LINKING SURVEY AND LINKEDIN DATA: UNDERSTANDING USAGE AND CONSENT PATTERNS

TAREK AL BAGHAL [iD]*
ALEXANDER WENZ [iD]
PAULO SERÔDIO
SHUJUN LIU
CURTIS JESSOP
LUKE SLOAN

Linking social media and survey data allows researchers to create novel metrics, track survey respondents longitudinally, and capture real-time intricate data. LinkedIn is a social networking platform with a strong emphasis on employment and business interactions which has not been explored in a survey data linkage context previously. Using a nationally representative panel survey, we explore the feasibility of linking survey and LinkedIn data. Our analysis focuses first on understanding the demographic profile of LinkedIn users in the UK, as understanding this context shows the coverage of who can be asked to consent. We then explore consent outcomes, assessing the impact of question placement and wording on participants' willingness to link data, and identifying other factors associated with LinkedIn consent. Our findings reveal that a notable proportion of respondents have a LinkedIn account, indicating a higher usage rate compared to previous research. Employment status, education level, and income are key determinants of having a LinkedIn

TAREK AL BAGHAL is Deputy Director of Understanding Society and Professor of Survey Methodology at University of Essex, Colchester, UK. ALEXANDER WENZ is Research Fellow at Mannheim Centre for European Social Research, University of Mannheim, Mannheim, Germany. PAULO SERÔDIO, Senior Research Officer at University of Essex, Colchester, UK. SHUJUN LIU is Research Associate at Cardiff University, Cardiff, UK. CURTIS JESSOP is Director of Attitudinal Surveys & the NatCen Panel at NatCen Social Research, London, UK. LUKE SLOAN is Professor at Cardiff University, Cardiff, UK.
*Address correspondence to Tarek Al Baghal, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom; E-mail: talbag@essex.ac.uk

account. Unlike previous studies, consent placement and wording does not have a significant influence on participants' willingness to link LinkedIn and survey data. However, social media posting frequency and possessing a university degree are associated with higher consent rates. Although LinkedIn users differ on certain demographic characteristics, the platform's unique focus on professional networking and career-related activities presents valuable opportunities for investigating employment and economic outcomes.

Statement of Significance

This article contributes significantly to the evolving landscape of social science by exploring the integration of survey and LinkedIn data. By delving into the feasibility of linking these two distinct data sources, the study addresses pressing questions surrounding data quality improvement and novel research avenues. The findings not only expand on the potential of combining social media and survey data, but also shed light on the utilization of LinkedIn, a platform with a pronounced professional focus. Revealing a higher prevalence of LinkedIn usage than previously estimated, this study highlights the changing dynamics of online interactions, particularly within the realm of employment and career-related activities. The investigation into consent dynamics and the identification of factors associated with LinkedIn usage and consent patterns underscore the nuanced interplay between participant behavior, consent, and data linkage. This research can guide future endeavors that seek to harness the synergies between survey and social media data, offering researchers novel insights into human behavior, social networks, and economic dynamics within an increasingly interconnected digital society.

---

1. INTRODUCTION

Linking social media and survey data has the potential to address novel research questions and improve data quality. Linked data allows researchers to generate new measures, follow survey respondents over time, collect detailed data in real-time, and make adjustments for survey nonresponse (Al Baghal et al. 2020). Several studies explore such linkages, with particular focus on ethics (Sloan et al. 2020), consent within the survey context (Al Baghal et al. 2020; Mneimneh 2022; Silber et al. 2022; Beuthner et al. 2023), and the

asymmetric nature of the two data sources (Al Baghal et al. 2021). Other research focuses on these linkages to identify the error in reported and actual social media use (Guess et al. 2019; Haenschen 2020; Henderson et al. 2021).

Most studies on these linkages use Twitter/X as the social media data source, with consent rates ranging from 27 to 53 percent (e.g., Karlsen and Enjolras 2016; Al Baghal et al. 2020; Henderson et al. 2021; Mneimneh et al. 2021; Mneimneh 2022). Twitter/X has been preferred largely due to the straightforward data access via its Application Programming Interface. However, other social media sources may be more useful for particular research questions. LinkedIn, a widely used social media platform centered on professional networking and career opportunities, boasts over 900 million users across over 200 countries and territories (LinkedIn 2023). Its extensive user base, exclusive content, and social networking features render it an invaluable data source to link to survey data for research on economic, employment, and social network topics. Studies have used LinkedIn data to explore how its use affects career benefits (Davis et al. 2020), to identify personality traits (Fernandez et al. 2021), and to examine whether connectedness improves start-up entrepreneurs' fundraising (Banerji and Reimer 2019). Connecting LinkedIn data with survey data would enable new research with these types of outcomes, for example, the effect of professional networks on income, job satisfaction, and life satisfaction. The longitudinal nature of social media data paired with panel surveys could also uncover how online and offline behaviors change.

We examine the feasibility of linking survey and LinkedIn data in a probability household panel in the UK by addressing the following research questions.

RQ1: What is the level of (a) LinkedIn usage and (b) consent for linking survey and LinkedIn data?

RQ2: Does the consent wording (with versus without importance statement) and the placement within the survey (early versus late) affect the consent to link survey and LinkedIn data?

RQ3: What are the respondent-level determinants of (a) LinkedIn usage and (b) consent to link survey and LinkedIn data?

2. BACKGROUND

Understanding whose data are available to link from LinkedIn or other data sources provides insight into its utility and is key to understand the scope and implications of consent. Like other social media platforms, LinkedIn may encounter undercoverage errors (Hsieh and Murphy 2017). The extent of undercoverage varies depending on the target population. For example, research focusing on active labor market participants is likely to face less undercoverage error than studies of the general population. When linking these data to surveys, non-users

in the sample are identifiable; given the wealth of data on both users and non-users, it is possible to identify the extent of these coverage errors.

LinkedIn's usage, though not as extensive as other social media platforms, shows interesting trends. Using a 2013 probability national sample of the United Kingdom, Blank and Lutz (2017) find that 9 percent of the population has a LinkedIn account, and these users are more likely to have higher income levels, be divorced, and have more experience using the internet. More recent nonprobability surveys from the United Kingdom's media regulator (Ofcom) show that around 16–17 percent of the population use LinkedIn, except in late 2020 and January 2021, showing usage at 27 percent (Ofcom 2019–2023). In the United States, LinkedIn usage shows a similar pattern, climbing from 27 percent in 2021 to 30 percent in 2023 (Pew Research Center 2024). The most recent Ofcom data show that in the United Kingdom, LinkedIn is more popular among men, aged 25–34 years, and those in managerial or professional roles (Ofcom 2019–2023).

Given the low consent rates for social media data linkage, identifying factors that can influence outcomes is key to improving consent rates, which helps reduces the risks of bias and increases statistical power and fieldwork efficiency. While this study focuses on the linkage between social media and survey data, the mechanisms of consent to link these data are not treated dissimilarly to other linkage consent research. Mneimneh (2022) developed a framework to understand social media and survey data linkage, which they base on frameworks from research on administrative and survey data linkage. Mechanisms have been similar in these different linkages; for example, lower privacy concerns are strong predictors of higher consent likelihood for both administrative data (Sakshaug et al. 2012) and social media linkage (Mneimneh 2022; Liu et al. 2024). As with administrative data linkage, consent to social media linkage rates vary by respondent characteristics. For example, in some surveys, women are less likely to consent than men (Al Baghal et al. 2020; Mneimneh 2022; Silber et al. 2022), while in others, there is no difference (Al Baghal et al. 2020; Mneimneh et al. 2021).

Survey design could also influence the linkage consent decision. Interviewer-administered surveys lead to higher consent rates across most requests (Jäckle et al. 2021), including social media linkage (Al Baghal et al. 2020). Placing the linkage request earlier in the questionnaire improves consent rates (e.g. Sakshaug et al. 2013), including in social media linkage requests (Beuthner et al. 2023). Wording highlighting the scientific value does not improve linkage consent in a variety of domains (Beuthner et al. 2023), and wording highlighting potential time savings have inconsistent effects (Sakshaug et al. 2013; Sakshaug and Kreuter 2014; Struminskaya et al. 2021). Comparisons of wording stating that respondents' survey data would be less valuable without linkage to wording stating their data would be more valuable with linkage similarly shows mixed results (Sakshaug et al. 2019; Kreuter et al. 2016).

3. DATA AND METHODS

This paper uses data from the fourteenth wave of the Innovation Panel (IP14), which is part of *Understanding Society*: The UK Household Longitudinal Study (UKHLS), conducted between May 19 and October 3, 2021 (University of Essex, Institute for Social and Economic Research 2021). The IP is a vehicle for methodological experimentation in a longitudinal survey design and is conducted annually. Interviews are attempted with all household members who are 16+ years of age. It uses a multi-stage probability proportionate to size sample of persons and households in Great Britain. A boost sample was added at IP14, with addresses selected using the same design, with an invitation sent by mail. Due to Covid-19 restrictions, all households were invited to complete the survey online; follow-ups with non-responding continuing sample households were attempted via telephone. The household response rate for the issued continuing sample was 62.7 percent and was 12.8 percent for the boost sample (AAPOR RR2). There were a small number of telephone interviews ($n = 265$), particularly among those asked the consent question ($n = 34$). Further, the IP14 boost sample could only respond online, as no telephone number was available. Therefore, only web responses are used for analysis.

We included questions on LinkedIn usage and for consent to link respondents' LinkedIn data to their responses. Two experiments were included to identify design impacts on consent. First, half of the respondents were asked questions earlier in the survey, and half asked later. IP14 includes significant routing, so exact question position within the survey differs by individual. However, those in the early condition received the question on average 5.2 minutes into their interview; those in the late condition received the consent question 33.1 minutes into their interview.

Second, to identify the impact of a statement of importance on consent, half of the consent requests included the additional sentences: "The data you would provide is key to this study. This data will enhance the understanding of your survey responses." This statement highlights the benefit linkage has, which was identified as an important motivator for consent to social media linkage in focus groups (Di Cara et al. 2020). The wording draws from Al Baghal and Lynn's (2015) findings, which demonstrate the efficacy of this statement in mitigating item nonresponse. Experimental conditions were allocated at the individual level for respondents in the continuing sample and at the household level for the IP14 boost sample. The consent question asked:

> We would like to know who uses LinkedIn, and how people use it. We are also interested in being able to link the information people have provided for this study to publicly available information from their LinkedIn accounts such as their employment or education history, their connections, or information about their employer.

Information collected from your LinkedIn account will be treated as confidential and protected in the same way as your interview data. Any LinkedIn information that would allow you to be identified will not be published.

The data you would provide is key to this study. This data will enhance the understanding of your survey responses. *(if in extra statement condition)*

Are you willing to tell us the name of your personal LinkedIn account and for your LinkedIn information to be linked with the information you have provided for this study?

We account for the complex design in variance estimation. However, some strata only contained one cluster, and so only clusters ($n = 151$) are used in variance estimation, which should provide more conservative standard errors. The preferred reporting items for complex sample survey analysis (Seidenberg et al. 2023) are presented in Appendix A. We conduct initial analyses exploring the experimental conditions alone. To understand the profile of those having a LinkedIn account and those consenting to data linkage, we estimate logistic regression models including experimental conditions (for consent) and respondent characteristics. Given the focus of LinkedIn, we include variables around labor force activity, including the respondents' employment status (in employment, unemployed, not in workforce), education (university/higher degree, professional or secondary qualifications, less than A-level) and personal monthly income. Full question wording is presented in Appendix B. The study derives monthly income from several survey questions and imputes missing income values following Little and Su's (1989) longitudinal method (Fisher and Hussein 2023). Given the skewed distribution of income, the natural log is used in the analysis.

Previous research shows more frequent social media use is also related to higher consent rates (Mneimneh 2022). We include social media behavior as indicated by whether they look and/or post to any social media frequently (at least several times a week compared to less frequently). Marital status (single, married, separated, widowed) is included as it is related to both having an account and linkage consent, although the reason for this is unclear (Blank and Lutz 2017; Mneimneh 2022). We control for age and sex, as these frequently are related to consent outcomes, but not in consistent ways (Jäckle et al. 2021). We include an indicator for whether or not a person is in the boost sample to identify potential sample effects. Models estimated use listwise deletion for missing data. Summary statistics are presented in Appendix C.

4. RESULTS

Overall, 26.2 percent ($n = 722$) of the sample has a LinkedIn account. There is no expectation that survey placement should affect reporting having an account, and no difference is found ($F(1,150) = 0.003$ $p = .96$). Table 1 shows the percentage of respondents reporting having a LinkedIn account that consented to link these data to their survey responses. Data are shown by whether the consent questions were asked early or late and if the importance statement was included.

The Rao–Scott χ^2 test is suggestive that the timing of asking for consent and including the statement might influence the likelihood of obtaining consent, although not significant at $p < .05$ ($p = .065$). Of those with an account, the consent rate is 6.5 percent higher for those asked early (44.2 percent) than those asked later (37.7 percent). This result is directionally consistent with past findings exploring consent for social media data linkage (Mneimneh 2022; Beuthner et al. 2023). Consent rates where the additional statement was presented are also 7.1 percent higher (44.7 percent) than where it was not (37.6 percent). A logistic regression (not shown) of the two conditions and the interaction of the two on consent show the interaction between location and inclusion of the statement is not significant ($F(1,142) = 0.08$, $p = .78$). However, the average marginal effects (AME) for the first-order effects of this model for being asked early (AME $= 0.066$, SE $= 0.036$, $p = .068$) and being provided the additional statement (AME $= 0.071$, SE $= 0.039$, $p = .067$) both trend toward higher levels of consent. Given these findings, all further analysis will focus on the experimental conditions separately alongside other covariates to better identify the individual contributions of the experimental conditions on consent rates.

To identify who has a LinkedIn account and who consents to link this social media account to their survey data, logistic regression models are estimated

Table 1. LinkedIn Consent to Linkage by Placement and Statement

	LinkedIn consent		
	No statement	With statement	Total (consent)
Asked early	40.3%	48.6%	44.2%
	191	*173*	*364*
Asked late	34.9%	40.8%	37.7%
	189	*169*	*358*
Total	37.6%	44.7%	41.0%
	380	*342*	*722*

NOTE.— Sample cell size n in italics. Rao-Scott $\chi^2_1 = 3.412$, $p = .065$.

Table 2. Odds Ratios and Confidence Intervals of Having a LinkedIn Account and Consent to Data Linkage

	Has LinkedIn Odds ratio (95% CI)	LinkedIn consent Odds ratio (95% CI)
Consent placement early	—	1.225
		(0.892, 1.681)
Additional statement	—	1.388
		(0.992, 1.943)
Boost sample	0.957	0.874
	(0.749, 1.222)	(0.638, 1.198)
Employment (Ref: Not in workforce)		
In employment	2.355*	1.203
	(1.741, 3.186)	(0.761, 1.903)
Unemployed	1.481	0.395
	(0.826, 2.655)	(0.107, 1.447)
Education (Ref: Less than A-level)		
University degree	5.584*	1.790*
	(4.065, 7.673)	(1.052, 3.045)
Professional/A-level	2.412*	1.689
	(1.736, 3.352)	(0.937, 3.046)
Income (log)	1.044*	0.971
	(1.015, 1.074)	(0.931, 1.012)
Age	0.986*	0.993
	(0.976, 0.995)	(0.979, 1.008)
Female	0.628*	0.854
	(0.519, 0.759)	(0.633, 1.152)
Marital status (Ref: single)		
Married	1.200	1.024
	(0.932, 1.544)	(0.666, 1.574)
Separated	1.129	1.303
	(0.768, 1.660)	(0.667, 2.548)
Widowed	1.378	1.521
	(0.737. 2.578)	(0.494, 4.682)
Social media usage		
Look frequently	2.027*	1.022
	(1.541, 2.666)	(0.621, 1.684)
Post frequently	0.869	1.605*
	(0.694, 1.086)	(1.155, 2.229)
n	*2582*	*689*

*$p < .05$.

using the covariates described above. For consent, the experimental allocations are also included. Table 2 presents the results of the models. As might be expected given LinkedIn's purpose, those in employment are significantly

more likely to have an account than those not in the workforce. A contrast test of coefficients shows that the difference in estimated likelihood of those in employment and those unemployed is not significant at $p < .05$ ($F(1,150) = 2.75, p = .099$).

Additional indicators show the effect of socioeconomic status on having a LinkedIn account. Education is strongly predictive of LinkedIn usage, with university graduates having 5.58 times greater odds of having an account than those with less than professional/A-level qualifications. Those with professional/A-level qualifications also are significantly more likely to have a LinkedIn account relative to those with less education. A coefficient contrast test shows that the estimate for university graduates having an account is significantly larger than for those with professional/A-level qualifications ($F(1,150) = 51.53, p < .001$). In line with the overall pattern on socioeconomic status, those with higher incomes also are more likely to have a LinkedIn account. Age and sex are predictive of LinkedIn usage, with younger and male respondents more likely to have an account. Frequent use of the internet to look at social media is related to having an account, but posting frequency is not significantly related. Marital status is not significantly related to having a LinkedIn account, contrary to previous research (Blank and Lutz 2017).

The model predicting consent among users shows fewer significant effects. Controlling for other factors, asking earlier in the survey does not have a significant impact on social media consent, unlike previous research (Mneimneh 2022). The additional statement may encourage consent, with those receiving this statement being an estimated 1.39 times more likely to consent, although it does not fall below $p < .05$ ($p = .056$). Education and social media posting behavior are the only other apparent characteristics associated with consent. Those with a university degree are more likely to consent to linkage than those with less than professional/A-level qualifications, although there does not appear to be a strong relationship between consent and having professional/A-level qualifications ($p = .081$). The difference in the effects between having a university degree and professional/A-level qualifications is not significant; however ($F(1,142) = 0.11, p = .75$). The effects of social media usage are reversed from whether someone has a LinkedIn account; those who post frequently on social media are significantly more likely to consent, while the effect of looking at social media frequently is not significant.

5. DISCUSSION

Linking survey data to multiple social media data sources could enhance both data sources. Recent studies focus on linking Twitter/X and survey data (Al Baghal et al. 2020; Mneimneh et al. 2021; Mneimneh 2022), but Twitter/X is only one source, and changes in data availability show that additional

sources should be considered. The experiment reported here provides evidence for the possibility of linking LinkedIn data to individuals' survey responses. This is the first identified study to explore linking representative survey data with this social media data source.

The first finding of note is on the relatively high usage of LinkedIn in our sample. Previous surveys have suggested usage of this social media site to be between 9 and 25 percent of the UK adult population. The percentage of users in this study is at the high-end of this range. This survey took place in 2021 during the Covid pandemic, which significantly affected the labor market (Blundell et al. 2022). Given LinkedIn's use to explore economic opportunities, it could be the increase was context-specific. The data show that those in employment are most likely to have a LinkedIn account, as expected. Those with higher education and income are more likely to have accounts, identifying the site's attraction to those with higher socioeconomic status. Men are also more likely to have a LinkedIn account.

There is weak evidence from the experiment conducted that asking for consent earlier and providing an additional statement of importance improve consent to linkage rates. While borderline significant, these findings are consistent with other research on consent (Jäckle et al. 2021) and the effects of similar statements on item nonresponse (Al Baghal and Lynn 2015). Few of the other tested characteristics are related to consent to linking LinkedIn data and survey responses. Only those with a university degree and those who posted social media content frequently are more likely to consent. That those more active on social media are more likely to consent may be a positive in that there may be more data to collect for these respondents. The lack of relationship to other respondent characteristics is a positive in this instance; significant effects suggest possible biases in the represented sample.

However, it is important to note that consent is only asked of those who have a LinkedIn account, and any representation is limited to that specific group. LinkedIn users are not representative of the general population; but this is the case for all social media platforms (Blank and Lutz 2017). The actual data obtained is further constrained to those providing consent. For example, we find those with a university degree are more likely to both have a LinkedIn account and to consent, meaning the potentially available linked LinkedIn and survey data will be from a significantly more educated subgroup than the general population.

The current study shows the possibility of asking for consent to link to LinkedIn data within a survey context; however, it does not explore the actual linkage of these data sources. The IP14 survey collected possible identifiers (e.g., name, employer) that could identify respondents' LinkedIn data. Initial review suggests only two respondents did not provide any account information. Future research should explore methods for linking to LinkedIn data using the collected identifiers and other data available from the survey records. This study only explored varying two survey design features to improve linkage

consent rates; alternatives may be more effective and continued study in identifying these are also needed.

Supplementary Materials

Supplementary materials are available online at academic.oup.com/jssam.

REFERENCES

Al Baghal, T., Sloan, L., Jessop, C., Williams, M., and Burnap, P. (2020), "Linking Twitter and Survey Data: The Impact of Survey Mode and Demographics on Consent Rates Across Three UK Studies," *Social Science Computer Review*, 38, 517–532.

Al Baghal, T., and Lynn, P. (2015), "Using Motivational Statements in Web Instrument Design to Reduce Item Missing Rates in a Mixed-Mode Context," *Public Opinion Quarterly*, 79, 568–579.

Al Baghal, T., Wenz, A., Sloan, L., and Jessop, C. (2021), "Linking Twitter and Survey Data: Quantity and Possible Biases," *EPJ Data Science*, 10.

Banerji, D., and Reimer, T. (2019), "Startup Founders and Their LinkedIn Connections: Are Well-Connected Entrepreneurs More Successful?," *Computers in Human Behavior*, 90, 46–52.

Beuthner, C., Weiß, B., Silber, H., Keusch, F., and Schröder, J. (2023), "Consent to Data Linkage for Different Data Domains—the Role of Question Order, Question Wording, and Incentives," *International Journal of Social Research Methodology*, 27, 375–388.

Blank, G., and Lutz, C. (2017), "Representativeness of Social Media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+, and Instagram," *American Behavioral Scientist*, 61, 741–756.

Blundell, R., Dias, M. C., Cribb, J., Joyce, R., Waters, T., Wernham, T., and Xu, X. (2022), "Inequality and the COVID-19 Crisis in the United Kingdom," *Annual Review of Economics*, 14, 607–636.

Davis, J., Wolff, H.-G., Forret, M. L., and Sullivan, S. E. (2020), "Networking via LinkedIn: An Examination of Usage and Career Benefits," *Journal of Vocational Behavior*, 118, 103396.

Di Cara, N. H., Boyd, A., Tanner, A. R., Al Baghal, T., Calderwood, L., Sloan, L. S., Davis, O. S. P., and Haworth, C. M. A. (2020), "Views on Social Media and Its Linkage to Longitudinal Data from Two Generations of a UK Cohort Study," *Wellcome Open Research*, 5, 44.

Fernandez, S., Stöcklin, M., Terrier, L., and Kim, S. (2021), "Using Available Signals on LinkedIn for Personality Assessment," *Journal of Research in Personality*, 93, 104122.

Fisher, P., and Hussein, O. (2023), "Understanding Society: The Income Data," *Fiscal Studies*, 44, 377–397.

Guess, A., Munger, K., Nagler, J., and Tucker, J. (2019), "How Accurate Are Survey Responses on Social Media and Politics?," *Political Communication*, 36, 241–258.

Haenschen, K. (2020), "Self-Reported versus Digitally Recorded: Measuring Political Activity on Facebook," *Social Science Computer Review*, 38, 567–583.

Henderson, M., Jiang, K., Johnson, M., and Porter, L. (2021), "Measuring Twitter Use: Validating Survey-Based Measures," *Social Science Computer Review*, 39, 1121–1141.

Hsieh, Y. P., and Murphy, J. (2017), "Total Twitter Error: Decomposing Public Opinion Measurement on Twitter from a Total Survey Error Perspective." in *Total Survey Error in Practice* (1st ed), eds. P. P. Biemer, E. Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, N. C. Tucker, and B. T. West, Hoboken, NJ: Wiley, pp. 23–46.

Jäckle, A., Beninger, K., Burton, J., and Couper, M. (2021), "Understanding Data Linkage Consent in Longitudinal Surveys," in *Advances in Longitudinal Survey Methodology*, ed. P. Lynn, Hoboken NL: Wiley, pp. 122–150.

Karlsen, R., and Enjolras, B. (2016), "Styles of Social Media Campaigning and Influence in a Hybrid Political Communication System," *The International Journal of Press/Politics*, 21, 338–357.

Kreuter, F., Sakshaug, J. W., and Tourangeau, R. (2016), "The Framing of the Record Linkage Consent Question," *International Journal of Public Opinion Research*, 28, 142–152.

LinkedIn (2023), About LinkedIn, Available at https://about.linkedin.com/.

Little, R., and Su, H. (1989), "Item Non-Response in Panel Surveys." in *Panel Surveys*, eds. D. Kasprzyk, G. Duncan, G. Kalton, and M. Singh, New York: John Wiley; Sons, pp. 400–425.

Liu, S., Sloan, L., Al Baghal, T., Williams, M., Jessop, C., and Serôdio, P. (2024), "Linking Survey with Twitter Data: Examining Associations among Smartphone Usage, Privacy Concern and Twitter Linkage Consent," *International Journal of Social Research Methodology*, 1–15.

Mneimneh, Z. (2022), "Evaluation of Consent to Link Twitter Data to Survey Data," *Journal of the Royal Statistical Society*, 185, S364–S386.

Mneimneh, Z., McClain, C., Bruffaerts, R., and Altwaijri, Y. (2021), "Evaluating Survey Consent to Social Media Linkage in Three International Health Surveys," *Research in Social and Administrative Pharmacy*, 17, 1091–1100.

Ofcom (2019–2023), "Adults' media use and attitudes report." Available at https://www.ofcom.org.uk/research-and-data/media-literacy-research/adults/adults-media-use-and-attitudes.

Pew Research Center (2024), "Social Media Fact Sheet." Available at: https://www.pewresearch.org/internet/fact-sheet/social-media/.

Sakshaug, J. W., and Kreuter, F. (2014), "The Effect of Benefit Wording on Consent to Link Survey and Administrative Records in a Web Survey," *Public Opinion Quarterly*, 78, 166–176.

Sakshaug, J. W., Couper, M. P., Ofstedal, M. B., and Weir, D. R. (2012), "Linking Survey and Administrative Records: Mechanisms of Consent," *Sociological Methods & Research*, 41, 535–569.

Sakshaug, J. W., Schmucker, A., Kreuter, F., Couper, M. P., and Singer, E. (2019), "The Effect of Framing and Placement on Linkage Consent," *Public Opinion Quarterly*, 83, 289–308.

Sakshaug, J. W., Tutz, V., and Kreuter, F. (2013), "Placement, Wording and Interviewers: Identifying Correlates of Consent to Link Survey and Administrative Data," *Survey Research Methods*, 7, 133–144.

Seidenberg, A. B., Moser, R. P., and West, B. T. (2023), "Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA)," *Journal of Survey Statistics and Methodology*, 11, 743–757.

Silber, H., Breuer, J., Beuthner, C., Gummer, T., Keusch, F., Siegers, P., Stier, S., and Weiß, B. (2022), "Linking Surveys and Digital Trace Data: Insights From Two Studies on Determinants of Data Sharing Behaviour," *Journal of the Royal Statistical Society*, 185, S387–S407.

Sloan, L., Jessop, C., Al Baghal, T., and Williams, M. (2020), "Linking Survey and Twitter Data: Informed Consent, Disclosure, Security, and Archiving," *Journal of Empirical Research on Human Research Ethics*, 15, 63–76.

Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., and Dolmans, R. (2021), "Sharing Data Collected with Smartphone Sensors," *Public Opinion Quarterly*, 85, 423–462.

University of Essex, Institute for Social and Economic Research [dataset] (2021). *Understanding Society: Innovation Panel, Waves 1-13, 2008-2020* [data collection] (11th ed), UK Data Service. SN: 6849. Available at http://doi.org/10.5255/UKDA-SN-6849-14.