with sponsored surveys. Again, some of these results are initially surprising, such as salient topics and the use of incentives associated with increasing nonresponse rate. If incentives are offered when response rates are already trailing, it may be too late to have the desired effect on response rates. We also see from the network that salient topics are less likely to be incentivised (i.e., negative association). It is possible that the saliency of the topic is not enough to entice respondents alone, without incentivisation. Sponsored surveys are associated with decreasing nonresponse, similar to Groves and Peytcheva (2008). On the whole, associations to nonresponse rate and nonresponse bias are weak (Table 3).

From Figure 6 it is clear that survey mode (node 11) is the most central in this network. Cross-referencing this to the parameters in Table 4 of the Appendix, we see that relative to paper surveys, face-to-face, mail, and telephone surveys are associated with a decrease in nonresponse rates. Owing to this, we decided to stratify our data and estimate a network for each survey mode to inspect how the interactions within them change to the overall model.

4.2.2 STRATIFIED NETWORK MODELS

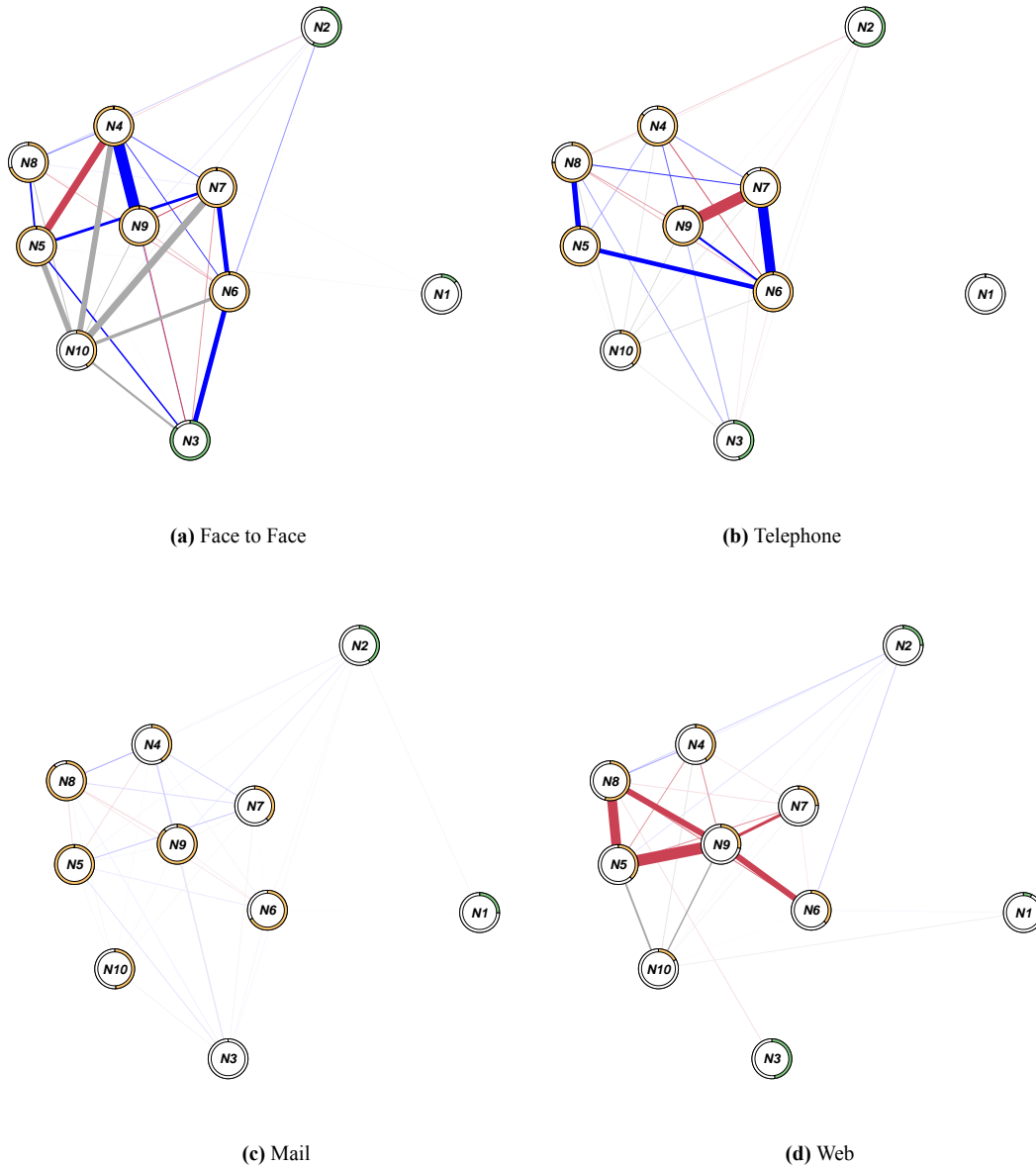


Figure 7: Stratified network models. Nodes are nonresponse bias (N1), nonresponse rate (N2), year (N3), topic saliency (N4), incentives (N5), reminder (N6), sponsored survey (N7), health/miscellaneous topic (N8), specific/general population (N9), question type (N10). Edges are positive (blue), negative (red), or undefined for over 2 categories (grey). Pie charts represent explained variance (green) and classification accuracy (yellow).

4.2.2.1 Network for face-to-face surveys A network graph on the associations between survey design characteristics, in face-to-face surveys, has 10 nodes, a mean edge-weight of 2.57, and moderate overall density: 24/45 (~53%) of all possible edges (Figure 7). Some of the associations between nonresponse rate and survey design characteristics remain in face-to-face surveys: positive connections to the use of reminders, surveys in specific populations, and health surveys; and negative connections to topic saliency. For the more surprising effects, as before, the conclusion is that some design choices (i.e.,

reminders) are not improving response rates in face-to-face surveys. On the contrary, salient topics *do* improve nonresponse rates for face-to-face surveys. Surveys involving special populations and health-related surveys and the associated increase in nonresponse rate is not that unexpected. Certain special populations can be hard to reach out-with their interest groups, and the primary target for health surveys are likely to be patients who due to illness are less able to respond. There is no relationship between year and nonresponse rate in face-to-face surveys. Face-to-face modes consistently benefit from higher response rates than other modes (Bowling 2005; Czajka and Beyler 2016; Rybak 2023), so it seems that they do not share in the general upward trend of nonresponse rate seen in the overall network.

Notably, the association between nonresponse rate and nonresponse bias is absent, giving the impression that the design characteristics in face-to-face surveys that lead to higher nonresponse rates are not in turn causing an increase in nonresponse bias. Recall from Table 2 that face-to-face surveys have the lowest nonresponse rate and (second) lowest bias, meaning face-to-face surveys are generally of a high standard.

4.2.2.2 Network for telephone surveys A network graph on the associations between survey design characteristics, in telephone surveys, has 10 nodes, a mean edge-weight of 0.99, and moderate overall density: 23/45 (~51%) of all possible edges (Figure 7). The telephone network carries structural similarities to the face-to-face network, except the key connections are of opposite sign: decreasing nonresponse occurs when the topic is salient, when reminders are used, when the survey is sponsored and when it is health related. These design characteristics are effective in telephone surveys, but evidently not in face-to-face surveys. Continuing with the assumption that health surveys, for instance, target people are ill, a telephone call might be more convenient to the patient (i.e., because they have access to their phone). Nevertheless, telephone surveys exhibit declining response rates over the years.

Although nonresponse rates and nonresponse bias are higher in telephone surveys (see Table 2), there are few strong drivers of this in the network apart from year. A unique disadvantage of telephone surveys is that they require immediate cooperation from the respondent (arguably this is true for face-to-face, but a copy of the survey can be dropped off). If the respondent is caught at a bad moment, disengaging from the survey is as simple as ending the phone call. These transient influences are not captured in our data set, but are an interesting thought for future work as they correspond to Brick and Tourangeau (2017)'s random propensity model of nonresponse.

4.2.2.3 Network for mail surveys A network graph on the associations between survey design characteristics, in mail surveys, has 10 nodes, a mean edge-weight of 0.16, and high overall density: 36/45 (~80%) of all possible edges (Figure 7). A considerably sparse network comprising several weaker connections relative to the previous networks, there are still some direct associations between nonresponse rates and survey design characteristics such as the (publication) year, when the topic is salient, when incentives are used and when the population is specific. The network also shows that web surveys are less likely to use reminders when the topic is salient or when specific populations are of interest. Perhaps the response rate in web surveys are more dependent on features like reminders, changing the relationships between the nonresponse rate and other characteristics. A reliance on reminders is intuitive, for instance, if (e)-mail surveys are hidden among more important

deliveries (e.g., household bills, appointment letters). In fact, the use of reminders *is* associated with decreasing nonresponse rates. Likewise, sponsored surveys and health surveys lead to lower nonresponse rates, likely for similar reasons to telephone surveys (i.e., mail health surveys are more accessible to patients). Sponsored mail surveys could be official in capacity (i.e., coordinated national surveys). These factors together are responsible for the lower overall bias (Table 2). A faint, direct, association is present between nonresponse rate and nonresponse bias (despite both being low comparative to other modes), eloquently demonstrating the complexity of accounting for this relationship with many design features.

4.2.2.4 Network for web surveys A network graph on the associations between survey design characteristics, in telephone surveys, has 10 nodes, a mean edge-weight of 0.99, and high overall density: 14/45 (~38%) of all possible edges (Figure 7). Web surveys experience higher nonresponse rates and greater nonresponse bias than other modes (Table 2). Yet, there are fewer direct associations to survey design characteristics in the network model. Nonresponse rate does increase somewhat with the use of reminders, incentives, and for health-related surveys. An increase in nonresponse bias is also present when the survey question is about an “observable” characteristics of the respondent, probably due to the anonymity offered by online surveys that enable respondents to give less accurate to certain questions (e.g., for social desirability purposes).

A stark contrast in the web network is the strong, negative associations between survey topic, saliency, the use of incentives, reminders, and survey sponsorship; suggesting that these features are far less prevalent for online surveys and are a possible explanation for the high nonresponse rate (Table 2), but not necessarily for the large nonresponse bias.

4.2.3 NODE CENTRALITY

Node centrality varies across the different survey modes: the topic saliency (node 4) is most central to the face-to-face and mail networks, reminders (node 6) are most central to the telephone network, and the population type (node 9) is most central to the web network (Figure 8). Face-to-face surveys are almost exclusively interviewer-led, presenting opportunities to make the purpose of the survey more salient to the respondent (hence topic saliency becomes most central). The reasons for the centrality of topic saliency in mail surveys is less clear. It is possible that mail surveys are often specific to the recipient (e.g., healthcare follow-up surveys) or feel more obligatory (e.g., national surveys), adding to the saliency. Reminders are probably central in telephone surveys due to the ease at which telephone reminders can be administered in addition to the probable chance the respondent misses the first phone call. It is also unclear why the population type is so central to web surveys, but note that incentives are not far behind.

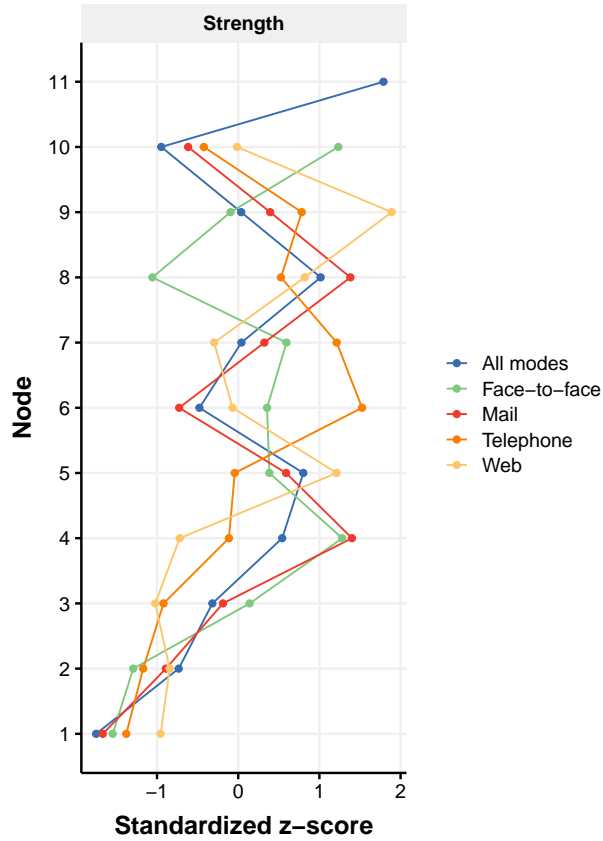


Figure 8: Centrality measure of node strength. Node 4 is most central for face to face and mail surveys. Node 6 is most central for telephone surveys. Node 9 is most central for web surveys.

Table 3: Estimated nodewise interactions for the network models of the absolute aggregated parameters.

Network		N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11
All modes	N1	0.00	0.16	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
	N2	0.20	0.00	0.25	0.00	0.00	0.08	0.41	0.00	0.23	0.15	0.63
	N3	0.03	0.56	0.00	0.70	0.67	1.36	0.53	0.89	1.19	0.95	1.59
	N4	0.03	0.00	2.88	0.00	0.49	0.68	0.18	1.47	1.01	0.27	0.76
	N5	0.07	0.15	0.60	0.23	0.00	1.59	0.61	0.77	0.22	0.44	2.14
	N6	0.04	0.25	3.12	0.33	1.53	0.00	0.05	0.40	0.55	0.27	1.43
	N7	0.00	0.40	2.10	0.28	0.64	0.00	0.00	0.75	0.32	0.31	0.72
	N8	0.00	0.19	2.87	1.56	0.79	0.74	0.82	0.00	1.71	0.13	1.03
	N9	0.00	0.32	3.19	0.93	0.15	0.74	0.09	1.55	0.00	0.26	0.71
	N10	0.04	0.14	0.62	0.11	0.29	0.28	0.33	0.00	0.24	0.00	0.47
	N11	0.03	0.86	3.89	0.50	1.30	1.69	0.41	1.81	1.03	0.46	0.00
Face-to-face	N1	0.00	0.00	0.00	0.00	2.33	0.00	0.00	0.00	0.00	0.00	-
	N2	0.00	0.00	5.10	1.31	0.00	0.96	25.15	15.59	2.04	1.37	-

Table 3: Estimated nodewise interactions for the network models of the absolute aggregated parameters. (*continued*)

Network		N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11
Telephone	N3	0.04	0.67	0.00	0.58	3.74	3.92	23.60	7.56	1.23	5.71	-
	N4	0.26	2.29	0.00	0.00	0.00	6.89	0.00	27.37	7.48	10.39	-
	N5	0.48	0.78	17.93	1.32	0.00	4.96	0.00	0.00	0.00	4.93	-
	N6	0.00	1.75	24.12	2.39	3.25	0.00	82.06	7.14	7.51	7.83	-
	N7	0.00	0.00	7.42	0.00	3.23	0.00	0.00	0.00	0.00	6.37	-
	N8	0.06	1.25	7.35	2.51	2.02	4.36	40.88	0.00	1.39	4.55	-
	N9	0.00	1.71	9.70	3.06	7.60	6.42	27.66	22.01	0.00	8.93	-
	N10	0.00	0.00	2.52	0.00	1.03	0.00	42.69	2.74	0.00	0.00	-
	N1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
	N2	0.00	0.00	3.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
Mail	N3	0.04	0.30	0.00	4.99	0.90	2.68	1.96	7.10	4.58	6.45	-
	N4	0.00	0.00	2.94	0.00	0.00	4.23	0.00	3.04	6.62	1.38	-
	N5	0.00	0.68	6.39	9.10	0.00	5.37	0.00	11.16	0.00	1.74	-
	N6	0.26	0.46	5.52	11.92	2.14	0.00	4.61	6.35	3.96	5.67	-
	N7	0.00	1.37	4.35	12.83	0.00	7.46	0.00	0.00	12.65	3.21	-
	N8	0.00	0.26	3.50	0.00	3.34	2.80	5.17	0.00	0.00	2.18	-
	N9	0.00	0.00	3.59	9.03	0.00	3.71	6.94	0.00	0.00	3.75	-
	N10	0.00	0.00	4.68	2.90	0.00	2.18	1.51	0.00	2.64	0.00	-
	N1	0.00	0.17	0.26	0.00	0.00	0.11	0.00	0.00	0.00	0.08	-
	N2	0.22	0.00	0.89	0.33	0.00	0.13	0.54	0.00	0.00	0.15	-
Web	N3	0.48	0.83	0.00	1.57	0.48	2.12	1.93	0.00	0.99	0.67	-
	N4	0.00	0.44	1.15	0.00	0.76	0.70	0.93	5.51	6.40	0.48	-
	N5	0.14	0.23	0.61	1.65	0.00	2.69	1.10	1.60	0.00	0.91	-
	N6	0.47	0.32	2.50	0.73	3.03	0.00	2.06	1.85	2.03	0.26	-
	N7	0.15	0.65	1.64	1.74	1.19	1.92	0.00	1.65	0.59	0.25	-
	N8	0.21	0.00	0.70	4.94	1.15	1.27	1.50	0.00	8.12	0.18	-
	N9	0.12	0.45	0.62	4.00	0.00	1.30	0.45	6.76	0.00	0.14	-
	N10	0.25	0.16	0.66	0.67	0.56	0.40	0.20	0.19	0.00	0.00	-
	N1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.44	-
	N2	0.00	0.00	2.29	0.00	0.00	0.87	0.00	0.00	0.00	1.26	-
	N3	0.00	3.11	0.00	53.81	3.23	0.00	3.62	3.97	1.78	6.18	-
	N4	0.00	0.64	1.72	0.00	6.91	1.51	7.39	3.43	0.00	5.76	-
	N5	0.00	0.72	1.18	25.26	0.00	0.00	20.70	3.92	5.65	4.61	-
	N6	0.16	0.00	0.00	15.51	0.00	0.00	6.63	0.72	1.61	0.00	-

Table 3: Estimated nodewise interactions for the network models of the absolute aggregated parameters. (*continued*)

Network	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11
N7	0.00	0.00	4.19	21.69	7.49	2.29	0.00	0.00	3.04	2.13	-
N8	0.00	0.00	2.58	18.32	0.00	2.64	0.00	0.00	3.75	4.15	-
N9	0.00	0.00	1.98	0.00	0.00	3.25	13.16	5.42	0.00	3.55	-
N10	1.68	0.28	1.55	19.85	7.47	0.00	13.90	0.00	2.29	0.00	-

Note: Node names match the network graphs; interactions should be read as rows predicting columns.

4.2.4 ROBUSTNESS OF THE NETWORK

As the parameters of these network models are *estimated* from data (i.e., not observed), it is crucial to assess the robustness of the models. Parameter estimation is vulnerable to low sample sizes and class imbalance, which are features of our stratified networks. We assessed robustness in two ways: **(1)** by bootstrapping the accuracy of edge-weights; and **(2)** bootstrapping the stability of centrality indices.

The bootstrapped edge-weights are presented in Figure 9, for each network model: overall, face-to-face, telephone, mail, and web. Unsurprisingly, the edge-weights are most accurate for the networks with larger sample sizes (overall, mail). In the face-to-face, web, and telephone networks edge-weights are often estimated to be zero, indicating the absence of an edge in all bootstrapped samples. Moreover, the shaded region reveal that some edge-weights have considerably wide confidence intervals and there are occasions when the bootstrapped and original sample mean markedly diverge. Although the sample size within each network is sufficient for estimation, the class imbalance over the categories of some survey design characteristics are the likely culprits. In part, this is a genuine effect of survey design, where certain design choices are common to a mode. However, the `mgm` package cannot handle missingness and further casewise deletion during estimation exacerbates the class imbalance. Future work should aim to collect enough data to have more balanced classes.

Figure 10 presents the correlation of the bootstrapped centrality indices with the centrality indices of the original sample. Correlations remain high even as the percentage of subsamples dropped from the data increase, meaning stability is excellent overall and never drops below the minimum recommendation (black line). Ergo, we are confident that the centrality indices estimated in the original samples is not dependent on particular observations (subsamples) of the data and are representative of the population of surveys.

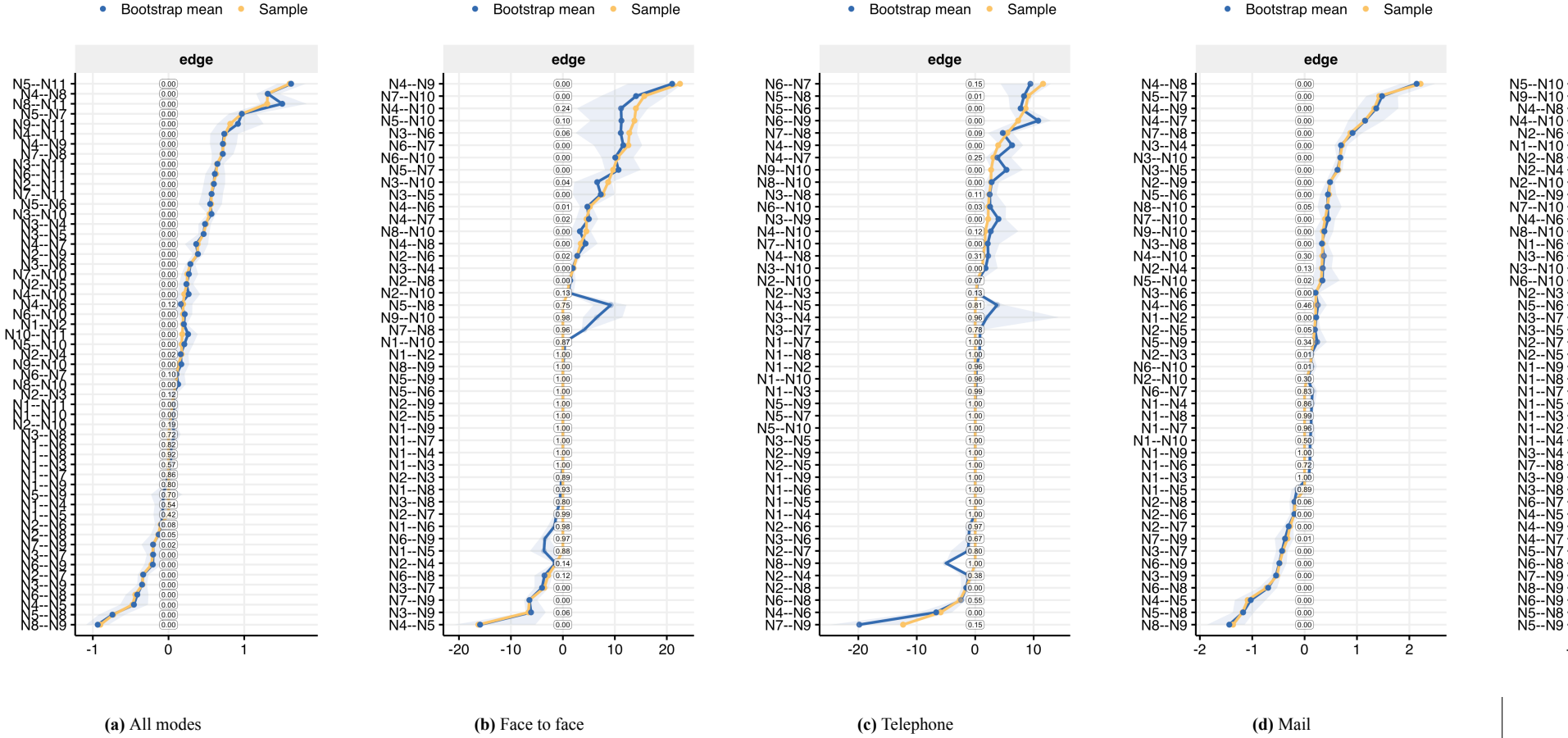


Figure 9: Nonparametric bootstrap samples ($n=1000$) of network edges and their confidence intervals (blue line and shaded region) plotted against the original sample edges (yellow line). Several edges are always estimated to be zero and the non-zero edges are less accurate in networks with lower sample sizes (face-to-face, telephone, and web). X-axis: edges; y-axis: standardized coefficients.

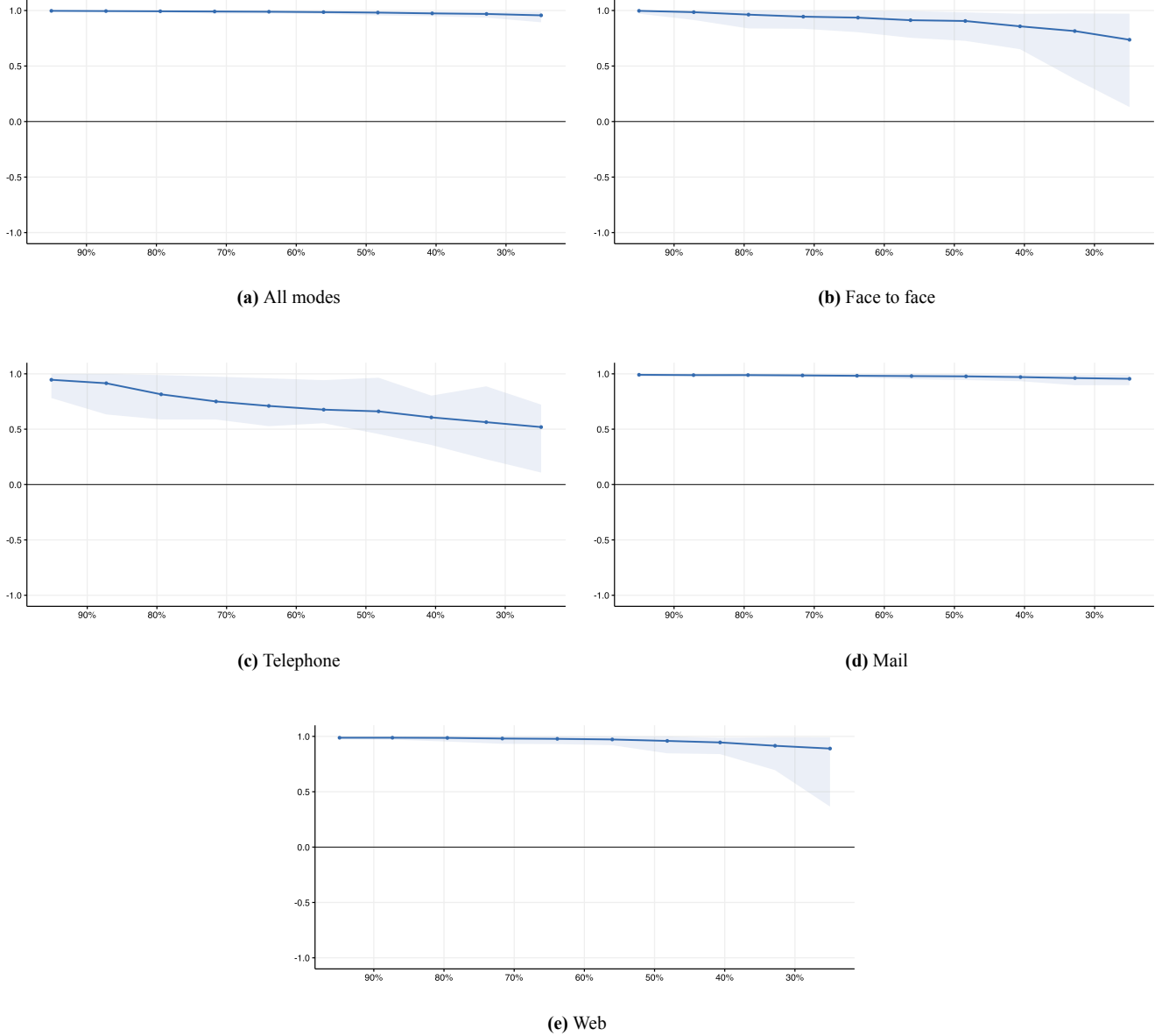


Figure 10: Case-dropping bootstrap samples ($n=1000$) of centrality indices (node strength) correlated with the centrality indices in original sample. Node strength is a stable measure across all networks. X-axis: percentage cases dropped, y-axis: correlation.

5 DISCUSSION

In this paper, we studied the impact of survey design characteristics on the relationship between the nonresponse rate and nonresponse bias using a novel network modelling method. Rising nonresponse rates have triggered several changes in how we design, conduct, and analyze survey data - and these choices are often the most challenging that survey researchers face. A long-held belief is that more expensive survey designs produce more accurate estimates through mitigating nonresponse. We demonstrate that the consequences of survey design are far from straightforward: each design choice affects another and these relationships are cyclic. The merit of our approach is in seeing these dynamics at once, enabling us to generate causal

hypotheses supported by our substantive knowledge of the survey literature and climate. If we base our theories solely on the theoretical link between nonresponse rate and nonresponse bias (Bethlehem 1988), the consequences appear clear. Yet, in our study, as in previous studies Beullens and Loosveldt (2012), we find no consistent or strong evidence that the nonresponse rate is a good indicator of nonresponse bias.

Consistent with the original meta-analysis by Groves et al. (2006), we find that nonresponse rates and nonresponse bias occur in an inconsistent manner and study-level characteristics have varied predictive power across survey modes. We demonstrated that survey mode is highly influential, and only by decomposing our results by survey mode did we find that survey characteristics operate differently over them. For instance, design choices that we make in good faith may not be having the desired effect given some other design choices we forego (e.g., using reminders but not incentives may not have an overall positive effect on response rates). While some of the decisions about survey design may be made due to the ease of implementation (e.g., reminders are low cost in online surveys), researchers should consider if they are necessary more than other design features especially if budgets are restricted. It should be stated that most of the networks do not implicate that survey design characteristics are hugely important for nonresponse bias. On the other hand, most can be related to nonresponse rates in some way to a less (i.e., mail) or more (i.e., face-to-face) degree. This could mean that the current state of survey design is pretty good. It could equally mean that the search of mechanisms of nonresponse rate and nonresponse bias continues. An interesting focus for further research is the more transient, fluctuating, circumstances that surround the *individual* respondent, such their personality, mood, backgrounds, and restrictions on their time.

An interesting observation we make, is that survey mode is directly associated with nonresponse rates *and* nonresponse bias, making it *common cause*. Referring to the causal models presented in Figure 1}, we see that common causes can induce spurious relations between two variables. Indeed, we only observe an effect between nonresponse rate and nonresponse bias comparable to the literature in the first network model. It (mostly) disappears when mode is stratified, suggesting that the relationship between nonresponse rates and nonresponse bias is conflated by mode and consequently in most studies that do not consider a comprehensive range of modes. Note that mail surveys present an exception where this relationship is intact. We note two possibilities: **(1)** as mentioned, the connection is spurious; and **(2)** the effect of nonresponse rates on nonresponse bias is specific to mail surveys. The latter is interesting to consider, since response rates are rather high in mail surveys (~42%) but nonresponse bias is low (~9%). If the effect is specific to mail surveys, it contradicts how the view that nonresponse rates are indicative of nonresponse bias. Still, it is possible that certain design features of mail surveys that boost response rates only do so for particular groups in the desired sample leading to biased estimates. We propose two potential scenarios where this could arise: **(1)** Reminders also improve response rates in mail surveys, as does survey sponsorship and health-related surveys. Sponsor bias can occur when respondents have preconceived opinions about the sponsoring organisation, causing them to respond in a more positive or negative way. **(2)** Health surveys are typically conducted in patient populations, meaning the topic is intrinsically salient to them and motivating them to respond. If the health survey is conducted at a time when some patients are most unwell, the response rate is high only among relatively healthy patients. In both cases, nonresponse declines but nonresponse bias increases.

Our study is based on a much larger, more diverse set of survey characteristics, and is the most extensive meta-analysis on the topic (Daikeler et al. 2020; compared to Groves et al. 2006; and Rybak 2023). In addition, this is the first practical implementation of a statistical method that is better aligned with the causal theories of nonresponse rate and nonresponse bias in surveys (Fay 1986; Groves et al. 2000; Kohler et al. 2022; Mercer et al. 2017). Let this only be the beginning: we encourage survey researchers to contextualize their findings within a causal perspective. This will lead to better survey design and more reliable survey estimates.

Some limitations to our study warrant discussion. First, the coding strategy depends on the information that is explicitly given in each study. There are cases where this information is ambiguous or absent (e.g., if a study did not mention incentives, we assumed that it was not a feature and coded it as “no”). It is possible that on occasion this approach was wrong. Moreover, the coding of “topic saliency” was a (somewhat subjective) decision made by the team of coders based on the stated topic, purpose, and sample. Also, the data collection and coding procedure is ongoing and, as such, inter-rater reliability has not been calculated. The coders did, however, discuss any disagreements or ambiguities and made a final decision together. Nevertheless, the coding strategy is open to human error and that error is not quantified here.

Secondly, as with any method, the parameters of the network model are estimated from the data and are not observed. Unlike other methods, this is a first application of network modelling to this topic and we have no “benchmark” study with a similar method to compare our results to (i.e., the estimated parameters could be wrong but the extent to which they could be wrong is vague). We feel assured that the main finding that nonresponse rates, nonresponse bias, and survey design are only weakly related corresponds with the traditional methods in the literature Rybak (2023). Furthermore, the results are moderately robust in terms of edge accuracy and very robust in terms of centrality stability.

Lastly, the issue of class imbalance impedes the quality of our sample and estimates based on highly imbalanced classes are unreliable (e.g., only 6.7% of telephone surveys mention incentivisation). Class imbalance also reduces the statistical power to detect statistical effects and it is possible that many of the (weak) edges in the networks models are a result of this. This is compounded by the conservative nature of the estimation procedure implemented in the `mgm` (i.e., both LASSO regularization and EBIC model selection). It is possible that with much larger sample sizes we may find stronger effects. We encourage the reader to keep these limitations in mind when interpreting the findings.

6 CONCLUSION

A novel method to an antiquated problem demonstrates that the contribution of nonresponse rate to nonresponse bias is small, but nuanced. Network visualisations expose us to this nuance in a way that is not possible with more traditional modelling frameworks like meta-regression. In a single model, we can see the intricacies of how survey design features are related to nonresponse rate and nonresponse bias in a way that is conducive to generating causal hypotheses. The numerity of interactions are beyond the scope of a single paper, but we hope to inspire critical thought about these network structures among survey researchers and to open a debate about the use of causal inference in survey research. Our results are contingent on what

is featured in the network models. There is scope for other survey design characteristics to be included (e.g., survey length, country, duration of fieldwork) in addition to characteristics concerning the individual (e.g., personality, demographics) and the wider survey climate.

7 REFERENCES

- AAPOR (2022), “[Best Practices for Survey Research - AAPOR](#),” *American Association for Public Opinion Research*.
- Becker, R. (2017), “Gender and Survey Participation An Event History Analysis of the Gender Effects of Survey Participation in a Probability-based Multi-wave Panel Study with a Sequential Mixed-mode Design,” *methods, data*, 29 Pages. <https://doi.org/10.12758/MDA.2021.08>.
- Beebe, T. J., Jacobson, R. M., Jenkins, S. M., Lackore, K. A., and Rutten, L. J. F. (2018), “Testing the Impact of Mixed-Mode Designs (Mail and Web) and Multiple Contact Attempts within Mode (Mail or Web) on Clinician Survey Response,” *Health Services Research*, 53, 3070–3083. <https://doi.org/10.1111/1475-6773.12827>.
- Bethlehem, J. (1988), “Reduction of nonresponse bias through regression estimation,” *Journal of Official Statistics*, 4, 251–260.
- Bethlehem, J. (2010), “Selection Bias in Web Surveys,” *International Statistical Review*, 78, 161–188. <https://doi.org/10.1111/j.1751-5823.2010.00112.x>.
- Bethlehem, J., and Cobben, F. (2008), “Indicators for the Representativeness of Survey Response,” in *Proceedings of Statistics Canada Symposium*, ed. B. Schouten, p. 8.
- Beullens, K., and Loosveldt, G. (2012), “Should high response rates really be a primary objective?” *Survey Practice*, 5, 1–5. <https://doi.org/10.29115/SP-2012-0019>.
- Beullens, K., Loosveldt, G., Vandenplas, C., and Stoop, I. (2018), “Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts?” *Survey Methods: Insights from the Field (SMIF)*. <https://doi.org/10.13094/SMIF-2018-00003>.
- Biemer, P. P. (2010), “Total Survey Error: Design, Implementation, and Evaluation,” *Public Opinion Quarterly*, 74, 817–848. <https://doi.org/10.1093/poq/nfq058>.
- Bowling, A. (2005), “Mode of questionnaire administration can have serious effects on data quality,” *Journal of Public Health*, 27, 281–291. <https://doi.org/10.1093/pubmed/fdi031>.
- Brick, J. M., and Tourangeau, R. (2017), “Responsive Survey Designs for Reducing Nonresponse Bias,” *Journal of Official Statistics*, 33, 735–752. <https://doi.org/10.1515/jos-2017-0034>.
- Brick, J. M., and Williams, D. (2013), “Explaining Rising Nonresponse Rates in Cross-Sectional Surveys,” *The ANNALS of the American Academy of Political and Social Science*, 645, 36–59. <https://doi.org/10.1177/0002716212456834>.
- Chen, S., Witten, D. M., and shojaie, A. (2015), “Selection and estimation for mixed graphical models,” *Biometrika*, 102, 47–64. <https://doi.org/10.1093/biomet/asu051>.
- Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B., and Wenz, A. (2020), “A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research,” *Journal of Survey Statistics and Methodology*, 8, 4–36. <https://doi.org/10.1093/jssam/smz041>.

- Curtin, R., Presser, S., and Singer, E. (2000), "The Effects of Response Rate Changes on the Index of Consumer Sentiment," *Public Opinion Quarterly*, 64, 413–428. <https://doi.org/10.1086/318638>.
- Czajka, J. L., and Beyler, A. (2016), "Declining Response Rates in Federal Surveys: Trends and Implications," Office of the Assistant Secretary for Planning; Evaluation.
- Daikeler, J., Bošnjak, M., and Lozar Manfreda, K. (2020), "Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates," *Journal of Survey Statistics and Methodology*, 8, 513–539. <https://doi.org/10.1093/jssam/smz008>.
- Draugalis, J. R., Coons, S. J., and Plaza, C. M. (2008), "Best Practices for Survey Research Reports: A Synopsis for Authors and Reviewers," *American Journal of Pharmaceutical Education*, 72, 11.
- Elliott, M. N., Edwards, C., Angeles, J., Hambarsoomians, K., and Hays, R. D. (2005), "Patterns of Unit and Item Nonresponse in the CAHPS Hospital Survey," *Health Services Research*, 40, 2096–2119. <https://doi.org/10.1111/j.1475-6773.2005.00476.x>.
- Epskamp, S., Borsboom, D., and Fried, E. I. (2018), "Estimating psychological networks and their accuracy: A tutorial paper," *Behavior Research Methods*, 50, 195–212. <https://doi.org/10.3758/s13428-017-0862-1>.
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012), "Qgraph : Network Visualizations of Relationships in Psychometric Data," *Journal of Statistical Software*, 48. <https://doi.org/10.18637/jss.v048.i04>.
- ESS (2016), *The European Social Survey (ESS) – a research instrument for the social sciences in Europe*.
- Fay, R. E. (1986), "Causal Models for Patterns of Nonresponse," *Journal of the American Statistical Association*, 81, 354–365. <https://doi.org/10.2307/2289224>.
- Fincham, J. E. (2008), "Response Rates and Responsiveness for Surveys, Standards, and the Journal," *American Journal of Pharmaceutical Education*, 72, 43.
- Gary, P. R. (2007), "Adjusting for Nonresponse in Surveys," in *Higher Education: Handbook of Theory and Research*, Dordrecht: Springer.
- Greenland, S., and Robins, J. M. (1986), "Identifiability, exchangeability, and epidemiological confounding," *International Journal of Epidemiology*, 15, 413–419. <https://doi.org/10.1093/ije/15.3.413>.
- Groves, R. M. (2006), "Nonresponse Rates and Nonresponse Bias in Household Surveys," *The Public Opinion Quarterly*, 70, 646–675.
- Groves, R. M., Couper, M. P., Presser, S., Singer, E., Tourangeau, R., Acosta, G. P., and Nelson, L. (2006), "Experiments in Producing Nonresponse Bias," *Public Opinion Quarterly*, 70, 720–736. <https://doi.org/10.1093/poq/nfl036>.
- Groves, R. M., and Peytcheva, E. (2008), "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis," *Public Opinion Quarterly*, 72, 167–189. <https://doi.org/10.1093/poq/nfn011>.
- Groves, R. M., Singer, E., and Corning, A. (2000), "Leverage-saliency theory of survey participation: Description and an illustration," *Public Opinion Quarterly*, 64, 299–308. <https://doi.org/10.1086/317990>.
- Gummer, T. (2019), "Assessing Trends and Decomposing Change in Nonresponse Bias: The Case of Bias in Cohort Distribu-

- tions,” *Sociological Methods & Research*, 48, 92–115. <https://doi.org/10.1177/0049124117701479>.
- Haslbeck, J. M. B., and Waldorp, L. J. (2015), “Structure estimation for mixed graphical models in high-dimensional data,” arXiv.
- Haslbeck, J. M. B., and Waldorp, L. J. (2020), “Mgm : Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data,” *Journal of Statistical Software*, 93. <https://doi.org/10.18637/jss.v093.i08>.
- Hedlin, D. (2020), “Is there a ‘safe area’ where the nonresponse rate has only a modest effect on bias despite non-ignorable nonresponse?” *International Statistical Review*, 88, 642–657. <https://doi.org/10.1111/insr.12359>.
- Hernán, M. A., and Robins, J. M. (2006), “Estimating causal effects from epidemiological data,” *Journal of Epidemiology and Community Health*, 60, 578–586. <https://doi.org/10.1136/jech.2004.029496>.
- Hua, T. T. (2023), “The Link Between Survey Response Rates and Nonresponse Bias: Theory, Simulations, and Empirical Evidence From the Household Pulse Survey,” 20.
- Keeter, S., Miller, C., Kohut, A., Groves, R. M., and Presser, S. (2000), “Consequences of Reducing Nonresponse in a National Telephone Survey,” *The Public Opinion Quarterly*, 64, 125–148.
- Koch, A., Halbherr, V., Stoop, I. A. L., and Kappelhof, J. W. S. (2014), *Assessing ESS sample quality by using external and internal criteria*, Mannheim: European Social Survey, GESIS.
- Kohler, U., Jabkowski, P., and Kołczyńska, M. (2022), “Against all odds: On the robustness of probability samples against decreases in response rates,” 44.
- Lee, M. H. (2011), “STATISTICAL METHODS FOR REDUCING BIAS IN WEB SURVEYS,” PhD thesis, Canada: Simon Fraser University.
- Lee, S., Brown, E. R., Grant, D., Belin, T. R., and Brick, J. M. (2009), “Exploring Nonresponse Bias in a Health Survey Using Neighborhood Characteristics,” *American Journal of Public Health*, 99, 1811–1817. <https://doi.org/10.2105/AJPH.2008.154161>.
- Lipps, O. (2009), “Attrition of Households and Individuals in Panel Surveys,” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1367371>.
- Lipps, O. (2010), “Effects of different Incentives on Attrition and Fieldwork Effort in Telephone Household Panel Surveys,” *Survey Research Methods*, 4, 81–90. <https://doi.org/10.18148/srm/2010.v4i2.3538>.
- Little, R., J. A., and Rubin, D. B. (2002), *Statistical Analysis with Missing Data*, Wiley Series in Probability and Statistics, Hoboken, NJ: John Wiley & Sons, Ltd.
- Lohr, S. L. (2021), *Sampling: Design and Analysis*, New York: Chapman; Hall/CRC. <https://doi.org/10.1201/9780429298899>.
- Luiten, A., Hox, J., and Leeuw, E. de (2020), “Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys,” *Journal of Official Statistics*, 36, 469–487. <https://doi.org/10.2478/jos-2020-0025>.
- Lundström, S., and Särndal, C.-E. (1999), “Calibration as a method for deriving nonresponse adjusted weights.” *Journal of Official Statistics*, 15, 305–327.

- Maslovskaya, O., and Lugtig, P. (2022), “Representativeness in Six Waves of Cross-National Online Survey (CRONOS) Panel,” *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185, 851–871. <https://doi.org/10.1111/rssa.12801>.
- McGonagle, K. A. (2020), “The Effects of an Incentive Boost on Response Rates, Fieldwork Effort, and Costs across Two Waves of a Panel Study,” *Methoden, daten, analysen*, 14, 241–250. <https://doi.org/10.12758/mda.2020.04>.
- Mercer, A. W., Kreuter, F., Keeter, S., and Stuart, E. A. (2017), “Theory and Practice in Nonprobability Surveys: Parallels between Causal Inference and Survey Inference,” *Public Opinion Quarterly*, 81, 250–271. <https://doi.org/10.1093/poq/nfw060>.
- Pearl, J. (2009), *Causality. Models, Reasoning, and Inference*, New York: Cambridge University Press.
- Petersen, M. L., Porter, K. E., Gruber, S., Wang, Y., and Laan, M. J. van der (2012), “Diagnosing and responding to violations in the positivity assumption,” *Statistical methods in medical research*, 21, 31–54. <https://doi.org/10.1177/0962280210386207>.
- R Core Team (2021), *R: A language and environment for statistical computing*, Vienna, Austria: R Foundation for Statistical Computing.
- Rybak, A. (2023), “Survey mode and nonresponse bias: A meta-analysis based on the data from the international social survey programme waves 1996–2018 and the European social survey rounds 1 to 9,” *PLOS ONE*, 18, e0283092. <https://doi.org/10.1371/journal.pone.0283092>.
- Sakshaug, J. W. (2022), “Reducing Nonresponse and Data Linkage Consent Bias in Large-Scale Panel Surveys,” *Forum for Health Economics and Policy*, 25, 41–55. <https://doi.org/10.1515/fhep-2021-0060>.
- Schoeni, R. F., McGonagle, K. A., and Andreski, P. (2012), “Response Rates in National Panel Surveys,” *The ANNALS of the American Academy of Political and Social Science*, 645, 60–87. <https://doi.org/10.1177/0002716212456363>.
- Schoot, R. van de, Bruin, J. de, Schram, R., Zahedi, P., Boer, J. de, Weijdemans, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., and Oberski, D. L. (2021), “An open source machine learning framework for efficient and transparent systematic reviews,” *Nature Machine Intelligence*, 3, 125–133. <https://doi.org/10.1038/s42256-020-00287-7>.
- Singer, E. (2011), “Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications,” *Journal of Official Statistics*, 27, 379–392.
- Southern, D. A., Lewis, S., Maxwell, C. J., Dunn, J. R., Noseworthy, T. W., Corbett, G., Thomas, K., and Ghali, W. A. (2008), “Sampling ‘hard-to-reach’ populations in health research: Yield from a study targeting Americans living in Canada,” *BMC Medical Research Methodology*, 8, 57. <https://doi.org/10.1186/1471-2288-8-57>.
- Struminskaya, B., and Gummer, T. (2022), “Risk of Nonresponse Bias and the Length of the Field Period in a Mixed-Mode General Population Panel,” *Journal of Survey Statistics and Methodology*, 10, 161–182. <https://doi.org/10.1093/jssam/smab011>.
- Toepoel, V., and Schonlau, M. (2017), “Dealing with nonresponse: Strategies to increase participation and methods for post-

- survey adjustments,” *Mathematical Population Studies*, 24, 79–83. <https://doi.org/10.1080/08898480.2017.1299988>.
- Tourangeau, R. (2017), “Presidential Address,” *Public Opinion Quarterly*, 81, 803–814. <https://doi.org/10.1093/poq/nfx031>.
- Tourangeau, R., Groves, R. M., and Redline, C. D. (2010), “Sensitive Topics and Reluctant Respondents: Demonstrating a Link between Nonresponse Bias and Measurement Error,” *Public Opinion Quarterly*, 74, 413–432. <https://doi.org/10.1093/poq/nfq004>.
- Van de Schoot, R., De Bruin, J., Schram, R., Zahedi, P., De Boer, J., Weijdem, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Tummers, L., and Oberski, D. (2022), *ASReview: Active learning for systematic reviews*, Zenodo. <https://doi.org/10.5281/zenodo.6591802>.
- Yang, E., Baker, Y., Ravikumar, P., Allen, G., and Liu, Z. (2014), “Mixed Graphical Models via Exponential Families,” in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 1042–1050.

APPENDIX

Tables 4 to 8 contain the interactions disaggregated over the nodes in the network, including their sign (+/-). Row and column names follow the same pattern as the node labels presented in the main paper, with .0, .1, .2, .3, and .4 representing the levels. For binary variables, .0 always represents "no" and .1 always represents "yes". For the survey mode (N11), .0, .1, .2, .3, and .4 represent "paper", "face-to-face", "mail", "web", and "telephone" modes respectively. For the question type, .0, .1, and .2 represent questions relating "demographic", "observable", and "unobservable" characteristics respectively. For survey topic, .0 and .1 represent "miscellaneous" and "health" surveys.

Table 4: Disaggregated interactions between all node levels in the overall network.

	Predict Columns														Predict Rows																																		
	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2	N11.0	N11.1	N11.2	N11.3	N11.4	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2	N11.0	N11.1	N11.2	N11.3	N11.4			
N1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N4.0	-	-	-	-	-	-	0.36	-	-0.20	-	-0.46	-	-1.36	-	-0.74	-	-0.28	0.00	-	-0.21	0.22	-0.17	1.76	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.1	-	-	-	-	-	-	-0.36	-	0.20	-	0.46	-	1.36	-	0.74	-	0.28	0.00	-	0.21	-0.22	0.17	-1.76	-	-	-	-	-	0.59	-0.59	-0.19	0.19	-0.32	0.32	-1.26	1.26	-0.71	0.71	0.00	0.82	0.00	-1.05	0.69	0.00	0.48	-2.30	-		
N5.0	-	-	-	-	-	-	-	-	-0.53	-	-1.15	-	0.77	-	-	-	0.31	0.22	-	-2.35	-1.53	-1.17	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N5.1	-	-	-	-	-	-	-	0.53	-	1.15	-	-0.77	-	-	-	-0.31	-0.22	-	2.35	1.53	1.17	0.00	-	-	-	-	-	-	-	-0.60	0.60	-0.79	0.79	0.69	-0.69	-	-	0.23	0.00	-0.03	-5.12	2.82	0.70	0.00	-1.03	-	-	-	
N6.0	-	-	-	-	-	-	-	-	-	-	-0.07	-	0.40	-	0.19	-	0.10	-0.25	-	0.20	-0.75	0.16	-0.92	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N6.1	-	-	-	-	-	-	-	-	-	0.07	-	-0.40	-	-0.19	-	-0.10	0.25	-	-0.20	0.75	-0.16	0.92	-	-	-	-	-	-	-	-	-	-	-0.14	0.14	0.48	-0.48	0.26	-0.26	0.00	-0.14	0.49	0.00	-0.57	1.53	-0.33	1.30	-	-	
N7.0	-	-	-	-	-	-	-	-	-	-	-	-	-0.79	-	0.21	-	-0.35	-0.32	-	0.62	-0.15	-0.82	-0.55	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N7.1	-	-	-	-	-	-	-	-	-	-	-	-	0.79	-	-0.21	-	0.35	0.32	-	-0.62	0.15	0.82	0.55	-	-	-	-	-	-	-	-	-	-	-	-0.63	0.63	0.20	-0.20	-0.43	0.00	0.05	0.00	-2.07	-0.27	0.49	0.15	-	-	
N8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.89	-	0.00	0.14	-	1.72	0.78	1.68	1.26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N8.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.89	-	0.00	-0.14	-	-1.72	-0.78	-1.68	-1.26	-	-	-	-	-	-	-	-	-	-	-	-	0.92	-0.92	0.00	0.12	-0.25	4.07	-0.67	1.22	-0.26	0.00	-	-	
N9.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00	-0.25	-	1.56	0.67	0.72	0.15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N9.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00	0.25	-	-1.56	-0.67	-0.72	-0.15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.03	0.00	0.40	2.45	-1.38	0.00	-0.06	0.40	-	-	-
N10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N10.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N10.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N11.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N11.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N11.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N11.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N11.4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 5: Disaggregated interactions between all node levels in the face-to-face network.

	Predict Columns																Predict Rows																			
	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2
N1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.0	-	-	-	-	-	-	2.34	-	-3.06	-	-1.35	-	-4.14	-	-7.38	-	0.00	-0.96	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.1	-	-	-	-	-	-	-2.34	-	3.06	-	1.35	-	4.14	-	7.38	-	0.00	0.96	-	-	-	-	-	30.19	-30.19	-7.3	7.3	-7.76	7.76	-2.74	2.74	-37.71	37.71	0.00	0.00	82.88
N5.0	-	-	-	-	-	-	-	-	-	-	-8.91	-	-16.31	-	-	-	10.52	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N5.1	-	-	-	-	-	-	-	-	-	-	8.91	-	16.31	-	-	-	-10.52	0.00	-	-	-	-	-	-	-	-	-10.45	10.45	0.00	0.00	-	-	0.00	-16.65	50.18	
N6.0	-	-	-	-	-	-	-	-	-	-	-18.54	-	3.78	-	0.00	-	0.00	-14.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N6.1	-	-	-	-	-	-	-	-	-	-	18.54	-	-3.78	-	0.00	-	0.00	14.41	-	-	-	-	-	-	-	-	-6.67	6.67	1.89	-1.89	4.25	-4.25	0.00	-16.01	25.49	
N7.0	-	-	-	-	-	-	-	-	-	-	0.00	-	3.64	-	-7.21	5.95	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N7.1	-	-	-	-	-	-	-	-	-	-	0.00	-	-3.64	-	7.21	-5.95	-	-	-	-	-	-	-	-	-	-	-	-	-0.82	0.82	9.27	-9.27	0.00	26.42	-47.81	
N8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.74	-	1.50	0.80	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N8.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.74	-	-1.50	-0.80	-	-	-	-	-	-	-	-	-	-	-	-	0.00	0.00	13.87	0.00	-9.79	
N9.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N9.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		

Table 6: Disaggregated interactions between all node levels in the telephone network.

	Predict Columns																Predict Rows																			
	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2
N1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.0	-	-	-	-	-	-	-3.92	-	5.30	-	-1.37	-	0.99	-	-2.30	-	-1.24	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.1	-	-	-	-	-	-	3.92	-	-5.30	-	1.37	-	-0.99	-	2.30	-	1.24	0.00	-	-	-	-	-	0	0	6.53	-6.53	-4.84	4.84	-1.72	1.72	-5.57	5.57	-0.25	8.14	0.00
N5.0	-	-	-	-	-	-	-	-	-7.61	-	-	-	-6.39	-	-	-	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N5.1	-	-	-	-	-	-	-	-	7.61	-	-	-	6.39	-	-	-	0.00	0.00	-	-	-	-	-	-	-9.64	9.64	-	-	-11.66	11.66	-	-	2.50	0.00	0.00	
N6.0	-	-	-	-	-	-	-	-	-	-	-13.57	-	1.76	-	-3.62	-	0.00	-4.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N6.1	-	-	-	-	-	-	-	-	-	-	13.57	-	-1.76	-	3.62	-	0.00	4.53	-	-	-	-	-	-	-	-	-	-9.64	9.64	3.49	-3.49	-11.06	11.06	0.00	-2.87	3.80
N7.0	-	-	-	-	-	-	-	-	-	-	-	-	-6.24	-	4.90	-	0.00	2.46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N7.1	-	-	-	-	-	-	-	-	-	-	-	-	6.24	-	-4.90	-	0.00	-2.46	-	-	-	-	-	-	-	-	-	-	-	-4.76	4.76	19.78	-19.78	0.00	1.67	-4.75
N8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.67	-	0.00	8.23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N8.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.67	-	0.00	-8.23	-	-	-	-	-	-	-	-	-	-	-	-	0.00	0.00	0.68	-3.24	0.00	
N9.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N9.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		

Table 7: Disaggregated interactions between all node levels in the network.

	Predict Columns																Predict Rows																			
	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2
N1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N4.0	-	-	-	-	-	-	1.12	-	-0.24	-	-1.48	-	-2.42	-	-1.56	-	0.30	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N4.1	-	-	-	-	-	-	-1.12	-	0.24	-	1.48	-	2.42	-	1.56	-	-0.30	0.00	-	-	-	-	-	1.05	-1.05	-0.19	0.19	-0.82	0.82	-2.01	2.01	-1.10	1.10	0.0	1.11	-0.47
N5.0	-	-	-	-	-	-	-	-	-0.47	-	-1.80	-	1.20	-	-0.14	-	0.65	0.22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N5.1	-	-	-	-	-	-	-	-	0.47	-	1.80	-	-1.20	-	0.14	-	-0.65	-0.22	-	-	-	-	-	-	-	-0.46	0.46	-1.05	1.05	1.17	-1.17	-0.19	0.19	0.6	0.00	0.00
N6.0	-	-	-	-	-	-	-	-	-	-	0.00	-	0.67	-	0.45	-	0.00	-0.15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N6.1	-	-	-	-	-	-	-	-	-	-	0.00	-	-0.67	-	-0.45	-	0.00	0.15	-	-	-	-	-	-	-	-	-	-0.08	0.08	0.78	-0.78	0.50	-0.50	0.0	-0.11	0.34
N7.0	-	-	-	-	-	-	-	-	-	-	-	-	-0.83	-	0.39	-	0.00	-0.59	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N7.1	-	-	-	-	-	-	-	-	-	-	-	-	0.83	-	-0.39	-	0.00	0.59	-	-	-	-	-	-	-	-	-	-	-0.92	0.92	0.26	-0.26	0.0	-0.19	1.27	
N8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.38	-	-0.23	0.17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N8.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.38	-	0.23	-0.17	-	-	-	-	-	-	-	-	-	-	-	-	1.35	-1.35	0.0	1.39	-0.63	
N9.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N9.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N10.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N10.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 8: Disaggregated interactions between all node levels in the web network.

	Predict Columns														Predict Rows																					
	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2	N1	N2	N3	N4.0	N4.1	N5.0	N5.1	N6.0	N6.1	N7.0	N7.1	N8.0	N8.1	N9.0	N9.1	N10.0	N10.1	N10.2
N1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
N2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.0	-	-	-	-	-	-	0	-	-	-	0.63	-	-1.05	-	0.00	-	-1.23	0.18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N4.1	-	-	-	-	-	-	0	-	-	-	-0.63	-	1.05	-	0.00	-	1.23	-0.18	-	-	-	-	-	7.15	-7.15	-	-	0.93	-0.93	-3.66	3.66	7.79	-7.79	0	7.75	-1.06
N5.0	-	-	-	-	-	-	-	-	-	-	5.91	-	10.80	-	10.96	-	-12.45	7.79	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N5.1	-	-	-	-	-	-	-	-	-	-	-5.91	-	-10.80	-	-10.96	-	12.45	-7.79	-	-	-	-	-	-	-	-	1.72	-1.72	16.46	-16.46	18.95	-18.95	0	5.66	-8.93	
N6.0	-	-	-	-	-	-	-	-	-	-	1.46	-	3.19	-	6.36	-	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N6.1	-	-	-	-	-	-	-	-	-	-	-1.46	-	-3.19	-	-6.36	-	0.00	0.00	-	-	-	-	-	-	-	-	0.00	0.00	6.74	-6.74	15.19	-15.19	0	0.00	-0.80	
N7.0	-	-	-	-	-	-	-	-	-	-	-	-	0.00	-	1.45	-	-2.86	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N7.1	-	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-1.45	-	2.86	0.00	-	-	-	-	-	-	-	-	-	-	2.15	-2.15	15.25	-15.25	0	0.00	0.00	
N8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N8.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N9.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N9.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
N10.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		

SUPPLEMENTARY INFORMATION

The following information can be sourced from Github: https://github.com/shandickson/thesis_masters \ R-file

data_preparation.Rmd containing code to perform the data preparation before analysis.

R-file **data_descriptives.Rmd** containing code to obtain the descriptive statistics, tables, and images seen in the manuscript.

R-file **data_networks.Rmd** containing code to perform the network analysis seen in the manuscript.

R-file **functions.R** containing all custom functions used in all of the above files.

R-file **data.RDS** containing the original data used in the analysis.