SHIFT
RESEARCH BRIEF | August 2025

Don’t Save the Demographics for Last:

No Breakoff Penalty for Placing Demographic Items Early in Web-Based Surveys

Introduction

The collection of demographic data on such core individual attributes as gender, race/ethnicity, and age is a crucial component of almost every social science survey, and social science surveys are one of the principal ways in which we collect information on such variables. Although the survey methods literature returns mixed results and instruction regarding the placement of demographic items (Dillman, Smyth, and Christian 2014), a more cautious strand of this work advocates for placing such questions at the end of survey instruments (Fink, Bourque, and Fielder 2003; Jackson 2016). Such scholars are concerned with respondents losing interest in the survey since they may regard demographic questions as uninteresting or irrelevant (Dillman 2007; Fink et al. 2003; Jackson 2016; Whitley and Kite 1996). They also warn that respondents may perceive demographic items as sensitive topics or too personal to share (Fink et al. 2003; Rea and Parker 2014; Sheatsley 2013).

It is important to revisit the issue of demographic item placement especially as it relates to the realities of contemporary survey data collection practice. There has been enormous growth in the use of non-probability sample web-based surveys over the past several decades (Couper 2000, 2017; Couper and Miller 2008), with data from such surveys employed in articles published in the leading journals in Sociology (Amengual and Bartley 2022; Bonikowski, Feinstein, and Bock 2021; Robbins, Dechter, and Kornrich 2022; Rosenfeld and Thomas 2012; Schneider and Harknett 2019), Political Science (Goren 2022; Huff and Kertzer 2018); and Economics (Hanspal, Weber, and Wohlfart 2021; Stantcheva 2023). But, such survey samples and methods pose a key challenge to the received wisdom on the placement of demographic items. Web-based surveys face a particularly acute version (Peytchev 2009, 2011) of the growing challenge of survey break-off and attrition (Massey and Tourangeau 2013b, 2013a). When demographic items are placed at the

Daniel Schneider
Harvard Kennedy School
dschneider@hks.harvard.edu

Kristen Harknett
UC, San Francsico
kristen.harknett@ucsf.edu

end of surveys that have high breakoff, the result can be high rates of item non-response on demographic variables. This can be a particular problem for non-probability web-based surveys that frequently post-stratify and weight their data on demographics (Amengual and Bartley 2022; Bonikowski et al. 2021; Goren 2022; Huff and Kertzer 2018; Robbins et al. 2022; Rosenfeld and Thomas 2012; Schneider and Harknett 2019; Thomas 2019; Wang et al. 2015; Witte, Amoroso, and Howard 2000). The result is a loss of statistical power in weighted analyses that are limited to respondents with complete cases on weighting variables as well as in estimates of gender, racial/ethnic, or age-based inequalities. However, it may be that the increase in item-non-response that comes from placing demographic items at the end of high-breakoff surveys is outweighed by the overall increase in survey breakoff that would be caused by placing demographic items early in the survey.

The magnitude and direction of this trade-off is an important empirical question for survey practice. However, despite the ambiguity surrounding the placement of demographic items, there is surprisingly little empirical research on this question, particularly as it translates to online surveys. Extant work finds little evidence that the inclusion of demographic items at all (Ziegenfuss et al. 2021) or their position in a survey affects survey breakoff (Giles and Feild 1978; Green, Murphy, and Snyder 2000), but this research focuses on mailed surveys and little of the evidence is recent.

We scrutinize this longstanding survey-design uncertainty in the context of web-based surveys. To do so, we leverage large-scale online surveys of hourly service-sector workers in the United States administered by the Shift Project to experimentally and observationally test the effect of demographic item placement on survey breakoff. We then estimate missingness for demographic items as well as the demographic composition of the survey sample across these different survey placements. Finally, we explore the potential tradeoff between item non-response on demographic questions placed late in the survey versus potentially higher overall breakoff when demographic items are placed early in the survey for statistical power through a data exercise estimating gender gaps across two sample outcomes.

We find no effect of demographic item placement on survey breakoff, with equal breakoff rates when demographic items are placed first, mid-way, and near the end of the survey instrument. We replicate this experimental finding in large-scale observational data. While there is no overall breakoff penalty to early placement of demographic items, there is a substantial decrease in missingness on demographic items when they are placed early versus late in the survey. We also show that while this does not affect the estimated gender composition of the survey sample, it does reduce the mean age of respondents and increases the share reporting being Hispanic. We then demonstrate how obtaining more complete demographic information generates more statistical power using an example of estimating gender gaps in wages as well as in having an on-call schedule. Altogether, these findings suggest certain demographic items may be more flexible in their placement than typically advised, particularly in online surveys.

Demographic Items Last: The Conventional Wisdom

Survey methodologists have long considered how question ordering influences respondents' engagement and accuracy, with many recommending that demographic items be placed late in a survey (Dillman 2007; Fink et al. 2003; Jackson 2016; Rea and Parker 2014; Sheatsley 2013; Whitley and Kite 1996). Dillman (2007), a leading voice in the survey design literature, draws from social exchange theory to recommend that researchers design questionnaires that encourage survey response by increasing perceived rewards from participation and fostering respondent trust. As such, demographic questions are often recommended to be placed later in a survey because of their reputation as uninteresting and sensitive.

Many survey-design experts warn that respondents often find demographic questions boring (Fink et al. 2003; Jackson 2016). To begin a survey with less-engaging material decreases the likelihood respondents finish it. Similarly, other researchers point out that personal demographics may not seem particularly relevant to the survey topic (Dillman 2007; Fink et al. 2003; Whitley and Kite 1996), so to place such questions at the beginning of the survey may discourage completion by weaking respondent's trust and engagement. It could also detract from introductory materials that orient respondents to the survey's subject matter.

Survey methodologists also point to the sensitive nature of demographic questions as another threat to survey completion and breakoff, cautioning that asking for personal information early in a survey may jeopardize respondents' trust (Fink et al. 2003; Rea and Parker 2014; Sheatsley 2013). Rea and Parker (2014) further justify placing demographic questions late in a survey by arguing that the questions building up to the demographic items can help build rapport and make these items seem less intrusive. They also contend that even if a respondent reacts negatively to being asked personal questions, placing such items at the end of a questionnaire ensures their previous responses are useable.

Empirical Evidence on Demographic Item Placement

A limited body of work has empirically tested how the placement of demographic questions shapes survey completion and breakoff, suggesting that the placement of these items is not as consequential as typically claimed. This is supported by experiments that randomly varied whether demographic questions were asked first or last in mailed paper questionnaires (Giles and Feild 1978; Green et al. 2000). Ziegenfuss et al. (2021) provide a more recent version of this test, finding no significant difference in response rates across the study arms of a mailed paper questionnaire where demographic items were (a) not asked, (b) placed at the end of the survey, or (c) asked on a separate sheet of paper. These studies also report that their various placements of demographic questions did not lead to biased data (e.g., Ziegenfuss et al. (2021) found consistent levels of disagreement between administrative and self-reported race/ethnicity data across their survey arms) or more item nonresponse (Giles and Feild 1978; Green et al. 2000; Ziegenfuss et al. 2021). In fact, Green et al. (2000) noted that demographics-at-the-end respondents had more missing demographic data than demographic-at-the-beginning respondents.

However, most studies examining how the placement of demographic questions impact survey data collection only did so using mailed paper surveys. Where demographic items are located might matter more in web-based formats where respondents can't see all the questions being asked at once. Moreover, the potential respondent pool for these studies were specialized and potentially had prior relationships with the survey administrators. For instance, Giles and Field (1978) mailed job satisfaction questionnaires to full-time faculty members of a large state university while Green et al. (2000) sent surveys about topics related to social work to registered members of the National Association of Social Workers. These specialized samples and use of only paper surveys limit these studies' generalizability.

Teclaw et al. (2012) begin to address some of these shortcomings by randomly varying whether demographic questions were asked first or last in an online survey offered to Veterans Affairs (VA) employees. Consistent with prior empirical work, their results reveal that placing demographic questions at the beginning of a survey boosted response rates for these items without altering the response rate for non-demographic items. They also showed that the mean responses for non-demographic items from demographics-first questionnaires were not significantly different than those from demographics-last ones. Although Teclaw et al. (2012) extend prior findings to web-based surveys, their potential respondent pool also consists of a specific occupational population–VA employees–using a work-related questionnaire. Do these results hold for surveys that social scientists rely on which sample broader populations who typically have no connection to survey administrators? Moreover, these studies have not explored how having access to more complete demographic data (because of placing these questions first) may lead to more statistical power.

Data

Shift Project Data Collection

We examine the issue of demographic item placement in online surveys using The Shift Project surveys. The Shift Project collects web-based surveys from service-sector workers, recruiting respondents through Facebook advertisements that target workers who are at least 18 years old, reside in the U.S., and are identified as employees at particular food service or retail employers. Advertisements appear on potential respondents' Facebook and/or Instagram feeds and offer a prize-based drawing incentive for completing a survey hosted on the Qualtrics platform. Detailed information on the data collection protocols is provided in Schneider and Harknett (2022) and the Shift Project data have been used in a number of published papers (Harknett, Schneider, and Irwin 2021; Schneider 2020; Schneider and Harknett 2019, 2021).

The Shift Project's approach has important similarities to increasingly commonly used alternative survey data sources in that it is fielded online and is a non-probability sample. However, it differs from data collected via such vendors as Dynata, Lucid, and Cint in that it does not recruit from an existing panel of respondents who frequently take surveys, but rather draws a fresh sample of respondents from the broad Facebook/Instagram user-base at each wave. This approach, treating Facebook/Instragram as a sampling frame with particularly rich individual-level targeting data, is increasingly commonly used in the social sciences (Zhang et al. 2020). We draw on two specific sources of data collected using these methods.

Experimental Sample

Following the survey design literature, demographic questions have always been placed near the end of Shift Project surveys. We directly tested how the placement of demographic questions shapes breakoff rates and data quality in an experiment that randomly assigned respondents to one of three survey conditions: (1) demographic questions asked at the beginning of the survey; (2) demographic questions asked in the middle of the survey; or (3) the original survey in which demographic questions were asked near the end.

However, we only varied the placement of survey items asking for respondents' race/ethnicity, age, and gender, leaving the rest of the demographic module in its original spot. We select these three questions not only because they are vital to the construction of survey weights that adjust for potential selection bias on unobservable characteristics (Schneider and Harknett 2022), but also because they are key dimensions along which inequalities in job quality exist (Bartel et al. 2019; Kristal et al. 2018; Levanon et al. 2009), a topic central to Shift Project research.

We fielded surveys in February 2023, recruiting a total of 1,475 employed respondents through three Facebook advertisements. We targeted employees from firms that had exhibited low, medium, and high rates of breakoff at the demographic module in the most recent wave of data collection prior to the experiment–Starbucks, Rite Aid, and Lowe's, respectively. The randomization was effective. 31.5% of respondents received the survey with race/ethnicity, age, and gender asked first (Arm 1), while 33.7% took the original survey with demographic items asked near the end (Arm 2) and 34.8% were asked their race/ethnicity, age, and gender in the middle of the survey (Arm 3).

Observational Sample

The results of the February 2023 survey experiment (reported below) immediately affected general survey practice at the Shift Project. We are then able to investigate how this experiment scales up by comparing the Fall 2022 wave of data collection, which placed demographics last in the survey following standard practice, with the Spring 2023 wave of the survey, that asked race/ethnicity, age, and gender first in the survey. This allows us to observationally test how demographic item placement influences breakoff rates by comparing Spring 2023 data (n = 31,581) to survey data collected in Fall 2022 (n = 21,495).

Methods

Survival Models of Breakoff

We estimate the effect of demographic item placement by comparing survey breakoff across experiment arms. To do so, we estimate cox proportional hazard models, where breakoff is defined as the event and progression across questions as time. In these models, we control for the survey "pathway" that respondents take, with the key difference being between union and non-unionized workers, who see slightly more questions about union attitudes and between non-parents and parents, who see additional questions on parenting. We also include fixed effects for the Facebook ads that link respondents to the survey since they target different workers at different service-sector firms. We present both survival curves and regression estimates.

We replicate these models using the observational data generated by comparing survey breakoff in the Fall 2022 wave of data collection, when the demographic items were placed at the end of the survey, and breakoff in the Spring 2023 wave of data collection, when the demographic items were placed first in the survey. This change was the direct result of the experiment fielded in February of 2023 and provides an illustration of how the findings from a more limited experimental test translate to broad based survey data collection. The observational data naturally do not provide as controlled a test as the experiment. Comparisons between Fall 2022 and Spring 2023 are meant to identify the effects of demographic placement but could be confounded by other survey changes. In particular, The Shift Project maintains a core of sampled firms across waves, but also alternates the inclusion of other firms, leading to imbalance in firm composition. Further, the images deployed in the recruitment advertisements vary in character from wave to wave. We estimate models that limit the sample, in turn, to only those respondents from the same set of firms, to only those respondents recruited using very similar images, and to only those respondents who satisfy both of these conditions. Just like in the experimental models, we include a control for survey pathway and Facebook ad fixed effects.

Demographic Non-Response

After assessing the effects of demographic item placement on survey breakoff, we then evaluate how demographic item placement impacts the procurement of respondents' gender identity, race/ethnicity, and age. Using both the experimental and observational data, we estimate linear probability models predicting the likelihood a respondent provides these key demographic measures.

Demographic Composition

We next estimate the effects of placement on the demographic composition of the sample in terms of gender identity, race/ethnicity, and age. We test for significant differences in race and gender composition between samples asked demographic questions first versus those asked them later using chi-square tests and for age differences using OLS regression, in both the experimental data and the observational data.

Improving Power

Finally, we draw on the observational data to illustrate the value of earlier demographic placement for preserving sample size in terms of statistical power, or the probability of correctly detecting statistical significance. Increased statistical power allows researchers to identify more fine-grained effect sizes, i.e., a study's minimum detectable effect size, and sample size is a fundamental component in obtaining adequate statistical power.

We perform a simple data exercise examining the variation in minimum detectable effect sizes across various hypothetical sample sizes generated from breakoff patterns observed in the observational data. We inspect this in analyses probing gender's association with wages and having an on-call schedule, specifically looking at differences between self-identifying men and women. We assume a standard significance level of 5% as well as 90% power and then input these generated sample sizes into Stata's "power oneslope" command along with summary statistics of these measures to obtain minimum detectable effect sizes for the different demographic item placements.

Results

Demographic Item Placement and Breakoff – Experimental Evidence

The February 2023 experiment provides the cleanest estimates of the effect of demographic item placement on survey breakoff. Figure 1a presents survival curves by treatment arm. The green line traces the survival/breakoff of respondents who were asked the demographic items at the end of the survey, the purple line does so for those asked demographic items towards the middle of the survey, and the blue line for those asked demographic items as the first question. For the survival curves for respondents asked demographics at the middle or end of the survey, the initial question they encountered asked about their employment status and subsequent questions asked about basic job characteristics. For those with demographics placed first, the three demographic questions preceded questions about employment.

Across the treatment arms, there is substantial breakoff from the survey as typical in non-probability, web-based surveys (Couper 2017; Massey and Tourangeau 2013b, 2013a; Peytchev 2009, 2011). While the total number of questions varies somewhat based on conditional display logic, nearly 50% of respondents who begin the survey have attrited by question number 18, out of a total of 126 questions.

In this high-breakoff context, moving demographic items to the beginning of the survey would be undesirable if it further increased breakoff. However, we find that is not the case. The curves show essentially no impact of demographic question placement on breakoff. In Model 1 of Table 1, we show the hazard of breakoff as a function of treatment arm. Neither of the coefficients are statistically significant–placement of the demographics at the beginning, middle, or end of the survey has no overall effect on breakoff.

Figure 1. Survey Breakoff by Demographic Item Placement Placement

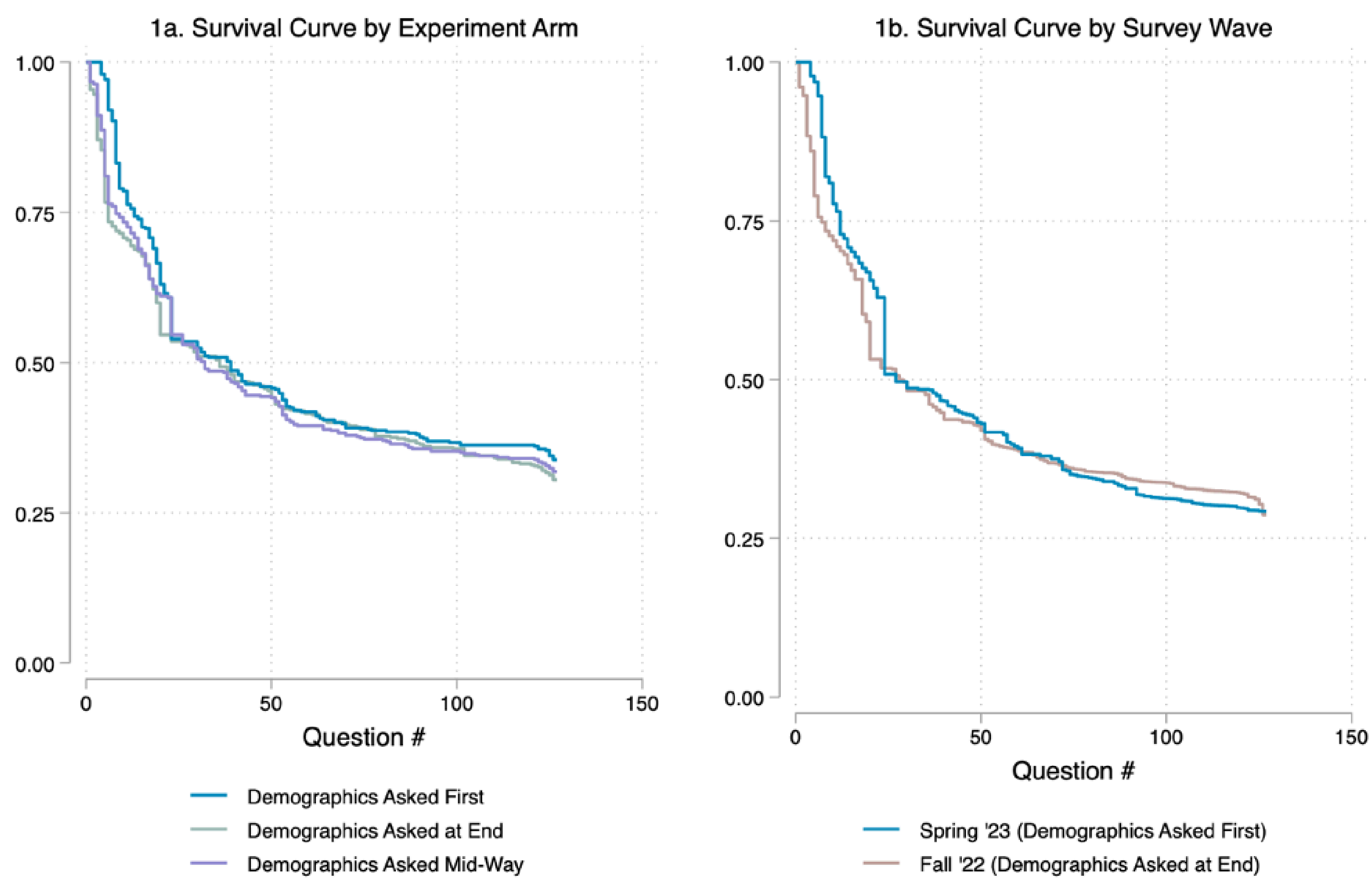

Table 1. Cox Proportional Hazard Model Estimates of Survey Breakoff by Demographic Item Placement

	(1)	(2)	(3)	(4)	(5)
	Experiment Results	Full Pooled Sample	Respondents From Firms Surveyed in Both Waves	Respondents From Ads Using Similar Images	Respondents From Ads Using Similar Images and From Firms Surveyed in Both Waves
Experiment Treatment (ref. Demographics Asked at End)					
Demographics Asked First	-0.128				
Demographics Asked Mid-Way	0.009				
Survey Wave (ref. Fall '22 – Demographics Asked at End)					
Spring '23 – Demographics Asked First		-0.376**	-0.123	-0.338***	0.401***
Observations	1475	53076	33356	42031	18449
Facebook Ad Fixed Effects	X	X	X	X	X
Survey Path Controls	X	X	X	X	X

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Demographic Item Placement and Breakoff – Observational Evidence

Figure 1b plots survival/breakoff curves comparing respondents in the Fall 2022 data, when demographics were asked at the end of the survey, and in Spring 2023, when demographics were asked at the beginning. Where the experiment is focused on respondents at 3 firms, these data include respondents from 101 large retail firms. The results, however, are very similar. There are high levels of breakoff from the survey, but no apparent differences in breakoff depending on the placement of demographic items.

Models 2-5 of Table 1 show the estimated hazard of breakoff from the Cox proportional hazard models. Model 2 presents the breakoff hazard estimated on the full sample of combined Fall 2022 (end of survey placement) and Spring 2023 (beginning of survey placement) data. These models control for both survey-recruitment-ad fixed effects and for the composition of survey paths in the sample. Model 2 reveals a statistically significant negative effect of placing demographics first on breakoff. In Models 3-5, we limit the sample to test the sensitivity of these results to aligning the Fall 2022 and Spring 2023 samples as closely as possible with respect to company composition and image similarity. In Model 3, we focus only on respondents in the sub-set of firms that were surveyed in both waves. Here, the coefficient is again negative, but not significant. In Model 4, we focus only on respondents who were recruited using similar recruitment images across both waves. Here, the coefficient is negative, similar in magnitude to Model 2, and statistically significant. Finally, in Model 5, we impose both sample restrictions and again estimate a significant and negative coefficient on placing demographic items first on breakoff.

Demographic Item Placement and Demographic Non-Response

While the position of demographic items in the survey has either null or small beneficial effects on breakoff, demographic placement has a huge effect on the provision of demographic data on age, race/ethnicity, and gender. Table 2 shows estimates of the share of respondents providing information on each of these crucial variables across the experimental arms (M1, M2, and M3) and from the most conservative observational sample, i.e., respondents from the same set of firms and recruited using very similar images (M4, M5, and M6).

Models 1-3 of Table 2 demonstrate that moving demographic items to the beginning of a survey results in more demographic data available for analysis. The coefficients for the treatment arms reveal a very large positive effect of placing demographics at the beginning or in the middle of the survey on securing non-missing responses for age, race/ethnicity, and gender. Intuitively, this effect is larger for placing demographics at the beginning, increasing the probability of securing respondent's age, race/ethnicity, and gender by a little more than 60 percentage points relative to asking demographics at the end of a survey. Models 4-6 show this pattern extends to the observational data–asking demographics first was associated with a nearly 60 percentage point increase in the probability of obtaining respondents' age, race/ethnicity, and gender relative to asking demographics last.

Demographic Item Placement and Demographic Composition

We also examine how the placement of demographic items influences the demographic composition of the sample. Table 3 illustrates differences in the gender

Table 2. Estimates of Demographic Data Non-Response by Demographic Item Placement Position

	Experiment Results			Respondents from Ads Using Similar Images & From Firms Surveyed in Both Waves		
	(1)	(2)	(3)	(4)	(5)	(6)
	Age Nonmissing	Race-Ethnicity Nonmissing	Gender Nonmissing	Age Nonmissing	Race-Ethnicity Nonmissing	Gender Nonmissing
Experiment Treatment (ref. Demographics Asked at End)						
Demographics Asked First	0.603***	0.610***	0.618***			
Demographics Asked Mid-Way	0.221***	0.226***	0.220***			
Survey Wave (ref. Fall '22 – Demographics Asked at End)						
Spring '23 – Demographics Asked First				0.581***	0.554***	0.570***
Observations	1475	1475	1475	18449	18449	18449
Facebook Ad Fixed Effects	X	X	X	X	X	X
Survey Path Controls	X	X	X	X	X	X

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

and race/ethnic composition of the sample as well as the average difference in age between respondents who were asked demographics first and those who were asked demographics last. We exclude the experimental treatment of asking demographics mid-way for simplicity and because all their differences were insignificant.

Although there are no significantly different compositions of gender-identity or racial-ethnic identity in the experimental data, Table 3 reveals that the Spring 2023 sample (demographics asked first) had a Hispanic population that was 2 percentage points greater (a 17% greater share Hispanic, with the difference significant at $p < .001$) than that of Fall 2022 (demographics asked last). The differences across the other racial-ethnic categories, while significant, were smaller. Finally, in both the experimental and observational data, the sample of respondents who were asked demographics first were, on average, younger, by 4 years, than those asked demographics last.

Demographic Item Placement and Statistical Power

The placement of demographic items also has substantial consequences for statistical power. Table 4 demonstrates how the sample size recovery resulting from asking demographics first provide more power to detect potential gender disparities in hourly wages and having an on-call schedule. For example, given the observed breakoff rates of these particular survey items in Fall 2022–when demographics were asked

Table 3. Differences in Demographic Composition Between Demographics Asked First and Demographics Asked at the End

	Experimental				Observational			
	Demog. First	Demog. at End	diff.		Demog. First	Demog. at End	diff.	
Gender (%)								
Male	27.5	30.3	-2.8		32.7	32.6	0.1	
Female	63.6	63.1	0.5		64.1	64.1	0.0	
Nonbinary	8.9	6.7	2.2		3.1	3.3	-0.2	
N	451	195			8757	3746		
Race/Ethnicity (%)								
White, Non-Hispanic	74.7	78.9	-4.2		73.2	74.4	-1.2	***
Black, Non-Hispanic	5.6	3.6	2.0		5.7	5.2	0.5	***
Hispanic	11.9	7.7	4.2		14.1	12.0	2.1	***
Other or 2+ Race	7.8	9.8	-2.0		7.1	8.3	-1.2	***
N	447	194			8686	3680		
Age (years)	32.8	37.1	-4.3	***	37.1	41.6	-4.5	***
N	448	198			8654	3704		

last–and Spring 2023–when demographics were asked first–if 500 individuals started both surveys, analyses investigating gender and wages using Fall 2022 data would only be able to detect an effect size of 0.215 standard deviations (or wage differences of around $1.30) while Spring 2023 analyses would be powered to detect an effect size of 0.145 standard deviations (or wage differences as small as $0,87). As seen in Table 4, Fall 2022 gender wage-gap analyses would be able to reach a similar level of statistical sensitivity as this if 1000 individuals started the survey, in which case they would be able to detect an effect size of 0.152 standard deviations.

In other words, a researcher studying gender wage-gaps using Fall 2022 survey data would want to increase their initial sample by more than double to achieve similar levels of statistical power as a researcher using Spring 2023 data. A similar pattern emerges for identifying a relationship between gender and having an on-call schedule. For example, analyses leveraging Spring 2023 data that had 500 initial survey respondents would be equipped to detect an effect size of 0.154 standard deviations while identical analyses using Fall 2022 data would not reach a similar level of power unless they had 1000 initial survey respondents, powering them to detect an effect size of 0.151 standard deviations. Table 4 shows that the loss of power is of less consequence when sample sizes increase by an order of magnitude. However, even for these larger samples, loss of power due to breakoff on demographics may be consequential for analyses that seek to estimate interaction effects or for subgroup analyses.

Table 4. Differences in Minimum Detectable Effect Sizes Across Differential Survey Breakoff Observed in Demographics Asked First vs. Demographics Asked at the End

	Minimum Detectable Effect Size of Gender Gaps (in SDs)			
	Wages (N)		**On-Call (0.89*N)**	
Total Number of Survey Respondents	Demographics Last (0.46*N)	Demographics First (N)	Demographics Last (0.52*N)	Demographics First (N)
500	0.215	0.145	0.214	0.154
1000	0.152	0.103	0.151	0.109
5000	0.0676	0.0459	0.0674	0.0486
10000	0.0478	0.0324	0.477	0.0344
20000	0.0338	0.0229	0.0337	0.0243
30000	0.0276	0.0187	0.02751	0.0198

Note: Gender gaps only reflect male-female differences for simplicity. Multipliers in parentheses reflect breakoff rates for the respective survey items/order.

Discussion

Demographic information is without question among the most-commonly collected data elements in surveys and among the most-often included information in descriptive and regression analyses. Demographic data is essential for characterizing a survey sample and is often of keen analytic interest for revealing age, race, and gendered disparities. Demographic information is commonly included as a regression covariate and as a key data element in post-stratification and weighting as well as in procedures to impute missing data. For all these reasons, optimizing the collection of demographic information is of great importance in survey research.

With online surveys playing an increasingly large role in survey data collection, our paper provides empirical insight into the placement of demographic variables on online surveys. Although more measured approaches to survey design recommended saving demographic questions for the end of the survey–because these questions were either too dull or too sensitive–in an experimental test, we find on the contrary that placing demographics first has benefits and does not have the feared downside of increasing survey breakoff.

Using survey data collected by The Shift Project, and an embedded question order experiment, we show that placing age, gender, and race/ethnicity variables at the start of the survey did not increase breakoff but rather slightly reduced breakoff before finishing the survey. By randomly assigning survey respondents to receive demographics questions as the first survey questions, in the middle of the survey, or at the end of the survey, we generated evidence from a rigorous comparison across these three experimental conditions. We were able to marshal further evidence from comparing our Spring 2023 survey wave, which asked demographics first, with our Fall 2022 survey wave, which asked demographics last. The results of this comparison were consistent with the experimental evidence, increasing our confidence in the finding that leading with demographics does not come at the cost of greater survey breakoff in online, non-probability surveys.

Further, placing demographic questions as the first questions in our survey yields some major benefits, stemming from vastly increasing the share of the survey sample for which demographic information was available. The gains in demographic information were substantial, increasing the share providing demographics among those who began the survey by 60 percentage points. Larger samples of demographic information provide a multitude of downstream benefits that enhance research. For example, we demonstrated how the larger sample sizes that result from asking demographics first provide greater power for analyses that seek to examine disparities in outcomes between self-identifying men and women that, in the case of Shift Project data, would necessitate researchers using surveys asking demographics last but striving for the statistical power generated by demographics-first surveys to approximately double their initial sample of survey respondents, which can quickly increase costs. Also, placing demographics first maximizes sample size and power when applying survey weights, given that demographic variables are frequently used when constructing post-stratification weights to adjust sample composition to be representative of the target population. Further benefits include allowing for an analysis of survey break off that captures differential breakoff rates by age, gender, and race/ethnicity, sharpening estimates of the demographic composition of the survey sample and improving the imputation of missing data because demographic information is available for almost the entire sample who began the survey.

Our results also showed that leading with demographics changed and sharpened the descriptive information on our sample composition in potentially important ways. In particular, maximizing responses to the demographics questions by asking them first revealed that the survey sample was younger and more Hispanic than it appeared when questions were asked late in the survey. One interpretation of this is that younger workers and Hispanic workers had higher survey break off rates than their counterparts. This information would be obscured if demographic information were collected at the end of the survey.

Although the findings in this paper are generated by a rigorous experimental research design and complemented with large-scale observation data, some features of our survey sample should be kept in mind when considering the generalizability of the results. Our survey sample recruited workers employed at large retail and food service companies in the United States. These workers tended to be in their 20s through 40s and were largely low-wage workers. It is possible that leading with demographic information for a different target population could yield different benefits and costs. Furthermore, our study only examined the placement of demographic items central to Shift Project research questions: race-ethnicity, gender, and age. Therefore, those seeking to use a web-based survey with a markedly different target population and/or demographic characteristics of interest may want to conduct their own survey experiment and can consider drawing on the analytic comparisons deployed here in their replication.

The rise of web-based surveys has provided rich, new opportunities for data collection that would not have been feasible using other modes of data collection. However, this increased reliance on web-based surveys necessitates revisiting conventional wisdom and putting it to the test. In this paper, we found that the more cautious guidance of placing demographic items later in the survey did not hold. The discovery is an important one in the context of high rates of survey break off in web-based surveys and the essential nature of demographic information for the research process. Given that demographic items are often high priority items, needed for multiple essential research purposes, it is very good news that placing them first and maximizing the sample size with non-missing information on these key items does not come at the expense of increased survey breakoff.

Authors

Dylan Nguyen *is a Sociology PhD Candidate at Harvard University*
dnguyen@hks.harvard.edu

Daniel Schneider *is the Malcolm Wiener Professor of Social Policy and Professor of Sociology at Harvard University*
dschneider@hks.harvard.edu

Kristen Harknett *is Professor of Social Behavioral Sciences in the School of Nursing at the University of California-San Francisco*
kristen.harknett@ucsf.edu

Connor Williams *is the data analyst for The Shift Project*
cwilliams@hks.harvard.edu

Acknowledgements

Research reported in this publication was supported by National Institutes of Aging (R01AG066898). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Aging. We also gratefully acknowledge support from the W.T. Grant Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of these foundations.

References

Amengual, Matthew, and Tim Bartley. 2022. “Global Markets, Corporate Assurances, and the Legitimacy of State Intervention: Perceptions of Distant Labor and Environmental Problems.” American Sociological Review 87(3):383–414. doi: 10.1177/00031224221092340.

Bartel, Ann P., Soohyun Kim, Jaehyun Nam, Maya Rossin-Slater, Christopher Ruhm, and Jane Waldfogel. 2019. “Racial and Ethnic Disparities in Paid Family and Medical Leave Access and Use: Evidence from Four Nationally Representative Datasets.” Monthly Labor Review.

Bonikowski, Bart, Yuval Feinstein, and Sean Bock. 2021. “The Partisan Sorting of ‘America’: How Nationalist Cleavages Shaped the 2016 U.S. Presidential Election.” American Journal of Sociology 127(2):492–561. doi: 10.1086/717103.

Couper, M. P., and P. V. Miller. 2008. “Web Survey Methods: Introduction.” Public Opinion Quarterly 72(5):831–35. doi: 10.1093/poq/nfn066.

Couper, Mick P. 2000. “Web Surveys: A Review of Issues and Approaches.” The Public Opinion Quarterly 64(4):464–94.

Couper, Mick P. 2017. “New Developments in Survey Data Collection.” Annual Review of Sociology 43(1):121–45. doi: 10.1146/annurev-soc-060116-053613.

Dillman, Don A. 2007. Mail and Internet Surveys: The Tailored Design Method. 2nd ed. Hoboken, NJ: John Wiley & Sons Inc.

Fink, Arlene, Linda B. Bourque, and Eve P. Fielder. 2003. The Survey Kit: How to Conduct Self-Administered and Mail Surveys. Vol. 3. 2nd ed. Thousand Oaks, California: Sage Publications, Inc.

Giles, William F., and Hubert S. Feild. 1978. “Effects of Amount, Format, and Location of Demographic Information on Questionnaire Return Rate and Response Bias of Sensitive and Nonsensitive Items.” Personnel Psychology 31(3):549–59. doi: 10.1111/j.1744-6570.1978.tb00462.x.

Goren, Paul. 2022. “Pliable Prejudice: The Case of Welfare.” American Journal of Political Science 66(4):961–76. doi: 10.1111/ajps.12652.

Green, Robert G., Katrina D. Murphy, and Shelita M. Snyder. 2000. “Should Demographics Be Placed at the End or at the Beginning of Mailed Questionnaires? An Empirical Answer to a Persistent Methodological Question.” Social Work Research 24(4).

Hanspal, Tobin, Annika Weber, and Johannes Wohlfart. 2021. “Exposure to the COVID-19 Stock Market Crash and Its Effect on Household Expectations.” The Review of Economics and Statistics 103(5):994–1010. doi: 10.1162/rest_a_01011.

Harknett, Kristen, Daniel Schneider, and Véronique Irwin. 2021. “Improving Health and Economic Security by Reducing Work Schedule Uncertainty.” Proceedings of the National Academy of Sciences 118(42). doi: 10.1073/pnas.2107828118.

Huff, Connor, and Joshua D. Kertzer. 2018. “How the Public Defines Terrorism.” American Journal of Political Science 62(1):55–71. doi: 10.1111/ajps.12329.

Jackson, Sherri L. 2016. Research Methods and Statistics: A Critical Thinking Approach. 5th ed. Boston, MA: Cengage Learning.

Kristal, Tali, Yinon Cohen, and Edo Navot. 2018. “Benefit Inequality among American Workers by Gender, Race, and Ethnicity, 1982–2015.” Sociological Science 5:461–88. doi: 10.15195/v5.a20.

Levanon, Asaf, Paula England, and Paul Allison. 2009. “Occupational Feminization and Pay: Assessing Causal Dynamics Using 1950-2000 U.S. Census Data.” Social Forces 88(2):865–91. doi: 10.1353/sof.0.0264.

Massey, Douglas S., and Roger Tourangeau. 2013a. “Introduction: New Challenges to Social Measurement.” The ANNALS of the American Academy of Political and Social Science 645(1):6–22.

Massey, Douglas S., and Roger Tourangeau. 2013b. “Where Do We Go from Here? Nonresponse and Social Measurement.” The ANNALS of the American Academy of Political and Social Science 645(1):222–36. doi: 10.1177/0002716212464191.

Peytchev, Andy. 2009. “Survey Breakoff.” Public Opinion Quarterly 73(1):74–97. doi: 10.1093/poq/nfp014.

Peytchev, Andy. 2011. “Breakoff and Unit Nonresponse Across Web Surveys.” Journal of Official Statistics 27(1):33–47.

Rea, Louis M., and Richard A. Parker. 2014. Designing and Conducting Survey Research. 4th ed. San Francisco, CA: Jossey-Bass.

Robbins, Blaine G., Aimée Dechter, and Sabino Kornrich. 2022. “Assessing the Deinstitutionalization of Marriage Thesis: An Experimental Test.” American Sociological Review.

Rosenfeld, Michael J., and Reuben J. Thomas. 2012. “Searching for a Mate: The Rise of the Internet as a Social Intermediary.” American Sociological Review 77(4):523–47. doi: 10.1177/0003122412448050.

Schneider, Daniel. 2020. “Paid Sick Leave in Washington State: Evidence on Employee Outcomes, 2016–2018.” American Journal of Public Health 110(4):499–504. doi: 10.2105/AJPH.2019.305481.

Schneider, Daniel, and Kristen Harknett. 2019. “Consequences of Routine Work-Schedule Instability for Worker Health and Well-Being.” American Sociological Review 84(1):82–114. doi: 10.1177/0003122418823184.

Schneider, Daniel, and Kristen Harknett. 2021. “Hard Times: Routine Schedule Unpredictability and Material Hardship among Service Sector Workers.” Social Forces 99(4):1682–1709. doi: 10.1093/sf/soaa079.

Schneider, Daniel, and Kristen Harknett. 2022. “What’s to Like? Facebook as a Tool for Survey Data Collection.” Sociological Methods & Research 51(1):108–40. doi: 10.1177/0049124119882477.

Sheatsley, Paul B. 2013. “Questionnaire Construction and Item Writing.” in Handbook of Survey Research, edited by P. H. Rossi, J. D. Wright, and A. B. Anderson. Academic Press.

Stantcheva, Stefanie. 2023. “How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible.” Annual Review of Economics 15(1):205–34.

Thomas, Reuben J. 2019. “Online Exogamy Reconsidered: Estimating the Internet’s Effects on Racial, Educational, Religious, Political and Age Assortative Mating.” Social Forces 1257–86. doi: 10.1093/sf/soz060.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2015. “Forecasting Elections with Non-Representative Polls.” International Journal of Forecasting 31(3):980–91. doi: 10.1016/j.ijforecast.2014.06.001.

Whitley, Bernard E., and Mary E. Kite. 1996. Principles of Research in Behavioral Science. Mayfield Publishing Company.

Witte, James C., Lisa M. Amoroso, and Philip E. N. Howard. 2000. “Research Methodology: Method and Representation in Internet-Based Survey Tools– Mobility, Community, and Cultural Identity in Survey2000.” Social Science Computer Review 18(2):179–95.

Zhang, Baobao, Matto Mildenberger, Peter D. Howe, Jennifer Marlon, Seth A. Rosenthal, and Anthony Leiserowitz. 2020. “Quota Sampling Using Facebook Advertisements.” Political Science Research and Methods 8(3):558–64. doi: 10.1017/psrm.2018.49.

Ziegenfuss, Jeanette Y., Casey A. Easterday, Jennifer M. Dinh, Meghan M. JaKa, Thomas E. Kottke, and Marna Canterbury. 2021. “Impact of Demographic Survey Questions on Response Rate and Measurement: A Randomized Experiment.” Survey Practice 14(1):1–11. doi: 10.29115/SP-2021-0010.