

Misclassification error when identifying job stayers in the Current Population Survey

Chase Eck¹

August 2017

Abstract

I evaluate the degree of misclassification error in Card and Hyslop's (1997) method for identifying job stayers in the Current Population Survey (CPS). Benchmarking to job tenure data from the biennial CPS job tenure supplement, I find that Card and Hyslop's method, which relies on industry and occupation codes, misclassifies approximately one-third of hourly workers. However, I also find that these errors do not substantially distort estimates of downward nominal wage rigidity. The reason is that industry and occupation miscoding is essentially random.

Keywords: wage rigidity, job stayers, CPS, misclassification

JEL Codes: J30, J31, E24

¹ Department of Economics, University of Arizona, 1130 East Helen Street, McClelland Hall 401, Tucson, Arizona 85721-0108, email: eckes1@email.arizona.edu. I would like to thank Gary Solon for his excellent comments and suggestions as well as Jean Roth for providing dictionary files for the National Bureau of Economic Research (NBER) Current Population Survey (CPS) archive. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1. Introduction

The Current Population Survey (CPS) has been one of the main U.S. data sources for studying job stayers' year-to-year wage changes, especially in research on downward nominal wage rigidity. An apparent limitation of the CPS for such research is that most months of the survey do not identify which workers have stayed with the same employer for the last year. An influential study by Card and Hyslop (1997, henceforth CH) dealt with this limitation by assuming that workers who reported the same two-digit occupation and industry in each year had stayed with the same employer.

CH acknowledged, "Many of the observed industry or occupation switches are presumably attributable to misclassification errors."² In such cases, the CH method may misclassify job stayers as changers. I will refer to this sort of misclassification as Type I error. The CH approach also could generate Type II error -- misclassifying workers who changed employer but stayed in the same industry and occupation as job stayers.

A more recent paper by Elsby, Shin, and Solon (2016, henceforth ESS) addressed this issue by using only Current Population Surveys that included the job tenure supplement, which in recent years has been administered every other January. Using the tenure supplement's question about how long respondents had been in their current job, ESS could study year-to-year wage changes of workers that explicitly reported they had been in their job for over a year. The advantage of the ESS approach is that it presumably achieves more accurate identification of job stayers. On the other hand, it can be applied to only 1 CPS out of every 24. In contrast, the CH method can be used for all 24.

Treating the ESS method as a benchmark, this note uses the tenure supplement months of the CPS to evaluate the CH method. In those months, I examine the overlap in classification between the two methods, and I study the degree to which their differences in classification yield different results with respect to nominal wage rigidity, a common application. I first replicate the ESS study and then apply the CH method to the same data. Assuming the ESS job stayers more accurately represent the full set of true stayers, I can identify the errors and degree of sample loss associated with the CH method. The CH method fails to identify a little more than a third of job stayers due to Type I error. When I compare the results from the two methods in their original context, studying nominal wage rigidity, I find that the two methods yield broadly similar estimates.

2. Method

As a baseline, I first replicate the ESS results for 2007-2008, 2009-2010, and 2011-2012 using the January CPS files hosted at the National Bureau of Economic Research. Following the ESS procedure, I matched survey records from each January CPS including a tenure supplement back one year using household IDs and line numbers. Per Madrian and Lefgren (2000), I then eliminated false matches by taking out observations for which race and gender changed or the

² Section 6 of Bound, Brown, and Mathiowetz (2001) reviews the substantial literature on industry and occupation classification errors. Additional discussions are in Kambourov and Manovskii (2013) and page 279 of Polivka and Miller (1998).

age change was less than -1 or more than 3. I then subsetted the sample to observations in the outgoing rotation groups (4th and 8th interview months), the groups for which wage variables are reported.

Following ESS, I dropped all observations for which hourly wage was imputed, removed observations with top-coded wages, and restricted the sample to those between the ages of 16 and 64. I then calculated the change in each observation’s log nominal wage and generated identifiers for whether each observation is a job stayer based on the ESS method (ESS stayer) or the CH method (CH stayer). Table 2, to be discussed below, summarizes my estimates of downward nominal wage rigidity using each method and compares closely to ESS’s Table 5.³ For simplicity, I report results only for hourly workers.

3. Results and Discussion

After replicating the ESS study, I extended the sample period to include 2013-2014 and 2015-2016, and examined the difference between ESS and CH stayer methods. I chose this sample period as it contains the Great Recession as well as several years after, ensuring that the results aren’t specific to downturns or recoveries.

Table 1 shows the unweighted number of hourly workers in each year by classification type. In each year, about a third of ESS changers are classified as stayers by CH. These Type II errors seem fairly inconsequential, however, in the sense that more than 90 percent of CH stayers also are ESS stayers. Type II errors do not loom larger because there are relatively few ESS changers in the sample. In each year, approximately 1000 (approximately 85%) of the longitudinally matched hourly workers are ESS stayers. To understand this pattern, it is useful to keep in mind that the CPS samples housing units, not households. The matching procedure therefore implicitly eliminates households that moved. Consequently, the sample is comprised solely of non-movers, a population that may contain a particularly large proportion of job stayers.

<i>Year</i>	<i>ESS Stayer</i>		<i>ESS Changer</i>	
	CH Stayer	CH Changer (Type I error)	CH Stayer (Type II Error)	CH Changer
<i>2007-2008</i>	712	394	64	117
<i>2009-2010</i>	712	376	39	86
<i>2011-2012</i>	567	365	39	90
<i>2013-2014</i>	585	351	42	114
<i>2015-2016</i>	530	348	69	117

Table 1: Unweighted counts of hourly workers by stayer type and year.

As for Type I errors, more than a third of ESS stayers are classified as changers by CH. To a first approximation, then, CH stayers are a subgroup of ESS stayers. As a result, in a January with a tenure supplement, the ESS method delivers a larger sample of job stayers than the CH

³ My results are slightly different for two reasons. First, I used the tenure supplement weights instead of an average of the outgoing-rotation-group weights. Second, I corrected an error in the ESS code that inadvertently removed never-married individuals from the 2007-2008 and 2009-2010 samples.

method. In the bigger picture, however, the CH method can deliver a much larger sample because it can be applied to every month.

Finally, I examined how the different job stayer classification methods impact the estimates of nominal wage rigidity as measured by the fraction of job stayers who experienced no or negative nominal wage changes. As reported in Table 2, I find that there are only small differences in estimates of negative or no wage change. This finding suggests that, while CH stayers are mostly a subgroup of ESS stayers, they are not substantially different. That is, misclassification of industry or occupation is essentially random among hourly workers.

<i>Year</i>	<i>Zero nominal wage change</i>			<i>Negative nominal wage change</i>		
	ESS	CH	Both	ESS	CH	Both
<i>2007-2008</i>	17.2%	16.8%	16.0%	19.0%	18.7%	19.7%
<i>2009-2010</i>	20.0%	20.0%	19.3%	23.3%	25.1%	24.5%
<i>2011-2012</i>	19.2%	19.4%	19.1%	25.5%	24.0%	25.9%
<i>2013-2014</i>	20.7%	20.9%	19.9%	21.5%	21.9%	23.0%
<i>2015-2016</i>	18.8%	19.1%	17.6%	18.2%	18.2%	18.9%

Table 2: Fraction of hourly job stayers zero or negative nominal wage changes by classification method.

4. Conclusion

The CH method largely identifies a particular subset of ESS stayers. The subset produced by the CH method yields broadly similar estimates of nominal wage rigidity, a central application of identifying job stayers. This result suggests that the misclassification of industry or occupation among hourly workers in the CPS is essentially random. Consequently, it is likely that the two methods yield similar results in applications involving longitudinal matches of the CPS. The CH method, then, may be preferred in applications that require data from multiple months or need a larger sample size. The ESS method is preferred if the study is restricted to January in a given year as it provides a more accurate accounting of job stayers and a larger sample in that month.

References

- Bound, John, Charles Brown, and Nancy Mathiowetz. "Measurement error in survey data." In James J. Heckman and Edward Leamer (eds.), *Handbook of econometrics*, vol. 5, pp. 3705-3843. North-Holland, 2001.
- Card, David, and Dean Hyslop. "Does inflation 'grease the wheels of the labor market?'" In Christina D. Romer and David H. Romer (eds.), *Reducing inflation: Motivation and strategy*, pp. 71-122. University of Chicago Press, 1997.
- Elsby, Michael W. L., Donggyun Shin, and Gary Solon. "Wage adjustment in the Great Recession and other downturns: Evidence from the United States and Great Britain." *Journal of Labor Economics* 34, no. 1, pt. 2 (2016): S249-S291.
- Kambourov, Gueorgui, and Iourii Manovskii. "A cautionary note on using (March) Current Population Survey and Panel Study of Income Dynamics data to study worker mobility." *Macroeconomic Dynamics* 17, no. 1 (2013): 172-194.
- Madrian, Brigitte C., and Lars John Lefgren. "An approach to longitudinally matching Current Population Survey (CPS) respondents." *Journal of Economic and Social Measurement* 26, no. 1 (2000): 31-62.
- Polivka, Anne E., and Stephen M. Miller. "The CPS after the redesign: Refocusing the economic lens." In John Haltiwanger, Marilyn E. Manser, and Robert Topel (eds.), *Labor statistics measurement issues*, pp. 249-286. University of Chicago Press, 1998.