How to Catch an Outlier: A Robust Method for Hours and Earnings Estimation in the Current Employment Statistics Survey September 2019

Paige Schroeder[1]

U.S. Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington, DC 20212

Abstract

Estimates of labor statistics are susceptible to being heavily influenced by certain businesses responding to surveys of employment, hours worked, and earnings. Winsorization can be used to identify and treat influential microdata, and is a method that the Current Employment Statistics State and Area program uses to improve efficiency in employment estimates. However, for the average weekly hours and hourly earnings estimates, there is not currently a robust method to identify outliers. This is mainly due to differences in estimation techniques. While employment estimates target a population total, average weekly hours and hourly earnings estimates use ratios of two totals. Subsequently, measuring the influence of a report on the latter estimates is not as straightforward as it is for employment because of the differing complexity in estimation techniques. This paper presents four different influence functions that were developed for catching outliers in the hours and earnings data. Each one is evaluated via simulations to determine the most efficient estimator. While some influence functions performed better than others, all proved better than having no robust method at all.

Keywords: influential observations, Current Employment Statistics, hours and earnings, robust estimation

[1] Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

1. Introduction

Outliers in establishment surveys can heavily influence estimates of labor statistics, especially in smaller domains. A technique known as Winsorization can be used to improve efficiency in estimators by reducing the weight of influential observations. A variation of that technique is used for sub-national employment estimates produced by the Current Employment Statistics (CES) survey. CES State and Area publishes employment, hours, and earnings estimates for all 50 states, plus the District of Columbia, Puerto Rico, and the Virgin Islands. It also produces them for over 450 metropolitan areas and divisions, from broad to detailed industry levels. As the area and industry levels become more detailed and the domains become smaller, certain respondents can influence the estimates even more so. The employment estimator uses a method robust to outliers, which can be respondents with large survey weights, large over-the-month changes in employment, or a combination of the two. However, there is not currently a robust method for the hours and earnings estimators. For respondents reporting hours and earnings data, influential observations may also have higher overall levels of total hours worked or total payroll in addition to large survey weights and over-the-month changes in those quantities. This paper will go over proposed methods of measuring influence of hours and payroll microdata on average weekly hours and hourly earnings series. With each proposed influence function, evaluation methods and their results will be analyzed to determine how each perform.

2. CES Background

There are around 142,000 businesses and government agencies that the CES surveys each month, collecting their employment, payroll, and paid hours data. In order to be included in the sample each month, the business must report positive data for both the current month as well as the prior month. Businesses reporting both are considered to be part of what is called the matched sample.

To estimate all employment AE for the current month c, CES uses a weighted-link-relative formula [2]:

$$\widehat{AE}_c = \widehat{AE}_p * \frac{\sum_i w_i ae_{c,i}}{\sum_i w_i ae_{p,i}} + B/D \tag{1}$$

Using the matched sample, this takes the weighted sum of all employees ae in the current month divided by the weighted sum of all employees in the previous month p and multiplies that by the previous month's employment estimate. A forecasted birth-death factor B/D is added on to account for expected residual change from business births and deaths that month.

To estimate average weekly hours AWH and average hourly earnings AHE, a weighted-link-and-taper formula is used:

[2] A full list of variables and abbreviations used in this paper is available in the appendix.

$$\widehat{AWH}_c = 0.9\widehat{AWH}_p + 0.1\frac{\sum_i w_i wh_{p,i}}{\sum_i w_i ae_{p,i}} + \left(\frac{\sum_i w_i wh_{c,i}}{\sum_i w_i ae_{c,i}} - \frac{\sum_i w_i wh_{p,i}}{\sum_i w_i ae_{p,i}}\right) \quad (2)$$

$$\widehat{AHE}_c = 0.9\widehat{AHE}_p + 0.1\frac{\sum_i w_i pr_{p,i}}{\sum_i w_i wh_{p,i}} + \left(\frac{\sum_i w_i pr_{c,i}}{\sum_i w_i wh_{c,i}} - \frac{\sum_i w_i pr_{p,i}}{\sum_i w_i wh_{p,i}}\right) \quad (3)$$

where wh is the reported weekly hours and pr is the reported payroll for each respondent i. The difference link accounts for the difference of the ratios for the current month and previous month. The taper is the level to which the difference link is applied and is a weighted combination of the previous month's estimate and the previous month's sample mean in the current matched sample. This estimator ensures that month-to-month changes are driven by the difference between current and previous month means within the matched sample and that the overall level of the estimates track the overall sample average over time.

Besides the differences in estimation methods, the two types of statistics also differ because the hours and earnings series have a smaller sample size due to item response rates that are much lower than for employment. This can cause more variability in the data. And unlike employment estimates, which are benchmarked to administrative counts from the Quarterly Census of Employment and Wages, there is not a census that can be used to benchmark hours and earnings estimates, so the true population values are unknown.

3. Measuring Influence

3.1 Measuring Influence for Employment

To identify and treat outlying reports in employment, a version of Winsorization[3] is run on the matched sample. The influence each observation has on the weighted-link-relative is measured by using weighted residuals, which are simply how much each observation k deviates from the link-relative:

$$wResidual_k = w_k\left(ae_{c,k} - \frac{\sum_i w_i ae_{c,i}}{\sum_i w_i ae_{p,i}} * ae_{p,k}\right) \quad (4)$$

Influential reports are ones that have much larger weighted residuals than the other reports in the matched sample. To determine how much larger the residual would need to be compared to the others, a cut-off value is calculated using the distribution of those residuals in that series. For all observations above that cut-off value, an initial adjustment factor is calculated to bring the residual(s) back closer to the cut-off value.

When applied in estimation, the final weight adjustment factor is determined using the following rules:

[3] See Kokic and Bell, 1994

1. For certainty units (units selected with probability=1):

$$final\ adjustment = \begin{cases} 0, & initial\ adjustment < 0.5 \\ 1, & initial\ adjustment \geq 0.5 \end{cases}$$

2. For non-certainty units:

$$final\ adjustment = \begin{cases} 0, & initial\ adjustment \leq 0.3 \\ 1, & initial\ adjustment > 0.4 \\ initial\ adjustment^{4}, & otherwise \end{cases}$$

When the final adjustment is equal to one, no weight adjustment factors are applied. When the final adjustment equals zero, that observation is completely removed from the link-relative and designated as an atypical unit. Otherwise the observation is kept in the link-relative but a weight-adjustment factor known as a downweight is applied to it.

To calculate the employment estimate when atypicals and/or downweights are identified, the formula becomes:

$$\widehat{AE}_c = \left[\widehat{AE}_p - \sum_j ae^*_{p,j}\right] * \left[\frac{\sum_i w_{i*} r_i * ae_{c,i}}{\sum_i w_{i*} * r_i * ae_{p,i}}\right] + \sum_i ae^*_{c,j} + B/D \qquad (5)$$

where r is a downweight associated with report i, * is all atypical sample data, with report j being atypical in the current month.

3.2 Measuring Influence for Hours and Earnings

Formulating an influence function for hours and earnings is not as straightforward as with employment because we are now targeting population totals of two variables. Looking at just over-the-month changes in the reported averages for hours or earnings does not capture how that report will actually affect the estimate. This is because the totals for each variable are estimated first before calculating the averages. For average weekly hours, total weekly hours are estimated first and then divided by total employment after. For average hourly earnings, total earnings are estimated first, and then divided by the estimated total weekly hours.

The average weekly hours formula with atypicals and downweights applied is:

$$\widehat{AWH}_C = 0.9\widehat{AWH}_P - .9\left(\left(\frac{\sum_i w_i r_i wh_{p,i} - \sum_i w_i wh^*_{p,j}}{\sum_i w_i r_i ae_{p,i} - \sum_i w_i ae^*_{pj}}\right)\left(\frac{\widehat{AE}_P - \sum_i ae^*_{p,j}}{\widehat{AE}_P}\right) + \frac{\sum_i wh^*_{p,j}}{\widehat{AE}_P}\right) +$$
$$\left(\left(\frac{\sum_i w_i r_i wh_{c,i} - \sum_i w_i wh^*_{c,j}}{\sum_i w_i r_i ae_{c,i} - \sum_i w_i ae^*_{c,j}}\right)\left(\frac{\widehat{AE}_C - \sum_i ae^*_{c,j}}{\widehat{AE}_C}\right) + \frac{\sum_i wh^*_{c,j}}{\widehat{AE}_C}\right) \qquad (6)$$

[4] If $initial\ adjustment * sample\ weight < 1$, then $final\ adjustment = \frac{1}{sample\ weight}$

Average hourly earnings[5] is the same format but with payroll pr in place of weekly hours wh, wh in place of employment ae, and the estimate of total weekly hours $\widehat{WH}$ in place of the employment estimate $\widehat{AE}$.

The four proposed influence functions for the hours and earnings series are listed below for Average Weekly Hours. Average Hourly Earnings is in the same format like above.

Version 1:

$$Residual_k = \frac{1}{\widehat{AE}_c}\left(wh_{c,k} - ae_{c,k} * \frac{\sum_i w_i wh_{c,i}}{\sum_i w_i ae_{c,i}}\right) - \frac{0.9}{\widehat{AE}_p}\left(wh_{p,k} - ae_{p,k} * \frac{\sum_i w_i wh_{p,i}}{\sum_i w_i ae_{p,i}}\right)$$

Version 2:

$$Residual_k = -\left[\frac{\left(\frac{\sum_i w_i wh_{p,i} - w_k wh_{p,k}}{\sum_i w_i ae_{p,i} - w_k ae_{p,k}}\right)\left(\widehat{AE}_P - ae_{p,k}\right) + wh_{p,k}}{\widehat{AE}_P}\right] + \left[\frac{\left(\frac{\sum_i w_i wh_{c,i} - w_k wh_{c,k}}{\sum_i w_i ae_{c,i} - w_k ae_{c,k}}\right)\left(\widehat{AE}_c - ae_{c,k}\right) + wh_{c,k}}{\widehat{AE}_c}\right]$$

Version 3:

$$Residual_k = -0.1\widehat{AWH}_P - 0.9\left[\frac{\left(\frac{\sum_i w_i wh_{p,i} - w_k wh_{p,k}}{\sum_i w_i ae_{p,i} - w_k ae_{p,k}}\right)\left(\widehat{AE}_P - ae_{p,k}\right) + wh_{p,k}}{\widehat{AE}_P}\right] + \left[\frac{\left(\frac{\sum_i w_i wh_{c,i} - w_k wh_{c,k}}{\sum_i w_i ae_{c,i} - w_k ae_{c,k}}\right)\left(\widehat{AE}_c - ae_{c,k}\right) + wh_{c,k}}{\widehat{AE}_c}\right]$$

Version 4:

$$Residual_k = \widehat{AWH}_C - \widehat{AWH}_{C,-k}$$

[5] Full equation for AHE with atypicals and downweights applied is in the appendix.

where

$$\widehat{AWH}_C = 0.9\left(\widehat{AWH}_P - \frac{\sum_i w_i wh_{p,i}}{\sum_i w_i ae_{p,i}}\right) + \frac{\sum_i w_i wh_{c,i}}{\sum_i w_i ae_{c,i}}$$

$$\widehat{AWH}_{C,-k} = 0.9\left(\widehat{AWH}_P - \frac{\sum_i w_i wh_{p,i} - w_k wh_{p,k}}{\sum_i w_i ae_{p,i} - w_k ae_{p,k}}\right) + \frac{\sum_i w_i wh_{c,i} - w_k wh_{c,k}}{\sum_i w_i ae_{c,i} - w_k ae_{c,k}}$$

Which simplifies somewhat to:

$$Residual_k = 0.9\left(\frac{\sum_i w_i wh_{p,i} - w_k wh_{p,k}}{\sum_i w_i ae_{p,i} - w_k ae_{p,k}} - \frac{\sum_i w_i wh_{p,i}}{\sum_i w_i ae_{p,i}}\right) + \frac{\sum_i w_i wh_{c,i}}{\sum_i w_i ae_{c,i}} - \frac{\sum_i w_i wh_{c,i} - w_k wh_{c,k}}{\sum_i w_i ae_{c,i} - w_k ae_{c,k}}$$

Version 1 was the simplest and the most similar to the basic hours and earnings formulas that did not have weight-adjustment factors applied. Versions 2 and 3 were more similar to the hours and earnings formulas with weight-adjustment factors applied, with version 3 also adding in a small weight on the previous month's estimate. Version 4 took out each observation one at a time (similar to a jackknife procedure) and measured the difference of what the estimate would have been with and without that observation included.

In addition to the influence functions, several other modifications were also tested, which included:

1. Using the same guidelines as in employment (see section 3.1) when it came to calculating the final adjustments
2. Only applying downweights to non-certainty[6] units, instead of making any atypical:

$$final\ adjustment = \begin{cases} initial\ adjustment, & initial\ adjustment \leq 0.4 \\ 1, & otherwise \end{cases}$$

3. Decreasing the cut-offs for certainty units:

$$final\ adjustment = \begin{cases} 0, & initial\ adjustment < 0.4 \\ 1, & initial\ adjustment \geq 0.4 \end{cases}$$

4. Adjusting the formula for calculating the cut-off point so it was closer to the next largest residual.

Combinations of each modification were made, as well.

[6] Guidelines for certainty units remained the same, as these units should not partially represent other respondents

4. Evaluations

As stated earlier, the hours and earnings estimates do not have a census to which they can be benchmarked and so other evaluation methods were needed for evaluating the influence functions. Three of the methods used and their results are described below.

4.1 Using Extreme Observations

One of the first evaluation methods conducted was using what was called "extreme observations." First, data were pulled for what was considered to be a representative Metropolitan Statistical Area (MSA) with an average hourly earnings estimate of $25/hour, that was fairly constant over time and had no obvious outliers in the sample data. Using a given month, each influence function was run on that series to get distributions of the weighted residuals. Next, 9 different respondents were chosen from the CES survey (not belonging to that MSA) that had reported earnings data that would be expected to be outliers. These reports either had large survey weights, large over-the-month-changes in average hourly earnings, and/or high average hourly earnings. Those respondents were added to the matched sample in that series one at a time, and then in groups, as well. The influence functions were then run again and the distributions of those residuals were compared. The goal was to see if those respondents had residuals that stood out from the rest of the observations in the matched sample.

Figure 1 has the original distributions of the residuals for each version using just the existing data in that series.

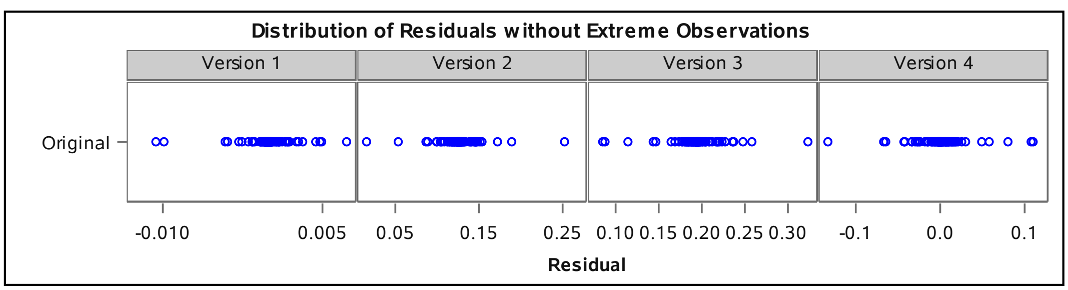

Figure 1: Distribution of residuals before adding any extreme observations

The first extreme observation added was a small employer with a high weight and a large over-the-month change in average hourly earnings. Figure 2 has what the distribution of residuals looked like before and after adding it. The red asterisk is the extreme observation.

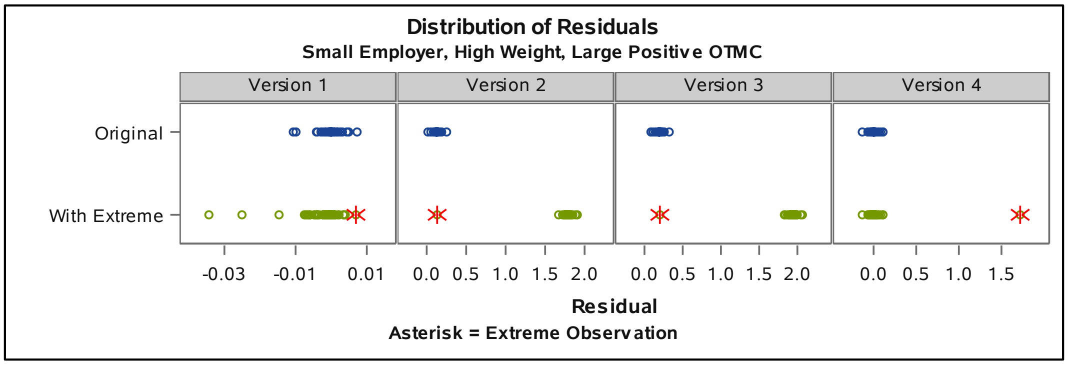

Figure 2: Distribution of residuals in same series, before and after adding an extreme observation

The extreme observation clearly stuck out from the rest of the observations in versions 2, 3, and 4. Whether the observation was to the left or right of the rest itself matters less than its distance from the rest of the distribution, only if it was clearly further away than the rest. In version 1, it was the largest right-sided residual but still not much larger than the rest.

The next extreme observation shown below in Figure 3 was a large employer with a small weight, high pay, and just a slight decrease in average hourly earnings.

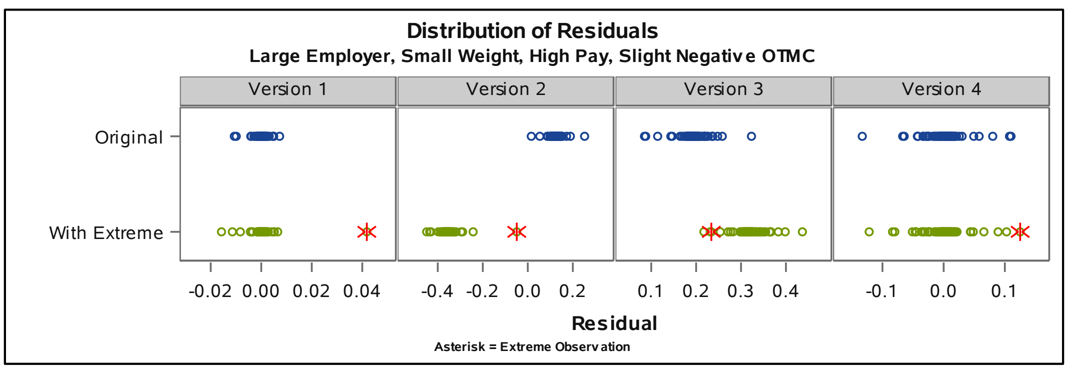

Figure 3: Distribution of residuals in same series, before and after adding an extreme observation

In figure 3, the extreme observation added here had a much larger residual compared to the rest of the observations in versions 1 and 2. Version 4 still had the largest right-sided residual but not as far away from the rest, and in version 3, the extreme observation's residual was not even the largest in this case.

The last observation added was a doctor's office with few employees, a large weight, high pay, and a slight over-the-month change in average hourly earnings.

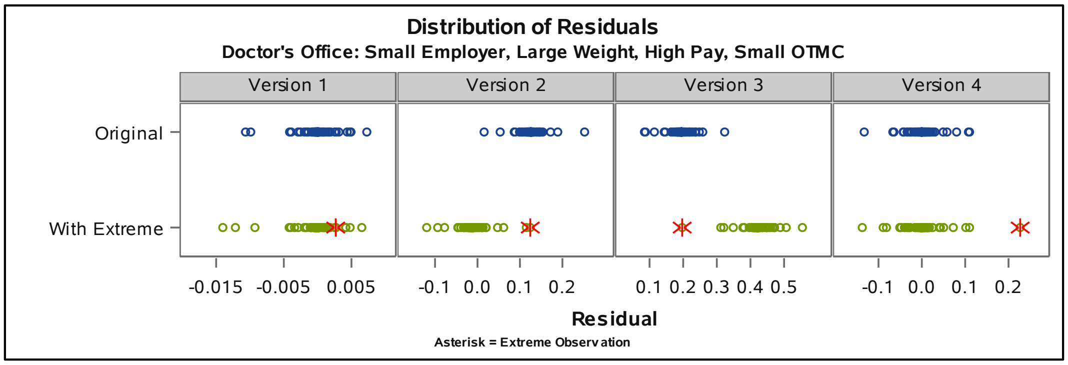

Figure 4: Distribution of residuals in same series, before and after adding an extreme observation

As seen in figure 4, versions 2, 3, and 4 performed fairly well when the doctor's office was added. Again, version 1 failed to distinguish it from the rest of the observations.

After examining the residuals, the Winsorization procedure was run to determine if observations qualified for a weight-reduction factor. Modifications to the Winsorization process affected whether they qualified, as well. While not shown, version 1 failed to pick up many of the observations as outliers; it also identified the most non-influential or non-extreme observations. Versions 2 and 4 performed the best, in that they picked up the most extreme observations as eligible for some type of weight-reduction factor.

4.2 Simulated Populations

The next evaluation method conducted was creating simulated universes from which to sample. Four different universes were constructed using seven years' worth of CES microdata, designed to match the proportion of establishment size classes in the frame from which CES samples, the Quarterly Census of Employment and Wages, which has administrative data for roughly 9 million unemployment insurance (UI) tax accounts. Each simulated universe for this evaluation was comprised of about 4,500 UI records, and simple random samples stratified by employment size were drawn from each. Two different sample sizes were used, one of size 50 and one of size 90, each being repeatedly drawn 100 times. After drawing each sample, the respondents were randomly assigned to one of either three different or two different groups, and then only the reports in each group were used to estimate each statistic. This was done to mimic the percentage of non-respondents in typical CES estimating cells. For the samples using a 33% response rate, the total number of replicates totaled 300 for each sample size, and for the samples with a 50% response rate, the total number of replicates was 200 for each sample size.

For each replicate, average weekly hours and hourly earnings were estimated over a twelve month time period for each robust version and modification, applying weight adjustment factors accordingly. As a comparison, the estimates were also calculated without any adjustments using formulas (2) and (3), meaning no version of robust was run; these unadjusted values will also be referenced as "version 0" or "v0." Several measurements of error were calculated, including the mean absolute error of the estimate, which was summed over all twelve months, the mean squared error of the over-the-month change, and the ratio of the root mean squared error of the over-the-month change for each robust version to the unadjusted root mean squared error of the over-the-month change. The absolute errors of the over-the-month changes were also calculated, and figures 5 and 6 display boxplots of these errors for all replicates by the different sample sizes and response rates.

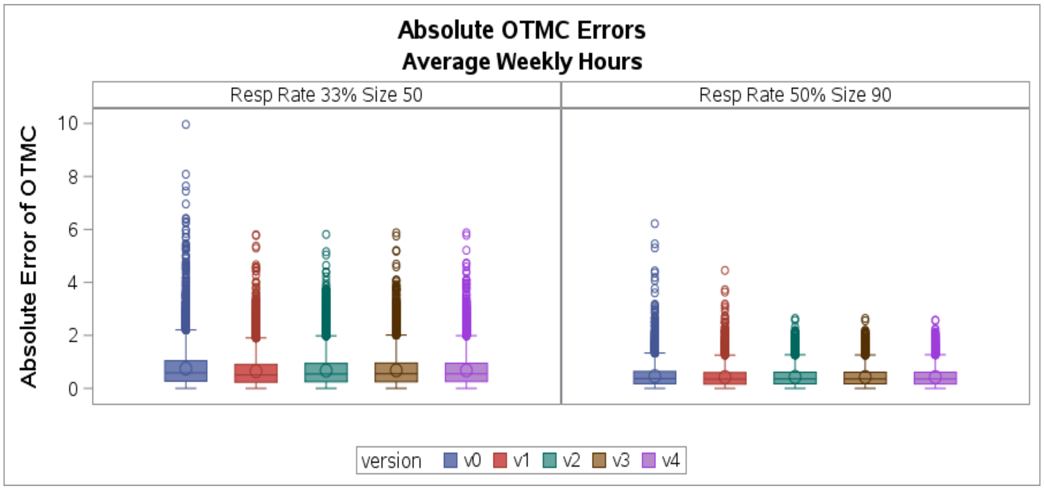

Figure 5: Boxplot of absolute OTMCs for average weekly hours by response rate and sample size

	Average Weekly Hours					
	Response Rate 33%, Sample Size 50			Response Rate 50%, Sample Size 90		
	Sum MAE of Estimate	MSE of OTMC	RMSE Ratio of OTMC	Sum MAE of Estimate	MSE of OTMC	RMSE Ratio of OTMC
Unadjusted	13.459	1.022	.	7.876	0.372	.
Version 1	**12.837**	**0.719**	**85.275**	**7.644**	0.308	92.843
Version 2	13.433	0.777	88.353	7.717	0.303	92.037
Version 3	13.970	0.785	88.878	7.801	0.300	91.573
Version 4	13.748	0.784	88.789	7.823	**0.298**	**91.369**

Table 1: Summary of errors for average weekly hours by response rate and sample size

Starting with average weekly hours, Figure 5 plots the absolute errors of the over-the-month changes using a 33% response rate and sample size of 50, as well as a 50% response rate and a sample size of 90. In some replicates, the unadjusted estimate had very large over-the-month change errors that were reduced when applying the weight-adjustment factors from the robust versions. This means that all of the robust versions identified outlier(s) in that replicate and applied weight adjustment factors accordingly to bring the replicate's estimate closer to the "true" population value. For average weekly hours, version 1 had the smallest errors, whereas with a higher response rate and larger sample, version 4 performed best under most measures.

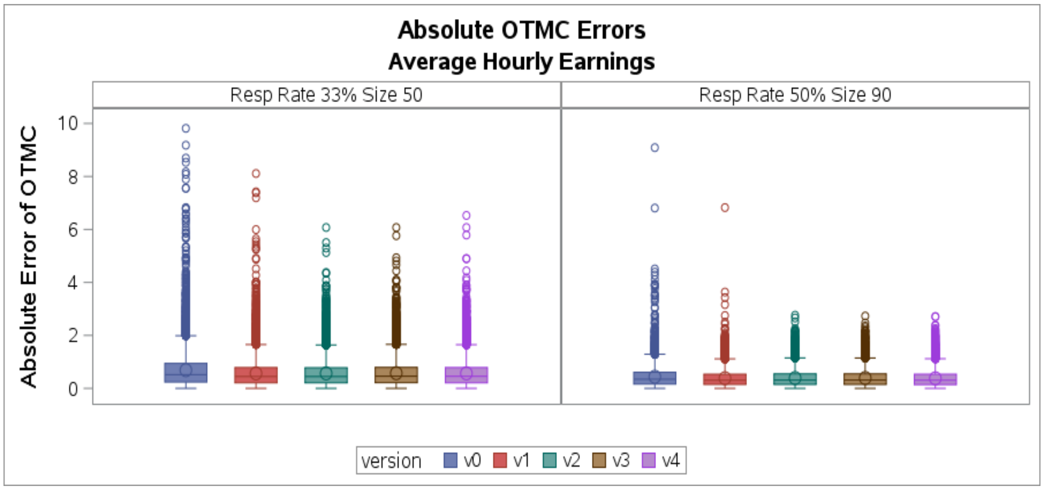

Figure 6: Boxplot of absolute OTMCs for average hourly earnings by response rate and sample size

	Average Hourly Earnings					
	Response Rate 33%, Sample Size 50			Response Rate 50%, Sample Size 90		
	Sum MAE of Estimate	MSE of OTMC	RMSE Ratio of OTMC	Sum MAE of Estimate	MSE of OTMC	RMSE Ratio of OTMC
Unadjusted	24.831	0.928	.	15.182	0.363	.
Version 1	**22.941**	0.572	79.944	14.571	0.258	86.081
Version 2	23.200	**0.564**	**79.196**	14.821	0.260	85.580
Version 3	23.123	0.580	80.204	**14.345**	0.263	86.101
Version 4	22.937	0.567	79.476	14.442	**0.247**	**83.772**

Table 2: Summary of errors for average hourly earnings by response rate and sample size

For average hourly earnings, version 2 performed best under the low response and sample size specifications, while version 4 once again had the smallest errors when the response rate was higher and the sample size larger.

4.3 Results for Published Series

Lastly, results are presented for the evaluations based on published CES series. Each variant of atypical and downweight identification method was applied to real matched samples and respective modified estimates were created for the period starting in 2012 up through 2018. Currently, the time series were evaluated based on reasonableness, although other methods of analysis are also being conducted. Particular attention was paid to series where subject matter experts had previously raised concerns about estimate quality. The figures below present some of the series that had quality concerns. The solid lines are the currently published data, and the dashed lines are what the time series would look like with each version of robust being applied.

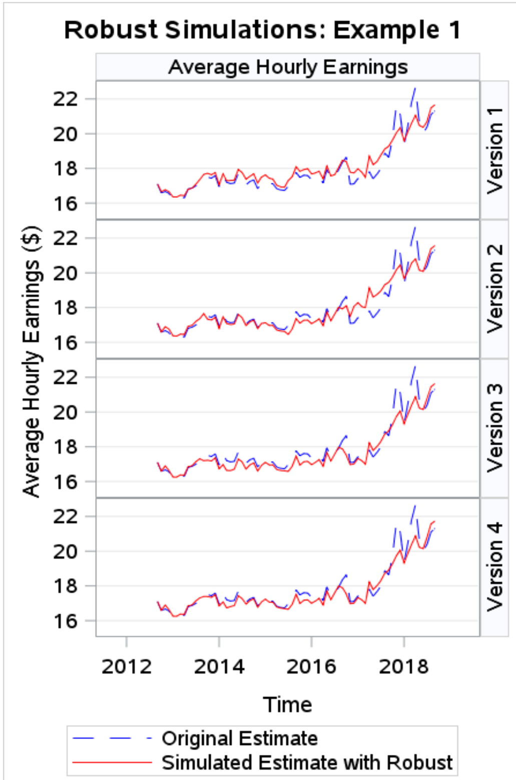

Figure 7: Average Hourly Earnings, example 1

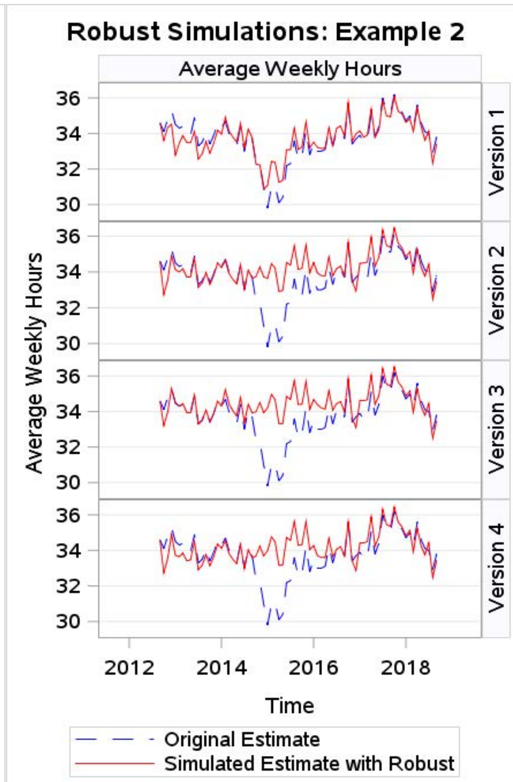

Figure 8: Average Weekly Hours, example 2

The first example in figure 7 is for average hourly earnings. This was a series with quality concerns regarding the large movements in 2018, which seemed unprecedented. When the estimate was simulated using the actual matched sample data for each version of robust, all versions lessened the over-the-month changes so they weren't as extreme. Version 2 also smoothed out the drop in 2017, as well.

The next example shown in figure 8 is the currently published average weekly hours for another series, which shows a large dip in 2015 that does not return to its previous levels for several months. When the estimate was simulated using each robust version, the drop was smoothed out in versions 2, 3, and 4. Version 1 lessened that drop but did not smooth it out completely.

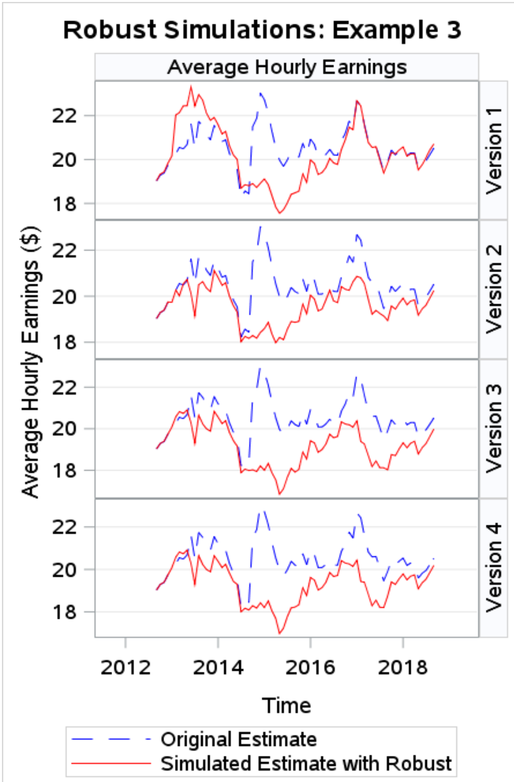

Figure 9: Average Hourly Earnings, example 3

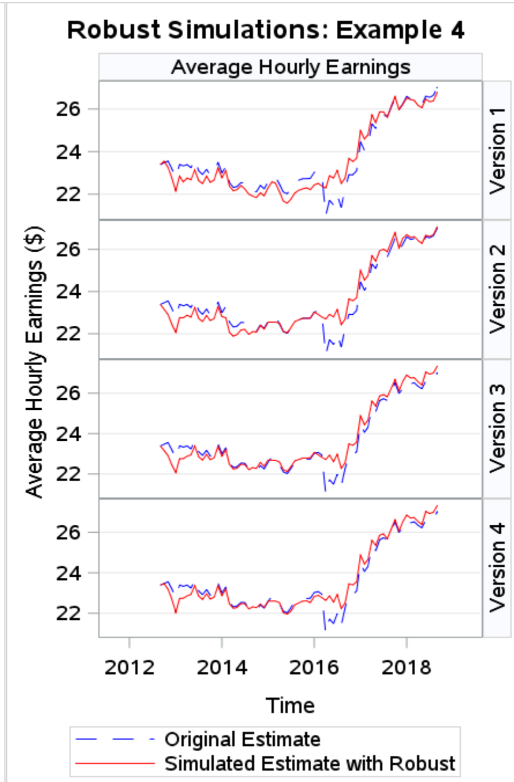

Figure 10: Average Hourly Earnings, example 4

Moving to the third example, figure 9 shows a published average hourly earnings series with a large increase in late 2014, and a similar but less extreme increase in 2017. Each version greatly reduced or prevented that increase, but versions 1, 3, and 4 still showed a decrease (though not as extreme) several months later from when it was returning to its previous levels. Version 2 was able to smooth out that decrease, though. Also to note is in version 1, the simulation actually increased the estimate in 2013, making the jump there much larger and less reasonable.

In the last example, Figure 10 shows each version smoothing out a temporary drop in average hourly earnings in 2016. However, each simulation also caused a larger drop at the beginning of 2013, which did not seem reasonable either.

5. Conclusions

All robust estimation methods considered are an improvement to having no robust method at all. In identifying the extreme observations that were added to the series, version 1 did not perform as well as the others. When creating estimates from a simulated population, all versions produced fairly similar, positive results; the average absolute errors of the over-the-month changes for all versions were less than they were without robust. Then, when estimates were simulated with actual matched sample data for each different influence function, there was a dramatic improvement in many series that had quality concerns. Again, version 1 did not perform as well as the others here, though.

The final version, along with any modifications, is still being decided upon. The performances for average hourly earnings are being more closely examined, though, since that is not bounded like average weekly hours is and is more likely to have more outliers in the microdata. After deciding on the final version, the robust hours and earnings procedure is intended for use in production of CES State and Area estimates, similar to the process for employment estimation.

6. References

Bureau of Labor Statistics [cited September 30, 2019], “Employment, Hours, and Earnings from the Establishment Survey,” *BLS Handbook of Methods* (chap. 2), Washington, D.C.: U.S. Department of Labor. Available at https://www.bls.gov/opub/hom/pdf/ces-20110307.pdf.

Gershunskaya, J. and Huff, L. (2004). “Outlier Detection and Treatment in the Current Employment Statistics Survey,” Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kokic, P. N. and P.A. Bell (1994). “Optimal winsorizing cutoffs for a stratified finite population estimator,” Journal of Official Statistics, 10, 419-435.

West, S., Kratzke, D., and Grden, P. (1997). “Estimators for Average Hourly Earnings and Average Weekly Hours for the Current Employment Statistics Survey,” Proceedings of the Section on Survey Research Methods, American Statistical Association.

Acknowledgments to Steven M. Mance and Kate Hubbard for the help with developing the influence functions presented in this paper.

7. Appendix

7.1 Variable definitions and abbreviations

Variable	Description
c	Current month
p	Previous month
$\widehat{AE}$	Estimated employment for all employees
$\widehat{AWH}$	Estimated average weekly hours
$\widehat{AHE}$	Estimated average hourly earnings
$\widehat{WH}$	Estimated total weekly hours
B/D	Net birth/death factor
ae	Reported all employees
w	Weight associated with a CES report
i	A CES report in the matched sample
j	A CES report in the matched sample where the current month is atypical
$*$	Atypical sample data
r	A downweight (or weight-adjustment factor not equal to 0) associated with a CES report
wh	Reported weekly hours
pr	Reported weekly payroll

7.2 Average Hourly Earnings with atypicals and downweights applied

$$\widehat{AHE}_C = 0.9\widehat{AHE}_P - .9\left(\left(\frac{\sum_i w_i r_i pr_{p,i} - \sum_i w_i pr^*_{p,j}}{\sum_i w_i r_i wh_{p,i} - \sum_i w_i wh^*_{pj}}\right)\left(\frac{\widehat{WH}_p - \sum_i wh^*_{p,j}}{\widehat{WH}_p}\right) + \frac{\sum_i pr^*_{p,j}}{\widehat{WH}_p}\right)$$
$$+\left(\left(\frac{\sum_i w_i r_i pr_{c,i} - \sum_i w_i pr^*_{c,j}}{\sum_i w_i r_i wh_{c,i} - \sum_i w_i wh^*_{c,j}}\right)\left(\frac{\widehat{WH}_c - \sum_i wh^*_{c,j}}{\widehat{WH}_c}\right) + \frac{\sum_i pr^*_{c,j}}{\widehat{WH}_c}\right)$$

7.3 More simulations