Economics Bulletin

Volume 38, Issue 2

Misclassification error when identifying job stayers in the Current Population Survey

Chase S. Eck
University of Arizona

Abstract

I evaluate the degree of misclassification error when identifying job stayers in the Current Population Survey using the method proposed by Card and Hyslop (1997). I find that even though their method misclassifies approximately one-third of hourly workers, these errors do not substantially impact estimates of downward nominal wage rigidity, a common application. However, I find that misclassification error influences the results in studies that require estimating the share of the population that are job stayers and changers and their wage changes, such as studies of wage adjustment, and develop a correction for it.

I would like to thank Gary Solon for his excellent comments and suggestions, an anonymous referee for suggesting the application to studies of wage adjustment, and Jean Roth for providing dictionary files for the National Bureau of Economic Research Current Population Survey archive. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Citation: Chase S. Eck, (2018) "Misclassification error when identifying job stayers in the Current Population Survey", *Economics Bulletin*, Volume 38, Issue 2, pages 1077-1082
Contact: Chase S. Eck - eckcs1@email.arizona.edu.
Submitted: September 04, 2017. **Published:** May 25, 2018.

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 and wage adjustment in general. 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."[1] 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. As ESS note, their results are broadly consistent with CH's results for these months.[2] I go further by examining the overlap in classification between the two methods. Additionally, I study the degree to which misclassification error impacts studies of estimates of nominal wage rigidity and wage adjustment. 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. However, when I apply these results to a study of wage adjustment the methods provide substantially different results.

2. Replication 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

[1] 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).

[2] Daly et. al (2012) also study nominal wage rigidity. They find smaller spikes at zero relative to ESS using the CH method and the ratio of reported weekly earnings to reported weekly hours as their wage measure for non-hourly workers.

one year using household IDs and line numbers. Following Madrian and Lefgren (2000), I then eliminated false matches by taking out observations for which race and gender changed or the 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. Nominal Wage Rigidity

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.

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

	ESS Stayer			ESS Changer		
Year	CH Stayer	CH Changer (Type I error)	% Type I error	CH Changer	CH Stayer (Type II error)	% Type II error
2007-08	712	394	36%	117	64	35%
2009-10	712	376	35%	86	39	31%
2011-12	567	365	39%	90	39	30%
2013-14	585	351	38%	114	42	27%
2015-16	530	348	40%	117	69	37%

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.

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

	Zero nominal wage change			Negative nominal wage change		
Year	ESS	CH	Both	ESS	CH	Both
2007-08	17.2%	16.8%	16.0%	19.0%	18.7%	19.7%
2009-10	20.0%	20.0%	19.3%	23.3%	25.1%	24.5%
2011-12	19.2%	19.4%	19.1%	25.5%	24.0%	25.9%
2013-14	20.7%	20.9%	19.9%	21.5%	21.9%	23.0%
2015-16	18.8%	19.1%	17.6%	18.2%	18.2%	18.9%

In addition, I compared the wage-change distributions produced by each classification method. As shown in Figure 1 the two methods produce broadly similar distributions. These findings suggest 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. Consequently, studies which focus on within group characteristics will be largely unaffected.

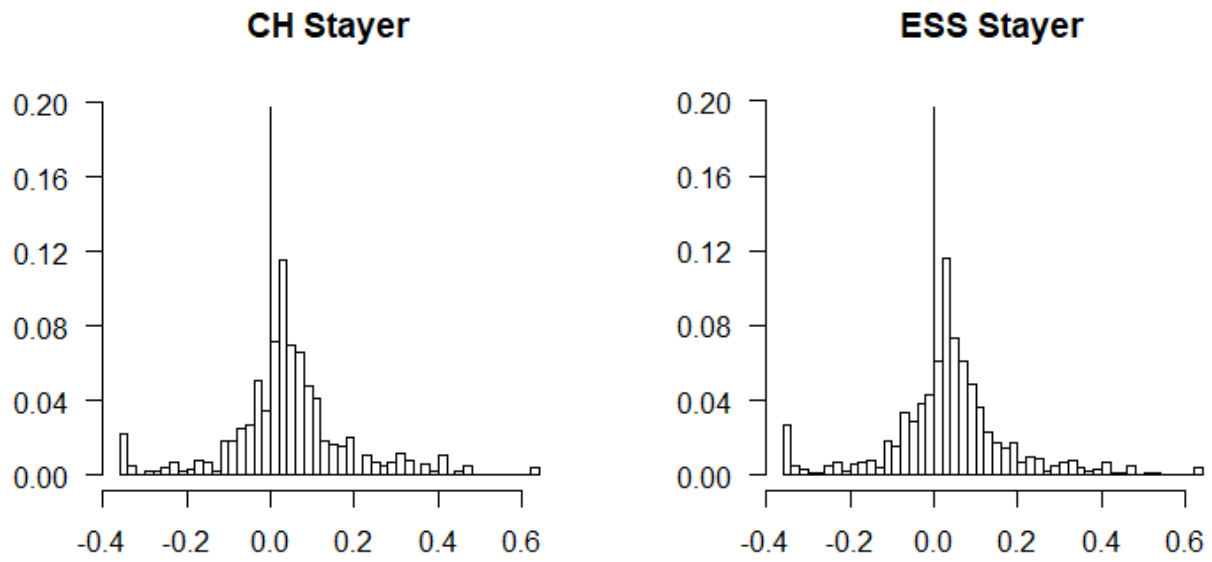

Figure 1: Distribution of change in log nominal hourly wages for US hourly workers from 2011-2012 by stayer type.

4. Wage Adjustment

Proper classification of job stayers and changers is also crucial to understanding drivers of wage adjustment over the business cycle. This application differs from wage rigidity because it focuses on comparing job stayers and changers rather than calculating the within group wage change distributions. Because classification error will change the proportion of stayers and changers in the population they will necessarily impact any attempt to decompose changes in broad measures of wage adjustment into changes driven by stayers and changers.

To demonstrate this point, I examine the potential impact of misclassification error on the decomposition employed by Daly and Hobijn (2016, henceforth DH) in their study of real wage cyclicality. They classify workers into job changers and stayers using the CH method with some additional information as in Fallick and Fleischman (2004) and Nagypál (2008). Additionally, they make use of dependent interviewing questions to expand their definition of job changing to include within employer job changes. (Daly, Hobjin, and Wiles, 2011). DH then use a modified shift-share analysis to decompose changes in the aggregate real wage into the changes due to workers in different labor states. DH find that wage changes by job stayers and job changers account for the vast majority of aggregate real wage changes and that job changers contribute more than job stayers due to their higher share of the population.

In their decomposition, each labor type can contribute to the aggregate real wage change through its share of the population and its median wage change. Following the results above, we would not expect misclassification error to impact the estimates of the median wage change of each group. Consequently, misclassification error may affect their results only through its impact on the estimates of the share of each group. The DH decomposition hinges on the share of stayers and changers. As the authors note: "[The relative importance of changers] owes to the fact that the share of full-time employed who change jobs, C , is larger than the share who stay in the same job S, while the earnings changes, or shifts, of C and S are similar

Table 3: Unweighted counts of misclassified observations in the CPS.

	ESS Stayers			ESS Changers		
Year	Total CH Stayers	Mis-classified	Misclassification rate	Total CH Changers	Mis-classified	Misclassification rate
2007-08	776	64	8%	511	394	77%
2009-10	751	39	5%	462	376	81%
2011-12	606	39	6%	455	365	80%
2013-14	627	42	7%	465	351	75%
2015-16	599	69	12%	465	348	75%
Average	671.8	50.6	8%	471.6	366.8	78%

Since DH include within employer job changes in their definition of job changers I am unable to distinguish between misclassification error and job changers who changed roles but not employers. However, as an illustrative exercise, suppose that DH had defined a job changer as a worker who changed employers, as in CH and ESS. I can then estimate how misclassification error would have impacted the share of job changers and stayers in their decomposition. Under the assumption that the additional variables they use do not introduce more measurement error, my estimates provide an upper bound on measurement error in their paper. I can then apply a correction for the misclassification using the following procedure.

I estimate how many stayers are changers and vice versa using the observed errors in the replication above. These results are reported in Table 3. I then apply the misclassification rate to the proportion of stayers and changers reported in Table 2 in Daly and Hobijn (2016) to estimate the share of full-time workers that were misclassified as job stayers and changers. Finally, I estimate the share of job stayers and changers that would have resulted from using the ESS method by adding the fraction of the population that are true positives for a given type (stayer or changer) to the fraction that are false positives of the other type. These results are presented in Table 4. This method may be used to correct for misclassification error in studies which rely on accurately estimating the share of the population that are job stayers and job changers.

After the correction, stayers now make up about 75% of the population and changers make up about 13.8% of the population. These share estimates are similar to the ESS shares presented above in Table 1. Consequently, the contribution of stayers is much higher than changers after the correction. These results also imply that changers are responsible for more of the variance of real wages.

Table 4: Estimated shares of job stayers and changers after correcting for misclassification error.

	DH Calculated share	Estimated false positives (pp)	Estimated true positives (pp)	Estimated ESS shares
Stayers	40.7	3.1	37.6	75.3
Changers	48.4	37.7	10.7	13.8

5. 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 that focus on estimating within group statistics. 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.

In studies that compare the two groups, such as studies of wage adjustment, misclassification may significantly impact the results through biased estimates of each group's share of the population. These relative shares often drive the results in studies that seek to estimate the relative impact of job changers and stayers making it necessary to correct for misclassification. In this case, it is possible to use the CH method to get estimates for each month and then correct for misclassification error using the procedure above, if you define a job changer as a worker who changes employers. This approach has the advantage of constructing a more complete time-series while mitigating the impact of misclassification error.

References

Bound, John, Charles Brown, and Nancy Mathiowetz (2001) "Measurement error in survey data" in *Handbook of econometrics*, vol. 5 by James J. Heckman and Edward Leamer, Eds., North-Holland, 3705-3843

Card, David, and Dean Hyslop (1997) "Does inflation 'grease the wheels of the labor market?'" in *Reducing inflation: Motivation and strategy* by Christina D. Romer and David H. Romer, Eds., University of Chicago Press, 71-122

Daly, Mary C., and Bart Hobijn (2016) "The intensive and extensive margins of real wage adjustment" Federal Reserve Bank of San Francisco

Daly, Mary C., Bart Hobijn, and Brian Lucking (2012) "Why has wage growth stayed strong?" *FRBSF Economic Letter* 10, no. 2

Daly, Mary C., Bart Hobijn, and Theodore Wiles (2012) "Dissecting aggregate real wage fluctuations: Individual wage growth and the composition effect" Federal Reserve Bank of San Francisco Working Paper 2011-23

Elsby, Michael W. L., Donggyun Shin, and Gary Solon (2016) "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, S249-S291

Fallick, Bruce, and Charles A. Fleischman (2004) "Employer-to-employer flows in the US labor market: The complete picture of gross worker flows" FEDS Working Paper No. 2004-34

Kambourov, Gueorgui, and Iourii Manovskii (2013) "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, 172-194.

Madrian, Brigitte C., and Lars John Lefgren (2000) "An approach to longitudinally matching Current Population Survey (CPS) respondents" *Journal of Economic and Social Measurement* **26**, no. 1, 31-62

Nagypál, Éva (2008) "Worker reallocation over the business cycle: The importance of employer-to-employer transitions" Manuscript, Northwestern University

Polivka, Anne E., and Stephen M. Miller (1998) "The CPS after the redesign: Refocusing the economic lens" in *Labor statistics measurement issues* by John Haltiwanger, Marilyn E. Manser, and Robert Topel, Eds., University of Chicago Press, 249-286