Measures of Variance Across CPI Populations November 2019

Joshua Klick, Owen Shoemaker
U. S. Bureau of Labor Statistics
2 Massachusetts Avenue NE, Washington DC 20212
klick.joshua@bls.gov, shoemaker.owen@bls.gov

Abstract
The Consumer Price Index (CPI) estimation is divided into a lower level calculation of basic indexes measuring price change within item/area cells, and then an upper level aggregation of price change across cells for a target population. CPIs are currently calculated for three populations: the urban population represented by the CPI-U, the wage earner and clerical worker population represented by the CPI-W, and the elderly population represented by the CPI-E. Basic level indexes and weights serve as inputs to upper level estimation, where the weights vary by population. However, the same basic indexes are used across populations. This paper evaluates weight and price changes across populations to define differences of estimates, and then evaluates how the respective population differences impact the measure of price change. In order to measure significant differences between the CPI-U and other populations, this paper proposes a new standard error measurement methodology. The methodology uses a Jackknife approach at the area level of geographical aggregation comparable to the special (SRC) item categories published variance estimates as an alternative to the published Stratified Random Groups (SRG) method replicate based variance estimates.[1]

Key Words: Consumer Price Index, CPI-U, CPI-W, CPI-E, Variance, Jackknife

1. Background

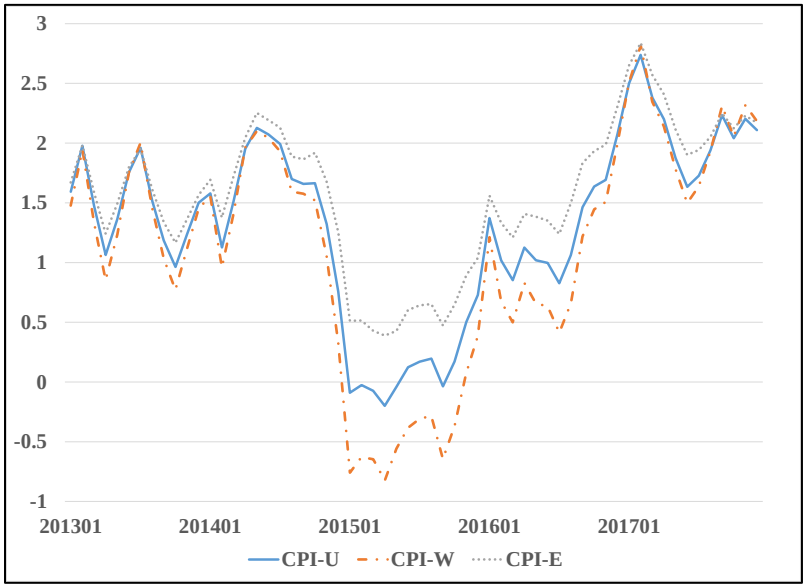
The CPI is defined as, "a measure of average change in prices paid by urban consumers for a fixed market basket of goods and services."[1] The Office of Management and Budget (OMB) designates the CPI as a Principal Federal Economic Indicator that describes the general direction of the economy, and includes a publication schedule for the year.[2] There are a number of CPI products that measure consumer inflation, a rise in the level of prices, or consumer deflation, a decline in the level of prices. This paper will focus on population based products: the CPI-U for the urban population (U), the CPI-W for the wage earner population (W), and the CPI-E for the Experimental Elderly population (E). Use of these CPI products by the Federal Government is summarized below:[3]

[1] See Bureau of Labor Statistics Handbook of Methods Chapter 17.
[2] See Office of Management and Budget, Statistical Policy Directive No. 3.
[3] See Office of Management and Budget, Federal Registrar/Vol. 84, No. 88/ Tuesday May 7, 2019/Notices.

Table 1. Summary of CPI Population Based Products used by the Federal Government

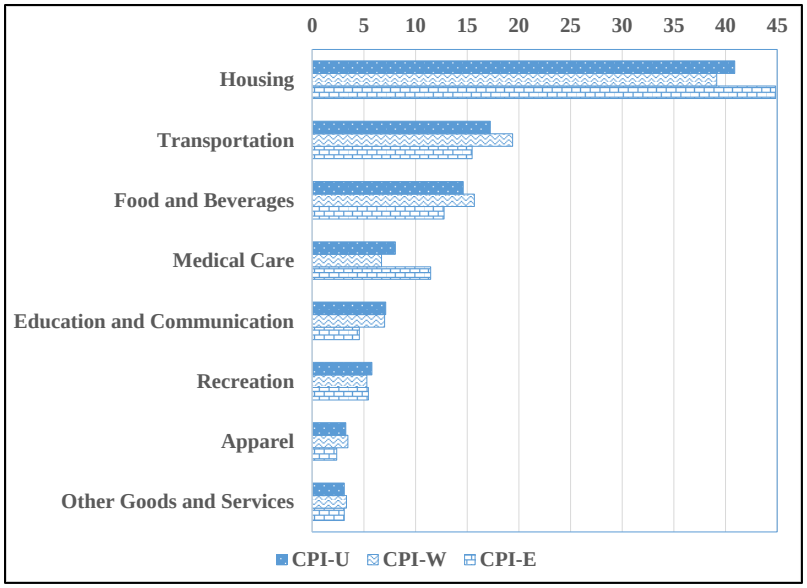
CPI-U	Official poverty thresholds, Treasury inflation-indexed securities, Deflate nominal values in a variety of measures
CPI-W	Social Security benefits, Federal transfer payments
CPI-E	Experimental index not used for official purposes

Why does the CPI publish indexes for multiple populations if the market basket is fixed and the price change is based on urban consumers? Because price change is different across populations due to the respective weights from the Consumer Expenditure survey.[1] Graph 1 displays the 12 month change for the three populations. While there are differences, there are currently no measures of standard errors across populations to understand whether these differences are statistically significant. Examples of item differences across populations are summarized in Graph 2.For example, the budget share of the urban population for recreation items is greater than the wage earner and elderly populations, the wage earner budget share for transportation items is greater than the urban and elderly shares, and the elderly shares for medical care are higher than the urban and wage earners.[4] The Lowe price index formula published by the CPI assumes consumers exhibit fixed quantity preferences in response to price increases, and so if prices increase therefore the index increases. Price change affects populations differently depending on the relative importance of their respective budget share. The question for this project is whether the CPI US All Items price change for the urban population is significantly different than wage earner and elderly subgroup populations.



Graph 1. CPI-U, CPI-W, and CPI-E 12 month change

[4] See CPI Biennial Expenditure Weight Fact Book, 2013-2014. Internal BLS report dated 2016.



Graph 2. 2013-2014 Biennial expenditure weight relative importance across populations-percent of all items

The following reviews CPI index construction, and differences across populations. Next, the crude Jackknife variance estimation methodology, based on area level geographical strata, is compared to the published variance estimates from the Sample Random Group (SRG) methodology. This paper concludes by discussing limitations and applications, which includes describing proposed circumstances when Jackknife variance estimates are an acceptable alternative to SRG.

2. Overview

2.1 Components of CPI products

There are four components that define the above CPI products: consumption goods and services, a geographic sample, a defined population, and an index aggregation method. The first three components depend on eligibility, and the last component depends on the index estimation formula used for aggregation.

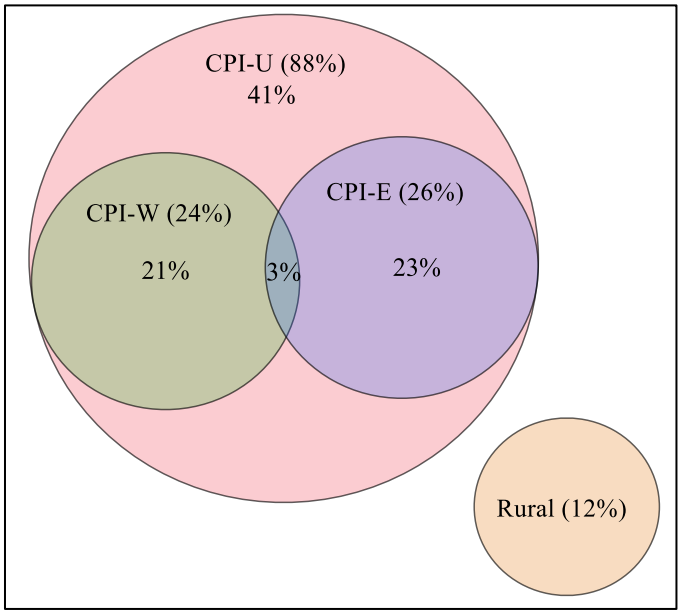
1. Consumption goods and services that define the CPI elementary item strata consist of general purchases that occur on a daily basis and exclude taxes and investments. In order to exclude investments related to owner occupied housing and related services, the CPI uses a rental equivalence value, which is a common methodological approach across national level statistical agencies.
2. The geographic sample that defines the CPI elementary area strata are based on a selection of 87 urban PSUs from the 1990 Census sample design, which aggregate to 38 areas, for the time period of this paper January 2012 to December 2017. As of January 2018 the CPI geographic sample was revised to a 2010 Census sample design based on a selection of 75 urban PSUs, which now aggregate to 32 areas.
3. CPI products are based on U, W, and E eligible populations. The U reference population, which excludes rural households, farm and military households, nursing home residents and people in prisons. The W is a subset of the U where the majority of household income is from eligible wage earner occupations, and at least one household member was employed at least 37 weeks for an eligible wage earner occupation for the previous year.

The E is a subset of the U where the reference person or spouse of the household is 62 years of age or older.

4. The CPI aggregation method can be divided into a first stage elementary or basic level estimation, and second stage aggregate level estimation. At the first stage, price quotes and household expenditures are combined to calculate basic indexes. At the second stage, basic indexes are aggregated using the Lowe formula, also called the modified Laspeyres formula. This combination of elementary level price indexes and weights create aggregate indexes representative of a defined population. The W and E are reweighted versions of the U elementary indexes. The scope of this paper is limited to second stage aggregation for the CPI-U, CPI-W, and CPI-W.[5]

2.2 Differences across populations

Differences across populations can be summarized in terms of size, household characteristics, and expenditure weight relative importance as displayed in Graph 2. For the 2013-2014 period (the biennial expenditure period used to calculate CPIs from January 2016 to December 2017), there were approximately 126 million households as measured by the CE Interview survey. The proportion of households for U, W, and E populations is displayed in Graph 3. The U reference population was 88%, and the rural population not used was 12%. The W proportion was 24% and the E proportion was 26%, with a 3% overlap between these subgroup populations.



Graph 3. Proportion of CPI 2013-2014 CE Interview Households

Demographic and economic characteristics are summarized in Tables 2 and 3. In terms of demographic characteristics, the U has an average family size per household of 2.5, the W slightly larger at 2.9, and the E less at 1.8. As expected, the E exhibits the greatest average age of 71, the U at 50, and the W is the lowest at 44 due in part to the largest average number of children at 1 relative to the other populations.

[5] The proposed Jackknife variance methodology could be extended to additional CPI products such as the final Chained-CPI-U, initial Chained-CPI-U, or CPI-U Research Series which implements methodological improvements beginning in 1978 to create a continuous methodologically consistent series.

The U exhibits the largest average income at $67,000 and average expenditure at $49,411. The E has the smallest average income at $53,000, but the mean expenditure of $48,566 is only about $850 less than the U. The W displays an average income $3,000 greater than the E, but the least average expenditure at $42,095. The standard error of the E is the larger than the W, but the coefficient of variation is smaller than the W due to the higher average annualized expenditure.[6]

Table 2. Demographic Characteristics based on the 2013-2014 CE Interview

Average	CPI-U	CPI-W	CPI-E
Family Size	2.5	2.9	1.8
Age	50.6	43.9	71.4
Children	0.7	1.0	0.2
Elderly	0.4	0.1	1.1
Earner	1.3	1.7	0.6

Table 3. Economic Characteristics based on the 2013-2014 CE Interview in terms of dollars

	CPI-U	CPI-W	CPI-E
Rounded Mean Income	67,000	56,000	53,000
Annualized Expenditure			
Mean	49,411	42,095	48,566
Standard Error	556	679	731
Coefficient of Variation	1.13	1.61	1.51

3. Comparing SRG to Crude Jackknife Variance Estimation Methodology

The CPI publishes standard errors of the 12 month change for the U based on SRG methodology as displayed in the formulas in the Appendix, which is based on the cost weight to weight each of contributing components correctly. Given that the CPI updates aggregation weights for January of even years, a bridging factor is processed for even years so that the 12 month changes are equivalent.1

For indexes through December 2017, SRG employs Jackknife methodology and is considered a middling unbiased estimator because it measures variance within strata for self-representing PSUs, defined as 31 – 1 = 30 degrees of freedom, and across clusters for non-self-representing PSUs, defined as 27-1=26 degrees of freedom, for a total of 56 PSU degrees of freedom. The number of inputs for SRG variance estimates can be summarized as the sum of 83 replicates, 32 Areas, and one US aggregate for a total of 116 components.

In contrast, the proposed crude Jackknife methodology is an outer upwardly biased estimator because it measures variance across area level strata, which results in 38 – 1 degrees of freedom. A single area is deleted for each replicate consistent with SRC processing as described below:

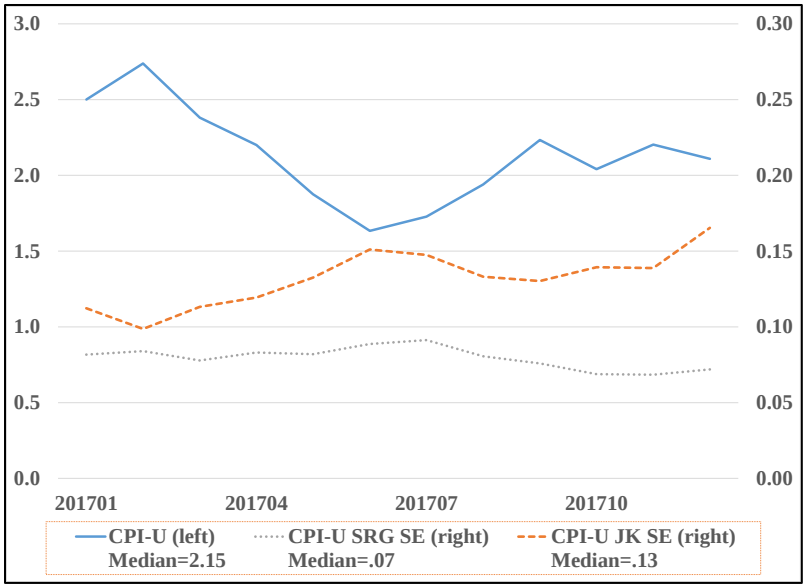
[6] Standard Error estimates based on CE balanced repeated replication method from the "Standard Errors in the 2016 Consumer Expenditure Survey" by David Swanson.

> BLS computes index series for 85 special (SRC) item categories, which are below the item stratum level and thus do not have accompanying replicate index values. (CE weights are produced only down to the item-stratum level in each index area.) The stratified random groups methodology requires a replicate structure. So, for these SRC items (such as butter or pork or new cars), an alternative variance estimation method is needed. Given the availability (at the regional and higher area levels) of independent estimates for these SRC items, the jackknife variance estimation methodology can be employed. Each area full sample cost weight can be subtracted from the all-area full sample cost weight to provide a jackknife replicate estimate. By taking the ratio of these replicate cost weight estimates at times t and t – k, subtracting 1, and multiplying by 100, one obtains the required jackknife replicate percent change value. (For the U.S. city average special item estimates, there are 38 independent index areas, and so 38 jackknife replicate estimates to work with.) (CPI Handbook of Methods 39)

The number of inputs for the crude Jackknife variance estimates can be summarized as the sum of 38 Areas and one US aggregate for a total of 39 components, or about one third of that required for SRG. In the formulas below to calculate the crude Jackknife variance estimate, item references are excluded.

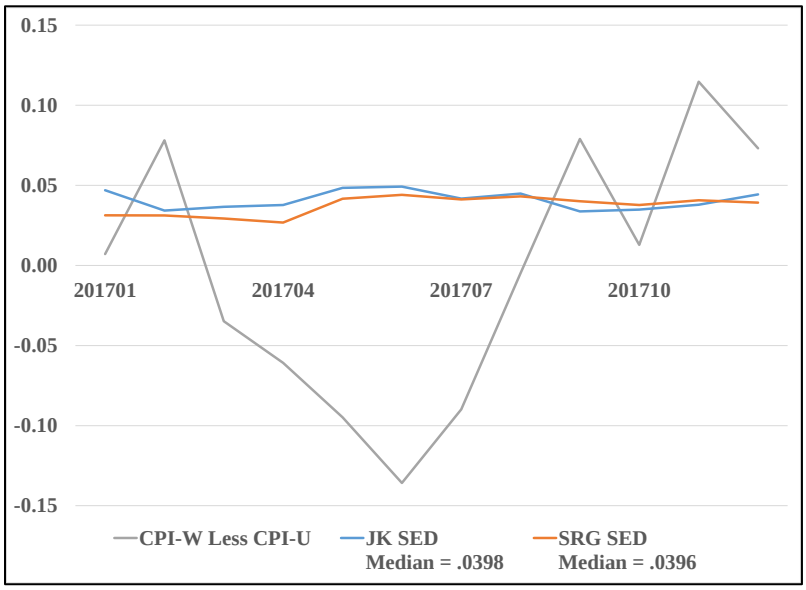
One advantage of this methodology relative to SRG is that Jackknife requires less elementary inputs as derived from second stage aggregation. This approach is used for SRCs for published variance estimates because currently the system calculates replicate indexes at the elementary item level, which are limited to calculating SRG variances estimates for elementary and aggregate item level indexes. For this research the crude Jackknife methodology is used to evaluate the US City Average All Items aggregate standard error difference of the U reference population relative to the W and E subgroup populations.

For a proof of concept consider that the 12 month change is about 2% for 2017 as displayed in Graph 4 referenced via the right vertical axis. The 12 month change is calculated to remove the index additive effects relative to the base period. The standard error referenced via the left vertical axis for SRG is trending level at median value of about 0.07%, compared to that of the crude Jackknife trending moderately upward with a median value of about 0.13%, indicative of the upward bias. These results provided motivation to extend the analysis of crude Jackknife in terms of analysis of significant differences of the resulting standard error difference across populations.

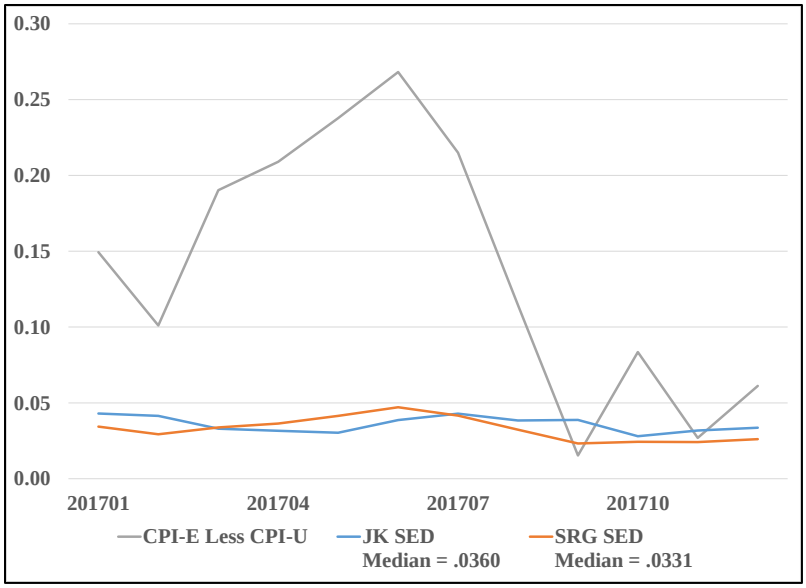


Graph 4. CPI-U 12 month change, with SRG and crude Jackknife standard error estimates

When evaluating the difference of W less U in Graph 5 and E Less U in Graph 6, and corresponding standard error difference values, the difference between crude Jackknife and SRG methods was further reduced relative to the gap displayed in Graph 4. Crude Jackknife standard error difference is biased, primarily upwards, to SRG for about ¾ of the months below for the W and E subgroup relative to U reference populations. Results below indicate that crude Jackknife is a rough approximation of SRG; however, these methods differ in terms of measuring variance across strata for the defined degrees of freedom.



Graph 5. CPI-W Less CPI-U 12 Month Change, with SRG and crude Jackknife standard error differences estimates

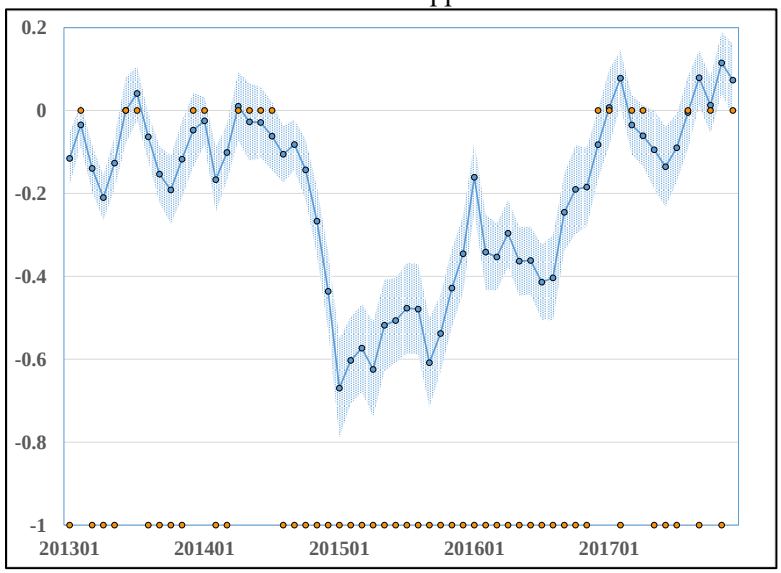


Graph 6. CPI-E Less CPI-U 12 Month Change, with SRG and crude Jackknife standard error differences estimates

4. Results

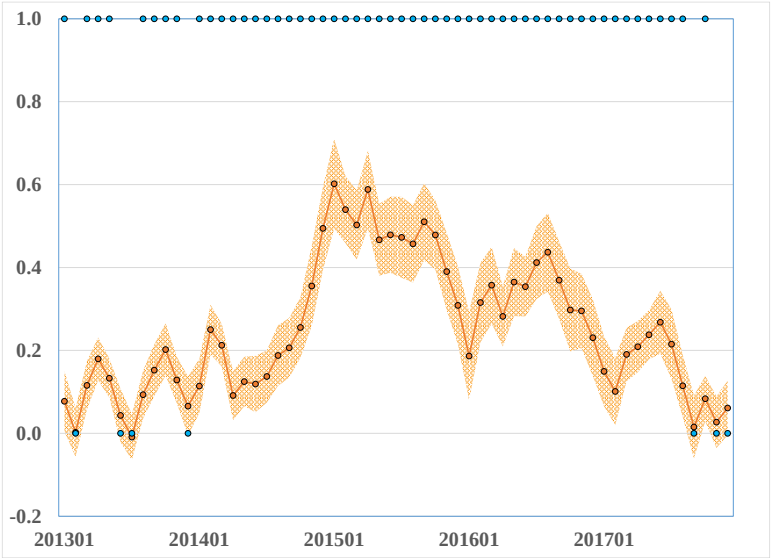
The 12 month change of the U is greater than the W for 51 out of 60 months as displayed in Graph 1. Six of the 9 months where the W is greater than the U occurred in 2017. A noticeable gap occurred in 2015 following a decline before and after this period. General trends of the all items 12 month change is similar to the energy category, and specifically gasoline, though at much higher and lower values.

The 12 month change of the W less the U is presented in Graph 7 below, where the significant difference flag is displayed as 0 not significant, and -1 as significant, and the 95% confidence interval are both based on the crude Jackknife standard error difference. As the 12 month change W less U moves away from 0 the more likely it will result in a significant difference. It is also worth noting that there are exceptions where the W is greater than the U, which may also result in a significant difference. Annualized results based on the median values are summarized in appendix A1.



Graph 7. CPI-W Less CPI-U 12 Month Percent Change with significant difference flag and 95% confidence interval of crude Jackknife standard error difference

The 12 month change of the U is less than the E for 59 out of 60 months as displayed in Graph 1. The 12 month change of the E less the U is presented in Graph 8 below, where the significant difference flag is now displayed as 0 not significant, and 1 as significant. A continuous significant difference occurs from February 2014 to August 2017. Annualized results based on the median values are summarized in appendix A2.



Graph 8. CPI-E Less CPI-U 12 Month Percent Change with significant difference flag and 95% confidence interval of crude Jackknife standard error difference

An overall summary of the above estimates are presented in Table 4, and median annual summaries are displayed in Tables 5 and 6. The W less U median was -.14, and the E less the U median was .21 consistent with Graphs 7 and 8. The median standard error difference between the W less U and E less U are nearly equal, with the latter slightly less.

One noticeable result is that the number of months of significant differences for the E less U is 10 greater than the W less U. For 2017, the W tracked the U more closely than the E, which produced more significant differences for the E. Decomposing the cause and effect of price change is beyond the scope of this paper, but for January 2017 as an example, the Transportation 12 month change for the E was higher than the U and W, potentially causing the resulting difference between the W and E.

Table 4. Summary of CPI population variance estimates 2013-2017

	CPI-U	CPI-W	CPI-E
Median:			
12 Month Change	1.488	1.373	1.641
Standard Error	0.130	0.130	0.140
Less CPI-U Difference		-0.141	0.214
Less CPI-U Difference Standard Error		0.043	0.038
Number of Months of Significant Differences		43	53
Percent of Months of Significant Differences		72%	88%

Table 5. CPI-U and CPI-W Median Annual Variance Estimates

YEAR	CPI-U	U-SE	CPI-W	W-SE	CPI-W Less CPI-U	SE Difference	RHO	Significant Differences
2013	1.488	0.100	1.394	0.113	-0.093	0.033	0.962	8
2014	1.661	0.104	1.565	0.095	-0.096	0.041	0.945	7
2015	0.049	0.198	-0.470	0.206	-0.520	0.054	0.969	12
2016	1.094	0.142	0.752	0.130	-0.341	0.045	0.950	11
2017	2.154	0.133	2.161	0.143	0.006	0.040	0.952	5

Table 6. CPI-U and CPI-E Median Annual Variance Estimates

YEAR	CPI-U	U-SE	CPI-E	E-SE	CPI-E Less CPI-U	SE Difference	RHO	Significant Differences
2013	1.488	0.100	1.578	0.097	0.091	0.030	0.957	8
2014	1.661	0.104	1.876	0.113	0.215	0.034	0.951	12
2015	0.049	0.198	0.559	0.211	0.509	0.047	0.977	12
2016	1.094	0.142	1.453	0.161	0.359	0.047	0.966	12
2017	2.154	0.133	2.200	0.138	0.046	0.036	0.964	9

5. Conclusion

Limitations of the crude Jackknife standard error variance methodology include: a. upwardly biased relative to the unbiased SRG variance estimates for individual population indexes and primarily upwardly biased relative to SRG for population difference, b. requires consistent elementary geography across the time periods of the 12 month change, and c. above results are limited to the US all items aggregate where preliminary variance estimates at lower levels are more volatile than SRG.

Acknowledging these limitations – crude Jackknife methodology provides an approximation of variance that is biased relative to SRG, where the bias is due to mismeasurement of variance across strata. For individual populations the crude Jackknife variance estimate is approximately double that of SRG. For differences across populations, the crude Jackknife variance estimate is approximately equivalent to that of SRG with a small upward bias for ¾ of months of 2017. Crude Jackknife is less complex than SRG, and is acceptable published measure of variance for SRCs, as well as would readily permit variance estimates for potential new index products based on lower levels of item specification such as Entry Level Item (ELIs), Clusters, Item Specifications, Item Brands, Outlets, etc. Crude Jackknife methodology also allows for efficient analysis of subgroup populations, particularly if analysis is limited to the US all items aggregate.

The results above provide a defined measure of bias of the crude Jackknife relative to SRG variance estimates in terms of standard error and standard error of difference estimates. Users will need to determine the amount of acceptable bias relative to the complexity and data available to determine if crude Jackknife is a feasible alternative to SRG. Potential applications of crude Jackknife beyond the above subgroup relative to reference populations include analysis of alternative definitions of the U reference population, analysis of alternative definitions of the W and E subgroup populations or other subgroup populations such as income quintiles. Related applications of crude Jackknife could extend

to evaluating the frequency of weight revisions, and evaluate modifying the frequency of weight revisions for respective subgroup populations.

Acknowledgements

Thanks to Anya Stockburger and BLS staff for helpful comments and questions. Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of the Labor Statistics.

References

Bureau of Labor Statistics (2019): BLS Handbook of Methods, Chapter 17, "The Consumer Price Index". Available at https://www.bls.gov/opub/hom/pdf/cpihom.pdf.

Bureau of Labor Statistics, "Fact Book Materials for the 2013-2014 Biennial Expenditure Weight Fact". Internal BLS report dated 2016.

Bureau of Labor Statistics, Fitzgerald, Jenny et. al. (2014). "Recommendations to Replace Stratified Random Groups Variance Methodology with Jackknife Variance Methodology in the CPI. Internal BLS Memorandum dated 2014.

ILO 2004, Consumer Price Manual: Theory and Practice, International Labour Organization, International Monetary Fund, Organization for Economic Co-operation and Development, Statistical Office of the European Communities, United Nations, The International Bank for Reconstruction and Development, The World Bank: Geneva.

Office of Management and Budget, Statistical Policy Directive No. 3. Available at https://www.bls.gov/bls/statistical-policy-directive-3.pdf.(1985)

Office of Management and Budget, Federal Registrar/ Vol. 84, No. 88/Tuesday, May 7, 2019/ Notices. Request for Comment on the Consumer Inflation Measures Produced by Federal Statistical Agencies. Available at https://www.govinfo.gov/content/pkg/FR-2019-05-07/pdf/2019-09106.pdf.

Schultze, C.L. and C. Mackie (eds.). 2002. "At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indices" (Washington, D.C.: National Academy Press).

Shoemaker, Owen. "10 Years of Comparative Results: Chained C-CPI-U vs. Regular CPI-U." American Statistical Association, Survey Methods Research Methods Section, Proceedings 2013. Available at https://www.bls.gov/osmr/research-papers/2013/pdf/st130020.pdf. (2013).

Shoemaker, Owen. "Variance Estimates for Price Changes in the Consumer Price Index January-December 2016". Available at https://www.bls.gov/cpi/tables/variance-estimates/2016.pdf. (2017).

Swanson, David. "Standard Errors in the 2016 Consumer Expenditure Survey". Available at https://www.bls.gov/cex/ce-se-2016.pdf. (2017).

Wolter, Kirk (2017). "Introduction to Variance Estimation" (Springer).

Appendix

A. Cost Weight for U, W, and E: $CW(A,f,t) = IX(A,f,t) \times AGGWT(A,f,t)$

CW =	Cost Weight
IX =	Index
AGGWT =	Aggregation Weight
A =	Area
f =	Full sample

t = Month

B. 12 Month percent change US level for U, W, and E:

For odd years: $PC(A,\ f,\ t,\ t-12) = \left(\frac{CW(A,f,t)}{CW(A,f,t-12)} - 1\right) \times 100$

For even years: $PC(A,f,t,t-12) = \left(\frac{CW(A,f,t)}{CW(A,f,t-12)} \times \frac{CW(A,f,t,OLD)}{CW(A,f,t-12,NEW)} - 1\right) \times 100$

PC = 12 month Percent Change
t-12 = 12 months prior
OLD = Previous biennial period
NEW = Current biennial period

C. 12 Month percent change Area level for U, W, and E:

For odd years: $PC(a,t,t-12) = \left(\frac{CW(a,r,t)}{CW(a,r,t-12)} - 1\right) \times 100$

For even years: $PC(A-a, f-r, t, t-12) =$

$$\left(\left[\frac{\left[CW(A,f,t)\times\frac{CW(A,f,t,OLD)}{CW(A,f,t-12,NEW)}\right]-\left[CW(a,r,t)\times\frac{CW(a,r,t,OLD)}{CW(a,r,t-12,NEW)}\right]}{CW(A,f,t-12)-CW(a,r,t-12)}\right] - 1\right) \times 100$$

r = Replicate

D. Variance of 12 month percent change for U, W, and E:

$$VAR = \left(\frac{N-1}{N}\right)\sum_{r=1}^{38}\left(\left[PC_{r,I} - PC_{f,I}\right]\right)^2$$

$$SE = \sqrt{VAR}$$

E. Comparing difference of W and E subgroup population (SUB) to U reference population:[7]

Variance difference: $\left(\frac{N-1}{N}\right)\sum_{r=1}^{38}\left(\left[PC_{SUB,r,I} - PC_{SUB,f,I}\right] - \left[PC_{U,r,I} - PC_{U,f,I}\right]\right)^2$

Standard Error Difference: $SE(SUB:U) = \sqrt{VAR(SUB:U)}$

Z-Score: $z(SUB:U) = \frac{|\mu_{diff(SUB:U)} - 0|}{\sigma_{(SUB:U)}}$

P-Value: $pval(SUB:U) = 2 \times \left(1 - probnorm\left(z(SUB:U)\right)\right)$

95% Confidence Interval: $95\%\text{CI}(SUB:U) = \mu_{diff(SUB:U)} \pm \left(\sigma_{(SUB:U)} \times 1.96\right)$

Underlying Correlation: $Rho(SUB:U) = -\frac{\left(VAR(SUB:U)^2 - \left(SE(SUB)^2 - SE(U)^2\right)\right)}{\left(2\times SE(SUB)\times SE(U)\right)}$

[7] Difference operator based on part on Shoemaker 2017, and Internal BLS Variance Monitor, although these references use SRG method for variance estimates.