Federal Reserve Bank of Boston

Research Department

Decembert 1, 2021

Current Policy Perspectives

Measuring the US Employment Situation Using Online Panels: The Yale Labor Survey

Christopher L. Foote, Tyler Hounshell, William D. Nordhaus, Douglas Rivers, and Pamela Torola

This report presents the results of a rapid, low-cost survey that collects labor market data for individuals in the United States. The Yale Labor Survey (YLS) used an online panel from YouGov to replicate statistics from the Current Population Survey (CPS), the government's source of household labor market statistics. The YLS's advantages include its timeliness, low cost, and ability to develop new questions quickly to study labor market patterns during the coronavirus (COVID-19) pandemic. Although YLS estimates of unemployment and participation rates mirrored the broad trends in CPS data, YLS estimates of those two rates were less accurate than for employment.

Chirstopher L. Foote is a senior economist and policy advisor in the Research Department at the Federal Reserve Bank of Boston. Tyler Hounshell is an economics pre-doctoral fellow at the Tobin Center for Economic Policy at Yale University. William Nordhaus is the Sterling Professor of Economics at Yale University. Douglas Rivers is a professor of political science at Stanford University and chief scientist at YouGov. Pamela Torola is an economics pre-doctoral fellow at Yale's Tobin Center.

The authors thank William Bannick of YouGov for outstanding assistance with this project. They also thank Alan Gerber, Matthew Shapiro, and Jason Faberman, who provided helpful suggestions in the design and execution of the project.

The survey received initial Yale IRB approval on April 10, 2020, and has received further approvals as it has been revised. Foote, Hounshell, Nordhaus, and Torola declare no financial conflicts of interest with the research. Rivers has a conflict of interest as an employee and shareholder in YouGov.

I. Introduction and Overview

Standard governmental information on the US labor market is derived from the Current Population Survey (CPS). In this report, we discuss the design and performance of rapid, low-cost collection of labor market data using an online panel, where results can be obtained within 24 hours of the end of the monthly reference week. The weekly Yale Labor Survey (YLS) was designed to measure the same statistics as the CPS and was collected from April 2020 through May 2021.[1] Like the CPS, the YLS asked a battery of questions concerning current and past employment, hours, and income. Unlike the CPS, the YLS was not based on a probability sample of the US population but instead relied on a large online panel of respondents maintained by YouGov, a firm specializing in online surveys.

Because the YLS drew upon an existing panel of potential respondents, it obtained responses inexpensively and quickly (within 24 hours). The YLS was also more flexible than the CPS. Although it drew its major questions from related questions in the CPS, the YLS included questions related to the unusual nature patterns of work and unemployment during the pandemic. It was able to develop new questions in the field quickly as labor market circumstances evolved. By relying on the online panel, however, the YLS had to surmount important sample-selection issues if it were to be useful for analysis of the US labor market. In this paper, we assess the YLS's performance, in part by comparing its results to CPS data.

The YLS began with some small pilot surveys during the week of March 29 through April 4, 2020; aside from two short hiatuses the survey was conducted regularly through late May 2021. This report covered 137,500 respondents in 109

[1] The CPS is a joint product of the Bureau of the Census and the Bureau of Labor Statistics (BLS). It is the source of official monthly household labor market statistics, such as the unemployment rate, labor force participation rate, and employment-to-population ratio.

waves over 55 weeks ending with the reference week of May 23 through 29, 2021. Fourteen weeks were also reference weeks for the CPS, so YLS results can be directly compared with CPS results for April 2020 through May 2021.

The YLS had three principal purposes. The first was to determine *whether it is feasible* to provide rapid turnaround estimates of complex socio-economic data such as the state of the national labor market. The second goal was to improve national economic policy and planning by *providing more timely estimates* of the state of the labor market. The third goal was to *test the accuracy of online panels,* which are a relatively new platform for performing population surveys. The following section discusses the extent to which these goals were met.

II. Major Results and Conclusions

II.A. Goals of the Study

Feasibility. Online surveys are promising because they can be conducted quickly and inexpensively. They draw from a specific group of people—those willing to take online surveys for modest compensation—so they are not guaranteed to be representative of the entire population. However, with careful selection and weighting of the observations, we attempt to remove as much selection bias in the YLS as possible.

Relative to feasibility, the project proves that a complicated online population survey can be collected both quickly and inexpensively. The YLS collected weekly labor market data and other population statistics for more than a year, with monthly sample sizes about one-tenth of those of the CPS. As we show below, results were broadly similar to those from the CPS, but these results were available in a matter of days and at less than 1 percent of the cost of the CPS. The first goal of feasibility, including low cost, was definitely achieved.

Timeliness. There is a heightened need for timely economic data in a time of unprecedented rapid developments. Unfortunately, there is a significant lag time between when government surveys are conducted and when their results are published. A clear example of a publication lag occurred when the pandemic shock hit the US labor market in March 2020. The monthly Employment Situation reports from the Bureau of Labor Statistics (BLS) cover labor market data during reference periods that include the 12th of each month. Thus, the reference week for the March 2020 CPS was March 8 through 14. The CPS was conducted during the following week, and the results were published on April 3.

The timing of the CPS turned out to be disastrous given the timing of the COVID-19 outbreak. The first state shutdown order came in California on March 19, 2020—the week after the March CPS reference week. Consequently, the March 2020 CPS did not show the deterioration of the labor market during the last half of March 2020. This deterioration was finally revealed by the April 2020 CPS, which was published in early May 2020—almost six weeks after the major employment shock took place. In fact, the initial YLS surveys in early April 2020 indicated that the US labor market was showing extreme stress, so the YLS provided information an entire month before the official government data did. It is clear, then, that the YLS showed the ability to provide important economic information on a virtually real-time basis.

Accuracy. The third aim of the YLS was to test whether any biases of online panels could be accounted for, in order to produce results similar to those of the CPS. An online panel is a set of individuals who have agreed to complete surveys through the internet. Panelists are recruited online and receive points or money for taking surveys. In the present context, the major advantages of online surveys are that they are inexpensive, can be run continuously, and can produce answers quickly—in a single day if the questions have already been coded into survey software.

Online surveys have become widely used in the last two decades, particularly in market research and election polling, but they seldom have been used to measure labor market activity. There are two types of online panels: opt-in and probability-based. (In the latter, panelists are randomly selected, though the combination of low cooperation rates and high levels of attrition result in response rates in the low single digits.) In both cases, quota sampling and weighting are used to compensate for selection bias. There is conflicting evidence about the relative accuracy of the different methods, and there is variation between different vendors (Gittelman et al. 2015, Kennedy et al. 2016). Election prediction provides the most credible measure of accuracy. Online opt-in panels have provided results similar to those of phone surveys. In the 2020 US presidential election, both approaches had problems. Opt-in panels outperformed traditional phone polls (Silver 2021), but neither method performed especially well (AAPOR 2021). However, employment and labor market participation are subject to different types of selection bias. Previous comparisons have not involved standard employment measures, nor have available labor market variables been used for sample selection, weighting, or estimation.

Because the labor market is tracked using the comprehensive and carefully crafted government CPS, and employment is measured using an independent establishment survey, we can obtain estimates of the accuracy of online demographic surveys by comparing the outcomes of the YLS with those of the CPS. As we discuss in the next section, evidence of the accuracy of online panels for labor markets is mixed. There appear to be persistent biases in reporting as well as episodic swings in the makeup of the respondent pool. Such swings could have resulted from social and political forces operating over a few weeks, but we have been unable to resolve the resulting discrepancies with our current approach.

A final conclusion is that the YLS succeeded in obtaining *independent estimates of the state of the labor market.* Virtually all existing complex demographic and economic surveys are conducted by the government, and they are expensive and difficult to duplicate. This study shows that it is possible to use alternative techniques to replicate the larger and more expensive demographic surveys.

II.B. Major Results

The YLS conducted studies of the labor market from April 2020 through May 2021. It succeeded in providing *independent estimates* of the state of the US labor market that parallel and largely replicate the estimates from the US government's CPS. The estimates are prepared weekly and are available less than a week after collecting the survey data. The survey questions are contained in Appendix K of the study.

Four main labor market series are compared. Two series are related to employment. One employment measure is the *work-for-pay ratio* (WFPR), which is our name for series that calculates the share of the population at work during the reference week; a second measure is the more familiar *employment-to-population ratio* (EPR), which includes both people at work and employed people absent from their regular jobs. We also compare YLS and CPS estimates of the *unemployment rate* (UR) and the *labor force participation rate* (LFPR), which, like the EPR, are defined in the standard way. The following are some key findings.[2]

The YLS was relatively successful at estimating employment status. The YLS successfully mirrored the CPS-reported drop in the EPR (from pre-pandemic levels of 63 percent) to around 52 percent in April. It also matched the subsequent rise to around 55 percent in June, the steady increase through October, and the leveling off through early 2021. Similarly, YLS estimates of WFPR tracked those of CPS from their April low (49 percent) to the more recent levels (around 57 percent).

Although the YLS generated an unemployment estimate that broadly tracked the UR from the CPS, the YLS estimate of the UR was consistently too high. Over the entire 14 months, the average UR from the YLS was 12 percent, while that of the

[2] These statistics use "final weights" version 1 to weight the respondents. For a discussion of weighting procedures, see Appendix E.

CPS was 8 percent. This overestimate was found among most demographic groups and time periods, even when looking at the data across 128 categories of race, gender, age, and education. Because YLS estimates of employment relatively closely matched those from the CPS, the YLS overestimate of the UR resulted in a YLS overestimate of the LFPR as well.

Patterns of labor market activity across major demographic groups generally mirrored patterns in the CPS. (a) Both the YLS and CPS found higher unemployment rates among Black and Hispanic respondents compared to white respondents throughout the whole period. The same is true for respondents under age 29 and older than 65 years. YLS respondents with college or post-graduate degrees showed much lower unemployment rates, another disparity mirrored in the CPS. (b) Estimates for major sectors in the YLS showed the great divide seen in the CPS between industries that were hard-hit (such as leisure activities) and those that fared well (such as financial services). (c) The errors in the employment-population ratio (EPR) are highest in the 65-plus age group. With a few exceptions, other YLS age groups are a close match to the CPS.

One of the remaining puzzles in the YLS is the consistent error in measuring unemployment over the last year—even after applying weights that reflect both demographic characteristics and past labor market status. While the source of this discrepancy has not been resolved, our analysis indicates that the problem arises partly from biases in retrospective measures of earlier labor force status. Because respondents tend to underreport earlier unemployment rates in their retrospective answers (relative to what they report contemporaneously), this leads to an overestimate of current unemployment rates because the baseline employed group contains people who should have been classified as unemployed. Our investigations did not find that other sources —such as difficulties in measuring search or layoff —were important contributors to bias.

The key finding of this study is that online surveys of complicated social, economic, and demographic characteristics of the population can be studied using online panels. However, it appears—at least for the questions involved in the labor market—that there are residual biases in reporting and/or sample selection that have not been identified and corrected.

The following two sections describe important aspects of the CPS and YLS, including sample selection, differences in questions across the two surveys, and the construction of sample weights. Section III provides a broad overview of these topics, and section IV goes into somewhat more detail. Results begin in section V.

III. Brief Description of Methods

III.A. Background on the CPS

The following is a description of the CPS, which is sponsored jointly by the Census Bureau and the Bureau of Labor Statistics (BLS). The Census Bureau administers the CPS using a probability-selected sample of about 60,000 occupied households.[3] Questions in the CPS concern labor market activities during the *reference week* that includes the 12th of the month. The fieldwork is typically conducted during the subsequent *survey week* that includes the 19th of the month.

The modern "activity-based" definition of unemployment dates back to the late 1930s, with refinements in that definition continuing through various revisions in the CPS. The core CPS questions separate the adult civilian non-institutional population (POP) into three groups: employed (E), unemployed (U), and not in the labor force (NILF). These three groups are exhaustive and mutually exclusive, so POP = E + U + NILF. Employed people are those who work for pay or profit (or

[3] The CPS is a survey of households and is often called the household survey. The other main government employment survey, conducted by the Current Employment Statistics (CES) program, gathers data from establishments. Monthly results from both the CPS and the CES are released on the same day, typically the first Friday of every month.

are temporarily absent from their jobs), while non-employed people must be actively looking for work or on temporary layoff to be counted as unemployed. The labor force (LF) is defined as E + U. This study examines four main labor force statistics: the work-for-pay ratio (WFPR), the closely related employment-population ratio (E/POP), the unemployment rate (U/LF), and the labor force participation rate (LF/POP).[4]

The CPS uses a complex design involving stratification and multistage selection of housing units. CPS initial contacts were in-person until the pandemic, with some follow ups by phone. Historically, the CPS response rate has been around 90 percent, but it has a declining trend. The average response rate for the 12 months ending in February 2020 was 83 percent, but the overall response rate declined to 65 percent in June 2020 and then recovered to 78 percent in January 2021 (US Bureau of Labor Statistics 2021a). It is not clear whether the weightings undertaken by the BLS and Census Bureau have adequately dealt with the massive non-response issues in recent months.

Before turning to a formal description of the YLS, we comment on the CPS as a formal point of comparison for our survey. The CPS is rightly considered to be the gold standard for household labor market surveys in the United States.

There is no doubt that the CPS is a valuable point of reference. However, in reality, the CPS is unlikely to ever measure unemployment with the same precision that is common in the physical sciences. For example, social surveys like the CPS regularly overestimate the fraction of people who vote in elections. Comparing survey data on voting to administrative data on voter turnout is useful because voter turnout is as close as we will ever come to an accurately measured population

[4] Measuring unemployment during the pandemic has been particularly challenging because the CPS was not designed with pandemic-induced lockdowns in mind. Particularly in the early months, the CPS incorrectly classified many unemployed people as "employed, but temporarily absent from work." Using microdata from the CPS, and following a method suggested in recent BLS publications, we created an alternative unemployment measure, *U3-alt,* to correct for the misclassification. This corrected unemployment rate is conceptually similar to the unemployment rate generated by the YLS. A description of the methods is found in Appendix C.

statistic (a point made clear by the very close 2020 US presidential election). According to survey experts, the CPS's regular November election supplement has regularly overestimated voter turnout on the order of 10 percentage points (Matthew DeBell et al. 2020). This shows us that even gold-standard surveys like the CPS cannot hope to attain the standards of measurement we have achieved for the gravitational constant or the mass of the electron.

A similar issue arises with respect to the impact of interviewer error on the discrepancy between YLS and CPS. Re-interview studies often find substantial errors in labor force measures, and some of the errors are introduced by interviews and re-interviews. To the extent that the YLS is anonymous and given to a population with experience in online panels, this is likely to impart a different kind of error from that associated with the US government-run CPS (Biemer and Forsman 2021).

III.B. The Yale Labor Survey

The YLS is designed to capture the major employment aspects of the CPS and illuminate unusual aspects of the labor market stemming from the COVID-19 pandemic that shook labor markets beginning in March 2020.

The main differences between the CPS and YLS involve both the *questions asked* in the two surveys and the *sample-selection and weighting methods*. The YLS's questions concerning labor market status are similar but not identical to those in the CPS, as explained below. Also, to better understand special features of the pandemic labor market, the YLS includes several COVID-related questions. Examples include questions asking whether respondents work at their regular workplaces or at home, whether they are paid by their employers even though they did not work, and whether they applied for or are receiving unemployment insurance. We also ask standard questions about recent hours of work, income, and when respondents held their last jobs.

The differences between the two surveys in their sample-selection and weighting methodologies are more important. The CPS is designed to be a probability-based sample of the US adult non-institutional population, and its statistical validity relies on its resembling a probability sample of the population to the greatest degree possible. By contrast, the YLS was administered by YouGov, a UK-based market research and survey firm, and uses an opt-in sample.

Additionally, the CPS and YLS differ in their sample sizes. After two pilot tests, the YLS was conducted weekly in waves of 1,500 to 5,000 respondents per week. The survey period covered in this report includes 137,500 observations over 55 weeks through the end of May 2021.

III.C. The YouGov Panel

Here we provide a brief overview of the YouGov panel and YLS methods, with additional detail provided in the next section. In contrast to the probability-based CPS, the YouGov panel from which weekly YLS samples were drawn is an *opt-in sample*; all interviews are conducted online among people who have previously agreed to complete YouGov surveys for compensation. This setup ensures rapid turnaround and low cost but also risks imparting selection bias on the resulting sample. One motivation for this study is to determine whether the results from a fast, inexpensive survey can provide useful insights before and between waves of slower, expensive, and more established surveys like the CPS.

To correct for sample bias, the YLS relied on adjustments that corrected for differences between panel participants and the US population. These adjustments were based on statistical models that are designed to improve the representativeness of the sample.

These adjustments involve two critical elements. The first is the procedure that drew a YLS sample as a subset of the YouGov panel. The YouGov panel turns out to be unrepresentative of the US population along many demographic characteristics (such as age, race, and education). But a more representative sample can be drawn from the YouGov panel through the application of appropriate sampling procedures. For the YLS, *quota sampling* was used to draw samples representative of US adults in terms of age, gender, education, and race. The sampling frame included 96 strata or cells, and respondents were selected from each cell approximately in proportion to the frequency of that cell in the February 2020 CPS. This month was chosen because the economy and the labor market were relatively stable, so we could match summary statistics from YLS respondents to corresponding averages in the US population.

The second critical element needed to make YLS results representative is the construction and application of *sample weights*. Quota sampling is intended to generate a sample that is broadly representative of the target population, but in practice it rarely generates samples that exactly match multiple population targets simultaneously. A "raking" procedure is therefore used to construct weights that align the YLS sample across six demographic characteristics (age, gender, education, race, marital status, and the presence of children).[5]

The use of quota sampling and the construction of sample weights ensure that YLS samples mirror the US population along important demographic characteristics. However, the YLS sample must also reflect the general labor market attachment of the US population. Accordingly, in addition to demographic information, we also used respondents' past labor market status in the construction of sample weights.

[5] See Appendix F for details of the raking procedure. The six demographic variables included in this procedure are collected either in the survey, the respondent's YouGov user profile, or both.

The labor market behavior of people who agreed to participate in the YLS survey may not be representative of the US population in terms of their labor market attachment. To see this, consider one cell of the panel—married white women with a college education aged 45–64 with no children. This group represents 1.29 percent of the YLS sample using the quota sampling and 1.28 percent after applying the post-stratification weights. These two proportions are virtually identical because, thanks to the quota sampling procedure, the proportion of the demographic group in relation to the whole sample is very close to the proportion of the demographic group in the US population. Ideally, rates of employment and unemployment in the sample group would mirror the corresponding rates of the same demographic group in the overall population. If such mirroring occurred for all demographic slices of a YLS sample, then the YLS could produce valid estimates of aggregate labor market data using only the quota-based demographic sampling and weights constructed from demographic data alone.

Unfortunately, in practice, the YLS sample did not closely match the representative sample generated by the CPS for the labor market. More specifically, employment and unemployment rates for narrowly defined demographic groups in YLS samples tended to be different from those rates for the same groups in the general population. Respondents tended to be unemployed more often than their population counterparts. As is shown in Appendix D, more than 90 percent of the 128 demographic cells over-reported unemployment relative to the CPS.

This bias stems from unobserved variables that affect YLS respondents' labor market behavior—variables that may include the respondents' work histories, health statuses, skills, and work attitudes, as well as local labor market conditions. For instance, the US government estimates that approximately 25 percent of Americans have a disability (Centers for Disease Control and Prevention 2020). According to the BLS, in 2020, 18 percent of people with a disability were employed compared to 62 percent without a disability (US Bureau of Labor

Statistics 2021b). The quota-sampling procedure does not include disability as a demographic variable, nor is disability one of the demographic variables used to construct the sample weights. The omission of disability from these two steps can therefore result in an unrepresentative sample, even after sample weights have been applied.

To address this problem and capture the complex set of unobserved labor market influences, our weighting procedure incorporates data on past labor market status as well as demographic information. When the weights are constructed, respondents' past labor market status is treated just like a demographic characteristic such as race or age. When past labor market status is included, a weighted sample from a survey taken in (say) February 2021 will match not only the demographic makeup of the US population, but also the rates of employment and unemployment in the previous month when past labor force status is measured (for example, December 2020 or January 2021).

In a sense, past labor market status creates a "quasi-panel," meaning that it allows incorporation of individual unobserved variables that are unchanged since the previous month. It is only a "quasi" panel because the earlier labor market status is a retrospective observation on the part of respondents and therefore subject to measurement error (such as recall or question error). To the extent that the retrospective labor market status is inaccurate or biased, this will tend to bias the weights and therefore the current estimates of labor market status. (See the discussion below and in Appendix G.)

We have incorporated prior labor market status into the sample weights in different ways as the project has evolved.

For the early months of the survey, labor market status was derived from answers provided by respondents from October 2019 through February 2020, collected by YouGov. Where these data were not available, the YLS asked a recall question about February 2020 labor market status and additional questions about current labor market status.[6]

As time passed, labor market status in February 2020 became less predictive of current labor market status. Starting in July 2020, therefore, we added retrospective questions about employment from February through June. We can use these questions to create "final weights" that reflect labor market activity in months closer to survey dates. The YLS final weights for the December CPS week, for example, used labor market status averaged from the October and November 2020 CPS microdata. These final weights roll forward over time as new CPS microdata become available. The weighting procedure is described in Appendix E.

An important question is whether this weighting procedure is likely to adequately address the sample selection issues inherent in this online survey. Comparisons of probability-based studies is an active area of research.[7] Studies of the relative accuracy of online non-probability surveys and probability-based surveys have mixed results. In any case, there is no systematic determination of which approaches are superior for which kind of population information (that is, pure demographic information, secondary information, and economic and social data). Moreover, many of the studies comparing different methods are relatively simple —asking questions such as "were you employed"—rather than the approach of the YLS, which involved multiple and overlapping questions. Finally, it is worth noting that even probability-based sampling—often considered the "gold standard" for survey research—has encountered major hurdles in recent years, as the willingness of randomly drawn respondents to participate in surveys has trended down. And

[6] Respondents were asked, "During the first two weeks of February 2020, did you do any work for pay or profit?" Those responding "Yes" were deemed employed in February 2020. Those responding "No" or "Not sure" by those currently employed were classified as unemployed in February 2020. Those who responded "No" and were not currently employed were allocated their current situation, with categories being one of employed, unemployed, retired, disabled, student, homemaker, and other.

[7] Potential adjustments include quotas, stratified random sampling from the panel, matching, post-stratification weighting, and propensity-score weighting. Our approach combines quotas and post-stratification weighting.

more recently, there were additional physical barriers to in-person interviews during the pandemic.

More details on the potential errors for the weights are discussed briefly in section VIII and Appendix J. The questions for the survey are contained in Appendix K. An example that works through the method of using prior labor force status is contained in Appendix G.

IV. Detailed Description of the Panel and Statistical Methods

This section provides additional detail on the source and selection of respondents for the study, sample weighting, calculation of standard errors, and assumptions needed for valid inferences.

IV.A. Source of Respondents

Respondents were drawn from YouGov's opt-in online panel, which resembles other access panels commonly used for market research and public opinion polling (Sudman and Wansink 2002). YouGov recruits participants using internet advertising campaigns (primarily Google Adwords, Facebook, and banner ads on popular websites; but also using co-registration, visitors to YouGov's home page, and referrals from existing panelists). After confirming their email addresses ("double opt-in"), the individuals provided personal and demographic information to become registered panelists. There is no well-defined sampling frame or established probabilities of selection for panelists. The panel is simply a pool of respondents available for conducting individual research studies. People who join online panels exhibit biases that are similar to those who answer random telephone surveys (for example, they are older and more likely to be white, and they have more schooling). Attitudinal studies have found that online panelists are early adopters, less traditional, and more environmentally concerned (Gittelman et al.

2020). Unlike in phone surveys, however, online panelists are approximately balanced on gender.

The issue of selection bias has become increasingly severe for both government and private surveys in recent years. We noted above that the CPS had a response rate of only 65 percent in June 2020, which is below the US government's statistical standard. Pew estimates that response rates in telephone surveys declined from 36 percent in 1997 to 6 percent in 2018 (Pew Research Center 2019).

Additionally, over time it has become increasingly difficult to reach target audiences. Most random-digit-dial phone surveys conducted today do not use random selection to choose respondents within a household. To reduce the number of women and older respondents in the sample, either explicit quotas or other procedures are employed to reduce selection bias. For example, the interviewer might first ask, "Out of all the people age 18 or older who are at home now, may I please speak to the youngest male?" If no male lives in the household, the interviewer might then ask, "May I speak to the youngest female?"

The major point here is that an accurate representation of the population can no longer assume that the responding sample has an equal probability of selection for all members of the target population. Rather, surveys must use procedures to weight individuals in the sample, and therein lies the modern art of survey research.

IV.B. Selection of Panelists for this Study

Samples for individual YouGov studies, like this one, are selected from the YouGov panelist pool that contains the target population (in this case, the US population 18 years and older). YouGov's panel is much larger than the sample size needed for any individual study, but the company conducts many studies simultaneously. At the time of this project, there were almost 200,000 active panelists.[8] YouGov uses quota sampling to select respondents from the panel for receiving invitations and an allocation algorithm to assign responding panelists for particular studies, which we describe now.

For the YLS, panelists were allocated to 96 quota cells, based on the cross-classification of their age (18–29, 30–44, 45–64, or 65+), gender (male or female), education (high school or less, some college, college degree, post-graduate degree), and race (white, Black, or Hispanic).[9] For each cell, a target number of respondents was selected that is proportional to the number of adults in the February 2020 CPS. For each panelist, probability of response was estimated based on past rates of participation and demographics. Panelists in each quota cell were randomly selected for invitations until the expected number of responses in each cell equaled the target number. The invitations did not describe the subject of the study, nor did they guarantee that the panelist would be assigned to any particular study.

Panelists who clicked on links in their email invitations were routed to one of the available studies according to an algorithm until the target number for the survey was reached or until the field period (say, 24 hours) ended. The algorithm assigned a value to each panelist for each study that the respondent qualified for. The value is based on the number of additional respondents needed to fill the respondent's quota cell, divided by the length of time remaining for fielding the survey.

As compensation for participating in this study, panelists received points that could be converted to cash after a minimum threshold was reached. For this study,

[8] An active panelist for this purpose is defined as a panelist who has completed a survey in the last month.

[9] YouGov includes "Hispanic" as an answer option for the question "What best describes your race?" The CPS asks separate questions about respondents' race and origin. In the CPS, we have grouped whites of Hispanic origin as Hispanic and Blacks of Hispanic origin with Blacks. Whites include any non-Hispanics who are not Black, including those identifying as Native American, Asian, Middle Eastern, and mixed race.

each respondent was awarded the equivalent of $0.50 in points. The median time to complete the survey was 9 minutes.[10]

IV.C. Weighting

Respondents were selected from YouGov's panel to join the study to be representative of all US adults in terms of four demographic variables (age, gender, education, and race). Due to non-response, the realized sample does not match the population targets exactly. We use *post-stratification weighting* to improve the representativeness of the sample. The post-stratification involves two sets of variables, demographic and labor market. We use a total of six demographic weighting variables: the four demographic variables used in the quota-based sampling (age, gender, education, and race) along with marital status and presence of children. Additionally, as noted above, we use variables to represent labor market status (LMS) to capture unobserved variables that represent an individual's labor market propensities. These were either February LMS in the early part of the survey or recent LMS in the later parts (see section III.C. above).

The purpose of weighting in this context is to adjust the sample to better represent the target population. Each respondent is assigned a positive weight, so that the fraction in each cell from the weighted sample matches the fraction of that cell from a census or other reliable estimate. The assumption is that by applying the same weights for computing means and proportions of other sample variables, this procedure will correct for differences in the characteristics between the sample and the target populations.

In the simplest case, both the sample and population can be partitioned into a set

[10] One interesting feature of the present survey is that respondents might have thought that they were working for pay because they were compensated for answering online surveys. As we note in the discussion of "nuggets" below, we correct for a misclassification of this group.

of mutually exclusive and exhaustive categories according to some characteristics. For example, if it is known that 52 percent of adults are female and 48 percent are male, while the sample is 60 percent female, weighting women by 52/60 and men by 48/40 will adjust the sample proportions to match the population proportions for gender. Cell weighting works well as long as the sample fractions in each category are not too small. If a particular age-race-education-gender category has zero people in the sample, it is not possible to use a (finite) multiplicative weighting to attain the population proportion.

The problem of zero-member cells limits the number of demographic characteristics that can be included in a quota-based sampling procedure. Consider a survey that must be balanced along multiple demographic characteristics (for example, age/education/race/region/gender). A naive approach would be to form a cross-classification using all characteristics and then do cell weighting using the full cross-classification. This high-dimensional plan fails in practice because the number of cells in a cross-classification grows quickly with the number of dimensions. For example, if there are four age categories (18–29, 30–44, 45–64, 65+), four education categories (high school or less, some college, college graduate, and post-graduate), three race categories (white, Black, Hispanic), four regions, and two genders, the cross-classification contains $4 \times 4 \times 3 \times 4 \times 2 = 384$ cells. If a sample cell is empty, it is impossible to set its weight as some positive number and match the corresponding population share. Even if there are only one or two sample observations in cells, the corrective weights can become large, making the resulting sample estimates unstable.

Therefore, for the YLS, we used quota-based sampling using only four categories (gender, race, education, age). We then constructed sample weights that further refined the sample along those characteristics and also incorporated marital status, the presence of children, and labor market status.

IV.D. Raking in the YLS

The general theory of raking (weighting) and its use in the YLS is discussed in Appendix F. The variables included, to begin with, six key demographic variables: age, gender, race, education, marital status and presence of children. To control for labor market attachment, LMS was added to these demographic characteristics, and several cross-classifications were also used.

Weights were computed for each day or week's sample. The weights are not exactly equal to the ratio of the population to sample proportion in each cell because we did not weight on all cross-classifications. It is impossible to match all cross-classifications with the daily samples because some cells in the full cross-classification were empty on particular days. An example that works through the method of using prior labor force status is contained in Appendix G.

IV.E. Statistical Properties

There are different methods for estimating the variance of sample means and proportions using raking weights. Little and Wu (1991 p. 90, eq. 19) provide an asymptotic variance formula under non-random selection. Unconditional variance estimates can be obtained by treating raking as a special case of calibration weighting (Chang and Kott 2008). Alternatively, Canty and Davison (1999) discuss bootstrapped variance estimates and confidence intervals, which are conceptually simpler if finite population corrections are not necessary.

Statistical inference is another important issue. The primary purpose of post-stratification weighting with opt-in samples is to reduce bias caused by self-selection and non-response. In principle, weighting can remove bias if panel selection and within-panel non-response are conditionally independent of the weighting variables. This is Rubin's "missing at random" condition (Little and Rubin 2019).

However, raking weights are based on a parametric response model that assumes that the log ratio of population proportions to sample selection probabilities obeys a main effects model without interaction (Little and Wureliably 1991, p. 87, eq. 5). That is, the only interactions relevant for selection bias involve variables whose population joint distribution is known. Nonetheless, even if raking does not eliminate all selection bias, it seems to perform reasonably well in practice when selection bias is not severe and sample sizes not too small. (A rule of thumb is to have at least 30 observations per cell.)

It is important to note, however, that raking can only remove bias that occurs because of nonrepresentative samples at the level of the post-stratified cells (for example, demographic and labor market). If there are biases in responses within the most detailed cells (for example, demographic characteristics and prior labor market status), then the weighting cannot remove that bias because it arises from unobserved variables.

In practice, post-stratification improves the estimates markedly when labor market variables are used but adds relatively little when only demographic variables are employed. The latter result is not surprising because quota sampling eliminates most of the demographic bias, but there is still a bias that can be removed by weighting on past labor market status.

V. Basic Labor Market Definitions and the COVID-19 Pandemic

V.A. Defining Labor Market Status

Like the CPS, our survey divided the US adult civilian non-institutional population into three groups: employed (E), unemployed (U), and not in the labor force (NILF). Because of survey limitations, we limited our analysis to the population 20 and over.[11] Here are the major definitions:

Employed people worked for either pay or profit during the reference week. We added to this group respondents who answered that they received pay even though they did not work during the reference week (as explained in Appendix A). *The work-for-pay ratio* (WFPR) measures the fraction of survey respondents who reported that they worked for pay or profit during the reference week. This fraction is adjusted for overreporting for those whose only jobs are answering online surveys.

Unemployed people are those who did not work for pay but were on temporary layoff or actively looking for work. In the YLS survey, the unemployment pool was comprised of: (a) Respondents who actively searched for work in the last 4 weeks and were available for work within 7 days, and (b) Respondents who were on layoff or furlough and expected to return to their jobs.[12]

People who are *not in the labor force (NILF)* are those who are neither employed nor unemployed.

V.B. Technical Note on Measuring Employment

The BLS has six "alternative measures of labor underutilization," denoted U1 through U6, which are published each month as part of the monthly jobs report. The standard unemployment rate is U3, defined as "total unemployed." The narrowest underutilization measure (U1) includes only the long-term unemployed, while the broadest (U6) is defined as "total unemployed plus all persons marginally

[11] People under 18 could not participate in the YLS because protection of human subjects requires parental consent. See Appendix B and footnote 14 for the effect of this limit on our choice of the 20+ population.

[12] Respondents could signal this expectation in two ways. One survey question asked non-working respondents about their present work situation, to which one possible answer was "laid off or furloughed from a job to which you expect to return." Respondents could also signal a job-recall expectation by answering yes to a separate question: "If you recently lost your job, have you been given any indication that you will be recalled to work within the next six months?"

attached to the labor force, plus total employed part-time for economic reasons, plus all people marginally attached to the labor force.” In January 2021, U1 was 3.4 percent, U3 was 6.3 percent, and U6 was 11.1 percent (seasonally adjusted).

The YLS attempted to replicate the headline measure, U3. However, during the early part of the pandemic, the BLS noted that it had probably misclassified many workers displaced by COVID-19 as “employed but absent from work,” when these workers should have been classified as unemployed. BLS calculations indicated that this misclassification probably lowered the reported unemployment rate (U3) by 5 percentage points in April 2020. Fortunately, improvements in the labor market and in CPS implementation reduced this error over time to around 0.6 percentage points by February 2021.

Because of the misclassification, YLS researchers used CPS microdata to construct an alternative measure of unemployment, moving workers classified as employed but absent from work for “other reasons” into the unemployment pool. The resulting measure, U3-alt, then allowed an apples-to-apples comparison with the unemployment rate in the YLS, where the classification error was less likely to occur. For a further discussion, see Appendix C. While the correction reduced the error in the calculation of labor force status in YLS in the early months, that improvement was smaller in the later months.

VI. Results of the Survey

VI.A. Overview

Table 1 and Figure 1 summarize results for the CPS survey months from April 2020 through May 2021. (For full results by month, see Appendix D.) We have direct comparisons for most months. Our estimates are limited to the population aged 20 and older (see Appendix B).[13] We show both the standard U3 measure of unemployment and the alternative concept, U3-alt, which includes an adjustment for classification errors in the survey as described in the last section.

The major conclusion is that the YLS estimates closely paralleled the labor market experience as described by the CPS. The estimates for employment were relatively accurate; those for unemployment tended to be slightly high, and consequently, the labor force participation rate was also higher than the CPS estimates.

Table 1 shows the average values and errors of each of the three major labor force categories for the 14 months, measured as a percentage of the adult population. The YLS captured the employment-population ratio closely over the period. However, it systematically overestimated unemployment, with a larger overestimate with the standard U3 than with U3-alt. The fraction of people not in the labor force was underestimated (that is, the participation rate is overestimated) largely because of the overestimate of unemployment.

TABLE 1. AVERAGE VALUES AND ERRORS FOR YLS AND CPS, APRIL 2020 TO MAY 2021

	Average monthly value Percent (%) of Population		
	Employed	Unemployed	Not in Labor Force
CPS	58.2	5.0	36.8
CPS-alt	57.2	5.9	36.8
YLS	57.0	7.4	35.6
	Average error Percent (%) of Population		
	Employed	Unemployed	Not in Labor Force
YLS - CPS	-1.1	2.4	-1.3
YLS - CPS-alt	-0.2	1.5	-1.3

[13] People under 18 are excluded from the sample because the protection of human subjects requires parental consent to participate in a survey. Although people aged 16–19 years have low labor force participation, they also have high unemployment rates, so there is a non-trivial difference between the 16+ unemployment rate and the 20+ unemployment rate. For the last two decades, the 16+ rate has been about ½ ppt higher than the 20+ rate, although this difference has trended lower since 2013.

Figure 1 shows monthly and weekly comparisons of the CPS and YLS for different concepts. The CPS-based estimates show both the official U3 rate and our constructed U3-alt rate.

Panels 1(a) and 1(b) show the persistent upward bias in the estimated unemployment rate, while panels 1(c) and 1(d) show that the survey was quite close to the CPS on the employment rate. The error in the unemployment rate increased after October 2020.

The alternative unemployment rate (U3-alt) is closer to the YLS unemployment rate than the standard U3 unemployment rate. The explanation is that many workers were mistakenly classified as employed in the CPS, whereas they were correctly classified as unemployed in the YLS. The difference between U3 and U3-alt declined over the period after April as pandemic-related absences for "other reasons" declined sharply.

The bottom line on the survey for the aggregates is that the YLS has proven remarkably accurate for employment but has consistently overestimated unemployment.

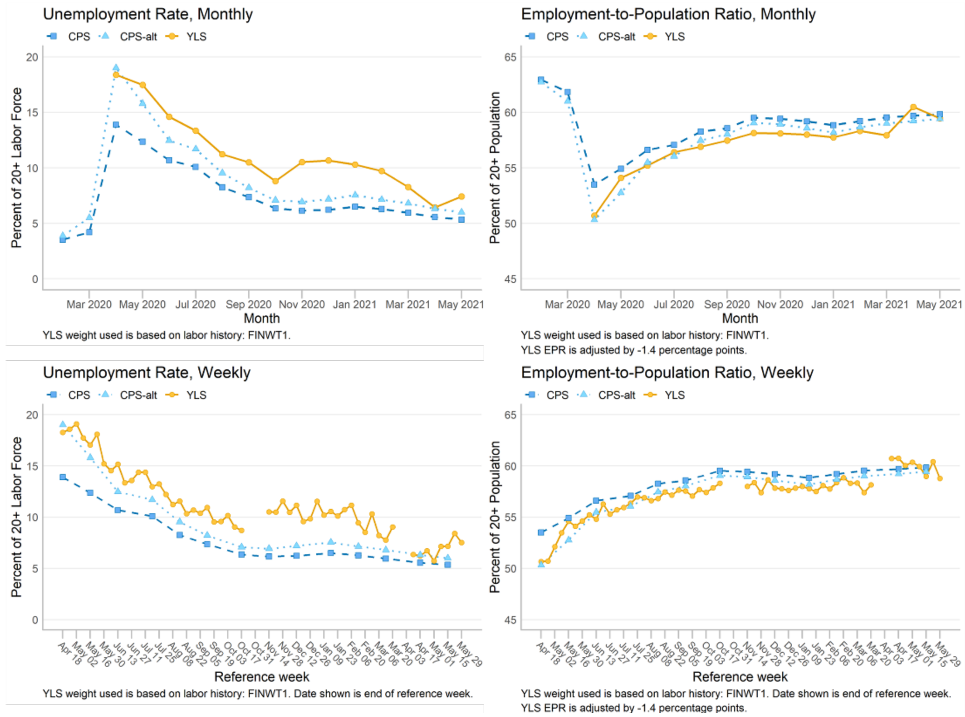

FIGURE 1(A) AND 1(B). UNEMPLOYMENT RATES BY WEEK AND MONTH, CPS AND YLS. FIGURE 1(C) AND 1(D). EMPLOYMENT-POPULATION RATIOS BY WEEK AND MONTH, CPS AND YLS

VI.B. Unemployment Rates for Major Groups

Next, we show the labor market status for different groups. Tables 2 and 3 provide the averages for the entire sample period. The underlying trends indicate, accurately, the following impacts.

Here are the major results: (a) The youngest age groups had the highest unemployment rates during the pandemic. (b) Among racial and ethnic groups, Black and Hispanic workers had the highest unemployment rates during the pandemic. (c) Among educated groups, lower educated groups had the highest unemployment rates during the pandemic. (d) Among occupations, those in service, construction, and transportation occupations were the most severely impacted. (e) Among industries, leisure and hospitality were the most severely affected.

Here are some results for demographic groups compared to the CPS. (a) The YLS tended to overestimate unemployment among females relative to males. (b) Among age groups, the YLS tended to overestimate unemployment, primarily among the oldest age group (age 65+). (c) There was no significant difference in estimates by racial groupings. (d) The YLS tended to overestimate unemployment among groups with lower education relative to those with higher education. (e) The YLS tended to overestimate unemployment for those with widowed and divorced as marital status.

Here are some results for economic groupings compared to the CPS: (a) Among regions, the YLS tended to overestimate unemployment in the South relative to the Northeast. (b) Among occupations, the YLS tended to overestimate unemployment in sales and underestimate in farming, transportation, and services. (c) Among industries, the YLS tended to overestimate unemployment dramatically in mining and information and overestimate in leisure and hospitality.

TABLE 2. AVERAGE MONTHLY UNEMPLOYMENT RATES FOR YLS AND CPS, DIFFERENT DEMOGRAPHIC GROUPS, APRIL 2020–MAY 2021

	CPS	CPS-alt	YLS
Gender			
Male	7.7	9.1	9.9
Female	8.1	9.7	12.7
Age			
20-29	11.6	12.7	14.3
30-44	7.1	8.3	10.1
45-64	6.8	8.3	9.9
65+	7.6	10.7	13.8
Race			
White	6.4	7.8	9.7
Black	11.5	13.1	14.5
Hispanic	10.3	11.7	14.1
Other	9.0	10.7	12.0
Education			
HS or less	10.6	12.2	15.0
Some college	8.9	10.6	13.1
College grad	6.0	7.4	8.2
Post grad	3.9	5.0	5.6
Marital status			
Married	5.7	7.1	8.0
Widowed	8.6	11.0	13.1
Divorced	8.2	10.0	14.2

Separated	11.1	13.2	13.1
Never Married	11.6	12.9	15.8

TABLE 3. AVERAGE MONTHLY UNEMPLOYMENT RATES FOR YLS AND CPS, DIFFERENT GROUPS, APRIL 2020–MAY 2021

	CPS	**CPS-alt**	**YLS**
Census Region			
Northeast	9.3	11.2	11.4
Midwest	7.1	8.3	10.4
South	7.0	8.4	11.5
West	9.1	10.7	11.5
Occupation			
Management, business, and financial	4.2	5.5	6.6
Professional and related	4.7	6.1	7.6
Service	13.5	15.6	14.7
Sales and related	8.7	10.4	13.0
Office and administrative support	7.4	8.3	9.3
Farming, fishing, and forestry	9.6	10.8	9.5
Construction and extraction	10.6	12.6	13.5
Installation, maintenance, and repair	6.6	7.4	10.0
Production	9.0	10.0	12.8
Transportation and material moving	11.7	13.3	11.3
Industry			
Agriculture, forestry, fishing, and hunting	5.4	6.5	7.9
Mining	14.1	15.2	6.3
Construction	8.7	10.7	11.2
Manufacturing	6.7	7.4	8.2
Wholesale and retail trade	8.3	9.6	10.8
Transportation and utilities	9.0	10.6	8.9
Information	8.2	9.9	7.3
Financial activities	4.1	5.1	5.4
Professional and business services	6.9	8.4	8.4
Educational and health services	5.5	6.8	8.0
Leisure and hospitality	20.6	23.0	26.3
Other services	9.7	13.1	13.4
Public administration	2.7	3.3	6.3

VI.C. Estimates Using Different Weights

We constructed different weighting models. Figure 2 shows the results using four different weights for the YLS and compares those with the results for the CPS. The four YLS weights are demographic weights and two sets of labor market weights (February and "final," which used recent months). These show the standard CPS version of U3 unemployment as well as our modified U3 version.

Several points are clear in the figures: (a) The demographic weights fare poorly in most cases. As explained above, this is likely due to unobserved variables that are important for labor market behavior, such as disability. (b) The February labor market weights do reasonably well in the early part of the period but diverge increasingly from the CPS in the later part of the year. The reason is that February status becomes increasingly obsolete as time passes. (c) The final labor market weights (reflecting labor market status in each of the cells in the last two months) track the actual CPS relatively closely. This is particularly true for the employment and work-for-pay data. However, the final weights tend to overestimate the CPS U3 throughout most of the period, although they are reasonably accurate at tracking U3-alt in the early months of 2020. (d) The results for the labor force participation rate (LFPR) are parallel to the results for unemployment, tending to overpredict because of the overestimate of the unemployment rate.

The clear conclusion of the data shown in Figure 2 is the critical importance of including labor market experience in raking the data. Demographic data alone do a relatively poor job in tracking the CPS.

A key finding is that the YLS matches the CPS more closely with the WFPR and EPR than with the UR and LFPR. Why is this so? Part of the UR discrepancy undoubtedly stems from the additional complications that arise when measuring unemployment as compared to employment. Measuring unemployment requires that the survey instrument not only discern whether a non-employed person is searching for a job but also whether this search is an active rather than passive one, since only active searches can lead directly to job offers. Additionally, the concept of "layoff" has evolved over time and is particularly ambiguous during a pandemic. When a restaurant shuts down in March 2020 and the employer tells workers that it will be only a short shutdown, does the worker consider this a temporary layoff? To the extent that interview surveys like the CPS and self-administered internet surveys like the YLS treat subtle labor market concepts differently, unemployment rates in the two types of surveys may differ.

Figure 2 also shows that errors in the YLS relative to the CPS were particularly large in the early stages of the pandemic and then during the spring of 2021. Part of the issue in spring 2021 was that a major discrepancy appeared between the responses of panelists not previously sampled for YLS ("new respondents") and those who had participated in an earlier wave ("repeats"). In early 2020, clearly, there were few repeats, but by the end of the study more than half of respondents were repeats. The reported labor market status of repeats was significantly different from new respondents, and that discrepancy tended to give large month-to-month errors in April and May 2021. In addition, the team experimented with different sampling of repeats, which (it was later discovered) resulted in large errors in those two months.

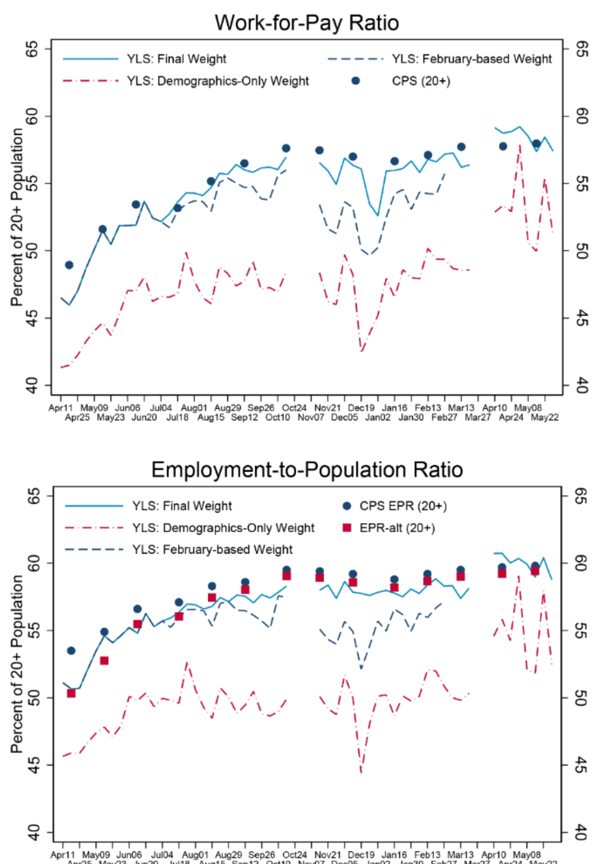

FIGURE 2(A) AND 2(B). COMPARISON OF ESTIMATES ON EMPLOYMENT FOR DIFFERENT WEIGHTS

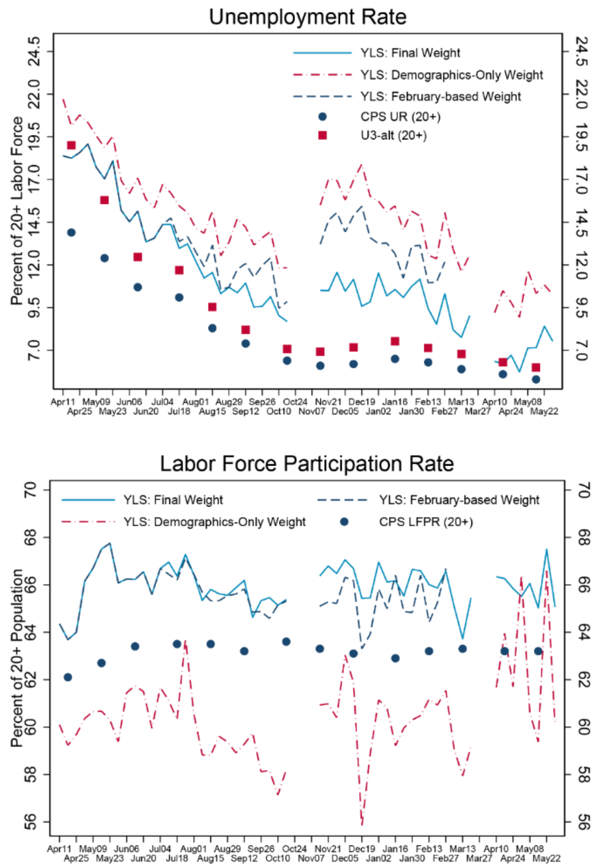

FIGURE 2C) AND 2(D). COMPARISON OF ESTIMATES ON UNEMPLOYMENT AND LABOR FORCE PARTICIPATION FOR DIFFERENT WEIGHTS

VII. Further Results

The YLS has many interesting findings for the pandemic period. This section examines a few of the nuggets from the surveys.

VII.A. Some Nuggets from the YLS

Results of the YLS contains many interesting nuggets that illustrate the impact of the 2020 pandemic on the labor market. Here are a few.

Work at home or the office? Many workers who normally worked outside the home started to work remotely as the effects of the pandemic spread. How large were the numbers? We asked where respondents worked during the survey year. We found that only 54 percent responded that they work entirely outside the home, while another 11 percent responded that they work both at home and outside the home.

Why absent from job? The YLS asked respondents about the reasons they were absent from work, adding pandemic-related questions to the normal questions in the CPS. It was interesting that only a small fraction of respondents listed childcare problems as a reason for absence. However, close to 33 percent listed, "I was temporarily absent from a job due to the coronavirus." Additionally, 8 percent said they were absent because of illness in their family, and 14 percent said they were absent because of their own illness.

When last worked? The survey asked when people lost their last jobs. About 1 percent of respondents who had ever worked replied that they last worked in each month from May 2019 through February 2020, though 6 percent lost their jobs in February 2020. There was a huge jump in job losses in March 2020, when about 25 percent of respondents reported losing their jobs. Since that time, the rate of job loss has averaged 2.7 percent per month, declining from 5.9 percent in April 2020 to 0.8 percent in May 2021.

Why lost job? The YLS asked people why they lost their jobs if they were employed prior to the onset of the pandemic. Of those who responded in April of 2020, 70 percent of workers said they lost their jobs because their firms reduced workers or hours because of COVID-19. That decreased sharply to 50 percent by August of 2020 and has steadily decreased to about 44 percent as of February 2021.

Answering internet surveys as a job. One of the issues with the YLS was that people might think that answering surveys represents "work for pay or profit." While this is a reasonable answer, we know that respondents are definitely biased toward people who respond to internet surveys, since they all do. To test the extent to which this might bias the responses, we directed respondents to not consider internet surveys as a job. Starting in wave 21, we queried respondents on this issue. With targeted questions, we determined that the fraction of the population that responded "yes" to the work-for-pay answer increased consistently by 1.4 percent from respondents whose only job was answering internet surveys. We were unable to reliably classify these individuals as unemployed or not in the labor force, but we noted that the number of employed is slightly overestimated in the survey and applied a -1.4 percent correction to our employment measures to diminish the bias.

Gender. There have been concerns that the CPS does not incorporate current views of gender. We therefore asked about both binary gender and a larger group of gender categories (N = 54,000). We found that 96.8 percent reported consistent binary gender on both the "Gender" and the "Gender7" question. Of the sample, 0.79 percent reported nonbinary gender and 1.11 percent reported inconsistent binary gender (all weighted values). This level of inconsistency is about the same as that of the traditional two-category gender question. The nonbinary gender groups tended to have higher labor force participation, but that was largely due to the younger age of that group.

Hours yesterday. One interesting calibration question was how many hours each respondent worked yesterday, which was asked of those who worked for pay. The mean response was 7.2 ($\pm$ 0.2) hours for those working for pay. This is 3.8 hours per adult when corrected for those not working for pay. The American Time Use Survey (ATUS) for 2019 reported an average of 3.6 hours per adult (US Bureau of Labor Statistics 2020). This is a remarkably close figure given the simplicity of the YLS question.

VII.B. Self-employment and Gig Economy Work

We also investigated the number of self-employed workers in the YLS sample. The CPS includes a "class of worker" characterization that includes wage and salary workers, self-employed workers, and unpaid family workers. Counting both the incorporated and unincorporated self-employed in the self-employment category, slightly more than 10% of employed people have been self-employed over the last several years in the CPS. This figure is close to the rates of self-employment in the YLS data, which we measured by asking people who worked for pay about their type of employer. We asked respondents whether they were "self-employed," and 13 percent of weighted respondents selected this answer, a figure slightly above the CPS self-employment rate.[14]

A separate YLS question asked workers to classify themselves using different categories than the class-of-worker question in the CPS. Respondents could note that they were working "for myself in my own firm," a "contract or gig economy worker," or "working for a wage or salary at a firm or other employer." For all waves that included this question, 15 percent of working respondents said they were working for themselves in their own firm, while another 11 percent said they were gig economy workers.

As other researchers attempting to measure the "gig economy" have found, it can be difficult to match workers' conceptions of their jobs with CPS concepts. Many workers might consider themselves gig-economy or self-employed when presented with one set of potential answers but call themselves wage-and-salary workers when presented with a different set of answers. Even the legal definitions and tax-law definitions regarding employment are complex and may not easily be understood by those who are not typical W2 employees.

VIII. Other Considerations

Many studies are forecasting US labor market characteristics, but few are tracking labor market responses in real-time. As of the date of the report, other than

[14] For more on the effects of high self-employment and multiple-jobholding rates in online surveys of the labor market, see Katz and Krueger (2019).

the CPS and the present study, we are aware of six other surveys that examine labor market dynamics in the COVID-19 period. The three main other studies published to date are by Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber (2020 CG&W), which relies on the Nielsen Homescan panel; a survey by Alexander Bick and Adam Blandin (2021 RPS), which relies on a Qualtrics panel; and a Census Bureau panel, the Household Pulse Survey (2021 HPS), which began April 23. The results of the other studies are summarized in Appendix H.

As with other surveys, there are several reasons why unemployment and participation estimates generated by YLS could differ from underlying population values. Often called "total survey error," these differences come from several sources: sampling error, nonresponse error, errors from differences in questionnaires and question wording, errors from interviewer vs. self-administered survey, and respondent error. Appendix J discusses our investigations into the sources of survey error. We conclude that sampling error is unlikely to be the major cause of the discrepancy between the YLS and the CPS. The most likely source is potential unrepresentativeness of the panel, even at the most detailed cell level.

REFERENCES

AAPOR (2021). AAPOR Report: Josh Clinton et al., *Task Force on 2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls,* American Association for Public Opinion Research, www.aapor.org, July 19, 2021.

Aaronson, Daniel, Scott A. Brave, R. Andrew Butters, Daniel Sacks, and Boyoung Seo (2020). "Using the Eye of the Storm to Predict the Wave of Covid-19 UI Claims." Federal Reserve Bank of Chicago, Working Paper no. 2020–10. https://doi.org/10.21033/wp-2020-10.

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh (2020). "Inequality in the impact of the coronavirus shock: Evidence from real time surveys." IZA—Institute of Labor Economics, Discussion Paper no. 13183. http://ftp.iza.org/dp13183.pdf.

Bernstein, Joshua, Alexander W. Richter, Nathaniel A. Throckmorton (2020). "COVID-19: A View from the Labor Market." Federal Reserve Bank of Dallas, Working Paper no. 2010 (April). https://doi.org/10.24149/wp2010.

Bick, Alexander, and Adam Blandin (2021). "Real-Time Labor Market Estimates During the 2020 Coronavirus Outbreak." Unpublished Manuscript, Arizona State University. https://sites.google.com/view/covid-rps/home.

Biemer, Paul P. and Gosta Forsman. "On the Quality of Reinterview Data with Application to the Current Population Survey." *Journal of the American Statistical Association* 87, no. 420 (1992): 915-23. Accessed March 5, 2021. doi:10.2307/2290627.

Brynjolfsson, Erik, John Horton, Adam Ozimek, Daneil Rock, Garima Sharma, and Hong Yi Tu Ye (2020). "COVID-19 and Remote Work: An Early Look at US Data." NBER, Working Paper no. 27344 (June). https://doi.org/10.3386/w27344.

Canty, Angelo J. and Anthony C. Davison (1999). "Resampling-Based Variance Estimation for Labour Force Surveys." *Journal of the Royal Statistical Society,* Series D, vol. 48 (September), pp. 379-391.

Castro, Miguel Faria (2020). "Fiscal Policy during a Pandemic." Federal Reserve Bank of St. Louis, Working Paper no. 2020-006G. https://doi.org/10.20955/wp.2020.006.

Centers for Disease Control and Prevention (2020). "Disability Impacts All of Us." https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html.

Chang, Ted and Phillip S. Kott (2008). "Using Calibration Weighting to Adjust for Non-response under a Plausible Model." *Biometrika*, vol. 95 (September), pp. 555-571.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber (2020). "Labor Markets during the COVID-19 Crisis: A Preliminary View." NBER, Working Paper no. 27017 (April). https://doi.org/10.3386/w27017.

DeBell, Matthew, Jon A. Krosnick, Katie Gera, David S. Yeager, and Michael P. McDonald. "The turnout gap in surveys: Explanations and solutions." *Sociological Methods & Research* 49, no. 4 (2020): 1133-1162.

Deming, W.E., and Frederick F. Stephan (1940). "On a Least-Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known." *The Annals of Mathematical Statistics*, vol. 11 (December), pp. 427-444.

Gittelman, Steven H., Randall K. Thomas, Paul J. Lavrakas, and Victor Lange (2015). "Quota controls in survey research: a test of accuracy and intersource reliability in online samples." *Journal of Advertising Research*, vol. 55 (December), pp. 368-379.

Foote, Christopher, William D. Nordhaus, and Douglas Rivers (2020). "The US Employment Situation Using the Yale Labor Survey." Cowles Foundation Discussion Paper 2243, July, New Haven, CT.

Ireland, C. Terrance, and Solomon Kullback (1968). "Contingency Tables with Given Marginals." *Biometrika,* vol. 55 (March), pp. 179-188.

Katz, Lawrence F. and Alan B. Krueger (2019). "Understanding Trends in Alternative Work Arrangements in the United States," *RSF: The Russell Sage Foundation Journal of the Social Sciences*, no. 5 (2019): 132-146.

Kennedy, Courtney, Andrew Mercer, Scott Keeter, Nick Hatley, Kyley McGeeney and Alejandra Gimenez (2016). *Evaluating Online Nonprobability Surveys*, Pew Research Center.

Little, Roderick J.A., and Donald B. Rubin (2019). *Statistical Analysis with Missing Data,* 3rd ed. Hoboken, NJ: Wiley.

Little, Roderick J.A., and Mei-Miau Wu (1991). "Models for Contingency Tables with Known Margins When Target and Sampled Populations Differ." *Journal of the American Statistical Association*, vol. 86 (March), pp. 87-95.

Lohr, Sharon L. (2010). *Sampling: Design and Analysis (Advanced Series),* 2nd ed. Boston: Brooks/Cole.

Pew Research Center (2019). "Response rates in telephone surveys have resumed their decline." https://www.pewresearch.org/fact-tank/2019/02/27/response-rates-in-telephone-surveys-have-resumed-their-decline/.

Pew Research Center (2020). "About Half of Lower-Income Americans Report Household Job or Wage Loss Due to COVID-19." https://www.pewresearch.org/social-trends/2020/04/21/about-half-of-lower-income-americans-report-household-job-or-wage-loss-due-to-covid-19/.

Şahin, Ayşegül, Murat Tasci, and Jin Yan (2020). "The Unemployment Cost of COVID-19: How High and How Long?" *Economic Commentary* (Federal Reserve Bank of Cleveland) (May 7), pp. 1-7. https://doi.org/10.26509/frbc-ec-202009.

Silver, Nate (2021). "The Death of Polling is Greatly Exaggerated," retrieved March 25, 2021 (https://fivethirtyeight.com/features/the-death-of-polling-is-greatly-exaggerated/).

Sudman, Seymour and Brian Wansink (2002). *Consumer Panels*, 2nd ed. Chicago: American Marketing Association.

US Bureau of Labor Statistics (2020). "CPS and COVID-19," February 2020, downloaded from https://cps.ipums.org/cps/covid19.shtml.

US Bureau of Labor Statistics (2020). "Table A-1. Time spent in detailed primary activities and percent of the civilian population engaging in each activity, averages per day by sex, 2019 annual averages." https://www.bls.gov/tus/a1-2019.pdf.

US Bureau of Labor Statistics (2021a). "Household survey response rates." https://www.bls.gov/osmr/response-rates/household-survey-response-rates.htm.

US Bureau of Labor Statistics (2021b). "Persons with a Disability: Labor Force Characteristics—2020." https://www.bls.gov/news.release/pdf/disabl.pdf

US Census Bureau (2019). *Current Population Survey Design and Methodology: Technical Paper 77*. https://www2.census.gov/programs-surveys/cps/methodology/CPS-Tech-Paper-77.pdf

US Census Bureau (2021). "Household Pulse Survey: Measuring Social and Economic Impacts during the Coronavirus Pandemic." https://www.census.gov/programs-surveys/household-pulse-survey.html.

Wozniak, Abigail, Joe Willey, Jennifer Benz, and Nick Hart (2020). "COVID Impact Survey: Version 3 [dataset]." Chicago: National Opinion Research Center.

Yeager, David S., Jon A. Krosnick, LinChiat Chang, Harold S. Javitz, Matthew S. 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*, vol. 75 (November), pp. 709-747.