# How Does the Dramatic Rise of CPS Nonresponse Impact Labor Market Indicators?*

Robert Bernhardt† David Munro‡ Erin L. Wolcott§

September 2023

**Abstract**

Within a decade, the share of households refusing to participate in the Current Population Survey (CPS) tripled. We show households that refuse one month but respond in an adjacent month account for an important part of the rise. Leveraging the labor force status of survey participants in the months surrounding their nonresponse, we find rising refusals suppressed the measured labor force participation rate and employment-population ratio but had little effect on the unemployment rate. Notably, nonresponse bias accounts for at least 10 percent of the reported decline in the labor force participation rate from 2000 to 2020.

**Keywords:** Current Population Survey, Unemployment Rate, Labor Force Participation Rate, Employment-Population Ratio, Non-interview, Survey Refusal, Bias.

**JEL Codes:** C83, E24, J64

*We thank Jim Hamilton, Hie Joo Ahn, Karel Mertens, Dan Bernhardt, David Wiczer, Larry Warren, and seminar participants from the Federal Reserve Bank of Dallas, U.S. Bureau of Labor Statistics, Council of Economic Advisors, University of Illinois Urbana-Champaign, St. Lawrence University, 2021 WEAI meetings, 2021 Liberal Arts Macro Conference, and 2021 Econ-Dynamics Workshop for their helpful comments.

†Department of Economics, University of Chicago, *rbernhardt@uchicago.edu*.

‡Department of Economics, Middlebury College, *dmunro@middlebury.edu* and GLO Fellow.

§Department of Economics, Middlebury College, *wolcott@middlebury.edu*.

# 1 Introduction

Since 2010, the share of occupied households in the United States not responding to the Current Population Survey (CPS) increased from 8 to over 20 percent. Figure 1 shows the steady increase is from households refusing to participate in the survey. We focus on Type A unit nonresponse which is where an occupied housing unit does not respond. This is distinct from Type B and Type C unit nonresponse which is where a housing unit is unoccupied, and it is distinct from item nonresponse which is where a household responds to the survey but the interviewee fails to answer a specific question. Although Figure 1 reports nonresponse rates through July 2021, we do not focus on the recent, temporary spike in nonresponse for "other" reasons related to the Census Bureau suspending in-person interviews in April 2020 because of COVID-19. (BLS, 2020; Rothbaum and Bee, 2021).[1,2]

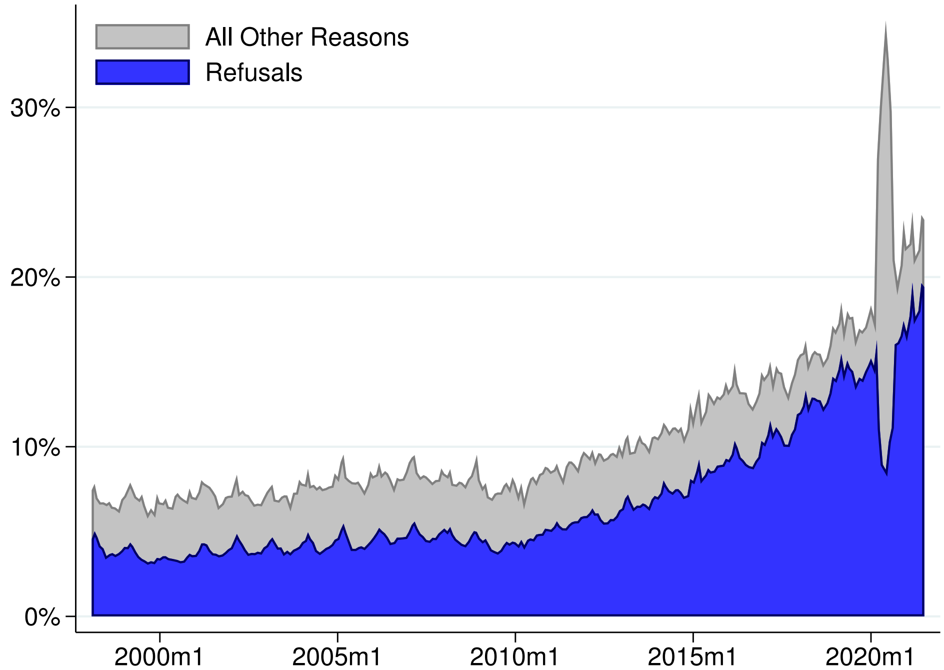
Headline labor market statistics are calculated from the CPS. These missing households raise questions about the accuracy of key labor market indicators used to monitor the United States economy and calibrate economic models. We document that the increase in missing observations from household nonresponse is *not* random. It has biased the labor force participation rate and employment-population ratio down but has had little discernible effect on the unemployment rate. We offer a correction method to adjust for nonresponse. Our correction method is an improvement over current methods because it imputes the labor force status of missing observations instead of weighting the sample primarily to make it demographically representative. In doing so, we find nonresponse bias is larger than previously thought and is growing.

The major challenge researchers face studying nonresponse is that we cannot observe the characteristics of nonresponders to see whether there is selection into the nonresponding

[1] Refusals and "other" are sub-categories of Type A nonresponse.

[2] See Bick and Blandin (2023), Foote et al. (2021), and Faberman et al. (2022) for alternative survey data measuring the pandemic labor market.

Figure 1: CPS Nonresponse Rates of Occupied Dwellings




**Notes:** Authors' calculations using data from the CPS. The number of Type A non-interviewed households divided by the total number of interviewed and Type A non-interviewed households from January 1998 through July 2021. All Other Reasons include: no one home, unable to locate, temporarily absent, language barrier, and other.

group. But unlike purely cross-sectional surveys, the panel structure of the CPS provides some information about nonresponding households. The CPS surveys households eight times, separated by at least a month. Households have the choice to not respond during each of the eight survey months and we leverage this panel structure to learn more about nonresponders.

To glean information about the possible labor force status of nonresponders, it is important that we observe a large share of nonresponders at some other point in their panel life. We start by documenting that over the last decade, there has been a growing share of two types of nonresponses: (1) nonresponses from households that respond in an adjacent month and (2) nonresponses from households that do not respond (or are out of the survey) in an adjacent month. Because the first type contributes to at least a third of total nonresponses, and because we have information about these households during the months they *do* respond, nonresponses with adjacent information are the cornerstone of our analysis.

Among nonresponses with adjacent information, we classify their household as a household that: (A) leaves the survey after responding to a panel (drop-*out*) or (B) enters the survey after not responding to a panel (drop-*in*). Note, it is possible for a household to be both a drop-out and drop-in. If drop-outs, drop-ins, and consecutive responders were identical, we would worry less about selective attrition. We show, however, this is not the case in the CPS. Selective response behavior, especially since 2010, has artificially biased the sample away from individuals participating in the labor force.

We offer a correction method by leveraging the panel structure of the CPS. With a sample of consecutive responders, we calculate monthly flow rates between labor force statuses over time. We apply these flow rates to respondents in the months surrounding a nonresponse to fill in their missing observations with the likelihood they are employed, unemployed, and out of the labor force.

This correction method has little effect on the unemployment rate. However, the reported labor force participation rate and employment-population ratio are lower than our corrected time series. Using the raw counts of individuals in the CPS, we find the magnitude of our correction has grown by three-fold, to over a percentage point, between 2010 and 2020. This accounts for nearly 20 percent of the decline in the (unweighted) participation rate since the turn of the millennium. The unweighted series, however, is not the official labor force participation rate. The Bureau of Labor Statistics (BLS) weights the sample primarily to ensure it is demographically and geographically representative. Applying BLS weights to our corrected data is tricky because we do not want to overcorrect for nonresponse. After a careful multi-step process, we find the BLS weights correct for some but not all of the growing bias. Nonresponse bias still accounts for 10 percent of the decline in the official (weighted) participation rate since the turn of the millennium. More concerning, however, is if refusal rates continue to increase at the same rate—which they appear to be doing—

the bias is liable to increase in severity.

Our work contributes to a long-standing, yet still growing, literature seeking to understand the rise of nonresponse across household surveys (Harris-Kojetin and Tucker, 1999; Atrostic, Bates, Burt, and Silberstein, 2001; Brick and Williams, 2013; Schoeni, Stafford, McGonagle, and Andreski, 2013; Meyer, Mok, and Sullivan, 2015; Williams and Brick, 2018; Dutz, Huitfeldt, Lacouture, Mogstad, Torgovitsky, and van Dijk, 2021).

This paper complements several recent papers studying nonresponse in the CPS. Korinek et al. (2007), Bee et al. (2015), and Hokayem et al. (2015) study the effect of missing observations on measures of income but do not address measures of labor force status. Heffetz and Reeves (2019) show easy-to-reach and hard-to-reach respondents, as measured by the number of survey attempts, are systematically different. If nonresponders are more similar to hard-to-reach responders, low response rates impede survey accuracy. In concurrent work, Borgschulte, Cho, and Lubotsky (2022) hypothesize that the increase in refusal rates since 2010 is linked to anti-survey rhetoric among Republican or Tea Party supporters. The authors find inconclusive evidence for this hypothesis, but conclude that the political cycle has influenced response rates since the 1990s with individuals more likely to respond to the CPS when the sitting president aligns with their political party.

In work most close to ours, Ahn and Hamilton (2022) correct for several sources of bias, including rotation group bias, missing observations, and inconsistency between reported job-search durations and observed continuation probabilities. They also attempt to correct for both unit nonresponse and item nonresponse. Our paper differs in that we exclusively focus on adjusting for bias generated from rising unit nonresponse. Understanding how unit nonresponse, in particular, impacts important labor market indicators is of paramount and growing importance given the large and steady increase in survey refusals since 2010. In doing so, we find the labor force participation rate and employment-population ratio are most affected by nonresponse bias, while Ahn and Hamilton (2022) find the unemployment rate is most affected by all sources of bias and misclassification between unemployment and not in the labor force plays an outsized role.[3] Another important difference is that we apply the BLS weights to our corrected series in a way that minimizes the risk of overcorrecting the data.

Our work relates to a literature documenting the prevalence of rotation group bias in the CPS (Bailar, 1975; McCarthy, 1978; Solon, 1986; Halpern-Manners and Warren, 2012; Krueger, Mas, and Niu, 2017). Rotation group bias arises in a panel survey when, for instance, the unemployment rate calculated from households in the first month of the survey differs from the unemployment rate calculated from households in the second month.[4] Because of the notable differences across survey months, we condition on survey month when imputing labor market statuses for nonresponders.

The correction method we offer to account for rising nonresponse is similar to Abowd and Zellner (1985), Tucker and Harris-Kojetin (1998), Fujita and Ramey (2006), Nekarda (2009), and Ahn and Hamilton (2022) in that it conditions on survey participants' previous or future responses to learn about their missing responses.

The paper proceeds as follows. Section 2 describes the data. Section 3 illustrates the ways in which survey refusals are not random and depend on survey drop-in and drop-out behavior. Section 4 corrects for the bias from rising nonresponse; and Section 5 concludes.

# 2 Data

The Current Population Survey is a monthly survey conducted by the U.S. Census Bureau of about 60,000 occupied households (technically housing units) focusing on labor market,

[3]There is a separate literature studying misclassification of labor market variables in the CPS. For example, see Poterba and Summers (1986), Chua and Fuller (1987), Elsby et al. (2015), Kudlyak and Lange (2018), and Vom Lehn et al. (2021).

[4]Appendix A shows that rotation group bias for the labor force participation rate has risen alongside nonresponse rates, but the same cannot be said for the unemployment rate.

educational, and demographic variables. Most famously, it is used to compute the official unemployment rate, labor force participation rate, and employment-population ratio. The CPS uses a 4-8-4 rotating sample design, where selected households are surveyed for a total of eight months. Households are included in the sample for four consecutive months, excluded for eight months, and then surveyed again during the next four months, bringing the total number of survey months to eight. Each household, in a given month, is assigned a month-in-sample (MIS) number ranging from one to eight. The survey is designed so households are always entering and leaving the survey. By design, one eighth of households are surveyed in the first month, and one eighth are surveyed each month thereafter.

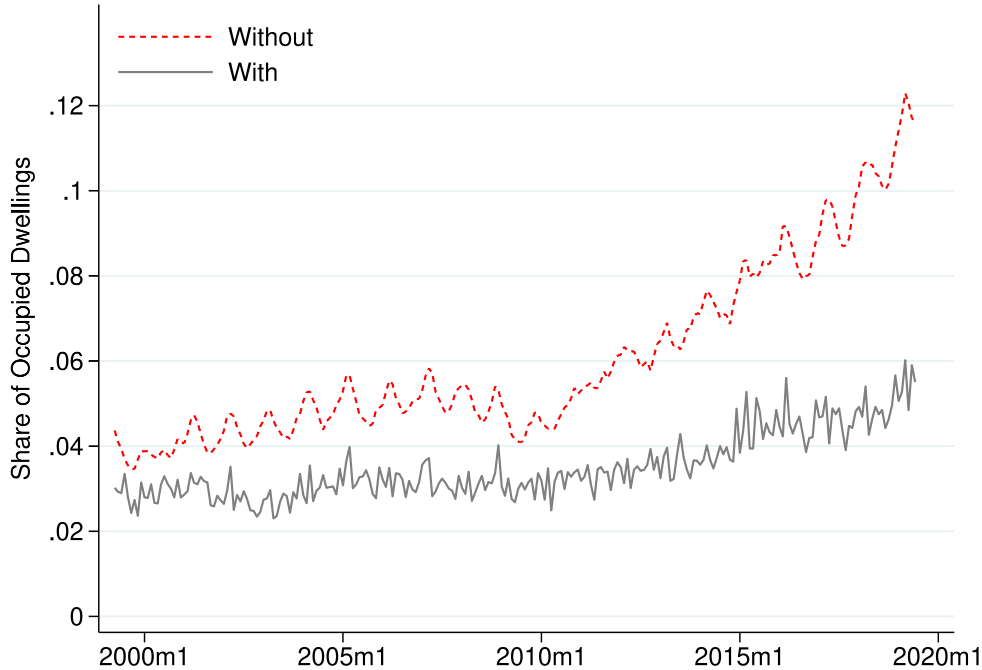
The CPS is a government survey but it is not legally required. Many households do not respond. It is important to note that a household is surveyed for eight months, counting nonresponse months. For example, if a household does not respond for the first two months but responds to all successive surveys, then the CPS will include two nonresponses and six responses for that household.

Our primary sample is the Current Population Survey microdata spanning January 1998 through December 2019 of individuals 16 years and older.[5] Each month of the data contains information on approximately 140,000 individuals in responding households, and approximately 10,000 to 15,000 in nonresponding households (Types A, B, C). For each of these observations we have household response indicators, and where available, personal demographic information and labor market data. We link households and individuals across the eight months of the panel using household, person, and month identifiers.[6] In total, the dataset includes about two million households.

[5]We begin the sample in 1998 because that is when the BLS began publishing the final composite weight (PWCMPWGT) which we use in Section 4.4. We end in December 2019 to avoid the unusual response rates during the pandemic.

[6]We do not put demographic restrictions when linking the panel, however, our results are robust to following Ahn and Hamilton (2022) and assuming a linked observation where the gender is not the same or the age differs by more than two years is a different person. If a responding household moves and is replaced by new occupants during the panel, the CPS records a different household ID.

Figure 2: Nonresponding Households With and Without Adjacent Information




**Notes:** Authors' calculations from the CPS. Lines are nonresponses as a share of total occupied dwellings in a given month. "Without" adjacent information are nonresponse where households either did not respond in an adjacent month or were out of the sample. "With" adjacent information are nonresponses where household responded in an adjacent month.

For our analysis, it is important to know whether nonresponses originate from households that respond in an adjacent month—who we have timely information about—or from households that nonrespond (or are out of the survey) in an adjacent month—who we have less timely information about. We refer to the former situation as nonresponses *with* adjacent information and the latter as nonresponses *without* adjacent information. Take, for example, a household that responds in the first month and then nonrespondes every month after. The nonresponse in MIS 2 has adjacent information from the response in MIS 1, but the nonresponses in MIS 3 through 8 do not.

Figure 2 plots the prevalence of nonresponses with and without adjacent information. Nonresponses *without* adjacent information (red dashed line) account for a larger share of possible responses than nonresponses *with* adjacent information (gray solid line). Another insight from Figure 2 is that both types of nonresponses have increased since 2010. The more pronounced uptick in the red line is driven by households who refuse all eight survey months and we know nothing about. Although nonresponses *without* adjacent information are more prevalent and have increased faster, in 2019, nonresponses *with* adjacent information still represented a third of all nonresponses.

In the following sections, we leverage these nonresponses with adjacent information to document non-random survey response behavior (conditioning on labor force status) and use their adjacent labor force information to correct three key labor market indicators.

# 3 Selective Response Behavior

Households that respond to only some months of the CPS give us a glimpse into what nonresponding households look like in the months they do respond. We document that there is selection. Nonresponses are more likely among people who are *in* the labor force during the months surrounding their nonresponse.

We classify nonresponding households with adjacent information into drop-outs and drop-ins. Drop-*outs* are households that respond in month $t$ but nonrespond in month $t+1$. Drop-*ins* are households that nonrespond in month $t-1$ but respond in month $t$. The CPS has seen a sizable share of both drop-outs and drop-ins since 1998. If drop-outs, drop-ins, and consecutive responders are identical, we would worry less about selective response behavior.[7] Unfortunately, this is not the case in the CPS. There are two margins of selection: (1) drop-outs and drop-ins collectively differ from consecutive responders and (2) drop-ins differ from drop-outs and there are more drop-ins that accumulate over the eight months of the panel.

[7]Nekarda (2009) shows that bias created from people physically moving out (Type B nonresponse) is small because the people moving in have similar characteristics. Our focus is on Type A nonresponse.

Table 1: Selective Response Behavior

Panel A: Drop-outs

| MIS $t$ Labor Force Status | MIS $t+1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 61.42%<br>(9,690,933) | 67.33%<br>(268,408) | -5.91 pp |
| Unemployed | 3.14%<br>(494,590) | 3.62%<br>(14,431) | -0.48 pp |
| NILF | 35.45%<br>(5,592,604) | 29.05%<br>(115,808) | 6.40 pp |

Panel B: Drop-ins

| MIS $t$ Labor Force Status | MIS $t-1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 61.27%<br>(9,668,014) | 66.63%<br>(339,031) | -5.36 pp |
| Unemployed | 3.06%<br>(483,627) | 3.27%<br>(16,619) | -0.21 pp |
| NILF | 35.67%<br>(5,628,464) | 30.10%<br>(153,154) | 5.57 pp |

**Notes:** Authors' calculations from linking households and individuals across month-in-samples (MIS) in the CPS, where for Panel A, $t \in \{1, 2, 3, 5, 6, 7\}$ and for Panel B, $t \in \{2, 3, 4, 6, 7, 8\}$. Data is aggregated over 1998-2019. Counts are in parenthesis. Each count is a person who lives in a household, where the household either responds or nonresponds. Nonresponses are all Type A. Panel A is the share and count of individuals with said labor force status in MIS $t$ and interview status in MIS $t+1$. Panel B is the share and count individuals with said labor force status in MIS $t$ and interview status in MIS $t-1$. The Response and Nonresponse columns add to 100%. The last column is the percentage point difference between the share of individuals who respond and nonrespond.

The Panel A of Table 1 reports the share and count of responders and drop-*outs* between MIS $t$ and $t+1$, averaging across all years and MIS from 1998 through 2019.[8] The first entry indicates the share of responders who were employed the month before. The next column indicates the share of nonresponders who were employed the month before, and the last column reports the difference. Only 61.42% of $t+1$ responders were employed in $t$, while 67.33% of nonresponders were employed in $t$, implying a $-5.91$ percentage point gap. The second row focuses on individuals who are unemployed. Here, too, the unemployed account for a smaller share of responders than nonresponders: there is a $-0.48$ percentage point gap. The third row focuses on individuals not in the labor force (NILF), and the pattern reverses. Responders are 6.40 percentage points more likely to be NILF than nonresponders. To summarize, drop-outs are more likely to be *in* the labor force than consecutive responders.[9]

The Panel B of Table 1 reports similar statistics as Panel A but this time for drop-*ins* between MIS $t-1$ and $t$. The first entry indicates the share of responders who were employed the month after. The next column indicates the share of nonresponders who were employed the month after, and the last column reports the difference. Only 61.27% of $t-1$ responders were employed in $t$, while 66.63% of nonresponders were employed in $t$, implying a $-5.36$ percentage point gap. The second row focuses on individuals who are unemployed. Here, too, the unemployed account for a smaller share of responders than nonresponders: there is a $-0.21$ percentage point gap. The third row focuses on NILF, and the pattern reverses once again. Responders are 5.57 percentage points more likely to be NILF than nonresponders. In other words, drop-ins are more likely to be *in* the labor force than consecutive responders.[10]

Taken together, both panels of Table 1 reveal that two margins put downward pressure on the reported labor force participation rate. The first is that drop-ins and drop-outs are both

[8] For both panels of Table 1, we only consider MIS where $t$ and $t+1$ (or $t-1$) are separated by a month.

[9] Table 1 does not use BLS weights. This is common in the nonresponse literature because the BLS only provides a "final" weight so it is difficult to distinguish the effects that nonresponse, sample design, and post-stratification have on these weights (Korinek et al., 2007). However, Appendix B.2 shows that using BLS weights has little impact on the findings of Table 1.

[10] Results of Table 1 hold if nonresponses are limited to refusals.

more likely to be *in* the labor force than consecutive responders. By definition, drop-ins and drop-outs respond less than consecutive responders and with "sticky" labor force statuses, this biases the sample away from labor force participation. The second margin at play is that there are more drop-ins than drop-outs.[11] Moreover, drop-ins contain a larger share of NILF than drop-outs: 30.10% relative to 29.05%. Accounting for the additional households that drop into the survey and are disproportionally NILF biases the full-sample participation rate downward relative to a participation rate calculated from only MIS 1 responses.

The results above highlight that whether someone continues to respond to the CPS, ceases responding, or commences responding after a nonresponse depends on their labor force status. Appendix B.1 shows this dependence has become stronger since 2010. This motivates our approach in the next section where we condition on a person's previous (and future) labor force status to impute missing observations. Our approach goes beyond conditioning on demographics, geography, and rotation group and finds that directly accounting for the labor force status of missing observations is important for measuring the labor market.

# 4 Correcting for the Bias

We now turn to correcting for the bias from nonresponse. Because the CPS is a panel, if a household responds at least once, we can infer information about their nonresponse from the month(s) they respond. Nevertheless, we need to make assumptions about the missing data and it is important to be explicit about these assumptions so readers can evaluate their plausibility. Our baseline correction estimates flow rates between labor force statuses and applies them to respondents' statuses in the month before or after a nonresponse.[12] This is our first major assumption: flow rates from responding and nonresonding households

[11] Atrostic et al. (2001) also points out the net number of CPS responders increases over month in sample.

[12] This is similar to logical imputation used in the Survey of Income and Program Participation (SIPP) for item nonresponse (not unit nonresponse) in that it makes an educated guess about a missing observation based on previous responses. See `https://www.census.gov/programs-surveys/sipp/methodology/data-editing-and-imputation.html`.

are identical. Our baseline correction does not account for all missing households, so as a secondary correction, we upweight individuals we infer information about to account for the rest of the sample. This is our second major assumption: nonresponding households with adjacent information have the same labor force characteristics as nonresponding households without adjacent information. We provide additional evidence for why these two assumptions are reasonable. In the last part of this section, we apply our correction methods and the BLS demographic weights so we can evaluate how rising nonresponse has biased the official statistics. All of our corrections suggest the bias has been growing since 2010.

## 4.1 Imputing Nonresponses with Adjacent Information

We measure aggregate flow rates between labor force statuses by focusing on individuals who respond for two consecutive months.[13] For this population, we calculate flow rates between three labor market statuses: employed ($E$), unemployed ($U$), and not in the labor force ($N$).[14] Let $z_i^s(t)$ represent the number of individuals who are in labor force status $i$ and MIS $s$ in month $t$ for $i = \{E, U, N\}$ and $s \in [1, 8]$. Conditioning on MIS is important because as we show in Appendix C, flow rates vary substantially by the MIS from which they are calculated. We then calculate two types of flow rates: forward flow rates and backward flow rates. Forward flow rates are the likelihood a respondent in labor force status $i$ at $t$ is in labor force status $j$ at $t+1$. Backward flow rates are the likelihood a respondent in labor force status $j$ at $t$ was in labor force status $i$ at $t-1$. Because flow rates vary over time, we calculate forward and backward flow rates for every combination of the three labor force statuses between 1998 and 2019. To preserve sample sizes, we do not condition on demographics, however Appendix H shows results are robust to further disaggregation.[15] To

[13]This group contains individuals from always responding households and individuals from sometimes responding households who have a nonresponse in a month other than the two in question.

[14]EE *flow* rates, for example, are distinct from employer-to-employer *transition* rates as in Fujita et al. (2020) because workers do not necessarily switch jobs.

[15]Appendix H shows results from Figure 3 do not depend on whether we condition on respondents being older or younger than 54, and separately on identifying as white or BIPOC. Conditioning on these demo-

further preserve sample sizes—and since we are interested in long-run trends of nonresponse bias—we calculate flow rates at an annual frequency based on the year of the first MIS.

Let $f_{ij}^s(t)$ be the forward flow rate between labor force status $i$ and $j$ at MIS $s$ and time $t$:

$$f_{ij}^s(t) = \frac{z_{ij}^s(t)}{z_i^s(t)}, \tag{1}$$

where $z_{ij}^s(t)$ is the number of individuals in labor force status $i$ and MIS $s$ at $t$ who move to labor force status $j$ and MIS $s+1$ at $t+1$. This forward flow rate is the share of individuals, for a given MIS, in labor force status $i$ who a month later are in labor force status $j$. Let $\bar{f}_{ij}^s(T)$ represent the average monthly forward flow rate for MIS $s$ calculated from individuals in calendar year $T$.[16]

Let $b_{ij}^s(t)$ be the backward flow rate between labor force status $j$ and $i$ at MIS $s$ and time $t$:

$$b_{ij}^s(t) = \frac{z_{ij}^{s-1}(t-1)}{z_j^s(t)}, \tag{2}$$

where $z_{ij}^{s-1}(t-1)$ is the number of individuals in labor force status $j$ and MIS $s$ at $t$ who came from labor force status $i$ and MIS $s-1$ at $t-1$. This backward flow rate is the share of individuals, for a given MIS, in labor force status $j$ who the month before were in labor force status $i$. Let $\bar{b}_{ij}^s(T)$ represent the average monthly backward flow rate for MIS $s$ calculated from individuals in calendar year $T$.[17]

By assuming flow rates of consecutive responders are the same as nonresponders, we can condition on the previous (and/or future) labor force status of nonresponders to impute their current labor force status.[18] Let $\mu_{kM\ell}^s(t)$ be a three-element row vector representing

graphic groupings results in a less than a 3 percentage point (or 10%) increase in the detected bias compared to the Flows Correction.

[16]Appendix C plots six annual forward flow rates averaged across all MIS for individuals in households that respond for two consecutive months.

[17]Appendix C plots six annual backward flow rates averaged across all MIS for individuals in households that respond for two consecutive months.

[18]We weaken this assumption in Appendix D by constructing flow rates from a sample of households that have at least one nonresponse, but where the nonresponse occurs outside the months in question. The idea

our correction probabilities for a Type A missing observation $M$ in MIS $s$ at month $t$ where statuses $k, \ell \in \{U, E, N, M\}$ are survey responses before and after the $M$ in question.

$$\mu^s_{kM\ell}(T) = \begin{cases} \left[\bar{f}^s_{kE}(T), \bar{f}^s_{kU}(T), \bar{f}^s_{kN}(T)\right] & \text{if } \ell = M, k \neq M \\ \left[\bar{b}^s_{E\ell}(T), \bar{b}^s_{U\ell}(T), \bar{b}^s_{N\ell}(T)\right] & \text{if } \ell \neq M, k = M \\ \left[\frac{1}{2}\left(\bar{f}^s_{kE}(T) + \bar{b}^s_{E\ell}(T)\right), \frac{1}{2}\left(\bar{f}^s_{kU}(T) + \bar{b}^s_{U\ell}(T)\right), \frac{1}{2}\left(\bar{f}^s_{kN}(T) + \bar{b}^s_{N\ell}(T)\right)\right] & \text{if } \ell \neq M, k \neq M \end{cases} \tag{3}$$

where, for example, $\bar{f}^s_{kE}(T)$ is the average forward monthly flow rate between $k \in \{E, U, N\}$ and $E$ with MIS $s \in \{2, 3, 4, 6, 7, 8\}$ in year $T$. Similarly, $\bar{b}^s_{E\ell}(T)$ is the average backward flow rate between $\ell \in \{E, U, N\}$ and $E$ with MIS $s \in \{1, 2, 3, 5, 6, 7\}$ in year $T$. Each missing observation is filled in with a three-element vector estimating the probability that the nonresponder is employed, unemployed, and not in the labor force.

The first component in Equation (3) uses forward flow rates and pertains to missing observations where the survey participant responded last month but failed to respond in the current month and failed to respond next month (or was out of the survey the next month), namely $\ell = M$. The second component in Equation (3) uses backward flow rates and pertains to missing observations where the survey participant failed to respond last month (or was not in the survey last month), namely $k = M$, and failed to respond in the current month but responded next month. The third component in Equation (3) uses both forward and backward flow rates and captures missing observations where the survey participant responds both last month $k \neq M$ and next month $\ell \neq M$, but not in the current month. Essentially, this missing observation is flanked by two non-missing, in-sample observations. To address the fact that we have two observations from which we can calculate the respondent's probabilistic labor force status, we use both sets of information by applying

is that people who sometimes respond are more similar to never responders than always responders. Under this alternative construction, Appendix D shows that our headline results hardly change, and if anything, the detected bias is larger.

forward and backward flow rates and averaging the results.[19]

## 4.2 Imputing Nonresponses without Adjacent Information

Equation (3) cannot account for all nonresponses. Missing observations with no response in an adjacent month are excluded. To address these missing observations, we apply sample weights such that respondents who satisfy one of the cases in Equation (3) are upweighted. Because these nonresponses *without* adjacent information account for up to two-thirds of nonresponses, correcting for them matters even more than the baseline correction.

This secondary procedure assumes households filled in by Equation (3) are identical (labor force wise) to households unaccounted for by Equation (3). This might be viewed as a strong assumption. To investigate its plausibility, we show in Appendix F that households that respond only sometimes—regardless of the number of times they nonrespond—are similar to each other, yet distinct from households that always respond. Therefore, nonresponders *with* adjacent information who are accounted for in Equation (3) are a good proxy for nonresponders *without* adjacent information who tend to live in households with many nonresponses.[20] This is analogous to hot-deck imputation where missing data is replaced with observations from similar units.

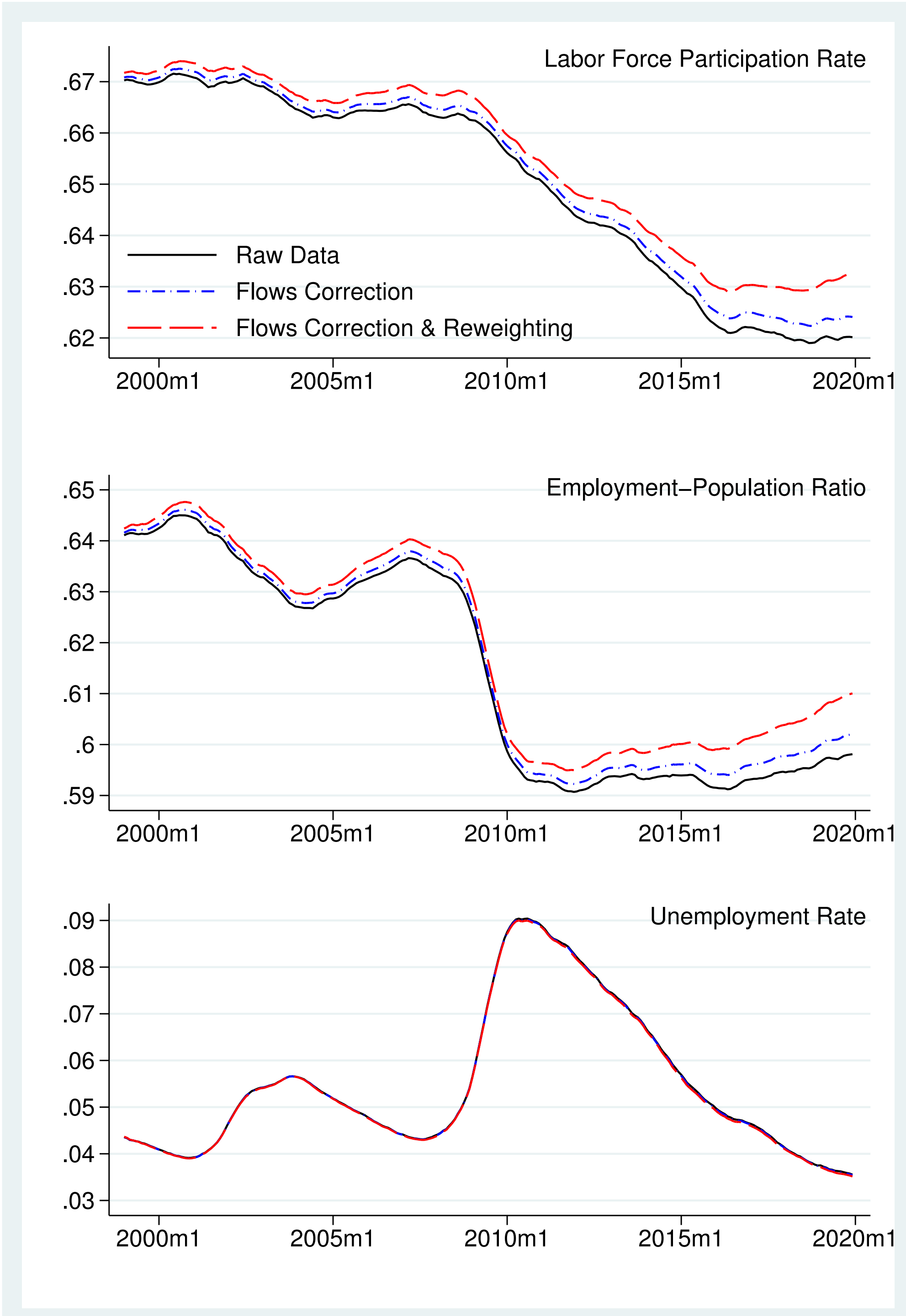
Figure 3 plots three estimates of three key labor market indicators from 1999 through 2019.[21] The black solid line represents rates calculated from the raw data without any adjustment. The blue dashed line imputes missing observations using Equation (3). The red dashed line further imputes missing observations by reweighting the sample to account for nonrespondents excluded from Equation (3). For the labor force participation rate and

[19] Appendix I shows that results are robust to using additional information for the third component of Equation (3) and filling in missing observations with flow rates computed from three consecutive responses.

[20] Appendix E shows that conditioning on the number of nonresponses a household registers throughout their survey life when implementing the reweighting procedure does not substantially change the results. If anything, it suggests the bias is larger.

[21] At this stage, none of the series in Figure 3 use the final composite weights provided by the BLS. We revisit this issue in Section 4.4.

Figure 3: Key Labor Market Indicators




**Notes:** Authors' calculations using data from the CPS for January 1999 through December 2019. All series are the 12-month historical moving average. The black solid line is the labor force participation rate calculated from the raw counts of respondents in the CPS. The blue dashed line adjusts for missing observations using our flows correction outlined in Equation (3). The red dashed line reweights our corrected series to account for missing observations that are excluded from Equation (3).

employment-population ratios, the dashed lines continually diverge from the solid line after 2010 which is exactly when nonresponse rates skyrocket. Notably, 20 percent of the decline in the labor force participation rate (as calculated from the raw counts of CPS respondents) since 2000 is from nonresponse bias. The unemployment rate looks different. All lines closely overlap, suggesting growing nonresponse has not discernably biased the unemployment rate. These findings line up with Section 3 where we document that most of the selective response behavior is between in and out of the labor force and not between employment and unemployment, and the latter is what the unemployment rate is calculated from.

## 4.3 Composition vs. Treatment

An important question relating to our correction approach is if a change in an individual's labor force status (i.e. "treatment") is itself a cause of nonresponse. We are unable to explore this question directly because it requires information of an individual exactly when they nonrespond, which, of course, we do not have. However, in this section, we discuss two examples of these "treatment" issues to highlight when our correction leads to an over or underestimate of the bias. We also highlight that the evidence we have suggests that our approach underestimates the bias.

**Example 1:**

Suppose a survey participant named Juan records the following statuses for the first three months: (1) employed, (2) nonresponse, (3) employed. While our approach for imputing Juan's missing observation for MIS 2 is relatively complicated, as a first approximation, let us assume we impute his status as employed. This imputation would increase the employment-population ratio for MIS 2 and—as we generally find in the paper—our "corrected" employment-population ratio would be larger than that calculated from the raw data. Suppose, though, that Juan's true (unobservable) labor force status in MIS 2 was NILF, and it was factors related to NILF that caused him to nonrespond. If we set Juan's missing observation to NILF instead of employed, this would mean the true employment-population ratio is lower than our "corrected" series. In this example, our correction approach would overestimate the employment-population ratio. For this to be the case, the likelihood someone responds to the survey must be negatively correlated with NILF. The evidence we have, however, suggests the opposite is true. In Table 1, individuals who reported NILF in MIS $t$ are *more* likely to respond to the survey in MIS $t+1$.

**Example 2:**

Suppose a survey participant named Jessica records the following statuses for the first three months: (1) NILF, (2) nonresponse, (3) NILF. As a first approximation let us assume we impute her status as NILF. This imputation would decrease the employment-population ratio for MIS 2 and our "corrected" employment-population ratio would be smaller than that calculated from the raw data. Suppose, though, that Jessica's true (unobservable) labor force status in MIS 2 was employed, and that factors related to employment caused her to nonrespond. If we set Jessica's missing observation to employed instead of NILF, this would mean the true employment-population ratio is higher than our "corrected" series. In this example, our correction approach would underestimate the employment-population ratio. For this to be the case, the likelihood someone responds to the survey must be negatively correlated with employment, which aligns with what we see in the data. In Table 1, individuals who are employed in MIS $t$ are *less* likely to respond to the survey in MIS $t+1$.

More convincing perhaps, in Appendix J we leverage *retrospective* questions in the CPS when someone responds to glean information about their labor force status in months they do not respond. Our results from this exercise provide evidence that nonresponse is simultaneously correlated with employment. There is also an intuitive reason why employment would cause nonresponse: working individuals have less time to answer the survey.[22]

[22]The average interview takes 10 to 15 minutes and depends on the number of household members.

Overall, the evidence we have—from information surrounding a nonresponse and from retrospective labor force questions—suggest employment causes households to nonrespond. Thus, our imputed labor force participation rate and employment-population ratio likely understate the true upward adjustment needed to fully correct the official statistics.

### 4.4 BLS Demographic Weights

An important part of the CPS are the weights provided by the BLS to ensure a representative sample. BLS weights target regional demographic information from the U.S. Census.[23] When aggregating individuals in the CPS to produce the official headline statistics, the BLS applies these weights to respondents. The analysis so far does not use BLS weights and instead uses the raw counts of individuals. This is common in the literature on nonresponse (e.g. Korinek et al. 2007) because applying BLS weights to imputed data can lead to a double correcting of nonresponders. For example, suppose we impute the labor force status of an individual who was missing in a given month using our Flows Correction. It is possible that the BLS weights for the responding population are, at least in part, also derived to account for those who are missing. If nonresponse behavior is strongly correlated with demographic characteristics, individuals who respond to the CPS but who also belong to a demographic group with a high nonresponse rate would be upweighted to match the BLS's demographic targets. If we impute missing individuals from this demographic group using our approach, but also apply the BLS weights to responders, we would end up overcorrecting for the nonresponders.[24] In this section, we apply BLS weights to our imputed data in a novel way and find they correct for about half of the rise in nonresponse bias.

[23] The BLS final composite weight corrects the data in two ways. First, it targets the population within a demographic-region cell. Second, it adjusts for rotation group bias within each cell. Because we aggregate the BLS weights to demographic-region cells—and at this level of aggregation, the weights are orthogonal to the rotation group correction—we do not apply the second correction. See U.S. Census Bureau (2019) for more details.

[24] Although not the focus of this paper, readers may be interested in how demographic characteristics of sometimes-responding households compare to that of always-responding households in order to think about the drivers of nonresponse. Appendix G shows that nonresponse is correlated with younger people and people of color.

To begin, we need the demographic population shares the BLS targets in their weighting approach. Using the demographic and region categories (age, sex, race, ethnicity, and state) described in the technical report by the U.S. Census Bureau (2019), we backward engineer the BLS targets. To save on computation, our demographic categories are broader than what the BLS targets, but consist of over 1,600 demographic cells, and when applied to the raw data generate a remarkably good match with the headline labor market statistics.[25]

We demographically adjust the raw data as follows. Using the BLS weights, we construct the share of the (BLS weighted) population in each demographic cell. Formally:

$$PopShare_{k,t}^{BLS} = \frac{p_{k,t}^{BLS}}{\sum_k p_{k,t}^{BLS}}, \tag{4}$$

where $p_{k,t}^{BLS}$ is the sum of the BLS weights for the population in demographic-region cell $k$ in month $t$ and $\sum_k p_{k,t}^{BLS}$ is the total (weighted) population across cells in month $t$.[26] Because these population shares are computed using the BLS weights, they should reflect the BLS's desired share of the population in each demographic category. We use these shares as our targets for the demographic shares of the population.[27]

We now apply this demographic weighting approach to our corrected data after we have imputed the labor force statuses of the nonresponders. Again, there are two key steps in our adjustment for nonresponders. The first is the Flows Correction where we fill in nonresponders who have a response in the preceding or subsequent month. For the demographic adjustment we also fill in these individuals' demographic and geographic information from the preceding and subsequent month. The second key step is the Reweighting Correction where we upweight these imputed nonresponders to account for the additional nonresponders who do not have responses in the preceding and subsequent month. Here, we assume

[25]See Appendix K.

[26]The demographic-region cells include 51 states including the District of Columbia; 4 age groupings, [16,30),[30,50),[50,70),[70,$\infty$); white and non-white; Hispanic and non-Hispanic; male and female.

[27]Appendix K compares the employment-population ratio constructed from our reverse-engineered BLS weights to the official series, and the lines are almost indistinguishable.

the demographics of households with adjacent information are similar to households without adjacent information. To demographically adjust this data, we now compute the share of the corrected population (responders and imputed nonresponders) in each demographic-region cell as:

$$PopShare_{k,t}^{Corrected} = \frac{p_{k,t}^{Corrected}}{\sum_k p_{k,t}^{Corrected}}, \tag{5}$$

where $p_{k,t}^{Corrected}$ is the sum of the population of responders and nonresponders for our corrected data who are in cell $k$ in month $t$ and $\sum_k p_{k,t}^{Corrected}$ is the total population in our corrected data in month $t$. We then adjust these demographic cells by:

$$AdjustmentFactor_{k,t}^{Corrected} = PopShare_{k,t}^{BLS} / PopShare_{k,t}^{Corrected}. \tag{6}$$

Another way to think about how much the BLS weights control for the nonresponse bias we correct for in the raw data is to look at the magnitude of $AdjustmentFactor_{k,t}^{Corrected}$. If the nonresponse behavior we correct for was perfectly described by differential response behavior across demographic cells, $AdjustmentFactor_{k,t}^{Corrected}$ would equal one for all $k$ cells, and the BLS weights would fully account for nonresponse bias. This is not the case.

Figure 4 illustrates how the BLS weights matter by plotting correction gaps for three labor market indicators. The dashed black line in each panel is the difference between our Flows Correction & Reweighting series and a key indicator derived from raw CPS counts. This line represents the magnitude of the correction from the raw data. For the labor force participation rate, the gap increased 104 basis points between 1999 and 2020. The solid blue line is computed analogously but where both the Flows Correction & Reweighting series and the raw count data have been demographically adjusted. This line illustrates the magnitude of the Flows Correction & Reweighting adjustment once demographics are appropriately controlled for. For the labor force participation rate, the gap has increased by 41 basis points between 1999 and 2020. In other words, the BLS weights manage to address 60 percent of the increase in nonresponse bias. The correction gaps for the employment-population tell a similar story, while the correction gaps for the unemployment rate hover around zero.

## 4.5 Discussion

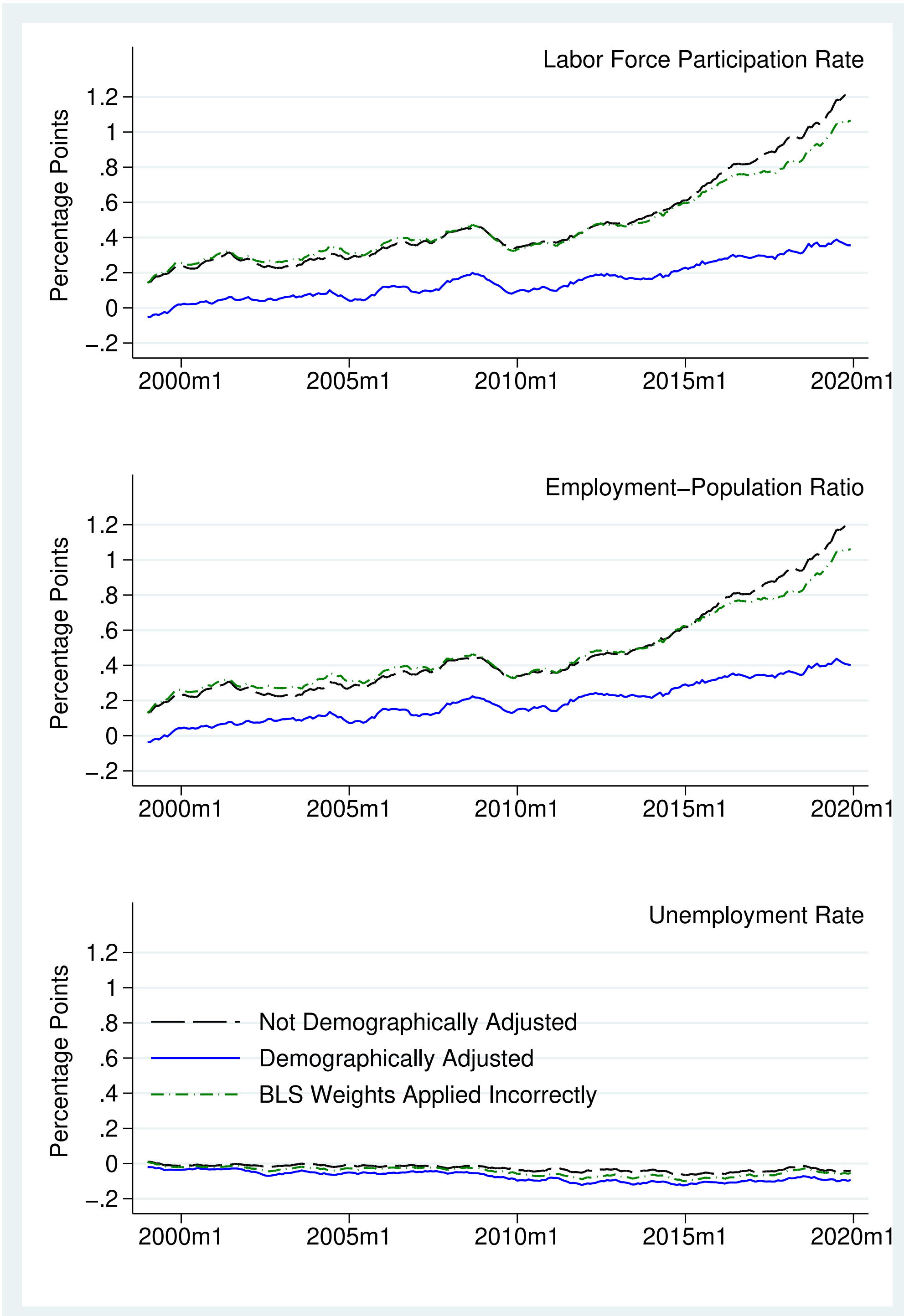
While the demographic adjustments provided by the BLS improve things, they do not eliminate rising bias for key labor market indicators. Our results align with the finding in Borgschulte et al. (2022) that about half of the recent rise in CPS refusal rates can be accounted for by demographic controls. We find nonresponse bias still accounts for a 40 basis point underestimate of the labor force participation rate and employment-population ratio, which is outside the BLS confidence bands.[28] This is approximately 10 percent of the reported decline in the official labor force participation rate since the turn of the millennium. On its own, 10 percent is sizable, but this finding is even more consequential considering that nearly half of the decline in the participation rate is from compositional changes in the population (Aaronson et al., 2012). This leaves only two percentage points of the decline to be accounted for by all other factors. The bias we document possibly cuts into the importance of other factors, such as skilled-biased technological change, behind falling participation (Abraham and Kearney, 2020; Wolcott, 2021).

These results also highlight a subtle yet important point related to work examining nonresponse in the CPS. Simply applying BLS weights to responders and nonresponders who have already been accounted for by other means (as in Ahn and Hamilton, 2022) can lead to overcorrecting. Refusals are only one of the many measurement issues Ahn and Hamilton (2022) impressively correct for in the CPS, and relative to misclassification issues, they play muted role. Since our focus is on missing observations from refusals, it is especially important to consider the interaction between imputed data and the BLS weights.

To highlight this point, we recompute the participation rate using our Flows & Reweight-

[28] In July 2023, 90 percent confidence bands were 23 basis points.

Figure 4: Correction Gaps




**Notes:** Authors' calculations using data from the CPS for January 1999 through December 2019. All series are the 12-month historical moving average.

ing corrected data from Section 4.2 but weight individuals by their BLS weight during aggregation. We then plot the gap between the BLS weighted key indicator of the corrected data and the official BLS statistic as a third line in Figure 4, denoted "BLS Weights Applied Incorrectly." This gap computes an erroneous bias because it applies the BLS weights as-is to the responding population even after nonresponding individuals have been filled in. The result is a gap that is substantially larger than our demographically adjusted Flows & Reweighting correction. Therefore, failing to adjust weights after missing individuals have been imputed risks overcorrecting for the missing population and overestimating the magnitude of nonresponse bias. In sum, an important contribution of the paper is highlighting the care that needs to be given when applying BLS weights to already corrected data. Further, we provide a methodology to address both nonresponse bias and demographic adjustment while avoiding overcorrecting.

# 5 Conclusion

How does the dramatic rise of nonresponse since 2010 impact labor market indicators? Rising nonresponse in the CPS has artificially suppressed the labor force participation rate and employment-population ratio but has had little discernible impact on the unemployment rate. We document that the rise in nonresponse is driven by households refusing to participate in the survey. We leverage the panel structure of the CPS to record the labor force status of nonresponding households in the months surrounding their nonresponse and use aggregate flow rates to impute missing observations. We offer a correction method for indicators calculated by both the raw counts of respondents in the CPS and respondents that have been demographically adjusted by the BLS. All methods point to the problem becoming worse since 2010. Although the BLS weights correct for some of the bias, they do not correct for all of it. Nonresponse appears to be a growing source of bias for key labor market indicators derived from the CPS.

## References


Aaronson, D., J. Davis, L. Hu, et al. (2012): "Explaining the decline in the US labor force participation rate," *Chicago Fed Letter*.

Abowd, J. M. and A. Zellner (1985): "Estimating Gross Labor-Force Flows," *Journal of Business & Economic Statistics*, 3, 254–283.

Abraham, K. G. and M. S. Kearney (2020): "Explaining the decline in the US employment-to-population ratio: A review of the evidence," *Journal of Economic Literature*, 58, 585–643.

Ahn, H. J. and J. D. Hamilton (2022): "Measuring labor-force participation and the incidence and duration of unemployment," *Review of Economic Dynamics*, 44, 1–32.

Atrostic, B. K., N. Bates, G. Burt, and A. Silberstein (2001): "Nonresponse in US Government Household Surveys: Consistent Measures, Recent Trends, and New Insights," *Journal of Official Statistics*, 17, 209–226.

Bailar, B. A. (1975): "The Effects of Rotation Group Bias on Estimates from Panel Surveys," *Journal of the American Statistical Association*, 70, 23–30.

Bee, C. A., G. M. Gathright, and B. D. Meyer (2015): "Bias from Unit Non-Response in the Measurement of Income in Household Surveys," *Unpublished Paper, August*.

Bick, A. and A. Blandin (2023): "Employer reallocation during the COVID-19 pandemic: Validation and application of a do-it-yourself CPS," *Review of Economic Dynamics*, 49, 58–76.

BLS (2020): "Frequently Asked Questions: The impact of the coronavirus (COVID-19) Pandemic on The Employment Situation for April 2020," *https://www.bls.gov/cps/employment-situation-covid19-faq-april-2020.pdf*.

Borgschulte, M., H. Cho, and D. Lubotsky (2022): "Partisanship and survey refusal," *Journal of Economic Behavior & Organization*, 200, 332–357.

Brick, J. M. and D. Williams (2013): "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys," *The ANNALS of the American academy of political and social science*, 645, 36–59.

Chua, T. C. and W. A. Fuller (1987): "A model for multinomial response error applied to labor flows," *Journal of the American Statistical Association*, 82, 46–51.

Cohany, S. R., A. E. Polivka, and J. M. Rothgeb (1994): "Revisions in the current population survey effective January 1994," *Emp. & Earnings*, 41, 13.

Dutz, D., I. Huitfeldt, S. Lacouture, M. Mogstad, A. Torgovitsky, and W. van Dijk (2021): "Selection in Surveys," *National Bureau of Economic Research*.

Elsby, M. W., B. Hobijn, and A. Şahin (2015): "On the importance of the participation margin for labor market fluctuations," *Journal of Monetary Economics*, 72, 64–82.

Faberman, R. J., A. I. Mueller, and A. Şahin (2022): "Has the willingness to work fallen during the covid pandemic?" *Labour Economics*, 79, 102275.

Foote, C. L., T. Hounshell, W. D. Nordhaus, D. Rivers, and P. Torola (2021): "Measuring the us employment situation using online panels: The Yale labor survey," *Journal of Economic and Social Measurement*, Preprint, 1–22.

Fujita, S., G. Moscarini, and F. Postel-Vinay (2020): "Measuring Employer-to-Employer Reallocation," *National Bureau of Economic Research*.

Fujita, S. and G. Ramey (2006): "The Cyclicality of Job Loss and Hiring," *Federal Reserve Bank of Philadelphia Working Paper*.

Halpern-Manners, A. and J. R. Warren (2012): “Panel Conditioning in Longitudinal Studies: Evidence from Labor Force Items in the Current Population Survey,” *Demography*, 49, 1499–1519.

Harris-Kojetin, B. and C. Tucker (1999): “Exploring the Relation of Economic and Political Conditions with Refusal Rates to a Government Survey,” *Journal of Official Statistics*, 15, 167.

Heffetz, O. and D. B. Reeves (2019): “Difficulty of Reaching Respondents and Nonresponse Bias: Evidence from Large Government Surveys,” *Review of Economics and Statistics*, 101, 176–191.

Hokayem, C., C. Bollinger, and J. P. Ziliak (2015): “The role of CPS nonresponse in the measurement of poverty,” *Journal of the American Statistical Association*, 110, 935–945.

Korinek, A., J. A. Mistiaen, and M. Ravallion (2007): “An Econometric Method of Correcting for Unit Nonresponse Bias in Surveys,” *Journal of Econometrics*, 136, 213–235.

Krueger, A. B., A. Mas, and X. Niu (2017): “The Evolution of Rotation Group Bias: Will the Real Unemployment Rate Please Stand Up?” *Review of Economics and Statistics*, 99, 258–264.

Kudlyak, M. and F. Lange (2018): “Measuring Heterogeneity in Job Finding Rates among the Non-Employed Using Labor Force Status Histories,” *Federal Reserve Bank of San Francisco Working Paper 2017-20.*

McCarthy, P. J. (1978): *Some Sources of Error in Labor Force Estimates from the Current Population Survey*, 15, National Commission on Employment and Unemployment Statistics.

Meyer, B. D., W. K. Mok, and J. X. Sullivan (2015): "Household Surveys in Crisis," *Journal of Economic Perspectives*, 29, 199–226.

Nekarda, C. J. (2009): "A Longitudinal Analysis of the Current Population Survey: Assessing the Cyclical Bias of Geographic Mobility," *Federal Reserve Board of Governors*.

Poterba, J. M. and L. H. Summers (1986): "Reporting Errors and Labor Market Dynamics," *Econometrica: Journal of the Econometric Society*, 1319–1338.

Rothbaum, J. and A. Bee (2021): "Coronavirus infects surveys, too: Survey nonresponse bias and the Coronavirus pandemic," *US Census Bureau*.

Schoeni, R. F., F. Stafford, K. A. McGonagle, and P. Andreski (2013): "Response Rates in National Panel Surveys," *The Annals of the American Academy of Political and Social Science*, 645, 60–87.

Solon, G. (1986): "Effects of Rotation Group Bias on Estimation of Unemployment," *Journal of Business & Economic Statistics*, 4, 105–109.

Tucker, C. and B. A. Harris-Kojetin (1998): "The Impact of Nonresponse on the Unemployment Rate in the Current Population Survey (CPS)," .

U.S. Census Bureau (2019): "Current Population Survey Design and Methodology Technical Paper 77," .

Vom Lehn, C., C. Ellsworth, and Z. Kroff (2021): "Reconciling Occupational Mobility in the Current Population Survey," *Journal of Labor Economics*.

Williams, D. and J. M. Brick (2018): "Trends in US face-to-face household survey nonresponse and level of effort," *Journal of Survey Statistics and Methodology*, 6, 186–211.

Wolcott, E. L. (2021): "Employment inequality: Why do the low-skilled work less now?" *Journal of Monetary Economics*, 118, 161–177.

# Online Appendix

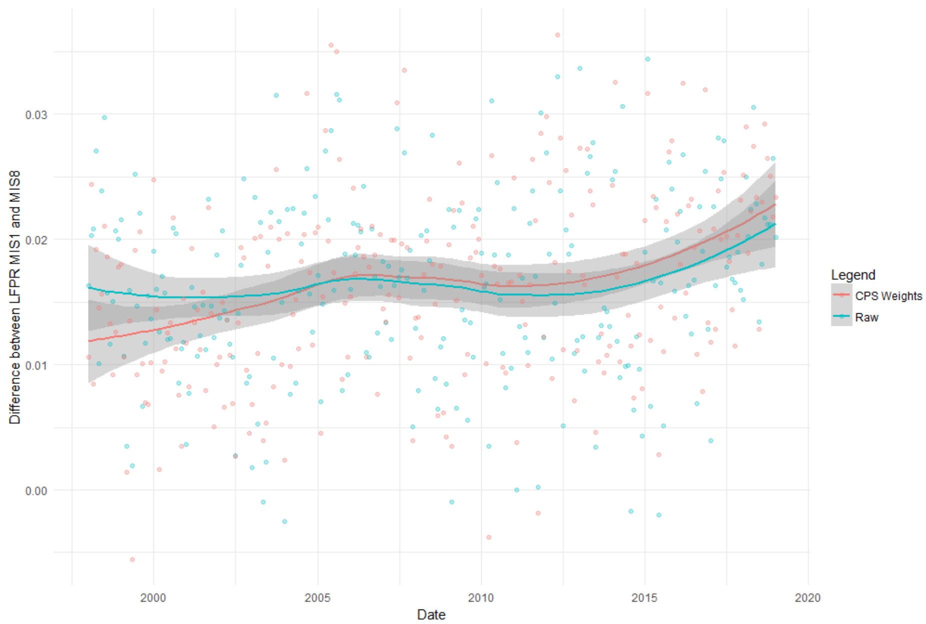
## How Does the Dramatic Rise of CPS Nonresponse Impact Labor Market Indicators?

Robert Bernhardt David Munro Erin L. Wolcott

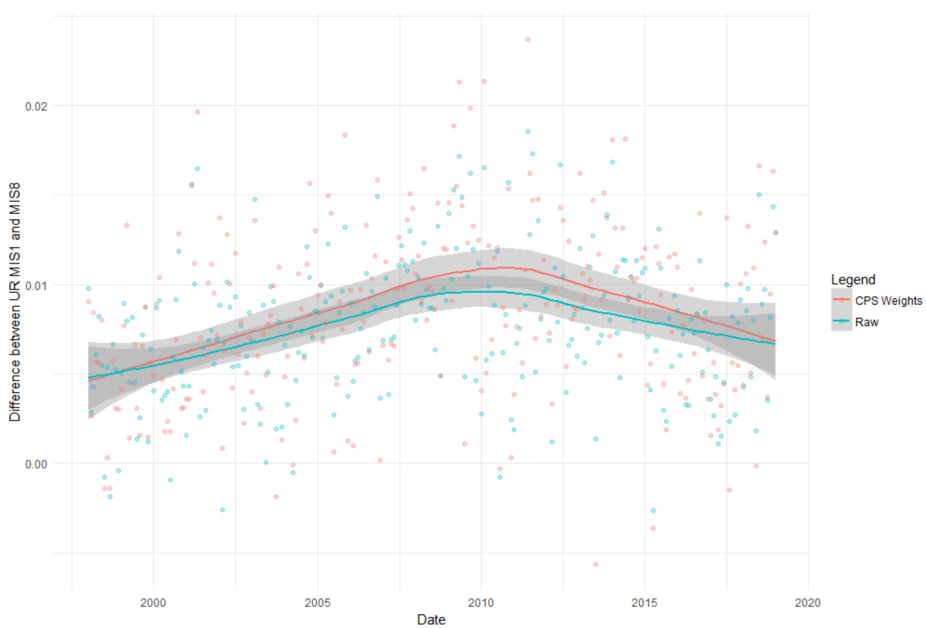
September 2023

# A Trends in Rotation Group Bias

Difference Between MIS 1 and MIS 8 Participation Rate




Difference Between MIS 1 and MIS 8 Unemployment Rate




**Notes:** Authors' calculations using data from the CPS. Difference in labor market indicators by the month-in-sample (MIS) the respondent was the in survey. The CPS weighted series applies the Composited Final Weight used to create BLS published labor force statistics. The raw series is unweighted.

# B Evidence of Selective Response Behavior

## B.1 Pre- and Post-2010

Table 1: Pre-2010 Selective Response Behavior

Panel A: Drop-outs

| MIS $t$ Labor Force Status | MIS $t+1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 63.07% (5,577,024) | 68.50% (117,705) | -5.43 pp |
| Unemployed | 3.04% (268,782) | 3.58% (6,149) | -0.54 pp |
| NILF | 33.89% (2,997,214) | 27.93% (47,988) | 5.96 pp |

Panel B: Drop-ins

| MIS $t$ Labor Force Status | MIS $t-1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 62.93% (5,565,327) | 67.63% (158,464) | -4.70 pp |
| Unemployed | 2.98% (263,740) | 3.24% (7,593) | -0.26 pp |
| NILF | 34.09% (3,015,163) | 29.13% (68,247) | 4.96 pp |

**Notes:** Authors' calculations from linking households and individuals across month-in-samples (MIS) in the CPS, where for Panel A, $t \in \{1, 2, 3, 5, 6, 7\}$ and for Panel B, $t \in \{2, 3, 4, 6, 7, 8\}$. Data is aggregated over 1998-2019. Counts are in parenthesis. Each count is a person who lives in a household, where the household either responds or nonresponds. Nonresponses are all Type A. Panel A is the share and count of individuals with said labor force status in MIS $t$ and interview status in MIS $t+1$. Panel B is the share and count individuals with said labor force status in MIS $t$ and interview status in MIS $t-1$. The Response and Nonresponse columns add to 100%. The last column is the percentage point difference between the share of individuals who respond and nonrespond.

Table 2: Post-2010 Selective Response Behavior

Panel A: Drop-outs

| MIS $t$ Labor Force Status | MIS $t+1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 59.32%<br>(4,113,909) | 66.44%<br>(150,703) | -7.12 pp |
| Unemployed | 3.26%<br>(225,808) | 3.65%<br>(8,282) | -0.39 pp |
| NILF | 37.42%<br>(2,595,390) | 29.90%<br>(67,820) | 7.52 pp |

Panel B: Drop-ins

| MIS $t$ Labor Force Status | MIS $t-1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 59.15%<br>(4,102,687) | 65.78%<br>(180,567) | -6.63 pp |
| Unemployed | 3.17%<br>(219,887) | 3.29%<br>(9,026) | -0.12 pp |
| NILF | 37.68%<br>(2,613,301) | 30.93%<br>(84,907) | 6.75 pp |

**Notes:** Authors' calculations from linking households and individuals across month-in-samples (MIS) in the CPS, where for Panel A, $t \in \{1, 2, 3, 5, 6, 7\}$ and for Panel B, $t \in \{2, 3, 4, 6, 7, 8\}$. Data is aggregated over 1998-2019. Counts are in parenthesis. Each count is a person who lives in a household, where the household either responds or nonresponds. Nonresponses are all Type A. Panel A is the share and count of individuals with said labor force status in MIS $t$ and interview status in MIS $t + 1$. Panel B is the share and count individuals with said labor force status in MIS $t$ and interview status in MIS $t - 1$. The Response and Nonresponse columns add to 100%. The last column is the percentage point difference between the share of individuals who respond and nonrespond.

## B.2 Impact of BLS Weights

Table 3: Selective Response Behavior with BLS weights

Panel A: Drop-outs

| MIS $t$ Labor Force Status | MIS $t+1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 61.09% | 66.95% | -5.86 pp |
| Unemployed | 3.33% | 3.78% | -0.45 pp |
| NILF | 35.57% | 29.28% | 6.29 pp |

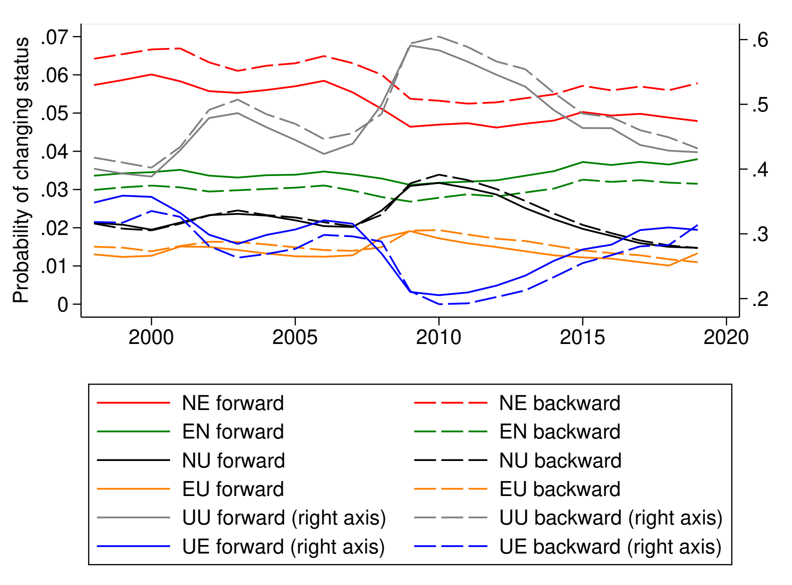
Panel B: Drop-ins

| MIS $t$ Labor Force Status | MIS $t-1$ Interview Status | | |
|---|---|---|---|
| | Response | Nonresponse | Difference |
| Employed | 60.95% | 66.26% | -5.31 pp |
| Unemployed | 3.26% | 3.41% | -0.15 pp |
| NILF | 35.80% | 30.33% | 5.47 pp |

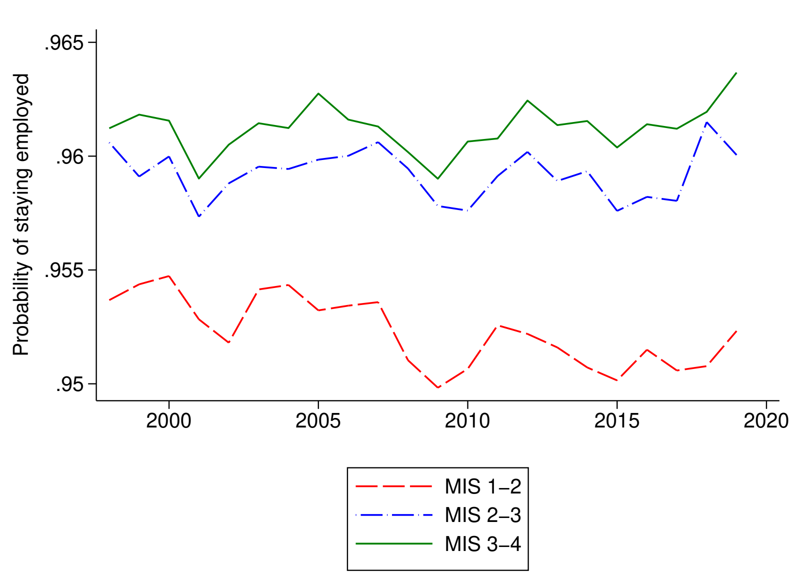
**Notes:** Authors' calculations from linking households and individuals across month-in-samples (MIS) in the CPS, where for Panel A, $t \in \{1, 2, 3, 5, 6, 7\}$ and for Panel B, $t \in \{2, 3, 4, 6, 7, 8\}$. Data is aggregated over 1998-2019. Numbers are calculated using BLS weights. Nonresponses are all Type A. Panel A is the share of individuals with said labor force status in MIS $t$ and interview status in MIS $t+1$. Panel B is the share and count individuals with said labor force status in MIS $t$ and interview status in MIS $t-1$. The Response and Nonresponse columns add to 100%. The last column is the percentage point difference between the share of individuals who respond and nonrespond.

# C Labor Force Status Flow Rates

Subset of Flow Rates Averaged Over MIS




Employment-to-Employment Flow Rates by MIS



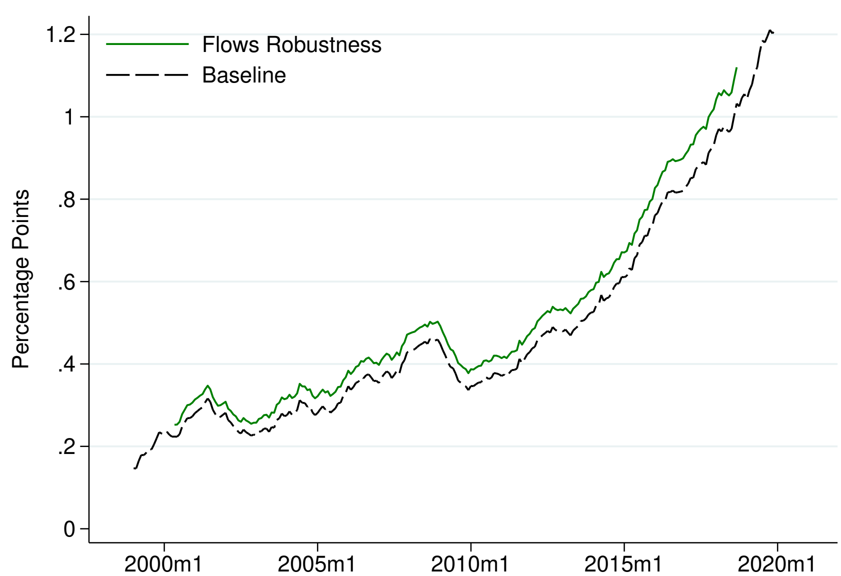

**Notes:** Authors' calculations using data from the CPS. Monthly flow rates are from individuals 16 years and older who move between unemployment (U), employment (E), and not in the labor force (N) from the full sample of consecutively responding households. Top panel plots a subset of forward and backwards flow rates. Bottom panel plots EE flow rates for three month-in-sample (MIS) pairs.

# D Robustness to Flows Assumption

In the corrections from Figure 3, flow rates are calculated from all individuals who register a response in the two months in question. This includes always responders and sometimes responders whose nonresponse(s) occurred in another month. To test robustness, we calculate flow rates using *only* households that have at least one nonresponse, but where the nonresponse occurs in months other than the ones used to compute the flow rates. We view this as a weaker assumption because we are using flow rates only calculated from sometimes responders to fill in missing observations for other sometimes responders.

With this weakened assumption to compute flow rates, we repeat the same adjustment procedure outlined in Sections 4.1 and 4.2. The figure below shows that our results barely change. The black dashed line plots the difference between the Flows Correction & Reweighting series and the Raw series from Figure 3. The solid green line plots this difference, but where the Flows Correction & Reweighting series only uses flow rates from sometimes-responding households. The Flows Robustness series is truncated at the beginning and end because it requires information on households linked over 16 months of the full survey duration. Notably, both lines start to exponentially increase in 2010. The Flows Robustness series reports an even higher prevalence of bias than the baseline specification. These results highlight that flow rates do not appear to be systematically different when computed from the restricted sample and that, if anything, doing so leads to an even larger computed bias.

Gap Between Corrected and Uncorrected Labor Force Participation Rate




**Notes:** Authors' calculations using data from the CPS. All series are 12-month historical moving averages. The Baseline is the Flows Correction & Reweighting.

# E Robustness to Reweighting Assumption

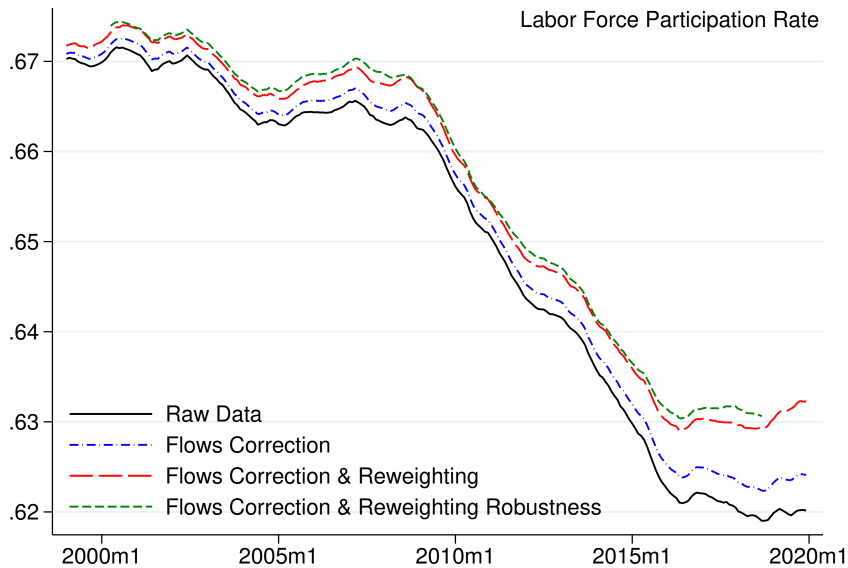
We assume nonresponders with adjacent information are the same, labor force wise, as-nonresponders without adjacent information. To test this assumption, we explore whether the number of times a household nonresponds matters for labor market indicators. The idea here is that nonresponders *without* adjacent information tend to come from households with a higher number of nonresponses. How labor market indicators vary by the number of household nonresponses indicates whether nonresponders with adjacent information (i.e. low nonresponse households, on average) are similar to nonresponders without adjacent information (i.e. high nonresponse households, on average).

Appendix F plots the average probability a respondent is employed (or NILF or unemployed) based on the number of times their household nonresponds. Notably, the employment rate of people in an always responding household (i.e. zero nonresponses) is much lower, at 60 percent, than the employment rate of people in a sometimes-responding household. Moreover, the employment rate of sometimes responders—whether a one-time or seven-time nonresponder—are similar ranging from 65 to 67.5 percent. In summary, the labor force characteristics of sometimes responders do not vary much regardless of the number of times they nonrespond, but are distinct from always responders. This suggests that assuming homogeneity among sometimes responders is reasonable. It is also worth noting that our reweighting assumption may be a conservative one. This can be seen from Appendix F. The probability of being employed for a person in a household with few nonresponses (e.g. 1-3 nonresponses) is lower than for a person in a household with many nonresponses (e.g. 4-7 nonresponses). Since most households without adjacent information have many nonresponses, using the sample average among sometimes responders from Appendix F to upweight these individuals would lead to lower employment rates than using the probability of, say, five-time nonresponding households.[1]

As an alternative reweighting assumption, the figure on the next page applies the labor force characteristics for nonresponders without adjacent information conditional on their total number of nonresponses. The idea is that we use information in Appendix F to fill in labor force characteristics for nonresponders without adjacent information. This alternative assumption allows for a richer relationship between labor force status and nonresponse behavior because it does not assume that individuals with and without adjacent information have identical labor force characteristics regardless of the number of times their household nonresponds. Consistent with the intuition outlined above, the figure below highlights that, if anything, our baseline reweighting assumption (red long dashed line) is slightly conservative because our reweighting assumption with richer information (green short dashed line) estimates a slightly larger bias for the labor force participation rate. We use our baseline adjustment in the paper because it is more conservative, is generated from a larger sample of households, and does not face truncation issues.

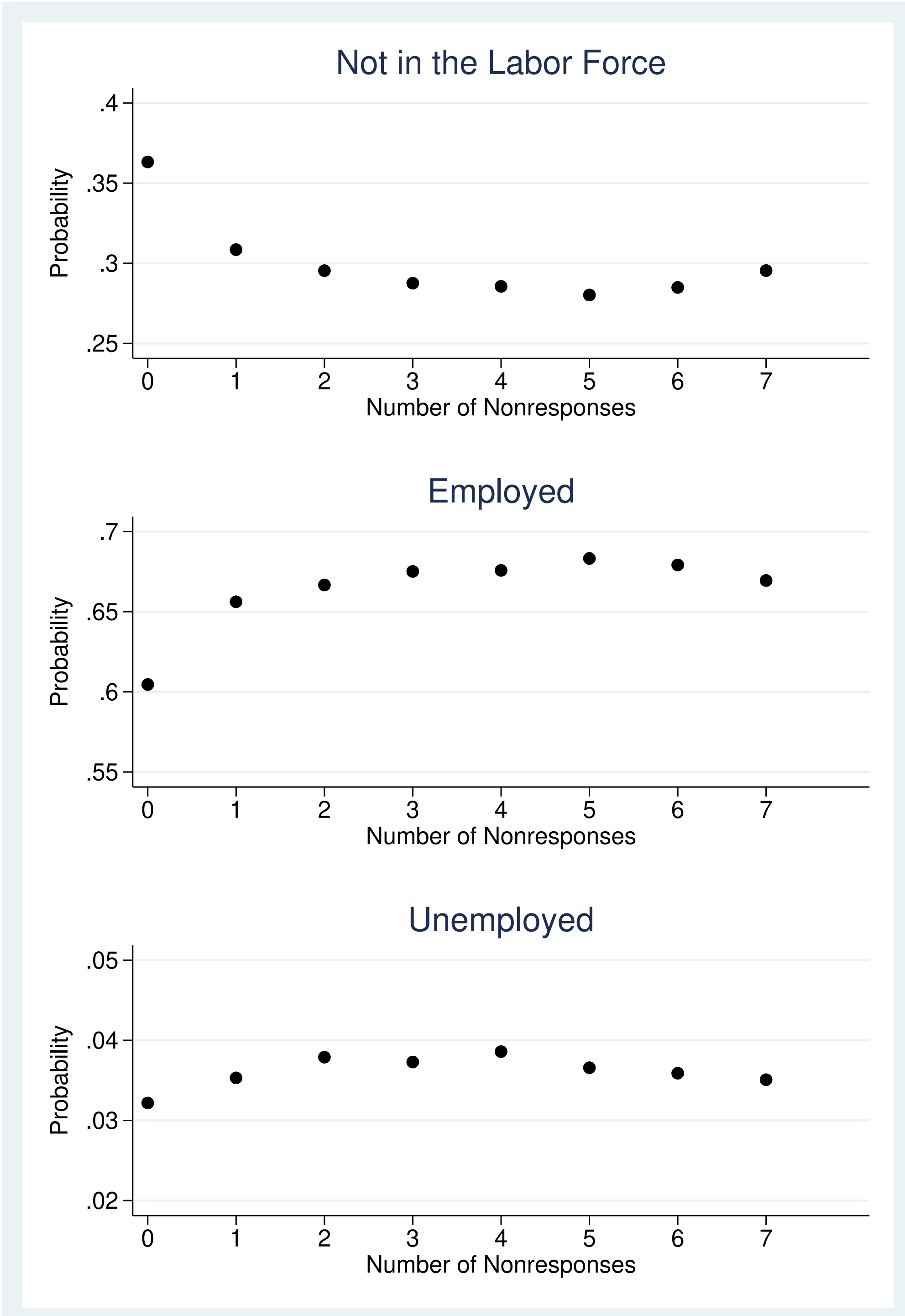
[1]Indeed, over the inner sample, the average number of nonresponses out of 8 for individuals with adjacent information is 2.3 and for individuals without adjacent information is 5.9.

Robustness Check for Reweighting Assumption




**Notes:** Authors' calculations using data from the CPS. The first three series run from January 1999 through December 2019. The Flows Correction & Reweighting Robustness series is truncated, running from April 2000 through September 2018, because respondents need to be observed for the span of the 16 month survey. All series are 12-month historical moving averages.

## F Labor Market Status by Number of Nonresponses




**Notes:** Authors' calculations using data from the CPS averaged over all months from 1998 through 2019. Probabilities are for individuals 16 years and older. Nonresponses are Type A unit non-interviews.

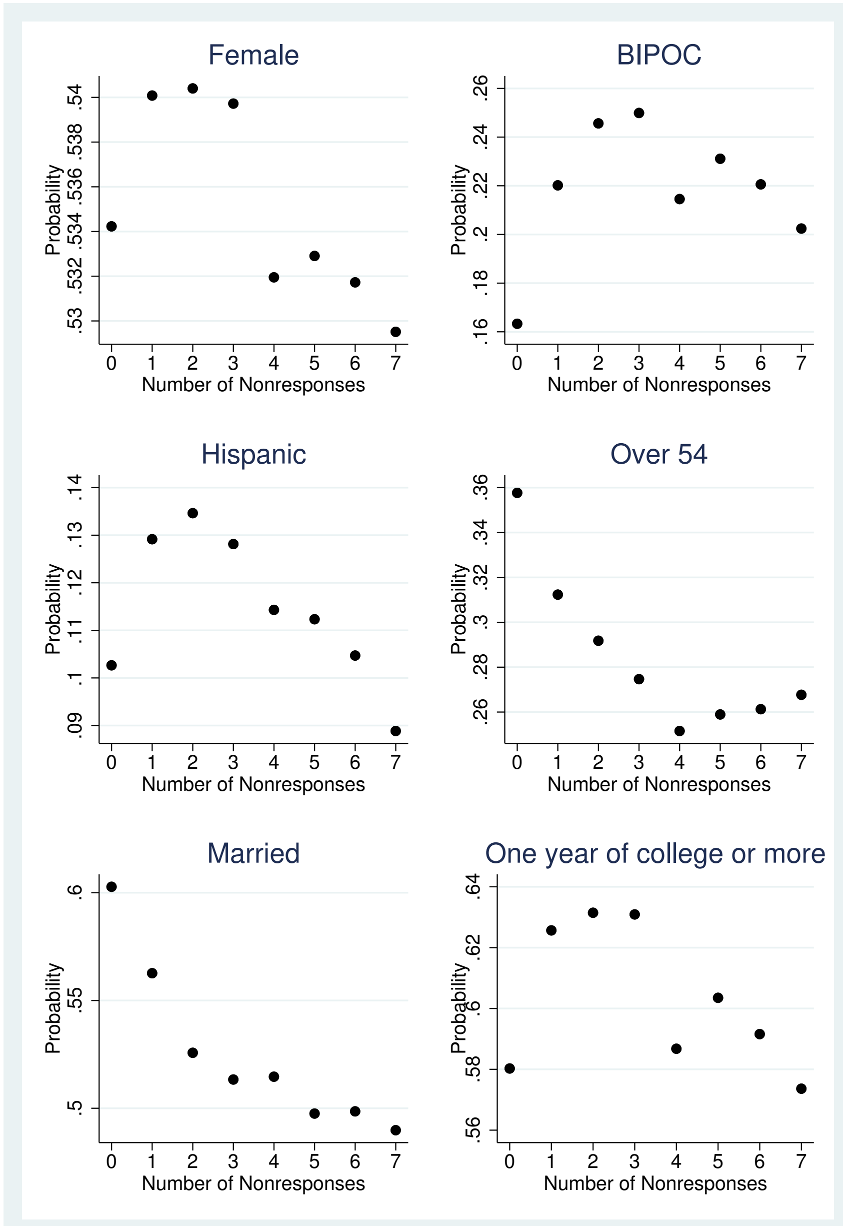
# G Demographics by Number of Nonresponses

Readers may be interested in knowing the characteristics of responders who respond to more months of the survey. Below is the results of a linear probability model where the left-hand side is an indicator equaling one if the respondent misses at least one month (i.e. sometimes responders) and equaling zero if the respondent answers all survey months. Women, respondents identifying as a race other than white, Hispanics, respondents under 55, the non-married, and respondents with some college are all more likely to be sometimes responders. We split the sample into two periods covering before and after the dramatic rise in nonresponse. This regression sheds light on the demographic drivers of nonresponse, but not on whether nonresponse biases the headline labor force statistics. The second question is the focus of this paper and demographics come into play in the second question by making sure the sample is nationally representative. See Section 4.4.

| VARIABLES | (1)<br>Sometimes Responder<br>1998-2009 | (2)<br>Sometimes Responder<br>2010-2019 |
|---|---|---|
| Female | 0.00169*** | 0.00664*** |
| | (0.000655) | (0.000827) |
| BIPOC | 0.0886*** | 0.0746*** |
| | (0.000987) | (0.00113) |
| Hispanic | 0.0581*** | 0.0402*** |
| | (0.00124) | (0.00136) |
| Over54 | -0.0607*** | -0.0875*** |
| | (0.000668) | (0.000840) |
| Married | -0.0482*** | -0.0465*** |
| | (0.000695) | (0.000859) |
| College | 0.0297*** | 0.0283*** |
| | (0.000670) | (0.000864) |
| Constant | 0.221*** | 0.304*** |
| | (0.000837) | (0.00107) |
| Observations | 1,515,433 | 1,164,855 |
| R-squared | 0.020 | 0.021 |

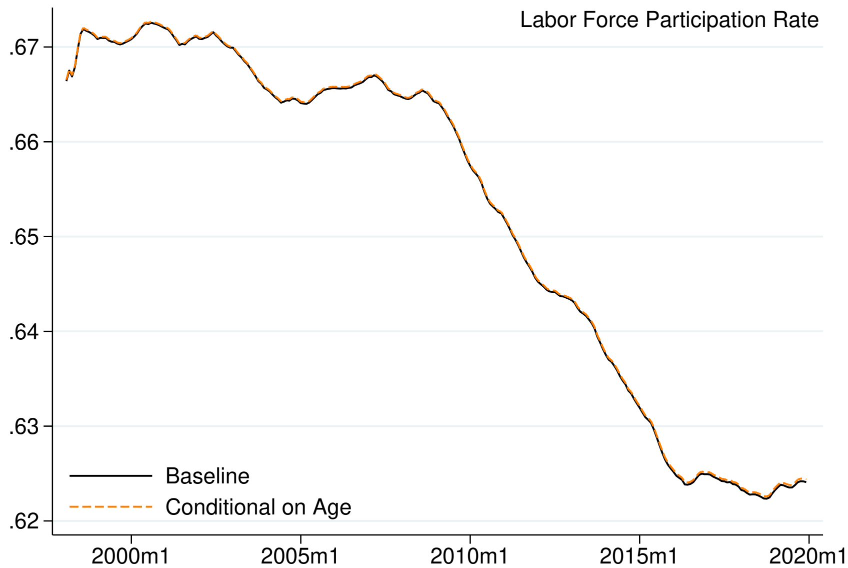
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

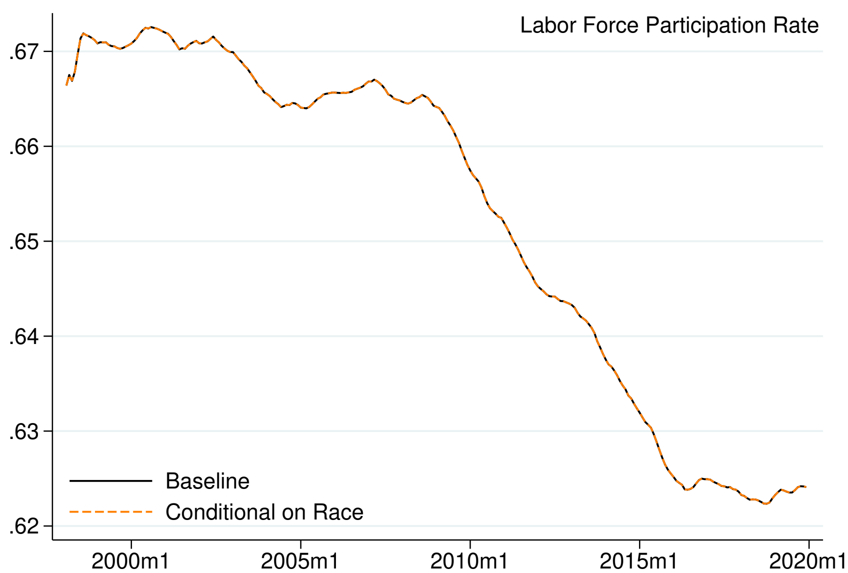
Another way to represent the demographics of partial vs. always responders is to plot the probability a respondent is part of demographic group by the number of times they nonrespond. Below is the result. Notably, always responders (represented by having zero nonresponses on the left) are systematically different from partial responders, especially for the characteristics that are the largest predictors of nonresponse, namely BIPOC and over 54.




**Notes:** Authors' calculations using data from the CPS. Each data point represents an average over 1998-2019.

# H Conditioning on Demographics

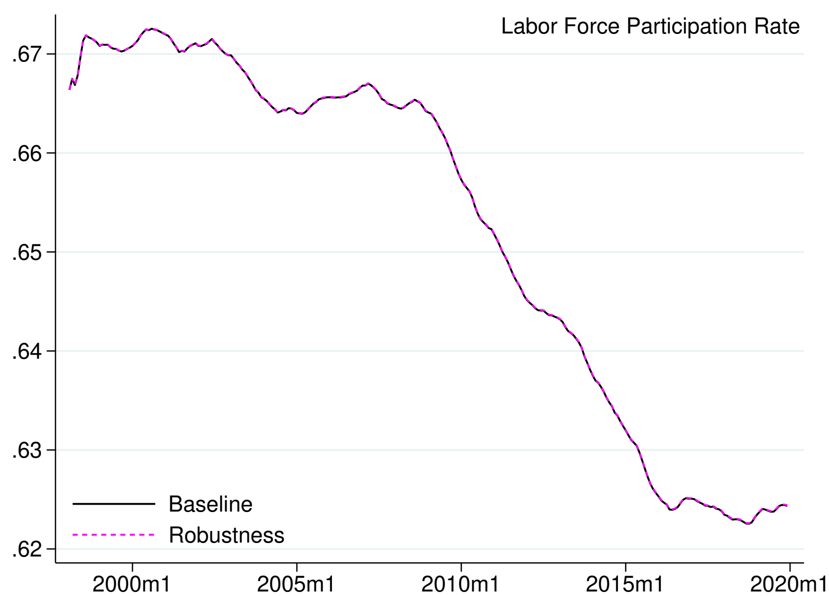



**Notes:** Authors' calculations using data from the CPS. The Baseline is the Flows Correction from Figure 3 of the main text. The dashed line in the top panel is conditional on whether the respondent is over or under age 54. The dashed line in the bottom panel is conditional on whether the respondent identifies as white or BIPOC. Figures show results are robust to disagregating flow rates by these demographic groupings.

# I Alternative for Nonresponse Flanked by Responses

The third condition in Equation (3) relates to nonresponses flanked by two responses. In the baseline specification, we apply forward and backward flow rates and average the results. In this robustness check, we compute these missing labor force statuses using probabilities calculated from individuals who respond three times in a row and have the same labor force statuses in the months the nonresponder is observed. To preserve sample size, we take the sample average of these probabilities (instead of computing them by year) and apply them to the missing observations flanked by two responses. All other missing observations—those taken care of in the first two conditions of Equation (3)—are treated the same as the baseline. The result is the dashed magenta line in the figure below. On average, the lines diverge by less than a basis point, suggesting this additional information does not alter our takeaways.




**Notes:** Authors' calculations using data from the CPS for January 1999 through December 2019. All series are the 12-month historical moving average. The Baseline is the Flows Correction from Figure 3 of the main text. The dashed magenta line uses more information to fill in missing observations that are flanked by responses in adjacent months.

# J Measuring Composition vs. Treatment

The lack of data implicit with a nonresponse makes examining empirical evidence for "treatment" issues difficult. One workaround is to leverage retrospective questions about work histories in the CPS. The retrospective question we consider in this appendix has two advantages: (1) it is asked every month, and (2) to our understanding is not a dependent interview question.[2] Respondents who are unemployed and report being available for work last week are asked when they last worked. They have the following options:"within the last 12 months," "more than 12 months ago," and "never worked." We label this question *LW* for "last worked."

We use this question to provide insight on the direction of possible "treatment" effects. Consider the following example to improve exposition. We examine respondents who are unemployed in MIS 1 and 3 but who register a nonresponse in MIS 2. We also condition on a response to the *LW* question in MIS 1 of either "more than 12 months ago" or "never worked." Thus, in the month before we observe a nonresponse, these individuals had no recent work activity. From this set of individuals, we calculate the share of them who also respond to the *LW* question in MIS 3 with "within the last 12 months." A change in their response to the *LW* question suggests they were working during the month they nonresponded. We do this for all rotation groups where there is adjacent information (i.e. MIS 2, 3, 6, and 7). Scenario (1) in the table on the next page outlines this procedure. The share of nonresponders who change their answer to the *LW* question is 21.5% which is our proxy for the share of nonresponders who were working during their nonresponse month.

One issue with treating this number as the share of nonresponses who were working in the nonresponse month is that it could be recall error. To quantify recall error, we look at individuals who were also unemployed in MIS 1 and 3, who switched their response to the *LW* question in the same way, but who responded to the survey and were coded as unemployed in MIS 2. The share of continuously unemployed individuals who switch their *LW* response is reported in Scenario (3) and is quite small: 0.3%, This suggests recall error is small.[3]

Using this approach, we can now speak to the direction of any "treatment" effects. To do this, we need a proxy for the likelihood that these nonresponders were employed during their nonresponse month (MIS 2 in this example) from compositional reasons alone. There is natural churn in labor force statuses, and this natural churn from U to E to U in our example here should be expected absent any additional "treatment" effects where changes in labor force status are "causing" the nonresponse. If our proxy for employment status using the change in the *LW* question looks quite different than the normal churn, that would be evidence of treatment effects. In our particular example, we calculate this "normal churn

[2]Unemployment duration is a dependent question where if a respondent's unemployment status was unchanged from the prior month, duration numbers would automatically increases by 4 or 5 weeks depending on the time between interviews (Chua and Fuller, 1987)

[3]It is possible that recall bias is worse for types of people who nonrespond. We examine this by looking at Scenario (2) for a subsample of respondents who nonrespond at least once. We find the ratio of households who provide an inconsistent answer to the retrospective question is 0.7% which is a similar magnitude to that reported for Scenario (2).

counterfactual" as follows. We find the total number of respondents who where unemployed in MIS 1 and 3, who respond to the $LW$ question in MIS 1 with either "more than 12 months ago" or "never worked", and who had any labor force status (U, E, or N) recorded in MIS 2. Of this total population, we compute the share who were employed in MIS 2. This share is 3.2% and is reported in scenario (3). This "normal" likelihood that unemployed individuals in MIS 1 and 3 are employed in MIS 2 is quite small relative to the 21.5% of the nonresponses who answer the $LW$ question in a way that suggests they were working during their nonresponse month. This suggests that there are treatment effects that go in the direction of employment "causing" nonresponse. As discussed in Section 4.3, Example 2 with Jessica, this treatment effect direction suggests we are underestimating the true bias in labor force metrics caused by nonresponse.

It should be strongly emphasized, however, that the nonresponders we are able to analyze with this retrospective question is a very small set of the total number of nonresponders and we caution against attempts to extrapolate these results to other nonresponders.

| Scenario | Description | MIS $t-1$ | MIS $t$ | MIS $t+1$ | Number of Households | Ratio |
|---|---|---|---|---|---|---|
| (1) | Share of M employed at $t$ | U ($LW > 12$) | M | U ($LW < 12$) | 155 | 21.5% |
| | | U ($LW > 12$) | M | U | 722 | |
| (2) | Testing for recall bias at $t$ | U ($LW > 12$) | U | U ($LW < 12$) | 127 | 0.3% |
| | | U ($LW > 12$) | U | U | 41,087 | |
| (3) | Testing if composition only | U ($LW > 12$) | E | U | 1,655 | 3.2% |
| | | U ($LW > 12$) | {U,E,N} | U | 52,134 | |

**Notes:** Authors' calculations using data from the CPS. Labor force statuses are represented by U (unemployed), E (employed), and N (not in the labor force). A missing observation from a household nonresponse is represented by M. For several of the month-in-samples (MIS), we condition on the retrospective question: when did you last work? $LW > 12$ indicates "more than 12 months ago" or "never worked." $LW < 12$ indicates "within the last 12 months."

# K Checking BLS Weights

One way to check how well the BLS backward engineered targets do, is to adjust the raw data with the demographic targets and construct labor force statics to compare against the official BLS statistics. We do this as follows. Using the raw individual counts, we compute the share of the population in each demographic cells as:

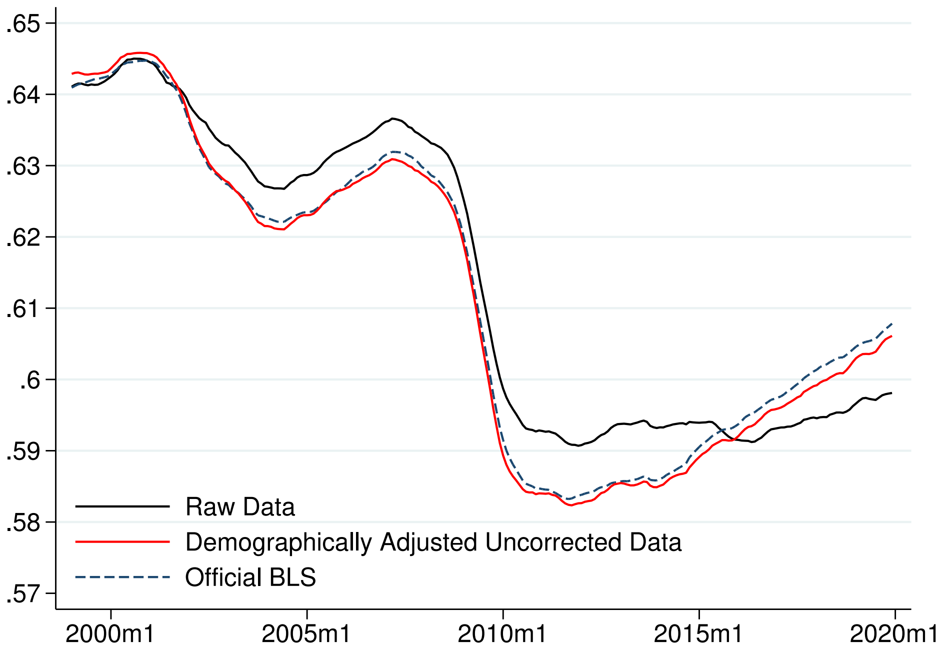
$$PopShare_{k,t}^{Raw} = \frac{p_{k,t}^{Raw}}{\sum_{k,t} p_{k,t}^{Raw}},$$

where $p_{k,t}^{Raw}$ is the sum of the (unweighted) population count in demographic cell $k$ in month $t$ and $\sum_k p_{k,t}^{Raw}$ is the total (unweighted) number of individuals in the survey in month $t$. We then up/down-weight individuals in these demographic cells with the following adjustment factor:

$$AdjustmentFactor_{k,t}^{Raw} = PopShare_{k,t}^{BLS} / PopShare_{k,t}^{Raw}.$$

In other words, if a certain demographic cell $k$ is underrepresented in the raw data, respondents in that group are upweighted so their population share matches those backward engineered from the BLS weights. To be clear, for this appendix, we only do this for individuals who respond to the survey. Using these adjustment factors, we then recompute the labor force statistics and compare them to the official BLS series. The figure on the next page plots the employment-population ratio. The black line is the same as the black line in the middle panel of Figure 3. It calculates the employment-population ratio from the raw counts of respondents in the CPS. The dashed line plots the official BLS series and the red line plots the demographically adjusted uncorrected data using the procedure above. There are a few points worth highlighting. The difference between the official headline statistic and that computed from the raw data is substantial, suggesting that BLS weights have an important impact on labor market indicators. The second observation is that our demographic adjustment (of the uncorrected data) gets us quite close to the official BLS series. This suggests that our quasi-headline series captures the vast majority of the adjustment resulting from the BLS weights.

Demographically Adjusted Employment-Population Ratio




**Notes:** Authors' calculations using data from the CPS for January 1999 through December 2019. All series are the 12-month historical moving average.