Review of the 2018 Consumer Price Index Geographic Revision November 2016

Authors: Steven Paben, William Johnson, John Schilp
Email: paben.steven@bls.gov
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue NE
Washington, DC 20212

Abstract: The Consumer Price Index (CPI) estimates the change in prices over time of the goods and services U.S. consumers buy for day-to-day living based on price quotes selected from multi-stage probability samples. The first stage being the selection of geographic areas. Every ten years the CPI updates its sample of geographic areas based on the latest Decennial Census to ensure the sample accurately reflects shifts in the United States population. This paper describes CPI's latest area redesign that will be used starting with 2018 indexes. Topics include area definitional changes, stratification, sample selection, and implementation issues.

Keywords: Area Redesign, Stratification, Multi-stage sampling

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

In the first stage of a multistage sample design the Consumer Price Index (CPI) selects geographic areas. In subsequent stages, the outlets where area residents make purchases, goods and services (items purchased), and area residents' housing units are selected. While samples of outlets, goods and services, and housing units within selected areas are rotated on a regular basis, the geographic sample has traditionally been rotated in all at once. The 2018 Area Revision will be the first time the CPI rotates in a new area sample over a period of years.

Historically, a new area sample had been selected and implemented following each Decennial Census. This selection is done jointly with the Consumer Expenditure Survey (CE) from which CPI obtains expenditure weights for the indexes as well as samples of items to select within outlets. However, the sample based on the 2000 Census was never implemented by the CPI. Due to the aforementioned practice of implementing the area sample all at once, field collection would need to continue in the old sampled areas while survey operations were starting up in the new areas. This practice caused a spike in field collection costs which required a special funding initiative in order to implement a new area sample. The solution to help offset this spike was to rotate the new area sample in over a period of years. Meanwhile, the CPI has remained on the area sample based on the 1990 Census that was introduced in 1998.

The general process used to select the new area sample was similar to the 1998 Area Revision[1]. The basic steps in this process include the following:

- Determine sample classification variables
- Construct primary sampling units (PSUs)
- Determine the number of sampled PSUs including number of self-representing PSUs

---

1 Williams, J.L., Brown, E.F., Zion, G.R., *The Challenge of Redesigning the Consumer Price Index Area Sample*, Proceedings of the Survey Research Methods Section, American Statistical Association (Vol. 1), 1993, 200-205.

- Determine stratification variables for the nonself-representing PSUs
- Allocate sample to nonself-representing PSUs and assign to strata
- Select a sample of the PSUs

While the basic process was similar, there were some notable methodology changes to the area sample design within each of the basic steps. First, the sample classification structure was changed for nonself-representing areas from four census regions x two size classes to nine census divisions. Second, the area definitions of the PSUs were updated to reflect the most recent Office of Management and Budget's (OMB) area definitions which consist of Core Based Statistical Areas (CBSAs).[2] Third, a decision was made to reduce the number of sampled PSUs in the CPI from the current 87 down to 75 in the new design. Fourth, there were changes to the stratification variables for the nonself-representing PSUs. Finally, there were changes made to the sampling process for selecting the sampled PSUs. The purpose of this article is to further explain all of the aforementioned methodological changes in the 2018 area revision, and to explain the plan to rotate to the new area sample over a four-year transition period.

Determine sample classification variables

In the CPI, geographic sample variables represent one dimension of the overall index classification structure. Specifically, the urban portion of the United States is divided into 38 geographic areas called index areas. Meanwhile, the set of all goods and services purchased by consumers is divided into 211 categories called item strata. This results in 8,018 (38 x 211) item-area combinations. There is concern that some of these basic indexes have a small number of price quotes which can lead to measurement bias. Prior research performed by BLS found a 0.2 to 0.3% per annum finite sample bias in the All Items CPI-U due to having basic item-area indexes with insufficient sample size. Since the magnitude of the bias is inversely related to sample size, an increase in the number of quotes per item-area cell will result in a proportionate reduction in the bias in the overall index.[3] With this area revision, there was a conscious effort to partly address this issue by reducing the number of index areas which would increase the average number of quotes per item-area cell.

In the 1998 sample design, areas were first classified by location, based on one of four census regions: Northeast, Midwest, South, and West. Then, each area was classified into one of three population size classes: self-representing areas (A-size), medium nonself-representing areas (B-size), and small nonself-representing areas (C-size). In the 2018 sample design, areas were first classified by location into one of nine census divisions: New England, Middle Atlantic, East North Central, West South Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific. The census divisions represent a further breakdown of census regions.[4] Then, each area was classified into one of two population size classes, self-representing or nonself-representing.

The main impetus for breaking the four census regions into the nine census divisions was to support more locally defined indexes. In order to maintain approximately the same number of classification groups for nonself-representing areas, it was decided to

---

[2] February 2013 (OMB Bulletin No. 13-01), https://www.whitehouse.gov/sites/default/files/omb/bulletins/2013/b-13-01.pdf.

[3] Bradley, R. (2007). *Analytical Bias Reduction for Small Samples in the U.S. Consumer Price Index*, Journal of Business & Economic Statistics, vol. 25, pp 337-346.

[4] **http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf**

collapse the B and C-size classes together.[5] There was some concern about letting medium and small size areas represent one another. However, it was decided to control the proportion of medium and small size areas within a census division by using a process called controlled selection (explained later in the Sample Selection section). A study on the effect of by-division indexes on the U.S.-All Items index as well as Regional CPI indexes found little difference on the index estimates or the 12-month standard errors.[6]

Given the change in classification variables from census region and size class to census division for the nonself-representing areas, the only way to reduce the number of index areas was to decrease the current number of self-representing areas from 31.

Construct Primary Sampling Units (PSUs)

Following each Decennial Census, the U.S. Office of Management and Budget (OMB) releases a new set of statistical definitions. The current definitions assign counties surrounding an urban core to geographic entities called Core Based Statistical Areas (CBSAs). The assignment is based on each county's degree of economic and social integration to the urban core as measured by commuting patterns. There are two types of CBSAs: metropolitan and micropolitan. A metropolitan CBSA has an urban core with more than 50,000 people and a micropolitan CBSA has an urban core between 10,000 and 50,000 people. CBSAs may cross state borders. Additionally, OMB defines Combined Statistical Areas (CSAs) which are combination of one or more CBSAs.

In the 1998 sample design, the CPI defined three PSU sizes (A-, B-, and C-sized). The B-sized PSUs were metropolitan statistical areas (MSAs) as defined by OMB in 1993, C-sized PSUs were urban parts of non-MSA areas, while the A-sized PSUs were a mixture of MSAs in which some were combined to maintain continuity with area definitions from the 1987 CPI Geographic Revision. Since CBSAs are the conceptual successor of earlier area definitions, using the metropolitan and micropolitan CBSA definitions for nonself-representing areas was a natural choice. However, there was an issue in determining whether or not the CSA definitions should be used for self-representing PSUs. In some cases the A-sized PSUs in the 1998 sample more closely resemble CSA definitions while others more closely resemble the new MSA definitions. The problem with CSAs is that in many cases they create a very large geographic area in which to conduct field operations and collect prices. Therefore, it was decided to strictly adhere to the new MSA definitions for self-representing PSUs.

Currently, BLS publishes the CPI for All Urban Consumers (CPI-U) which covers approximately 87 percent of the U.S. population. With the introduction of the CBSA concept to the CPI, the CPI-U coverage will increase to 94 percent of the U.S. population based on the 2010 Census. The area frame consisted of the 381 metropolitan areas representing approximately 85 percent of the population, and 536 micropolitan areas representing approximately 9 percent of the population.

Number of Sampled PSUs

For the area sample, the CPI has traditionally selected one PSU per stratum. The number of strata created determines the total number of PSUs in the sample. Determining the number of strata depends on a number of factors including total budget and the overall expected impact on the accuracy of the U.S.-All Items CPI. There are two components to

---

[5] Currently, there are 7 census region x size class groups for the nonself-representing areas as there was no sample allocated to the Northeast C-sized areas.

[6] Schilp, J. (2014). *Simulated Statistics for the Proposed By Division Design in the Consumer Price Index*, Proceedings of the American Statistical Association, Government Statistics Section.

accuracy that need to be taken into consideration: the expected impact on sampling variance and the expected impact on bias.

Currently, the CPI has 87 urban PSUs, while the CE has 91 PSUs (75 urban and 16 rural)[7]. With the 2018 area revision, the CPI will reduce the total number of PSUs in the CPI to 75. This reduction will eventually allow the CPI to collect prices in the same set of PSUs as the expenditure information collected from the CE survey.[8] By bringing the area samples for the two surveys into agreement, there will no longer be a need to make special weighting adjustments between the two surveys. Additionally, the reduction in PSUs should result in cost savings of implementing the new sample as we expect an increase in the percentage of overlapped areas and a decrease in the number of new areas. Most importantly, this decision will increase the average number of quotes per area in the CPI and, therefore, help address small sample bias of basic indexes. Of course, just maintaining the same number total of quotes in the CPI and reducing the number of PSUs would potentially increase the standard error of CPI estimates due to the loss of information in the PSU component of variance. Using CPI variance models, it was estimated the U.S.-All Items six-month standard error[9] would increase between 2% and 5%. These estimates were calculated with the assumption that the total number of quotes in the CPI would be maintained and the number of self-representing PSUs would be reduced from 31 to 23.

For determining the ideal number of self-representing PSUs given a 75 PSU design in the CPI, variance models again were used to simulate the U.S.-All Items six-month variance for different population size cutoffs. However, the range of the modeled six-month percent change standard errors for any cutoff value between a population of 2.0 and 3.0 million was extremely narrow. (The largest difference that was simulated was around 1%. To give some perspective, that is the difference between a standard error of .0657 and .0650). We believed this result allowed for some flexibility in determining the exact population cutoff within this range. It ultimately was decided to choose a population cutoff of 2.5 million. This population cutoff would result in 23 self-representing PSUs, 21 due to having a population greater than 2.5 million plus Anchorage, AK and Honolulu, HI. Anchorage represents all CBSAs in Alaska. Honolulu represents all CBSAs in Hawaii. Alaska and Hawaii are self-representing PSUs due to the fact that their geographic locations make price change within their consumer markets unique compared to other markets.

Given the 23 self-representing PSUs and the nine census divisions, this implies there will be 32 index areas x 211 item strata = 6752 basic indexes for the All U.S. index estimates under the new area design. This is an approximate 16% reduction in the number of basic indexes which we believe will help address the small sample bias issue.

Stratify nonself-representing PSUs

The goal of area stratification in the CPI is to decrease the overall sampling variance by grouping PSUs together which have similar characteristics that are highly correlated with price change and/or expenditure levels. In the 1998 sample design, four independent variables were used for stratifying the nonself-representing PSUs: normalized (centered and scaled by the range) longitude, the square of normalized longitude,

---

[7] The CE additionally covers rural areas while rural areas are out-of-scope for the CPI.

[8] The original 2000 Decennial based design had 86 urban PSUs for CE and CPI. CPI did not receive funding for its initiative in time to make the switch to the 2000 based design. CE did implement the 2000 Decennial based design but then had to cut 11 PSUs due to funding issues in 2006.

[9] The variance models are for six-month percent change standard errors since the main purpose of the models are for allocation of outlets and items where the samples are rotated every six months.

normalized latitude, and percent urban. Instead of simply repeating the stratification from the prior area revision, it was decided to re-analyze the effect of using demographic variables in the area stratification. This decision was primarily due to the introduction of the American Community Survey (ACS) which replaced the decennial long form.[10] The ACS allowed for rolling three-year estimates that would cover every community greater than a population of 20,000 and five-year estimates that would cover every community in the nation.[11] Whereas, previously it was only possible to get a snapshot of the demographic variables with the Decennial.

There were 52 main topics of social characteristics presented in the 2005-2007 three-year estimates of the ACS. It would have been difficult to investigate all of these variables when modeling CPI percent change estimates. Therefore, the first task was to limit the set of ACS variables to a manageable list, and the second task was to aggregate these statistics to the CBSA-level. This produced 23 variables that were used in the first stage of modeling.

The first step was to eliminate variables that were thought to have little explanatory power in the CPI. The guiding philosophy was to take variables from ACS that were thought to potentially have explainable economic effects on price change such as race, educational attainment, and property statistics. Housing variables were selected because "Shelter" makes up a large proportion of the CPI. The second task was to aggregate the county-level ACS data to the CBSA-level. When medians were available, the median statistics for all member counties in a CBSA were averaged. Examples include the average median property value or the average median income for a CBSA. Other CBSA statistics, such as an average or percent are calculated using a weighted average based on population of the requisite county averages. This yielded statistics such as "Average Median Household Property Value" or "Percent of CBSA Population who is Native American."

The final list of potential variables categorized by group were as follows:

- Housing - Average Median Household Property Value, Average Number of Vehicles per Household, Percent of Family Households, Percent of Occupied Housing Units, Percent Owner Housing Units, Housing Units per Square Mile
- Demographic - Population per Square Mile and Percent Urban
- Age - Percent of Population in their 20s, and Percent of Population between 35 and 44
- Race and Gender - Percent Male, Percent African American, Percent Native American, Percent Asian
- Income - Percent in Poverty, Average Median Household Income, Average Total Median Earnings
- Education[12] - Percent Less than 9th grade education, Percent with a 9th to 12th grade education, Percent with a High School Diploma, Percent with Some College, Percent with an Associate degree, Percent with a College Degree, Percent with a Graduate Degree.

To model these potential stratification variables, a series of non-overlapping All Items price relatives were developed for each PSU in the current area sample from the same time frame of the ACS demographics. The ACS variables investigated spanned from 2005 to 2007 and were released in December 2008. Unofficial B-sized and C-sized PSU All-Items CPI price relatives were produced; in addition to the existing A-sized PSU price

[10] An overview of the ACS data products can be found at: https://www.census.gov/programs-surveys/acs/

[11] The three-year product was subsequently discontinued in 2015.

[12] Education level for Age 25 and above

relatives, to be the responses in the modeling procedure. The 12-month price relatives used were for December to December for 2004-2005, 2005-2006, 2006-2007, and 2007-2008. This provided four responses for each PSU along with a set of covariates to be used in a repeated measures modeling procedure.

Backward elimination was used to limit the set of variables to those with small p-values (<.05). Anchorage and Honolulu were included because this model does not contain longitude and latitude. Anchorage and Honolulu were later removed from models which included latitude and longitude as they are obviously geographic outliers. The regression variables are listed as "effects" along with their resulting p-values (Pr > F). It is customary to remove the effect with the highest p-value first, rerun the model, which will change the p-values slightly and then repeat the steps of removing the highest p-value and running the model. This continues until only effects with small p-values (<.05) remain. This backward elimination process of removing one variable at a time was repeated until we arrived at the final model of all significant variables in Table 1.

Table 1. Initial proposed demographic model based on current 87 CPI PSUs

Effect	Numerator DF	Denominator DF	F-value	p-value
Time Period	1	81	173.42	<.0001
Census Region[13]	3	81	10.27	<.0001
Number of Households	1	81	12.51	0.0007
Household Property Value	1	81	11.51	0.0011

In order to enumerate the predictive power of a particular model, each one year All Items price relative in the time span was examined using linear regression. The first linear model used the one-year price relative December 2004 to December 2005 as the dependent variable, the second used the one-year price relative January 2005 to January 2006, the third February 2005 to February 2006 and so forth, until every one-year price relative was used. These linear models produce a distribution of R^2 statistics. The mean of this distribution was calculated to describe the predictive power of the model in question. This process was repeated for lower level one-year price indexes. The sub-indexes were for Energy, Local Services, Food and Beverage and Local Shelter. The backwards elimination process was followed for each of these sub-indexes as well. The R^2 statistic was again calculated to enumerate the predictive quality of the final set of variables on each sub-index. However, none of these proved to be highly predictable with R^2 statistics all under 0.3. Models were also evaluated by each of the four regions with little success. In addition to the variables from ACS, the stratification variables from the prior area revision were also considered: Longitude, Latitude, Longitude Squared, Latitude Squared and Percent Urban for each PSU in the current sample. However, these variables also had very little predictive power and were not always significant in the All Items model.

Once several final models were derived, the PSUs using these sets of variables were stratified with an "equal population" constraint. This constraint was that each stratum would have a population within 10% of the mean of all strata in the index area. This constraint was hampered by the "Number of Households" variable because it is highly

[13] Census region was used in lieu of Census division for two reasons. First, the current sample design was not designed to support division estimates. Second, the determination of the stratification model was complete prior to the decision to implement the census divisions.

correlated with population. Therefore, this variable was eliminated due to conflicting with the "equal population" constraint.

Meanwhile, since the area sample design is also to support the CE survey, ACS variables were additionally investigated for correlation with expenditure estimates. Average Median Household Income and Average Median Property Value were by far the best demographic variable predictors of consumer expenditures with an $R^2 = .65$ for the two-variable model investigated over the timeframe of 2005-2008.

Finally, three separate stratification models were investigated: a seven-variable model, a six-variable model and a four-variable model. The seven-variable model included Percent Urban, Income, Property Value, Longitude, Latitude, Longitude Squared and Latitude Squared. The six-variable model contained the same variables except for Percent Urban and the four-variable model eliminated the Longitude Squared and Latitude Squared terms as well as Percent Urban.

Finally, the four-variable model (Longitude, Latitude, Property Value and Income) was selected since it was believed that Longitude Squared and Latitude Squared added little value to the model. There do not seem to be any influential variables that can predict CPI price change over time and all regions. Table 2 shows the model prediction over time for each census region and sub-category investigated for the final four-variable model.

Table 2. R^2 for census regions and various sub-categories for final model

		Census Regions			
Index	Total	Northeast	Midwest	South	West
All Items	.158	.311	.224	.160	.330
Energy	.216	.529	.162	.261	.483
Food and Beverage	.164	.338	.174	.153	.392
Housing	.276	.469	.206	.236	.407
Local Services	.076	.201	.098	.207	.173

The final stratification model of Longitude, Latitude, Median Household Property Value and Median Household Income is a compromise between what are significant predictors for CE and the geographic variables used previously. The addition of the income and property value variables will greatly enhance the area stratification for CE greatly reducing the between PSU variance. While the introduction of the income and property value variables do not have significant explanatory power for price change, they do not harm the area stratification for CPI. In fact, producing more reliable consumer expenditure estimates only helps the CPI. Since this stratification model is very different from the prior stratification model, a sample overlap procedure will be used in order to maintain as many nonself-representing PSUs as possible with the 1998 area sample.

Allocate sample to nonself-representing PSUs and assign to strata

To allocate sample to the nonself-representing PSUs, the population for the self-representing PSUs was excluded for each census division. Table 3 presents the proportional to population size sample allocation for census divisions for the 2018 geographic area design. The 23 self-representing PSUs account for approximately 39 percent of the total U.S. population and about 42 percent of the CPI-U population. There are 52 nonself-representing PSUs representing the remaining 58 percent of the CPI-U population which includes both metropolitan and micropolitan areas.

Table 3. Census division distribution of selected sample units, 2018 revision

Census Divisions	Nonself-rep	Self-rep	Total
Total	52	23	75
1 Northeast	2	1	3
2 Middle Atlantic	4	2	6
3 East North Central	8	2	10
4 West North Central	4	2	6
5 South Atlantic	12	5	17
6 East South Central	6	0	6
7 West South Central	8	2	10
8 Mountain	4	6	10
9 Pacific	4	7	11

The next phase of the selection process was to assign the nonself-representing PSUs within each census division into strata based on a model of the four stratification variables (latitude, longitude, median household income, and median property value[14]). The primary objective of the PSU stratification is to minimize the between-PSU component of variance. That is, the PSUs within each stratum should be as homogeneous as possible. Additionally, within each census division, each stratum should have approximately the same population in order to minimize the variance. This is a type of constrained clustering problem. In the 1998 design, the Friedman-Rubin hill climbing algorithm was used to assign PSUs to strata. In the 2018 design, a new heuristic stratification algorithm was developed which used k-means clustering and zero-one integer linear programming.[15]

Select a sample of the PSUs

The final step of the selection process was to select one PSU per stratum. However, prior to making that final selection, we need to employ a couple of special selection procedures: the use of a sample overlap procedure and controlled selection.

Sample overlap methodology

In the 1998 design, a sample overlap procedure was used to select the nonself-representing geographic areas for CPI and CE. Sample overlap procedures increase the expected number of nonself-representing geographic areas that will be reselected in the new design. The use of an overlap procedure results in fewer new areas that need to be rotated in and fewer existing areas that need to be rotated out of the current sample, thus lowering the expected costs associated with moving operations to a new area design including the hiring and training of new field staff. Two different sample overlap procedures were considered: Perkins (1970) and Ernst (1986). The Perkins method is a heuristic procedure and was used in prior redesigns. The Ernst procedure uses linear programming. Due to the optimization of linear programming, the Ernst procedure will have a higher number of expected overlap PSUs and as a result will lower the overall cost of switching to the new area design for both surveys. Only nonself-representing metropolitan PSUs were deemed eligible for the procedure. For the micropolitan areas, there are issues with not having enough renters for the CPI Housing survey. Each

[14] Median Household Income and Median Property Value were derived from 2010 5-year ACS estimates for the final stratification model.

[15] King, S., Schilp, J., and Bergman, E. (2011). *Assigning PSUs to a Stratification PSU*. Proceedings of the American Statistical Association, Section on Survey Methods Research, 2235-2246.

micropolitan area must have enough renters for two samples of six panels each over the course of a decade between area redesigns. In the past, the CPI was sometimes forced to extend the area definition for the CPI Housing survey to include outlying rural counties in order to have enough renters. Attempting to retain micropolitan areas would only compound this problem. Table 4 shows the expected overlap for the two procedures compared to selecting a new set of nonself-representing areas independently.

Table 4. Expected sample overlap of PSUs by census division[16]

Census Divisions	PSU Design	Independent	Perkins	Ernst
Total	58	13.2	19.3	28.6
1 Northeast	2	0.5	0.7	1.0
2 Middle Atlantic	4	0.7	1.2	1.7
3 East North Central	8	2.9	3.6	4.8
4 West North Central	4	0.8	1.4	2.1
5 South Atlantic	14	3.2	4.5	7.0
6 East South Central	6	0.8	1.1	2.2
7 West South Central	8	2.1	2.6	4.4
8 Mountain	6	1.4	2.6	3.2
9 Pacific	6	0.9	1.6	2.2

Due to the fact that the Ernst (1986) procedure would yield a higher degree of overlap with the 1998 area sample, it was decided to use this procedure. The outcome of the procedure gives a new set of selection probabilities which were used to select the sample.

Controlled Selection

Controlled selection is a process of selecting a random sample of PSUs such that the probability of selecting certain preferred combinations of PSUs increases, and the probability of selecting non-preferred combinations of PSUs decreases. This is accomplished by controlling the interaction between the PSU selections in different strata. Since ultimately only one sample may be selected, there may be some important reasons for preferring some possible sample outcomes over others. This is usually because we judge that balance in the sample with respect to one or more additional factors will permit inferences to be made to the population with a greater degree of confidence. In other words, there is no reason why only information strictly from the strata should be used. If information is available on additional variables, then the population may be cross-classified, allowing an implicit stratification to be achieved with reference to each of these additional variables. However, multiple cross-classifications can ultimately overly constrain a sample such that it may be impossible to determine a possible sample that meets all of the constraints.

Controlled selection can be used to control on a variety of variables. In the 1998 design, the number of PSUs per state and the number of overlaps was controlled. For example, if Florida was expecting 2.3 PSUs from the South Region based on its share of population in the South, then controlled selection gave a 30% chance of getting 3 PSUs in Florida and a 70% chance of getting 2 PSUs in Florida. In this case, controlled selection eliminated the possibility of selecting samples with less than two PSUs in Florida or more than three PSUs in Florida.

[16] This analysis was done on 87 total PSUs with 58 nonself-representing PSUs prior to the decision to move to the final design of 75 total PSUs with 52 nonself-representing PSUs.

In the 2018 design, we were most concerned about controlling the number of micropolitan areas selected in the sample and controlling by state was deemed a secondary concern. The change from census region x size class to census divisions for nonself-representing index areas means that controlling the number of PSUs per state was deemed to be of less value.

Controlled selection is a computationally intensive process. The solution time of a controlled selection problem increases exponentially with the size of the problem, e.g. the number of strata. Small problems can take only seconds to run, but the solution time increases dramatically with an increase in the size of the problem. We could not solve a two-variable controlled selection problem for the South region using the software package, SOCSLP[17]. Therefore, only micropolitan status was used as the control variable and this was done at the census region level.

Sample outcome

After adjusting the sample selection probabilities using the Ernst (1986) sample overlap procedure and employing controlled selection for the micropolitan areas, one PSU per stratum was then randomly selected. In the final analysis, 33 of the 87 PSUs from the 1998 design will be dropped from the CPI. Although, two of these are due to treating the New York, NY CBSA as one PSU when previously it was treated as three PSUs. Meanwhile, only 21 of the 75 PSUs in the 2018 design will be considered new areas. Of the 21 new areas, 14 are metropolitan areas and seven are micropolitan areas.

New Area Design Implementation Plan

Once the new geographic sample was selected, a process to implement the new area sample into the four surveys used to construct the CPI was determined. The four surveys are: CE, Telephone Point of Purchase Survey (TPOPS), Commodity and Services (C&S), and the Housing Survey. In all previous CPI geographic revisions, the conversion process occurred all at once: that is, the administration of each survey switched from the old area sample to the new area sample in its entirety, albeit at different points in time. For example, for the 1998 revision, the CE switched to the new sample design in 1996; TPOPS was used to identify outlet frames in new PSUs during the 1995-1996 period; and data were collected in Housing and C&S by the fall of 1997, such that the CPI could be computed on the new area design by January 1998.

For the 2018 area revision, the CE already converted to the new sample in its entirety in 2015. However, for the other three surveys, the 21 new PSUs have been divided into groups whereby the new PSUs will be introduced over a four-year span. The benefit of this rotation process is to distribute the cost of dropping and adding new PSUs into Housing and C&S rather than incur a spike in data collection costs leading up to the point when the computation of the CPI converts to the new area structure. This section describes the constraints that influenced the rotation plan for TPOPS, C&S, and Housing, along with the timing of implementation.

It was decided to begin calculating indexes under the new area design with the first set of new PSUs and to let all dropping PSUs be considered as proxy candidates for new

---

[17] SOCSLP is written by Sun Wong Kim, Steven G. Herringa, and Peter W. Solenberger of the University of Michigan. The software is written in SAS and should remain useable in the future as SAS will continue to be supported at BLS. Details of the methodology used in SOCSLP is found in their paper “Optimizing Solution Sets in Two-way Controlled Selection Problems” at ftp://ftp.isr.umich.edu/pub/src/smp/socslp/socslp_paper.pdf

PSUs until the complete set of new PSUs could be rotated in. An ideal proxy for a given new PSU was considered to be one of the dropping PSUs within a new PSUs geographic stratum. Otherwise, a set of nearest neighbor rules were applied with the constraint that a nearest neighbor needed to be within 200 miles. If no eligible proxy existed, the new PSU was considered to be a "geographic hole" within the new area structure. There were eight new PSUs with no eligible proxy. Therefore, they were given priority in the rotation schedule.

In devising the rotation schedule, CPI determined the following field operational constraints: no more than six new PSUs could be rotated in per calendar year; and no more than two new PSUs could be rotated in any of the six BLS Regional Offices per calendar year. Since there were 21 new PSUs, the new PSUs would, therefore, be rotated in across four groups or waves over a four-year period. Six new PSUs would be introduced in each of the first three waves, and three new PSUs introduced in the final fourth wave. The timing of each successive wave begins exactly one year after each milestone for the prior wave. CPI will begin publishing under the new area design beginning with January 2018 indexes. This means Waves 2-4 will have new PSUs "proxied" under the new area design by a dropping PSU. All dropping PSUs that were not designated as a proxy will be dropped with Wave 1. There are three new PSUs that are considered geographic holes which are part of Wave 2 that will be entirely imputed for the first year under the new area design.

References

Bradley, R. (2007). *Analytical Bias Reduction for Small Samples in the U.S. Consumer Price Index*, Journal of Business & Economic Statistics, Vol. 25, pp 337-346.

Ernst, L.R. (1986). Maximizing the overlap between surveys when information is incomplete. *European Journal of Operational Research*, Vol. 27, 192 - 200.

Johnson, W., Paben, S.P., and Schilp, J. (2012). *The Use of Sample Overlap Methods in the Consumer Price Index Redesign*, Proceedings of the Fourth International Conference of Establishment Surveys, June 11-14, 2012, Montréal, Canada, American Statistical Association, CD_ROM.

King, S., Schilp, J., and Bergman, E. (2011). *Assigning PSUs to a Stratification PSU*. Proceedings of the American Statistical Association, Section on Survey Methods Research, 2235-2246.

Perkins, W. (1970). *1970 CPS Redesign: Proposed Method for Deriving Sample PSU Selection Probabilities Within 1970 NSR Strata*. Memo to Joseph Waksberg, U.S. Bureau of the Census.

Schilp, J. (2014). *Simulated Statistics for the Proposed By Division Design in the Consumer Price Index*, Proceedings of the American Statistical Association, Section on Government Statistics.

Williams, J.L., Brown, E.F., Zion, G.R. (1993). *The Challenge of Redesigning the Consumer Price Index Area Sample*, Proceedings of the Survey Research Methods Section, American Statistical Association (Vol. 1), 200-205.