**What's Not to Like?**
**Facebook as a Tool for Survey Data Collection***

Daniel Schneider
University of California, Berkeley

Kristen Harknett
University of California, San Francisco

* Daniel Schneider (Corresponding author): UC Berkeley, Department of Sociology, 480 Barrows Hall, Berkeley, CA 94720; djschneider@berkeley.edu. We gratefully acknowledge grant support from the National Institutes of Child Health and Human Development (R21HD091578), the Robert Wood Johnson Foundation (Award No. 74528) the U.S. Department of Labor (Award No. EO-30277-17-60-5-6), the Washington Center for Equitable Growth (Award No. 39092), the Hellman Family Fund, the Institute for Research on Labor and Employment, and the Berkeley Population Center. We received excellent research assistance from Carmen Brick, Paul Chung, Nick Garcia, Alison Gemmill, Tom Haseloff, Veronique Irwin, Sigrid Luhr, Robert Pickett, Adam Storer, Garrett Strain, and Ugur Yildirim. We are grateful to Liz Ben-Ishai, Annette Bernhardt, Michael Corey, Rachel Deutsch, Dennis Feehan, Carrie Gleason, Anna Haley-Lock, Heather Haveman, Heather Hill, David Harding, Julie Henly, Ken Jacobs, Susan Lambert, Sam Lucas, Andrew Penner, Adam Reich, Jennie Romich, Jesse Rothstein, Hana Shepherd, Stewart Tansley, Jane Waldfogel, and Joan Williams for very useful feedback. We also received helpful feedback from seminar participants at UC Berkeley Sociology, the Institute for Research on Labor and Employment, UC Berkeley MESS, The Washington Center for Equitable Growth, the Institute for the Study of Societal Issues, MIT IWER, UCSF, and UC Davis. This work was approved by the UC Berkeley Committee for the Protection of Human Subjects (2015-10-8014).

## ABSTRACT

In this paper, we demonstrate how Facebook targeted advertisements can be used to collect survey data. We illustrate the potential of survey sampling and recruitment on Facebook through the example of building a large employee-employer linked dataset as part of The Shift Project. We describe the workflow process of targeting, creating, and purchasing survey recruitment advertisements on Facebook. We address concerns about sample selectivity, and apply post-stratification weighting techniques to adjust for differences between our sample and that of "gold-standard" data sources. We then compare univariate and multi-variate relationships in the Shift data against the Current Population Survey and the National Longitudinal Survey of Youth-1997. Finally, we provide an example of the utility of the firm-level nature of the data by showing how firm-level gender composition is related to wages. We conclude by highlighting some unique strengths of the Facebook targeting advertisement approach, including the ability for rapid data collection in response to research opportunities, rich and flexible sample targeting capabilities, and low cost, and we suggest broader applications of this technique.

## INTRODUCTION

The virtues of probability sampling – in which samples are selected at random and sample members have a known probability of selection - are many and well appreciated. Foremost among these benefits is the ability to generalize from samples and draw valid inferences about populations. Importantly, realizing this benefit requires a sampling frame that accurately captures the target population and non-differential response to the survey invitation. For some hidden and hard-to-reach populations, sampling frames do not exist and probability sampling has never been an option, creating some impetus for developing tools for drawing inferences from non-probability sampling methods.

Further, increasingly, even for populations for which probability sampling has historically predominated, serious obstacles have arisen that add urgency to the task of finding alternative sampling and data collection techniques. For instance, the decline in the coverage of landlines has undermined the primary sampling frame for telephone surveys, and telemarketer fatigue and technology that facilitates call screening and call blocking have dramatically reduced survey response rates. As a result, response rates to non-governmental surveys have plummeted. For instance, at one of the nation's leading polling organizations, Pew Research, the response rate has dropped from 36% in 1997 to 9% in 2016 (Keeter et al., 2017). The natural concern is that the 1 in 10 individuals who do still respond to surveys could be substantially different from those who do not – that is, that probability samples cannot be seriously considered random samples of the population or at the least that total survey error is increasingly high.

In response to this urgent need for new data collection strategies, there have been important advancements in developing methods of non-probability sampling for addressing bias and yielding valid inferences (Zagheni and Weber, 2015). Recent research has shown that, using post-stratification weighting techniques to weight to gold-standard sources such as the Census on demographics, even surveys with very low response rates exhibit little evidence of bias on univariate statistics or bivariate relationships (Kohut, et al., 2012), with the main exception being measures of civic engagement (Keeter et al., 2017). This insight has led a new generation of survey researchers and statisticians to suggest that it is then worth revisiting the value of non-probability sample surveys (Goel et al., 2015; Wang et al., 2015).

Over the past several years, scholars have begun to take up this call, and have creatively harnessed data from such sources as Twitter, email, and Google searches to study migration, fertility, and other demographic processes (Billari, D'Amuri, and Marcucci. 2013; Reis and Brownstein, 2010; Zagheni and Weber, 2012; Zagheni et al., 2014). Researchers have also conducted online surveys using a range of online non-probability samples (Couper, 2017). These approaches appeal in part because they can be quickly implemented at generally very low cost (Goel et al., 2015; Nunan and Knox, 2011; Stern et al., 2014).

But, the research community has been divided on the scientific value of research using surveys collected from online non-probability samples. The weight of early research and discussion suggested that problems of under-coverage from limited and selective internet usage made this approach of limited use (Best et al, 2001; Bethlehem, 2010; Yeager et al., 2011). Research continues to point to problems both with point-estimates and relationships between variables in non-probability online opt-in panel surveys (Dutwin and Buskirk, 2017; Bruggen et al., 2016; Casler et al., 2013).

However, recent research has found more encouraging results for online non-probability samples recruited through websites and advertising such as through Mechanical Turk, the X-Box gaming console, and Google ad words. This work suggests that such internet-based samples can fairly closely resemble probability samples in terms of demographics (Stern et al., 2016) and, further, perform well when weighted, in terms of yielding results in line with benchmark samples that use more conventional probability sampling approaches (Goel et al., 2015; Clifford et al., 2015; Wang et al., 2015; Mullinix et al., 2015).

Yet, of all of non-probability web-based recruitment platforms employed to date, Facebook has the largest user-base, has broad global coverage, exhibits less selection than for opt-in panels, and validates respondents' identities. Some prior work has used snowball sampling on Facebook through affinity groups to collect surveys (Bhutta, 2012; Baltar and Brunet, 2012) and this work generally finds that associations from the resulting data resemble those estimated from standard data sets such as the General Social Survey (Bhutta, 2012). Other recent work in marketing (Nunan and Knox, 2011), medical research (Ramo and Prochaska, 2012; Thornton et al., 2016), and political science (Zhang et al., 2017; Samuels and Zucco, 2013) has begun to explore the use of Facebook advertisements to recruit respondents to surveys. These studies have generally attempted to use Facebook to develop samples meant to approximate the general population. Recently, demographers have demonstrated that the Facebook advertising platform can be used as a "digital census," and employed to estimate migrant populations by country and U.S. state (Zagheni, Weber, and Gummadi, 2017).

Building on insights from this recent research, we suggest that a unique benefit of sample construction on Facebook is the ability to use the detailed audience targeting capabilities that are at the heart of Facebook's advertising model to construct samples of otherwise difficult-to-sample populations (a point also alluded to in AAPOR, 2014 and Zagheni, Weber, and Gummadi, 2017). We suggest that one such population of particular academic and policy interest is the employees of specific named firms. Such employer-employee linked data would be valuable to economic sociologists, who are interested in understanding how firm-level characteristics such as ownership structure and unionization affect labor practices (Fligstein, 2001; Applebaum and Batt, 2014; Weil, 2009), to policy scholars, who are interested in assessing the impact of local and state labor laws that focus on specific large employers (Colla et al., 2014), and to economists who are concerned with measuring intra-industry variation in compensation (Lane et al., 2007; Andersson et al., 2005; Groshen, 1991a; Groshen, 1991b; Krueger and Summers, 1988).

However, this sort of employer-employee matched data has proven elusive to social scientists. Data sets that are commonly used to describe employees' job conditions such as the NLSY, PSID, or CPS do not allow a link to identifiable employers. Studies, such as the National Organizations Survey, that contain detailed data on firm practices do not contain data from multiple employees at a given firm. Restricted access employer-employee linked data such as the LEHD or the BLS's OES are limited by not publically identifying employers and by having a fairly circumscribed set of measures. An important constraint on this work has been the absence of a sampling frame of workers at a large set of specific companies and the enormous cost that would come with attempting to assemble such a sample from a general population survey.

We illustrate the potential of survey sampling and recruitment on Facebook through the example of building just this sort of employee-employer linked dataset for The Shift Project. We discuss the work-flow of using the Facebook advertising platform, describe the results of our data collection efforts, discuss useful strategies for post-stratification and weighting, and then compare key associations from our data with a range of survey data gathered using probability sampling methods. We then take up the important question of selection into the survey on unobservable attributes that cannot be easily accounted for with weights and propose an easily implemented test to gauge the significance of this problem. Finally, we exploit the firm-level structure of the data to estimate how firm-level gender composition is associated with wages. Our results show that this data collection approach yields data that are broadly consistent with gold standard probability samples at the national level, and open up rich opportunities for granular targeting of a variety of hard-to-reach populations.

## TARGETED ADVERTISING ON FACEBOOK

Using Facebook to collect survey data departs from traditional probability sampling and some have raised reasonable questions about such approaches (Groves, 2011; Smith, 2013). One potential concern arises from the sampling frame of Facebook users. In the recent past, both internet access and Facebook use has been confined to relatively narrow subgroups of the population, which tended to have relatively high socioeconomic status. However, internet access is now widespread in the United States among working aged adults. Recent estimates from the American Community Survey find that between 90-94% of working aged adults have a computer at home and 80-84% have broadband internet access at home (Ryan and Lewis, 2017). Among those who use the internet, the very large majority are active on Facebook – 79% overall and 86% of those 18-49 (Greenwood et al., 2016). The result is that 81% of Americans age 18-49 are now active on Facebook, far in excess of the percent of this population with landlines. Further, although people of color and low-income strata are less likely to have home computers and broadband access (Ryan and Lewis, 2017), Facebook use is nevertheless not especially stratified by demographic characteristics (Greenwood et al., 2016). In addition, unlike some online platforms, Facebook goes to some length to verify that each user account is associated with a unique identifiable person (Facebook, 2017).

Facebook has two other important advantages over both phone and address-based sampling. First, unlike phone and address based sampling, the Facebook profile is a portable and durable means of contact. Respondents can be reached by Facebook for survey recruitment whether at home or work, whether they have moved or have a long residential tenure, whether they change phone numbers or lose service. This represents a distinct advantage over conventional sampling frames.

Second, Facebook collects detailed data on the attributes of users that can be used by advertisers to target their campaigns quite precisely. Indeed, this capability is at the heart of Facebook's business model. These attributes include standard demographics such as age and gender, locational attributes, interests, as well as information on schooling and employment. This last field permits us to deliver advertisements that are targeted to users who work at specific firms. Given the goal of assembling a data set that includes large samples of workers at each of a large number of firms, this targeting capability is very valuable.

To illustrate, consider the effort that would be associated with assembling a sample of this type using traditional methods. Given that a large number of employers are unlikely to be persuaded to turn over lists of employees with contact information, one would need to begin with a nationally representative sampling frame (such as a purchased phone or address list) that would not contain any information on employer, screen on those in the labor force, then those currently employed, then those in the sector of interest, and then those at particular large companies. To take just one example, Walmart is far and away the largest private sector employer in the country with 1.4 million employees. However, that equates to just 0.044% of the 319,000,000 US adults. Given response rates of approximately 9% for non-governmental surveys (Keeter et al., 2017), that would entail attempting to contact approximately 505,00 adults by phone or mail to achieve a sample of 200 Walmart workers ((200/.09)/.0044). Given a survey, such as ours, that aimed to collect data from 200 workers at each of 40 large companies with collective employment of 6.9 million, one would need to contact approximately 4.1 million adults.

## DATA COLLECTION

Acting as an "advertiser," we use Facebook's audience targeting tools to purchase and place survey recruitment advertisements in the newsfeeds of Facebook users who work at specific companies. Each advertisement was targeted to employees of a specific company (or family of consumer-facing brands), in the 18-50 age range, who were located in the United States.

Facebook provides several options for the "marketing objective" of the campaign. Our default approach, selected after consultation with advertising specialists at Facebook, is to set the campaign aim as "traffic" and to optimize delivery on "link clicks." Advertisements appearing on Facebook must follow a fairly standardized design, but there are options within that framework. For instance, while every advertisement must link to a Facebook page, include a headline and advertisement text, an image, and may include a link to an external webpage, advertisers have substantial discretion in crafting the advertising text, in choosing the content of the image, and in using a single image as opposed to a carousel, a video, a slideshow, or a collection.

We used a simple template for all of our advertisements. Every advertisement included a single image drawn from licensed stock photography available at no charge on the Facebook advertising page. We selected images that seemed to most closely approximate an employee of the target company at work, matching on store environment and color and style of employees' uniforms. Every advertisement linked to an "[Author's University] Work & Family Study" Facebook page that itself included very little additional content. For the data reported on in our main analysis, every advertisement used the "headline" field to offer users the opportunity to enter a drawing for an Apple iPad. Finally, again for the data in our main analysis, every advertisement used the advertisement text field to include a standard recruitment message. This message took the form of "Working at <targeted employer>? Take a short survey and tell us about your job!" In Figure 1 we include sample advertisements that we have used to recruit workers to the survey.

Finally, Facebook offers various options for advertisement placement. Advertisers may opt to have their advertisements appear on Facebook (in the newsfeed and/or in the right-hand column on desktop), on Instagram, or on partner networks. All of our campaigns were placed on Facebook in the newsfeed and on Instagram. Users who click on the advertisement are routed to an electronic survey hosted by Qualtrics. The survey can be accessed on desktop or mobile devices. Users are asked to consent to participation and then begin the survey. In essence, Facebook serves as both the sampling frame and the recruitment channel.

## SURVEY DATA

Our survey includes five core modules. The first collects information on respondents' jobs including on job tenure, hourly wage, hours, benefits, and work scheduling practices. The second module collects information on respondent's household economic security including household income, public benefits use, and use of alternative financial services. The third module contains data on respondents' demographics. The fourth module assessed respondents' health and wellbeing including self-rated health, sleep quality, and depressive symptoms. The final module was asked of parents and included information on child wellbeing, parenting time, and childcare. The individual survey questions were drawn from existing large-scale surveys including the Fragile Families and Child Wellbeing Survey, the NLSY97, and the NHIS.

We fielded recruitment advertisements to Facebook users employed at 38 large retail firms, drawn from among the 100 largest retail firms by revenue in 2015 (National Retail Federation, 2015). We fielded these advertisements between September of 2016 and June of 2017. In total, our advertisements were shown to 3,270,228 Facebook users, including some who were shown one of our advertisements on more than one occasion. These advertisements generated 179,563 link clicks through to the introductory page of our survey at a total advertising and prize cost of $75,000. Then, 39,918 respondents contributed at least some survey data. In all, 5.3% of our advertisement views led to a clicked through to begin the survey and 22% of those individuals contributed some survey data (or 1.2% of all advertisement views), for an average cost of $1.88 per respondent.

Of the 39,918 respondents who contribute some survey data, we eliminate 6,468 respondents who report that they were not paid hourly. In addition, the survey included a data quality check that instructed respondents to select a specific option on a question. 96% of respondents who were presented with this item complied. However, this item was not asked of respondents who attrited early in the survey. The result is a sample of 32,142 respondents.

However, there was substantial attrition. Of the 32,142 respondents who began, 17,828 fully completed the survey, and among those, there was item non-response. We perform multiple imputation to account for this missing data. First, we impute data only for those respondents who completed the survey, but had item non-response. Second, we impute data for all respondents who completed the first survey module, including those who finished the survey with some item non-response and those respondents who attrited from the survey at various points.

Our final analysis sample for a single implicate using the first approach is 17,828 responses and for the second imputation approach is 29,722 responses, both distributed across 38 companies. Based on the first sample size, the average price per survey response was $4.21 and based on the second it was $2.52  With complete or imputed data for each respondent on 125 items, we estimate a per item cost of $0.034 and $0.02 – very similar to the cost that Goel et al. (2015) report for their survey using Amazon Turk – and at least 20 times cheaper than traditional RDD polling (Goel et al., 2015) on a per question basis and far more inexpensive even than that given the focus on employees of these 38 companies.  All of the analyses we describe below produced substantially similar estimates when using the imputations on the sample of 29,722 responses versus 17,828 responses.  For the sake of parsimony, we present only the analysis on those who completed the survey, with multiple imputation for item-non response (n=17,828).

**POST-STRATIFICATION AND WEIGHTING**

A concern with using a non-probability based sample, such as this one, is that respondents may differ from the target population. Sample over-representation on particular demographic attributes can be addressed using post-stratification and weighting of the survey data to a "gold standard" benchmark (Zagheni and Weber, 2015).

A key contribution of our application is to construct a survey sample that contains relatively large numbers of employees at each of several dozen employers.  This is valuable precisely because such data are not readily available from existing survey or administrative sources.  The consequence is that it is actually somewhat difficult to derive a good estimate of the demographic characteristics of our target population to use as a benchmark. Our solution is compare the demographics of our survey respondents against several candidate benchmark populations, none of which exactly capture our target population.

First, we pool data from the 2013-2015 American Community Surveys. We condition the ACS sample on respondents being age 18-55 and employed in industries in the retail sector (581, 591, 600, 601, 623, 633, 641, 642, 691) that are represented by the 38 companies. We exclude any of these respondents who report upper level managerial occupations. In total we have data on 482,608 ACS respondents who meet these inclusion criteria.

Second, we pool data from the 2010-2017 rounds of the Current Population Survey (CPS), focusing on the March Annual Social and Economic Supplement (ASEC). The ASEC is valuable because while the sample size is smaller than the ACS, the ASEC includes a measure of firm size that captures whether the respondent works at a firm with greater than 1,000 employers. While all of the firms in our data have substantially more than 1,000 employees, conditioning on this variable at least allows us to exclude the many retail workers who are employed at small non-chain firms from our analysis. Here too, we further condition the sample to those aged 18-55 who work in the relevant industries and occupations. In total, we have data on 32,221 CPS-ASEC respondents who meet these inclusion criteria.

Third, we extract data from the Facebook advertising platform on the demographics of users who work at each of the companies in our data. While the survey data provides us with the demographics of those who took the survey, we can get demographic information on the characteristics of all potential respondents from the Facebook sampling frame by drawing on the advertising platform. Further, while in ACS we can only generate a benchmark population of those in the comparable industry, and in the CPS-ASEC only of those in the comparable industry and at large firms, with the Facebook data, we can benchmark to the demographics of those at the very same company. The tradeoff is that we benchmark to those who are on Facebook, rather than to the broader population of all workers employed at those companies. Additionally, the demographic information available from the Facebook advertising platform is limited to respondents' age and gender.

We categorize our benchmark samples in terms of the matrix of demographic characteristics. We also create a variant that further stratifies by industry. For our ACS and CPS-ASEC benchmark samples, we stratify respondents into cells defined by age x race/ethnicity x gender x industry group. For these benchmarks, we categorize age into three bins (18-29, 30-39, or 40-55), race/ethnicity into four mutually exclusive bins (white, non-Hispanic; Black, non-Hispanic, Other or two-or-more races, non-Hispanic, or Hispanic), gender into two categories (male or female), and industry into nine groups (hardware, department stores, general merchandise, grocery, fast food, apparel, electronics, drug store, or other retail). For our Facebook benchmark, we construct a matrix of age x gender x 38 employer cells.

We then construct weights for each cell that are the ratio of the proportion of the benchmark sample in each cell to the proportion of our sample in that same cell. The intuition behind these weights is that when a particular subgroup is relatively larger as a proportion of the benchmark sample than it is in our sample, then this group will be up-weighted with a weight value that is greater than 1. Conversely, when a subgroup is relatively smaller as a proportion of the benchmark sample than it is in our sample, then this group will be down-weighted with a weight value less than 1.

Finally, we further account for variation in the number of employees who work at each of the firms in our data by adjusting the individual responses by company labor force size to correct for any over or under-representation of employees at particular companies in our survey data relative to the actual relative labor force of a given company (e.g. the share of respondents in our survey data who work at Walmart might either be too large or too small a percent of all respondents as compared to Walmart's share of total employment at the 38 companies in our data). To make this correction, we use detailed data on establishment-level employment from the Reference USA U.S. Businesses database, collapsing thousands of store-level records to generate total in-store employment at each of the 38 companies.

The result is a set of eight weights – (1) ACS by demographics, (2) ACS by demographics/industry, (3) ACS by demographics/industry with employer size correction, (4) CPS by demographics, (5) CPS by demographics/industry, (6) CPS by demographics/industry with employer size correction, (7) Facebook by demographics/employer, and (8) Facebook by demographics/employer with employer size correction.

Table 1 compares the unweighted demographics of our survey respondents (column 1) against each of these benchmarks – the ACS (column 2), the CPS-ASEC (column 3) and Facebook users (column 4). The table shows that the unweighted Shift sample is disproportionately female, young, and White, non-Hispanic compared with the broader population in the ACS and CPS samples. While we do not have information on race/ethnicity for Facebook users, for gender and age, our sample is more similar to the Facebook user population employed at these firms, though by no means identical.

The next set of columns tabulate the Shift data by gender, age, and race/ethnicity after applying the basic weights to the ACS, the CPS-ASEC, and to Facebook. We see that the weighting procedure clearly brings the Shift sample into alignment with these benchmarks in terms of gender, age, and race/ethnicity.

We also compare educational attainment and school enrollment in the unweighted Shift data against the ACS and CPS and then against the weighted Shift data. Here, we again see some discrepancies in educational attainment. However, most of the difference appears to be from those who have completed "some college" which is difficult to accurately assess. The share that reports a college degree is constant across the unweighted Shift, the ACS, the CPS, and the weighted Shift estimates. We also see that the estimate of school enrollment – 37% – in the unweighted Shift data is between the somewhat lower estimate in ACS (32%) and the higher estimate in CPS (46%).

## COMPARISON WITH NATIONAL SURVEYS

As previously mentioned, an important rationale for developing this method of survey recruitment using Facebook is to address a lack of available data. Although the employer-employee linked database that we have compiled is unique, we can make some comparisons of tabulations from our dataset to overlapping measures available in two widely used and carefully constructed probability sample national surveys: the National Longitudinal Survey of Youth (1997) (NLSY97) and the Current Population Survey (CPS). In particular, we estimate and compare (a) regression-adjusted wages, (b) job tenure, and (c) the relationship between job tenure and wages from the Shift data and from the NLSY97 and CPS data sources.

Both the NLSY and CPS surveys aim to assemble a representative sample of the United States population – the NLSY97 for the cohort born between 1980 and 1984 and the CPS for the non-institutionalized population over the age of 15. In contrast, the Shift survey aims to recruit respondents in a target population of retail workers under the age of 55 who are paid hourly and work at large firms. Additionally, the CPS has been fielded from 1962-2016 and the NLSY97 from 1997 - 2013, while the Shift data were collected in 2016 and 2017. Our first step then is to align the three samples as closely as possible. We select cases from the most recent rounds of the NLSY97 (2011 and 2013) and from the CPS (2010-2016). We next restrict both samples to respondents who are paid hourly and who work in the industries represented in our data (581, 591, 600, 601, 641, 623, 633, 642, and 691 in the Industry 1990 codes). The 2011 and 2013 rounds of the NLSY97 only include respondents between the ages of 26 and 34. But, the CPS includes respondents of a wide range of ages, and we restrict to those age 18-55 to align with the Shift data.

We then construct harmonized measures across the CPS, NLSY97, and Shift samples of several core variables: hourly wage (inflation adjusted using the CPI), job tenure, gender, age, and survey year. In total, we have 17,828 observations in the Shift data, 1,518 observations in the pooled

CPS, and 1,494 observations in the pooled NLSY-97. We apply the survey weights from the CPS or NLSY97 and we estimate the models on the Shift data using each of our constructed weights.

While the measures are harmonized, the samples from the CPS, NLSY97, and Shift are still not exactly comparable. First, the survey years differ – the NLSY97 data are available for 2011 and 2013, the CPS for 2010, 2012, 2014, and 2016 (when the job tenure module was asked), and the Shift data for 2016 and 2017. Second, the age range in the NLSY97 is much narrower than in the CPS and Shift. Third, the Shift data comes from employees of large firms, while the CPS and NLSY data are for the entire sector, regardless of firm size.

To make comparisons between these three data sets, we first estimate mean values of two key employment characteristics – hourly wage and tenure – after adjusting for age, gender, and indicator terms for year of survey. We compare the estimates from the Shift, CPS, and NLSY97 data. Table 2 shows the regression-adjusted mean wages and distribution of tenure by survey. We estimate wages with an OLS model as a function of tenure, age, gender, and year, and we estimate tenure with a multinomial logistic regression model as a function of wages, age, gender, and year. We estimate these models separately for each combination of survey x weight. Mean hourly wages are very similar across the three surveys - $10.31 in the CPS, $12.58 in NLSY97 and between $10.95 and $11.74 in the Shift data, depending on the weight. There are more substantial differences between the surveys in the distribution of tenure. Here, close to a third of CPS respondents have less than one-year of tenure as compared with 28% of those in NLSY and about a quarter of shift respondents. In contrast, a higher share of Shift respondents are estimated to have 1—2 years of tenure than the share of CPS or NLSY-97 respondents. In turn smaller shares of Shift respondents report 3-5 years of tenure as compared with CPS and NLSY-97. The share with 6 or more years of tenure is similar in NLSY-97 as in Shift, but lower in CPS. In no case do the numbers precisely agree across all three sources, but in no case are they substantially different, either.

Next, we examine whether the well-documented relationship between job tenure and wages varies across the three surveys. For each of the surveys, separately, (and separately for each of the 8 Shift weights), we regress wages on tenure, controlling for age, year, and gender.

Figure 2 presents the key coefficients from these models. Compared with having less than a year of tenure, we see in the left panel that those with 1-2 years of tenure receive a wage premium. The estimated size of this premium is fairly stable across the eight estimates using the eight weights from the Shift data – about $0.60. We see that this estimate also falls between the low estimate of essentially no return to 1-2 years of tenure in the CPS data and the estimate of about $1.30 in the NLSY-79 data. In the middle panel, we present the estimates of the return to waves of having 3-5 years of job tenure. Again, the estimated premium – about $1.80 is stable across the shift estimates and is somewhat higher than in the CPS and somewhat lower than in the NLSY-97. In both cases, the Shift estimates are closer to both the NLSY and the CPS estimates than these two data sources are to each other. The right-hand side panel presents the estimates of the returns to six or more years of tenure. Here, we see more variation in the estimated returns across the eight Shift estimates, ranging between $4.00 and $5.00, but again, essentially falling between the NLSY-97 and the CPS estimates. In Figure 3, we plot out the predicted wage values by tenure for each of the eight Shift estimates and then for the NLSY-97 and CPS. As we would expect given the coefficients, we see approximately parallel lines with a higher intercept for the NLSY-97 and a lower intercept for the CPS.

In sum, these comparisons of univariate statistics and multi-variate relationships between Shift and two high quality probability sample surveys are encouraging. On wages and tenure, Shift is no more different from the NLSY and the CPS than they are from each other. It is important to note though that Shift is not identical to either of the other surveys. But, by comparing against two probability sample surveys we see that no two of the surveys are identical to each other.

## TEST OF SELECTION ON UNOBSERVABLES

We post-stratify and weight our survey data to account for bias on observable demographic characteristics. And, we find that the weighted data can closely replicate established associations from the CPS and NLSY-97. However, it remains possible that our estimates could be biased by selection into the Shift survey on unobservables.

Here, we describe a test of the presence of such selection on unobservables that leverages the particular dynamics of advertising on Facebook and that would be available to anyone who used paid advertising to field a survey. We recruited respondents to the survey through paid advertisements on Facebook. We specified our target audiences and our advertisements were delivered to eligible users based on Facebook's advertisement placement algorithm. However, a unique feature of Facebook's paid advertisements is that users can engage with these paid posts in much the same way that they may engage with posts created by friends or institutions.

Facebook users can share the advertisement to their own timelines or those of their friends. The extent of this sharing can be gauged by the "social reach" of an advertisement in terms of the number of unique users who see the advertisement through social channels and in terms of the number of "social impressions" obtained through such channels. These may then generate "social clicks" in which users click through to the survey from a social share rather than from a paid placement.

Respondents who take our survey because their friends shared the content are likely to be different in meaningful ways than those who are targeted by our paid advertisements. Further, this social sharing may extend the reach of our advertisements beyond those who list their employer to those who do not list an employer, but whose employer is known to friends on Facebook. We leverage the fact that these forms of social engagement with our advertisements are then likely to shift the pool of respondents to the survey and introduce heterogeneity in the composition of the sample at the level of the recruitment advertisement. We compare those who came to the survey through advertisements that experienced high levels of social sharing with those who came through advertisements with little such social activity. Although the expected direction of potential bias is uncertain a priori, if unobserved characteristics bias our estimates, we should see a significant interaction between the extent of social sharing and job tenure on wage rates.

We assess the importance of such dynamics by sequentially interacting post shares, social impressions, social reach, and social clicks with job tenure to predict wages. We ask if there is any significant variation in the returns to tenure by whether respondents were recruited through highly shared recruitment advertisements or more circumscribed advertisements. Of 12 estimated interaction terms, we see that 2 are statistically significant. There is some evidence that return to at least six years of job tenure varies by social sharing. In Figure 4, we plot the estimate returns to tenure (coded into 1-2 years, 3-5 years, and 6+ years – all relative to less than a year) by the range of observed values for the four measures of social sharing. The lines are all flat for those with 1-2 years or 3-5 years of tenure. There is no evidence that the differential selection into the respondent pool induced by social sharing makes a difference for these estimates of the return to tenure. However, there is significant variation in the estimate of the return to 6 or more years of tenure by the number of social shares and the extent of social reach. Respondents who were recruited through these advertisements seem to have a smaller return to 6 or more years of tenure.

While significant, the variation is not large. The estimated return to six years or more of tenure in the preferred pooled models is $4.53. Here, at the lowest levels of social sharing the estimate is $4.70-$4.80 and at the highest levels it is $3.60-$3.70. By way of comparison, the estimated return to 6 or more years of tenure ranges from $4.96 in the NLSY-97 to $3.00 in the CPS. While there is some evidence of bias, the magnitude is substantially less than the difference between NLSY and CPS data sources.

## THE VALUE OF FIRM-LEVEL DATA: GENDER COMPOSITION AND WAGES

The prior section demonstrates how the Shift Project data produces estimates of wages, job tenure, and the relationship between tenure and wages that are broadly consistent with CPS and NLSY data sources and that do not show major bias from unobservables. However, one of the primary rationales for the Shift Project was to collect data at more granular levels, including samples of workers at particular named companies that is not readily available in standard data sources.

To give one illustration of how the Shift Project data on workers at named employers can be used to address research questions that cannot be addressed with existing data, we draw on the large existing literature on how the gender composition of jobs is associated with wages (England, 1992). Here, the leading theoretical explanation is that workers in female-dominated jobs are paid less precisely because work that is associated with women is devalued and so less well compensated, even though comparable in terms of job requirements to similar jobs that may be done mostly by men (Levanon, England, and Allison, 2009; England, et al., 1988). Sociologists, economists, and demographers have amassed a large body of evidence that workers employed in jobs that have larger shares of female incumbents are indeed paid lower wages (Levanon et al., 2009; Reskin and Bielby, 2005; England, 2005). Notably, this wage penalty is found for both women in female-dominated occupations and for men in such occupations (Budig, 2003).

However, this research has, in almost all cases, measured gender composition using occupations or the intersection of industry and occupation (Huffman and Velasco, 1997). This source of gender segregation is clearly important for the dynamics of gender inequality in wages. But, there are several other important sources of gender segregation as well. Reskin and Hartmann (1986) point out that in addition to occupational segregation, between-firm gender segregation may also importantly shape gender wage inequality – for instance as men are employed as waiters at fine dining establishments, but women as waitresses at coffee shops. Yet, very little existing research has examined the consequences of between-firm gender segregation for gender inequality in wages. Recent work using data from the Equal Employment Opportunity Commission examines how managerial gender is related to sex composition within firms (Huffman et al., 2010; Kurtulus and Tomaskovic-Devey, 2012), but the EEO file lacks data on wages. The work that comes closest to examining how firm-level segregation impacts wages is Tomaskovic-Devey's (1993) use of a unique 1989 survey of North Carolina employees in which respondents report on the gender composition of their co-workers. Tomaskovic-Devey (1993) finds that, drawing on this firm data, the percent female within a job is indeed negatively associated with wages. However, that research is limited to a single state, is more than 30 years old, and relies on a single reporter within each firm to gauge wages and gender composition.

Here, we show how, using the Shift data, we can examine how the gender composition of a relatively homogeneous set of service sector occupations are remunerated and whether this varies by the gender composition of the particular firm. We first generate firm-level measures of gender composition by taking the share of female respondents among all respondents at each of the 38 firms in our data, employing the Facebook weights discussed above. The percent female ranges from 23% at Gamestop to 92% female at Victoria's Secret with a mean (median) of 60% (59%) female across all 38 firms.

We next regress the hourly wage for male and female respondents (pooled) in our data on the gender composition of their employer. In a second model, we introduce controls for demographic and human capital characteristics that could plausibly confound this relationship – age, marital status, race/ethnicity, educational attainment, presence of children in the household, tenure on the job, and managerial status. The relationship between gender composition and wages could though be confounded by non-demographic and non-human capital factors. In particular, workers may accept lower wages in return for other compensating job features (Budig and England, 2001). In female-dominated occupations, these compensating differentials might be found in work schedules that would be less likely to conflict with care obligations. We control for this source of confounding by measuring work schedule type (regular day, regular night, regular evening, variable, split/rotating), week-to-week variation in work hours, number of weeks of advance notice, whether the employee works on-call shifts, whether the employee has had shifts cancelled, whether the employee has input into his/her work schedule, and a three-item scale measure of work-life conflict engendered by the employee's job. While measures such as the gender composition of firms are available in the LEHD, it would not be possible to control for this rich set of confounding factors in such administrative data. In a third model, we test if gender composition is similarly associated with men's wages and women's wages (as Budig (2003) finds). Finally, we investigate how the wage returns to tenure that we discussed previously may be moderated by occupational gender composition.

We present the results of these models in Table 3. Model 1 shows the unadjusted relationship between firm-level gender composition and wages and we see a large (-5.69) and statistically significant negative association. In Model 2, we see that, as we would expect, this estimate is substantially reduced after adjusting for worker-level factors – to about -2.31, but remains negative and statistically significant. In Model 3, we test if the wage penalty of working at a female-dominated firm is different for male and female workers. The interaction is small and not statistically significant. We then use the estimates in Model 2 to plot predicted wages by gender composition in Figure 5. We see that wages at the firms with the smallest share of female employees (Gamestop at 23% female) earn about $12.50 per hour as compared with $10.60 at the firm with the largest share of female employees (Victoria's Secret at 92% female). This difference is found after adjusting for the host of individual level characteristics described above. In plain terms, we find that the mostly female workforce selling women's underwear make about $2.00 per hour less than the mostly male workforce selling video games.

Model 4 provides some evidence for how these between-firm wage gaps take form over time. We see that there is a strong and significant interaction between job tenure and gender composition. We plot predicted wages by years of tenure across the percent of firm employees that are female in Figure 6. It is evident that at the male-dominated firms in our data there are substantial returns to tenure. Employees with 1-2 years tenure earn modestly more than those with less than a year of tenure and those with 3-5 years of tenure do better still. Those with the most tenure, six or more years, see substantial returns to their experience. In contrast, at female-dominated firms we simply see the absence of a career ladder – the returns to anything less than six years of tenure are non-significant and even long tenures of six years or more are associated with only a very modest wage gain. Interestingly, while the returns to 3-5 and 6 or more years of tenure are sharply graded, there is little evidence of an association between firm gender composition and wages among recent hires. We also test for interactions with respondent gender and do not find any evidence that these dynamics differ between men and women.

## CONCLUSION

We describe a new integrated approach to non-probability sampling and survey recruitment that leverages the powerful targeting capabilities of Facebook. Our intervention comes at a time when traditional probability sampling has been declared to be in crisis, beset by low response rates and worsening sampling frames. Important debate and testing continues on the viability of non-probability online surveys. But, out of this debate, there appears to be an emerging consensus that it remains important to continue to investigate the utility of nonprobability web-based surveys and that such approaches can have real value depending on the research objectives (Schonlau and Couper, 2017; AAPOR, 2010).

Here, we intervene to try to solve a problem that has long frustrated survey research – to build a sample of respondents, one must have a sampling frame. While researchers have found creative ways to build frames for the general population, it remains very difficult to sample respondents who are nested within organizational entities who may be reluctant or unable to share lists of employees, students, alumni, or members.

While marketers spend tens of billions of dollars a year using Facebook's targeting tools to try to build brand awareness and sell products to Facebook users, we show that these tools are valuable for survey research as well. We illustrate how targeted advertising on Facebook can be used to build an employee-employer matched data set where hundreds of employees at each of several dozen large firms are recruited and surveyed.

This approach to data collection has several advantages. First, as described above, it provides sampling frames that do not exist (or are very difficult to access) otherwise. Second, it allows for rapid data collection. Third, it is very low cost as compared to traditional survey approaches. We also show that these data can be easily weighted to the demographic attributes of similar target populations in such gold standard surveys as the ACS and the CPS, as well as to the eligible population of Facebook users. We then show that respondents in our data very closely resemble respondents in two large and widely used labor force studies – the CPS and the NLSY97 – on the key characteristics of wage and tenure. Indeed, there are relatively small differences in wages and tenure across the two studies and, to the extent that there are differences, the Shift data is no more different from the CPS and NLSY97 than they are from each other. We also show that multivariate relationships – between wages and tenure and wages – are very similar in the Shift data as in the NLSY97 and CPS.

For the example we illustrate here, these general benefits have tangible results. Because we can inexpensively target employees at a large number of firms, we deploy this method to build an ongoing national monitoring survey of employer management practices and job quality in the retail sector. The ability to rapidly implement a survey and quickly collect data allows us to use these tools to evaluate new local and state ordinances that regulate the employment practices of large retail firms.

This speed has allowed us to quickly collect pre-treatment data once laws are passed, but before they go into effect. Finally, the ability to collect large numbers of responses from employees nested within firms will permit us to examine how company-level attributes may shape the experience of low-wage work. To illustrate this potential, we estimated the relationship between the gender composition of particular employers and wages, showing that workers employed by service sector employers with a greater share of men in the workforce enjoy higher wages and higher returns to to job tenure compared with employers with more female workforces.

These tools are likely to be useful to researchers working in other areas as well where it can be difficult to recruit targeted samples or to access useful sampling frames. The approach is particularly well-suited to research questions that seek to nest multiple level-one observations within a set of level 2 units. While our case is companies, scholarship concerned with educational institutions (such as colleges, secondary schools, charters, etc…), military units, neighborhoods, or voluntary organizations might benefit from this approach.

## REFERENCES


AAPOR. 2010. "Report on Online Panels." *Public Opinion Quarterly* 74(4): 711-781.

AAPOR. 2014. "Social Media in Public Opinion Research: Report of the AAPOR Task Force on Emerging Technologies in Public Opinion Research." Report.

Andersson, Fredrik, Harry Holzer, and Julia Lane. 2005. *Moving Up or Moving On: Who Advances in the Low-Wage Labor Market*. New York: Russell Sage Foundation.

Applebaum, Eileen and Rosemary Batt. 2014. *Private Equity at Work*. New York: Russell Sage Foundation.

Baltar, Fabiola and Ignasi Brunet. 2012. "Social Research 2.0: Virtual Snowball Sampling Method Using Facebook." Working Paper.

Best, Samuel, Brian Krueger, and Clark Hubbard. 2001. "An Assessment of the Generalizability of Internet Surveys." *Social Science Computer Review* 19: 131–141.

Bethlehem, Jelke. 2010. "Selection Bias in Web Surveys." *International Statistical Review* 78 (2): 161–188.

Bhutta, Christine. 2012. "Not by the Book: Facebook as a Sampling Frame." *Sociological Methods & Research* 41(1): 57-88.

Billari, Francesco, Francesco D'Amuri. and Juri Marcucci. 2013. "Forecasting Births Using Google" Conference paper presented at the Annual Meeting of the Population Association of America.

Bruggen, E., J. van den Brakel, and J. Krosnick. 2016. "Establishing the Accuracy of Online Panels for Survey Research." Discussion Paper.

Budig, Michelle. 2003. "Male Advantage and the Gender Composition of Jobs: Who Rides the Glass Escalator?" *Social Problems* 49(2): 258-277.

Budig, Michelle and Paula England. 2001. "The Wage Penalty for Motherhood." *American Sociological Review* 66(2): 204-225.

Casler, Krista, Lydia Bickel, and Elizabeth Hackett. 2013. "Separate but Equal? A Comparison of Participants and Data Gathered via Amazon's MTurk, social media, and Face-to-Face Behavioral Testing." *Computers in Human Behavior* 29: 2156-2160.

Clifford, Scott, Ryan Jewell, and Philip Waggoner. 2015. "Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?" *Research & Politics* 2: 1-9.

Colla, Carrie, William Dow, Arindrajit Dube, and Vicky Lovell. 2014. "Early Effects of the San Francisco Paid Sick Leave Policy." *American Journal of Public* Health 104(12): 2453-2460.

Couper, Mick. 2017. "New Developments in Survey Data Collection." *Annual Review of Sociology* 43: 121-145.

Dutwin, David and Trent Buskirk. 2017. "Apples to Oranges or Gala versus Golden Delicious? Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples." *Public Opinion Quarterly* 81: 213-239.

England, Paula, George Farkas, Barbara Stanek Kilbourne, and Thomas Dou. 1988. "Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects." *American Sociological Review* 53(4): 544-58.

England Paula. 1992. *Comparable Worth. Theories and Evidence*. New York: Aldine de Gruyter.

Facebook. 2017. "What Names are Allowed on Facebook?" Accessed online at https://www.facebook.com/help/112146705538576.

Fligstein, Neil. 2001. *The Architecture of Markets: An Economic Sociology of Twenty-First-Century Capitalist Societies*. Princeton, NJ: Princeton University Press.

Goel, Sharad, Adam Obeng, and David Rothschild. 2015. "Non-Representative Surveys: Fast, Cheap, and Mostly Accurate." Working Paper.

Greenwood, Shannon, Andrew Perrin, and Maeve Duggan. 2016. *Social Media Update.* Pew Research Center.

Groshen, Erica. 1991a. "Sources of Intra-Industry Wage Dispersion: How Much do Employers Matter?" *Quarterly Journal of Economics* 106(3): 871-884.

Groshen, Erica. 1991b. "Five Reasons Why Wages Vary Among Employers." *Industrial Relations* 30(3): 350-381.

Groves, Robert 2011. "Three Eras of Survey Research." *Public Opinion Quarterly* 75(5): 861.

Huffman, Matt and Steven Velasco. 1997. "When More is Less: Sex Composition, Organizations, and Earnings in U.S. Firms." *Work and Occupations* 24(2): 214-244.

Huffman, Matt L., Philip N. Cohen, and Jessica Pearlman. 2010. "Engendering Change: Organizational Dynamics and Workplace Gender Desegregation, 1975-2005." *Administrative Science Quarterly* 55: 255-277.

Keeter, Scott, Nick Hatley, Courtney Kennedy, and Arnold Lau. 2017. *What Low Response Rates Mean for Telephone Surveys.* Pew Research Center.

Kohut, Andrew, Scott Keeter, Carroll Doherty, Michael Dimock, and Leah Christian. 2012. *Assessing the Representativeness of Public Opinion Surveys.* The Pew Research Center for the People and the Press.

Krueger, Alan, and Lawrence Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica* 56(2): 259–93.

Kurtulus, Fidan Ana and Donald Tomaskovic-Devey. 2012. "Do Female Top Managers Help Women to Advance? A Panel Study Using EEO-1 Records." *The Annals of the American Academy of Political and Social Science* 639: 173-197.

Lane, Julia, Laurie Salmon, and James Spletzer. 2007. "Establishment Wage Differentials." *Monthly Labor Review.*

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.

Mullinix, Kevin, Thomas Leeper, James Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2: 109-138.

Nunan, Daniel and Simon Knox. 2011. "Can search engine advertising help access rare samples?" *International Journal of Market Research* 53(4): 523.

Ramo, Danielle and Judith Prochaska. 2012. "Broad Reach and Targeted Recruitment Using Facebook for an Online Survey of Young Adult Substance Use." *Journal of Medical Internet Research* 14(1): e28.

Reskin, Barbara and Denise Bielby 2005. "A sociological perspective on gender and career outcomes." *Journal of Economic Perspectives* 19(1): 71–86.

Reskin, Barbara and Hedi Hartmann. 1986. *Women's Work, Men's Work: Sex Segregation on the Job.* Washington DC: National Academy Press.

Reis, Ben Y. and John S. Brownstein. 2010 "Measuring the Impact of Health Policies using Internet Search Patterns: The Case of Abortion" *BMC Public Health*, Vol. 10 No. 1, pp. 514.

Ryan, Camille, and Jamie M. Lewis. 2017. "Computer and Internet Use in the United States: 2015" *American Community Survey Reports, ACS-37*. Washington, DC: U.S. Census Bureau.

Samuels, David and Cesar Zucco. 2013. "Using Facebook as a Subject Recruitment Tool for Survey-Experimental Research." *Experimental Research.*

Schonlau, Matthais and Mick Couper. 2017. "Options for Conducting Web Surveys." *Statistical Science* 32 (2): 279-292.

Smith, Tom. 2013. "Survey-Research Paradigms Old and New". *International Journal of Public Opinion Research* 25(2): 218-229.

Stern, Michael, Ipek Bilgen, and Don Dillman. 2014. "The State of Survey Methodology: Challenges, Dilemmas, and New Frontiers in the Era of the Tailored Design." *Field Methods* 26(3): 284-301.

Thornton, Louise, Philip J. Batterham, Daniel B. Fassnacht, Frances Kay-Lambkin, Alison L. Calear, Sally Hunt. 2016. "Recruiting for health, medical or psychosocial research using Facebook: Systematic review." *Internet Interventions* 4 (1): 72-81.

Tomaskovic-Devey, Donald. 1993. "Gender and Race Composition of Jobs and the Male/Female, White/Black Pay Gaps." *Social Forces* 72(1):45-76.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2015. "Forecasting Elections with Non-Representative Polls." *International Journal of Forecasting* 31(3): 980-991

Weil, David. 2009. "Rethinking the Regulation of Vulnerable Work in the USA: A Sector-based Approach." *Journal of Industrial Relations.* 51(3): 411-430.

Yeager, David, Jon Krosnick, LinChiat Chang, Harold Javitz, Matthew Levendusky, Alberto Simpser, and Rui Wang. 2011. "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples." *Public Opinion Quarterly* 75 (4): 709-747.

Zagheni, Emilio, Ingmar Weber, and Krishna Gummadi. 2017. "Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants." *Population and Development Review* 43(4): 721–734.

Zahgeni, Emilio, and Ingmar Weber. 2015. "Demographic Research with Non-Representative Internet Data" *International Journal of Manpower* 36(1): 13-25.

Zagheni, Emilio and Ingmar Weber. 2012. "You are Where You e-mail: Using E-mail Data to Estimate International Migration Rates" *Proceedings of Web Science* (WebSci), pp. 348-351.

Zagheni, Emilio, Venkata Rama Kiran Garimella, Ingmar Weber, and Bogdan State. 2014. "Inferring International and Internal Migration Patterns from Twitter Data", *Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web* (WWW), pp. 439-444

Zhang, Baobao Matto Mildenberger, Peter D. Howe, Jennifer Marlon, Seth Rosenthal, and Anthony Leiserowitz. 2017. "Quota Sampling Using Facebook Advertisements Can Generate Nationally Representative Opinion Estimates." Working Paper.

## TABLES AND FIGURES

**Figure 1. Examples of Employer-Specific Survey Recruitment Advertisements Placed on Facebook.**

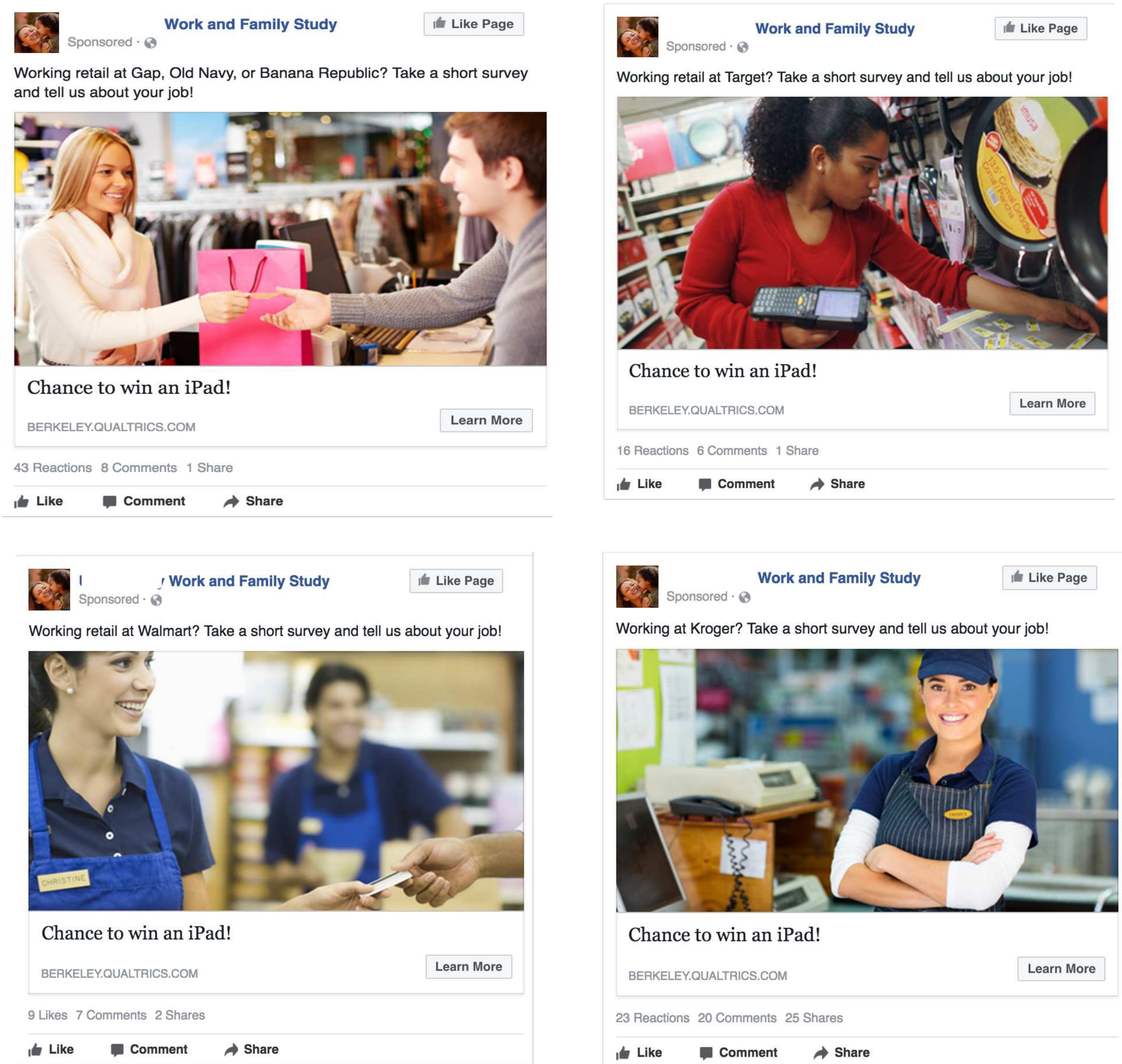

**Figure 2. Association between Job Tenure and Inflation Adjusted Hourly Wage in the CPS (2010-2016), NLSY97 (2011-2013), and Shift (2016-2017) surveys. Adjusted for age, gender, and survey year.**

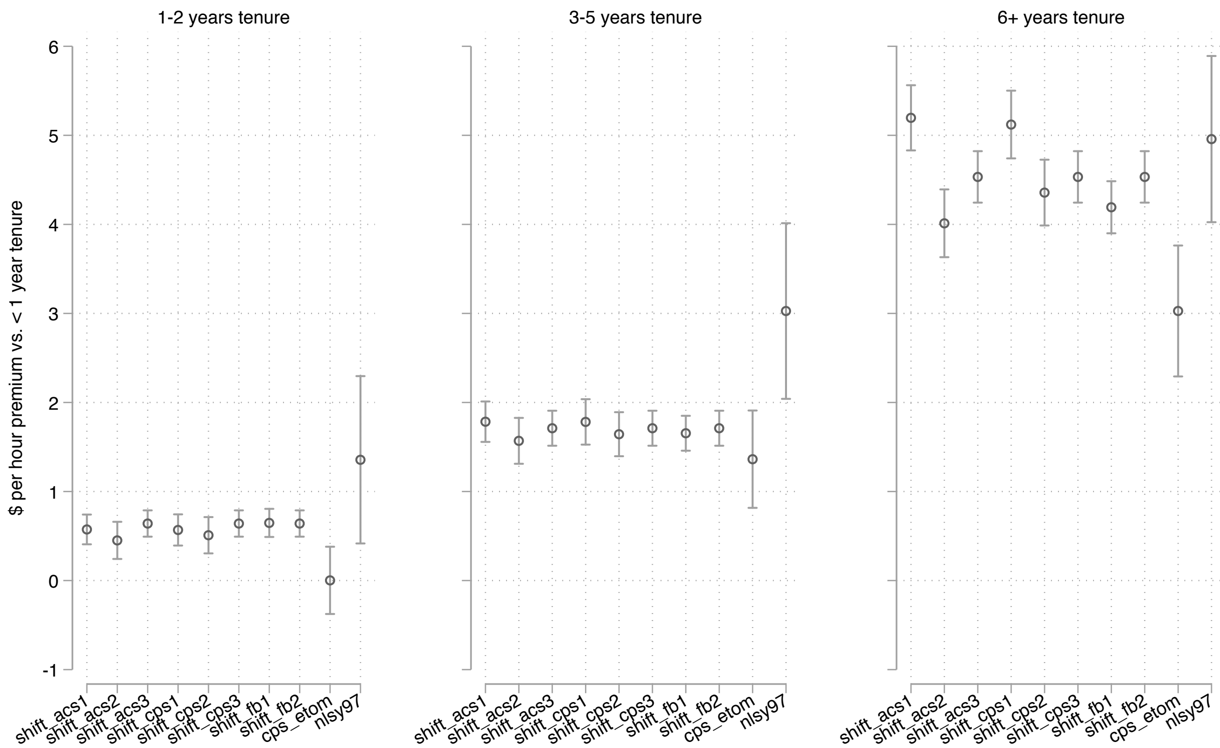

**Figure 3. Predicted Wages by Job Tenure in the CPS (2010-2016), NLSY97 (2011-2013), and Shift (2016-2017) surveys. Adjusted for age, gender, and survey year.**

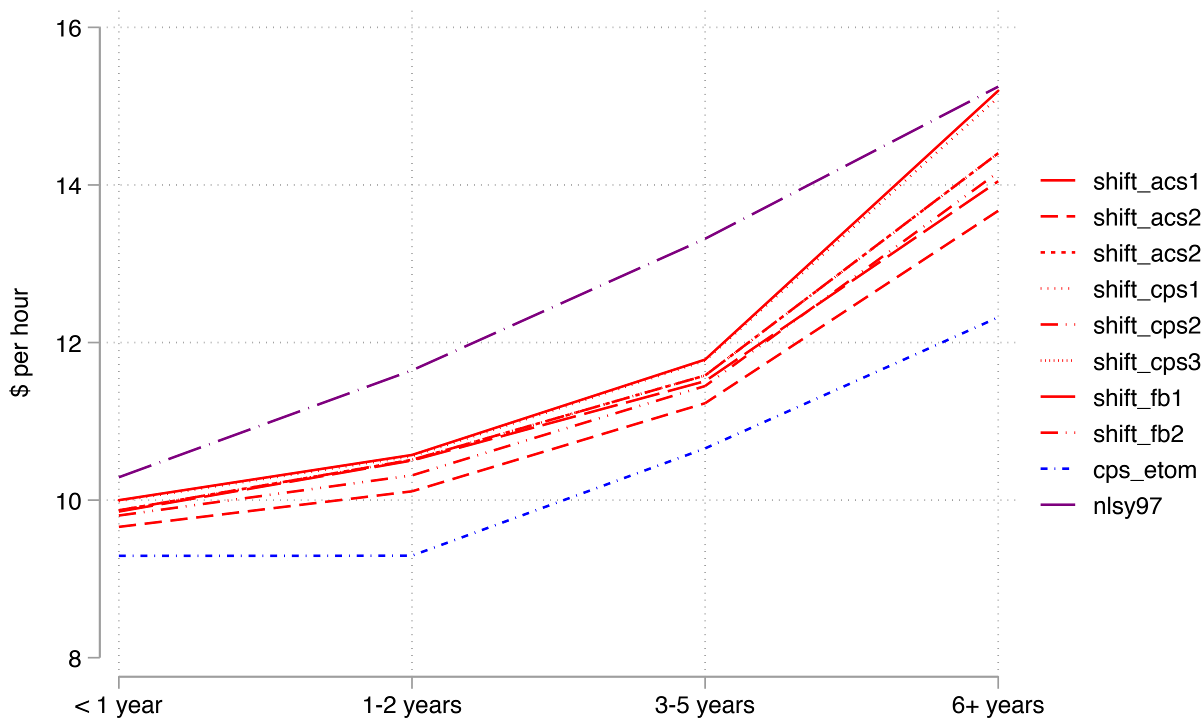

**Figure 4. Variation in Wage-Tenure Relationship in Shift Data by Advertisement-Level Social Sharing Activity**

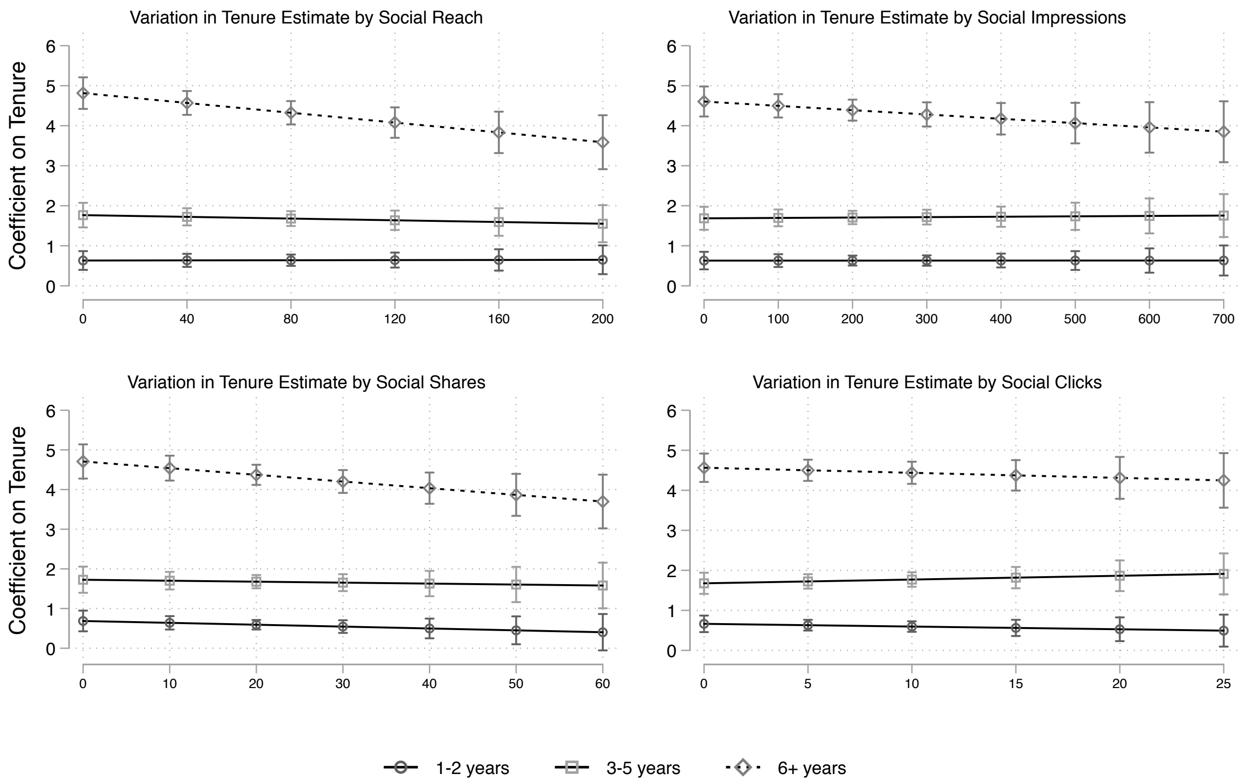

**Figure 5. Firm Gender Composition and Wages**

**Figure 6. Wage Returns to Tenure by Firm Gender Composition**

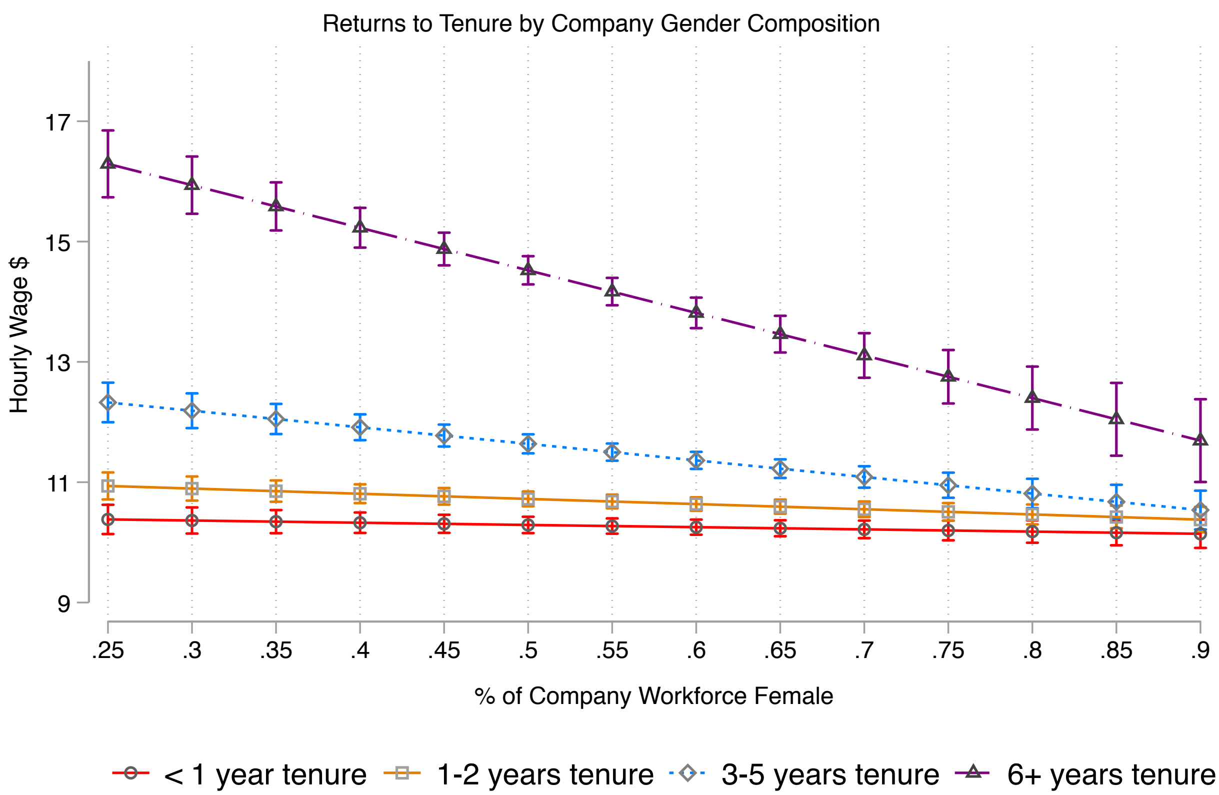

**Table 1. Comparison of Key Demographics in Shift, ACS, CPS ASEC, and Facebook**

| | Shift | ACS | CPS ASEC | Face book | Shift ACS_1 | Shift CPS_1 | Shift FB_1 |
|---|---|---|---|---|---|---|---|
| **Female** | 73% | 55% | 55% | 60% | 55% | 54% | 60% |
| **Age** | | | | | | | |
| 18-29 | 67% | 59% | 59% | 61% | 60% | 60% | 61% |
| 30-39 | 17% | 19% | 18% | 22% | 19% | 18% | 22% |
| 40-49 | 13% | 14% | 14% | 12% | 13% | 14% | 12% |
| 50-55 | 2% | 8% | 9% | 6% | 8% | 9% | 6% |
| **Race/Ethnicity** | | | | | | | |
| White, NH | 75% | 54% | 54% | -- | 54% | 53% | 76% |
| Black, NH | 4% | 15% | 17% | -- | 15% | 18% | 4% |
| Hispanic | 13% | 22% | 21% | -- | 22% | 21% | 12% |
| Other, NH | 8% | 9% | 7% | -- | 9% | 8% | 8% |
| **Education** | | | | | | | |
| HS or less | 39% | 51% | 54% | -- | 39% | 39% | 41% |
| Some Coll/AA | 52% | 40% | 37% | -- | 52% | 52% | 50% |
| BA or more | 9% | 9% | 9% | -- | 10% | 9% | 9% |
| **Enrolled in School** | 37% | 32% | 46% | -- | 33% | 33% | 33% |

**Table 2. Comparison of Mean Hourly Wage and Job Tenure in Shift, CPS, and NLSY97**

| | CPS | NLSY97 | Shift ACS_1 | Shift ACS_2 | Shift ACS_3 | Shift CPS_1 | Shift CPS_2 | Shift CPS_3 | Shift FB_1 | Shift FB_2 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mean Hourly Wage** | $10.31 | $12.58 | $11.74 | $10.95 | $11.47 | $11.71 | $11.24 | $11.47 | $11.35 | $11.47 |
| **Job Tenure** | | | | | | | | | | |
| < 1 year | 32% | 28% | 23% | 25% | 23% | 23% | 24% | 23% | 23% | 23% |
| 1-2 years | 25% | 21% | 33% | 34% | 33% | 33% | 34% | 33% | 33% | 33% |
| 3-5 years | 26% | 27% | 21% | 21% | 21% | 21% | 20% | 21% | 22% | 21% |
| 6+ years | 16% | 24% | 23% | 20% | 23% | 23% | 21% | 23% | 22% | 23% |

**Table 3. Firm Gender Composition and Wages**

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| % Female in Firm | -5.69 *** | -2.31 *** | -2.16 *** | -0.38 |
| Respondent Male | -- | 0.61 *** | 0.90 * | 0.60 *** |
| % Female * Female | -- | -- | 0.51 | -- |
| **Tenure** | | | | |
| < 1 Year | -- | (ref) | (ref) | (ref) |
| 1-2 Years | -- | 0.35 *** | 0.35 *** | 0.64 * |
| 3-5 Years | -- | 1.10 *** | 1.10 *** | 2.53 *** |
| 6+ Years | -- | 3.68 *** | 3.68 *** | 7.71 *** |
| **% Female * Tenure** | | | | |
| % Female * < 1 Year | -- | -- | -- | -- |
| % Female * 1-2 Years | -- | -- | -- | -0.46 |
| % Female * 3-5 Years | -- | -- | -- | -2.38 *** |
| % Female * 6+ Years | -- | -- | -- | -6.93 *** |
| Individual Controls | N | Y | Y | Y |
| N | 17828 | 17828 | 17828 | 17828 |