Examining Differences in the Labor Status in the Current Population Survey and the Census 2000[1]

Tamara Adams
U.S. Census Bureau
Washington, D.C.

1. Introduction

On the first Friday of every month, the Bureau of Labor Statistics releases the unemployment rate for the previous month. As one of the well-known economic indicators, the financial world eagerly awaits the estimates each month and the financial markets react to small changes in the unemployment rate each month. However, like all survey measures, the collection of the unemployment data each month from the Current Population Survey is subject to measurement errors due to nonresponse, respondent's errors, and the misunderstanding of questions.

If we examine the unemployment rate for March, April, and May, 2000, we see that the non-seasonally adjusted unemployment rate as measured by the Current Population Survey (CPS) was 4.3 percent, 3.7 percent, and 3.8 percent respectively. (www.bls.gov and Employment and Earnings, April, May, and June 2000). However, the unemployment rate measured from the Census 2000, taken from late March to early September was 5.8% (www.census.gov). These differences are larger than the standard errors for either of the surveys. The direction of the disparity is a recent phenomenon. From 1950-1980 the decennial census measured a lower unemployment rate than the CPS; the 1990 decennial census was the first to measure a higher unemployment rate than the CPS (Brown and Flaim, 1992). As we will see later, this may be related to the changing composition of the workforce and the inclusion of more women, temporary and seasonal work, and other non-traditional employment situations.

The above comparisons have been performed at a macro level. Using data from the CPS/Census Match conducted by the Census Bureau in the fall of 2002, we can examine labor force data at a micro-level in order to attempt to determine the phenomena in action. The micro-level data provides us with a unique opportunity to examine differences related to causal effects of change in labor force status between the CPS and the census, such as operational differences between the two data collection methods, questionnaire differences, and characteristics related to tenuous attachment to the work force.

2. The Current Population Survey and Census 2000 Long Form

2.1 The Current Population Survey

The Census Bureau conducts the Current Population Survey (CPS) for the Bureau of Labor Statistics each month. A nationwide sample of about 60,000 (since increased to 72,000) addresses is selected each month for interview. The primary purpose of the CPS is to measure the unemployment rate of the noninstitutionalized, civilian population of the U.S. who are at least 16 years old. The survey is a dual-mode survey with approximately 10% of the sample interviewed at a centralized computer-assisted telephone interview (CATI) facility. The remainder of the sample is interviewed by an interviewer using a computer-assisted personal interview (CAPI) instrument. A household is eligible for a phone interview only after their initial (usually face-to-face) interview in the CPS panel; the respondent provides a phone number, and the initial interviewer says that the household would be appropriate for a telephone interview (i.e., there are no language barriers or the interview was not difficult to obtain). Approximately 10 percent of interviews are completed via CATI each month. The nonresponse rates for February-May 2000 are all below 7 percent. (See U.S. Census Bureau, 2002 for more details on the CPS). For this analysis we consider those cases interviewed in February - May, 2000.

2.2 Census 2000 Long Form

The Census 2000 sampled approximately 1-in-6 households to complete the long form questionnaire (See Hogan, 1999 for more details on the long form sampling). In addition to population counts and demographic information collected for all households in the U.S., the long form also collects data on various topics, including labor force characteristics. The census, including the long form, is primarily a paper-and-pencil operation.

Most of the data are collected by mail-response with a self-administered paper questionnaire – approximately 72

[1]This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

percent of the mailout/mailback universe, which comprises 80 percent of the nation, completed a mail response questionnaire (Clark and Whitford, 2001). In addition, 18.8 percent of the nation was update/leave – an enumerator dropped off a questionnaire and the respondent completed the questionnaire and returned it by mail (Clark and Whitford, 2001). The remainder is primarily collected using a paper-and-pencil interviewer-administered questionnaire. The mode of collection varies with the date of the census interview – early respondents fill out the mail questionnaire, later respondents are interviewed. (See www.census.gov for more details on the long form data collection).

2.3 The CPS/Census Match

The Census Bureau conducted the CPS/Census Match in the fall of 2002. The basic March CPS and the Annual Demographic Supplement[2] were included, as were basic CPS month-in-sample 4 and 8 cases from February, 2000 and basic CPS month-in-sample 1 and 5 cases from April and May, 2000 (see U.S. Census Bureau, 2000 for more information on the rotation group structure of the CPS). Both addresses and persons were matched. No matches outside the state were sought. For information on the special weighting procedures associated with the CPS/Census Match, see Zbikowski (2004). For more information on the matching operations, please see Siegel (2004).

3. Literature Review and Discussion

Below, we will discuss various factors indicated by previous research that may contribute to the differences between the labor force measurement between the census and the CPS. These fall into several broad categories: reference period issues, respondent issues, data collection operation differences, questionnaire differences, and phenomena that are intrinsic to the labor force that make measurement difficult.

3.1 Reference Period Differences

Differences in the nature of their reference periods is one obvious source of inconsistency between the census and the CPS. Brown and Flaim (1992) hypothesized that reference period may have caused some of the macro-level differences between the 1990 census and the 1990 CPS unemployment data. First, the reference period is different for the two surveys – CPS collects data on a fixed reference period, the week containing the second Tuesday of each month (U.S. Census Bureau, 2002). However, the census questionnaire used "last week" as the reference period. Given that the census data were collected from January to September (the majority of the data were collected from March to May), "last week" referred to many different weeks. Therefore, we can expect some of the differences to be true differences – since the reference period shifted, some people's labor force status will also have shifted. Several studies (Bassi, Torelli, and Trivellato, 1998; Biemer and Bushery, 1999; McGovern and Bushery, 1999) show high rates of inconsistency for those classified as unemployed in the first measurement as compared to those classified as employed or not in the labor force in the first measurement. In other words, those captured as unemployed in the first measurement are more likely to change labor force status in the second measurement than those captured as employed or not in the labor force. In many studies, since the same measurement tool is used, there is a symmetry to the indices of inconsistency. Specifically, those measured as umeployed in the second measurement are more likely to have been classified differently in the first measurement as those who were captured as employed or not in the labor force in the second measurement. This is not as pronounced in this study since the measurement tools (i.e., the CPS and the census questionnaires) are not exactly the same. However, since we see changing of labor force status in surveys that have similar reference periods, a vastly different reference period cannot be the only factor affecting inconsistencies between the CPS and census measurements of labor force status.

3.2 Data Collection Operations

If we compare and contrast the data gathering operations, we see some operational differences. In general, these differences may lead to a difference in the data being collected. Figure 1 summarizes the differences between the data collection operations.

[2]Now referred to as the Annual Social and Economic Supplement (ASEC)

Figure 1 – Data Collection Difference Between the CPS and the Census		
	CPS	**Census**
Mode	Mostly personal visit with some CATI (approx. 10%)	Mostly paper mailout questionnaires with some personal visit
Interviewers	More experienced – worked on CPS for a longer time period	Less experienced – many hired for the census
Sponsorship	U.S. government survey	U.S. government survey with publicity. More well-known.
Automation	Fully automated (CAPI and CATI instruments)	Very little automation with complex skip patterns

3.3 Respondent Issues

Both the CPS and the census collect proxy data for labor force status – a household member other than the subject can provide data or someone outside the household can provide data. Although studies are not conclusive (Kojetin and Mullin, 1995; O'Muircheartaigh, 1991; Moore, 1988; Biemer and Bushery, 1999), data collected by proxies may tend to be more consistent in the CPS and the census, but the data collected may not be the "right answer". Since proxies outside the household are not collected until later in the census process during the enumerator data collection phases, mode effects and reference period effects may be further exacerbated by proxy data collection.

3.4 Questionnaire Differences

The questionnaires for the CPS and the census are quite different, especially in how they collect labor force data. Both questionnaires collect roster information and demographics before they collect the labor force data. Due to the multipurpose nature of the census long form, there are several differences between the CPS and the census in attempting to capture labor force data. In general these differences cause shortcomings that lead to ambiguity in the census; the CPS is able to have questions tailored towards more difficult specific situations than the census. For a summary of these differences, please see Figure 2.

Figure 2 – Differences in the Current Population Survey and Census 2000 Long Form Labor Force Questions		
Issue	**Current Population Survey**	**Census**
Self-Employment	Asks specific questions about self-employment and owning a business	Has only guidelines in small print under the first labor force item
Multiple Jobs	Has special question wordings and paths for people with multiple jobs	No allowance for multiple jobs
Absence from job	Asks about reasons for absence	Has only a short, non-exhaustive list of specific reasons for job absence
Retirement	Has a specific path for retirees	No path or questions for retirees
Vague Terms	Both the CPS and the census use terms that may be vague and subject to interpretation by respondents	

3.5 Labor Force Phenomena

Discrepancies between the CPS and the census classifications of the same individual's labor force status sometimes reflects true changes in status rather than measurement differences. For most people, labor force status is a constant within a year – they are either employed or they do not work. However, for some people, it is "weak and uncertain" (Bailar and Rothwell, 1984). That is, Bailar and Rothwell posit, for some people, labor force membership is an attitude – something that shifts from day-to-day. Like other attitudinal items, the reporting of labor force status would be very susceptible to change based on the questionnaire and the mode of the interview. Bailar and Rothwell (1984), in considering labor force reporting as an attitudinal item, report differences in labor force measurement by respondent (proxy or self), mode, and by interviewer.

In addition, the very concept that is being measured, "employment", can be a difficult phenomenon to understand. Both the CPS and the census rely on concepts such as "last week", "work", "job", "business", "layoff", and "looking for work" that may be vague. The respondents may define these terms differently than survey designers and data collectors (Campanelli, Martin, and

Rothgeb, 1991; Martin, Campanelli, and Fay, 1991; Norwood and Tanur, 1994), and respondents may tend towards overinclusion of "work" when asked to define their labor force status (Martin, Campanelli, and Fay, 1991). This would lead to definitional issues that would be further exacerbated by the attitudinal nature of employment.

4. Limitations

There are several limitations to this data analysis:

- Independence – The CPS and census, although operationally independent, often used the same respondents. The CPS respondent may remember the response given in the CPS interview and may repeat the same answer for the census (Biemer and Forsman, 1992). In addition, participation by a respondent in the CPS may affect their census data collection in much the same way that panel respondents are conditioned (Bassi, Torelli, and Trivellato, 1998). In the CPS reinterview, this can lead to an increase in the unemployment rate for the second interview (Chua and Fuller, 1987). We can attempt to control for this effect by using the census date of measurement.
- Rotation Group Bias – The CPS uses a 4-8-4 rotation scheme (U.S. Census Bureau, 2002). Rotation group bias, or estimates being correlated with the amount of time that a respondent spends in sample, may affect their responses (Bailar, 1975). Controlling by the month-in-sample may alleviate this effect on the analysis.
- Nonmatches – The CPS/Census Match does not have a complete matching (approximately 8% of CPS people do not match and this is differential across various population groups). The census could have omitted a given unit or the CPS may have erroneously listed a person. Either of these situations can cause nonmatches. However, the overall unemployment rate measured with and without the nonmatches is not statistically different.

5. Methods

Although we could compare unemployment rates for various indicators listed above, this would not take full advantage of the fact that we have micro-level data available to analyze the changes in labor force status between the CPS and the census. Instead, we will use a measure called the %inconsistent, where, if we had a 3x3 table as below in Figure 3, the %inconsistent would be as follows:

Figure 3 – Constructing the % Inconsistent

	Census			
CPS	Employed	Unemployed	Not In the Labor Force	Total
Employed	n_{11}	n_{12}	n_{13}	$n_{1.}$
Unemployed	n_{21}	n_{22}	n_{23}	$n_{2.}$
Not in the Labor Force	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{..}$

$$\%\ inconsistent = \frac{n_{12} + n_{13} + n_{21} + n_{23} + n_{31} + n_{32}}{n_{..}} * 100$$

Notice that in Figure 3, there are not columns for out-of-age range or those people considered to be in the Armed Forces by the census. We delete the people falling into these categories from the analysis. Those who fall out of the age range in either the CPS or the census are inconsistent due to their ages (i.e., their labor force data are not collected in one survey or the other), not collection of labor force status. Those in the Armed Forces are considered out of scope for the CPS.

We will use a two-tailed *t*-test, with a Bonferroni adjustment for multiple comparisons when necessary, to compare the %inconsistent from various subgroups.

We will use logistic regression modeling and multinomial logistic regression modeling. Both of these techniques model the logit of the probability of an event occurring. In our case, we will be modeling the following function for dichotomous logistic regression:

$$\log(\frac{p}{1-p})$$

where p is the probability of the event occurring (SAS Institute, 1999B). For multinomial logistic regression, where we have a nominal response variable, we will be modeling the generalized logit function:

$$\log\frac{\pi_{hij}}{\pi_{hir}}$$

where h,i,j are index variables for 3 response categories, π_{hij} is the probability of the given category, and π_{hir} is the probability of the last category to be used as a reference (SAS Institute, 1999A).

Variances are calculated using the CPS replicate design (U.S. Census Bureau, 2002) with a constant factor included in the replicate weights to account for the census long form sampling. In this method, we do not account for the sample design of the long form; however, we account for the differential weighting in the long form sample.

6. Results

For the purposes of analysis in this paper, we will use the 2000 basic CPS match (February-May). Using the links from the matches provided by the CPS/Census Match, we linked the basic CPS cases to the Census 2000 long form records. We will be treating the CPS cases that received the long form as a subsample of the CPS. If we include only records that have a final labor force status code for both the CPS and the census (i.e., are at least 16 years old in both the CPS and census data), we have 13,249 people in sample. If we examine Table 1, we see that the CPS unemployment rate for these cases is 3.9%. This is within standard error of the not seasonally adjusted CPS estimate of 4.3% for that month.

Table 1—March 2000 CPS Unemployment Rate and the Unemployment Rate from the Census 2000 for the Matched Cases in the CPS/Census Match

March 2000 CPS Unemployment Rate	3.9%
	(0.25)
Census 2000 Unemployment Rate	5.2%
	(0.25)

Note: These are for those cases in the CPS who received a long form. Standard errors in parentheses. These data do not include those who are out of age range or in the Armed Forces.

The census unemployment rate for this sample is 5.2%. These are significantly different at the α=0.05 level and represent quite different measurements of unemployment. If we examine Table 2, we see that 90.6% (58.08%/64.14%) of the CPS employed cases are also employed in the census, indicating a high rate of consistency for those classified as employed in the CPS. In addition, 86.2% (28.28%/32.79%) of the not in the labor force cases from the CPS are also not in the labor force in the census. Again, this is a relatively high rate of consistency. Proportionately, however, the largest inconsistencies are for those classified as unemployed in the CPS – only 42% (1.12%/2.65%) of those people are classified as unemployed in the census. Some of this change is expected. Given that the CPS data were gathered in mid-March and the census data were gathered from February through September, we would expect some true change. However, the change does not balance – the marginal totals for the census are not the same as those for the CPS.

For example, we examine the two cells which indicate conversions between employed and not in the labor force. There are 3.28% of people classified as not in the labor force in the CPS and employed in the census; there are 4.78% of people classified as employed in the CPS and not in the labor force in the census.

Therefore, there is a 1.5% net difference in the classification of people who are employed versus not in the labor force. Over the short time to which these two measurements refer (mid-March for the CPS compared to "last week" for the census), there should be a balance between these two estimates and the net difference should be zero.

Table 2 – Percentage of March 2000 Current Population Survey Records Matched to the Census Long Form by their Current Population Survey and Census Labor Force Statuses
(Weighted N= 207,875,749)

	Census Labor Force Status					
Current Population Survey Labor Force Status	**Out of Age Range**	**Employed**	**Unemployed**	**Not In Labor Force**	**Armed Forces**	**Total**
Out of Age Range	0.00	0.14	0.04	0.25	0.00	0.42
Employed	0.12	**58.08**	1.12	4.78	0.04	64.14
Unemployed	0.02	0.80	**1.12**	0.69	0.01	2.65
Not In Labor Force	0.10	3.28	1.11	**28.28**	0.01	32.79
Total	0.25	62.31	3.37	34.01	0.06	100.00

%inconsistent=12.5% (significant at the α=0.01 level)
Note: For illustrative purposes, we have included those out of age range in either survey and those in the Armed Forces. These records will not be included in the remainder of the analyses.

In order to tease out differences in the causal patterns for each of the types of inconsistencies, we used multinomial logistic regression. We performed multinomial logistic regression analyses for each of the three types of labor force classifications in the CPS: employed, unemployed, and not in the labor force. For each of these, we classified as our base variable those cases which were consistent – for example, those considered employed in the CPS and those who are also employed in the census. We modeled against the consistent cases, those cases which were inconsistent between the census and the CPS.

In Table 3, we provide a synopsis of the above findings.

7. Conclusions

There are no clear-cut answers to the question, "Why does the CPS labor force status measurement differ from the census labor force status measurement?" However, when we break down the different types of misclassification, we can start to see some patterns emerging.

The strongest predictor of inconsistency is the imputation of the census labor force status. Those cases with an imputed census labor force status have a much higher rate of inconsistency than other cases. This stands to reason since the census did not collect enough information to assign a labor force status without imputation.

Next, we see differences between the CPS and census based on date of census response. Those who respond later to the census tend to have more inconsistent answers than those who respond earlier. In addition, there may be differential effects of response date on various subgroups of people. There may also be effects of "last resort" data – last ditch efforts by the census to collect data may create more response error.

We also see differences based on demographic characteristics. We believe that different demographic groups are differentially attached to the workforce. This makes them differentially affected by date of census. We also see mode and respondent effects – we see more socially desirable answers in interviewer-administered interviews and more socially desirable answers from self-respondents. Finally we see more inconsistent answers for those who are *employed but absent* in the CPS and for those who own home businesses. Several previous studies

point toward the difficulty in measuring these concepts (Campanelli, Martin, and Rothgeb, 1991; Martin and Polivka, 1995; Rothgeb, Polivka, Creighton, and Cohany, 1991). The census questionnaire does not contain some of the specific questions that CPS does regarding these difficult-to-measure concepts.

If we examine the recent trend of the CPS unemployment rate being smaller than that of the census, we can gain some insight from this research. First, the largest drivers of the unemployment rate change will tend to be from those who are classified as employed in the CPS being classified as either unemployed or not in the labor force in the census. For those being classified as unemployed in the census, we see that there are census data collection effects. Specifically, there may be an interviewer effect due to social desirability bias. In addition, those who respond late tend to have inconsistent answers. Although this may be due to true change, it may also be due to "last resort" data collection efforts by the census. For those being classified as not in the labor force in the census, we see that those who are *employed but absent* are more likely to be considered not in the labor force by the census. The CPS questionnaire could be too sensitive to this classification or the census may be insensitive to this type of person. Given the research by Campanelli, Martin, and Rothgeb (1991), we suspect that the census is not sensitive enough. In addition, older people tend to be classified as not in the labor force by the census. Since the census does not have specific questions regarding retirement or employment during retirement, there may be misclassification on the part of the census. However, the CPS may be oversensitive to those who are retired in order to reduce respondent confusion and burden.

Future research should include additional modeling to tease out the different hypothesized constructs of data collection differences, true change, and questionnaire differences. In addition, one could more thoroughly analyze the differences in the questionnaires and paths taken during the interviews. Also, one could more carefully examine the interaction between all of the main effects. However, given the fluid, attitudinal nature of labor force status for some portions of the population, capturing this elusive characteristic will continue to be a challenge for survey designers.

8. References

Adams, Tamara, Byrne, Rosemary, and Ramos, Magdalena, "Results of Quality Assurance on the Accuracy and Coverage Evaluation Matching Operations", *ASA Proceedings of the Section on Survey Research Methods*, 2001.

Bailar, Barbara A., "The Effects of Rotation Group Bias on Estimates from Panel Surveys", *Journal of the American Statistical Association*, 70 , 23-30, 1975.

Bailar, Barbara A. and Rothwell, Naomi D., "Measuring Employment and Unemployment", *Surveying Subjective Phenomena*, Turner, Charles and Martin, Elizabeth (eds.), 1984.

Bassi, Francesca , Torelli, Nicola , and Trivellato, Ugo, "Data and Modelling Strategies in Estimating Labour Force Gross Flows Affected by Classification Errors", *Survey Methodology*, 24 , 109-122, 1998.

Biemer, Paul P. and Bushery, John M., "Estimating the Error in the Labor Force Data Using Markov Latent Class Analysis", *Proceedings of the FCSM Conference*, 1999. http://www.fcsm.gov/99papers/biemer.pdf

Biemer, Paul P. and Forsman, Gosta, "On the Quality of Reinterview Data with Application to the Current Population Survey", *Journal of the American Statistical Association*, Vol. 87, No. 420., pp. 915-923, Dec., 1992.

Brown, Sharon and Flaim, Paul, "Interpreting the Differences between BLS and Decennial Census Estimates of Unemployment", part of LAUS Technical Memorandum No. S-93-1, Internal Bureau of Labor Statistics Documentation, November 18,1992.

Bushery, John M. , Brick, J. Michael , Severynse, Jacqueline , and McGuinness, Richard A., "How Interview Mode Affects Data Reliability", *ASA Proceedings of the Section on Survey Research Methods*, pp. 600-604, 1996.

Campanelli, Pamela C., Martin, Elizabeth A., Rothgeb, Jennifer M., "The Use of Respondent and Interviewer Debriefing Studies as a Way to Study Response Error in Survey Data" *Statistician*, Vol. 40, No. 3, Special Issue: Survey Design, Methodology and Analysis (2), pp. 253-264, 1991.

Chua, Tin Chiu , and Fuller, Wayne A., "A Model for Multinomial Response Error Applied to Labor Flows", *Journal of the American Statistical Association*, 82 , 46-51, 1987.

Clark, Jon R. and Whitford, David C., *Quality of Census 2000 Processes*, DSSD Census 2000 Procedures and Operations Memorandum Series, # B-3, February, 28, 2001.

Hogan, Howard. "Long Form Sampling Specifications for Census 2000", DSSD Census 2000 Procedures and Operations Memorandum Series, # LL-5, Nov. 17, 1999. http://landview.census.gov/dmd/www/pdf/LL-5.pdf

Kojetin, Brian A. , and Mullin, Paul, "The Quality of Proxy Reports on the Current Population Survey (CPS)", *ASA Proceedings of the Section on Survey Research Methods*, 1995

Martin, Elizabeth A., Campanelli, Pamela C., Fay, Robert E., "An Application of Rasch Analysis to Questionnaire Design: Using Vignettes to Study the Meaning of `Work' in the Current Population Survey", *Statistician*, Vol. 40, No. 3, Special Issue: Survey Design, Methodology and Analysis (2), pp. 265-276, 1991.

McGovern, Pamela D. and Bushery, John M., "Data Mining the CPS Reinterview: Digging into Response Error", *Proceedings of the FCSM Conference*, 1999. http://www.fcsm.gov/99papers/mcgovern.pdf

Moore, Jeffrey C., "Self/proxy Response Status and Survey Response Quality: a Review of the Literature", *Journal of Official Statistics*, 4 , 155-172, 1988.

Norwood, Janet L. and Tanur, Judith M., "A Review Measuring Unemployment in the Nineties", *Public Opinion Quarterly*, Vol. 58, No. 2., pp. 277-294, Summer, 1994.

O'Muircheartaigh, Colm. "Simple Response Variance: Estimation and Determinants", *Measurement Errors in Surveys*, ed. Biemer, et al. 1991.

SAS Institute, SAS/*STAT User's Guide, Version 8, Volume 1*, SAS Publishing, 1999

SAS Institute, SAS/*STAT User's Guide, Version 8, Volume 2*, SAS Publishing, 1999

Siegel, Paul, "Accuracy of Data for Employment Status as Measured by the CPS- Census 2000 Match ", Census 2000 Testing, Experimentation, and Evaluation Program., Report B.7.

US Census Bureau, *Current Population Survey Design and Methodology*, Technical Paper 63RV, March 2002.

Zbikowski, Andrew, "Some Issues in Weighting the Current Population Survey - Census 2000 File", *ASA Proceedings of the Section on Survey Research Methods*, 2004, In Publication.

Table 3 – Synopsis of Multinomial Models

Indicator	Effect
Demographic Characteristics	After controlling for mode and date of census response, we are still seeing effects of demographic characteristics. Mode and true change may differentially interact with various levels of the demographic groups.
Labor Force Status	
Employed but absent	Those *employed but absent* in the CPS are more likely to be classified as not in the labor force in the census. May be due to the lack of appropriate questions in the census.
Not in the labor force	If a person is *not in the labor force – other* in the CPS, they are more likely to have inconsistent classifications than those disabled or retired. The vagueness of the category may make it difficult to capture properly in the current questionnaires (either census or CPS).
Home business ownership	People tend to have inconsistent classifications in the CPS and census if someone in the household owns a business. May be due to the lack of a specific question in the census concerning home businesses.
Proxy data	Self-respondents are more likely to report socially desirable answers of "employed".
Imputed Census Labor Force Status	The strongest covariate by far is if the census imputed the labor force data. In all models, it indicates a tendency towards inconsistency.
Mode effects	There is a mode effect in the census—more socially desirable answers in enumerator questions. May be due to a distancing effect of mail surveys.
Date of census response	Later census responders tend to have more change.