Modeling Variances to Determine Sample Allocation for the Current Population Survey

Reid Rottach[1], Greg Erkens[2]
[1]US Census Bureau, 4700 Silver Hill Road, Washington DC 20233
[2]Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington DC 20212

Abstract
The Current Population Survey employs a two-stage rotating panel design, with each month's sample being made up of eight replicate second-stage samples. Panel correlations for state and national unemployment estimates are modeled, allowing us to make predictions for sample allocation which account for proposed changes in the sample design and changes in the underlying population. We use the panel correlations to estimate variance components. Calibration and composite estimators are considered.

Key Words: weighting, annual sampling, rotating panels

1. Introduction

The Current Population Survey (CPS) is a household survey jointly sponsored by the Bureau of Labor Statistics and the US Census Bureau. It is the source of monthly estimates of labor force characteristics, such as unemployment, employment, and not-in-labor-force for different demographic groups. The official estimate of the unemployment rate, which is a principal economic indicator, comes from the CPS.

We are currently redesigning the sample for the survey, which will begin phasing in starting April of 2014. The CPS has a multi-stage sample design and all stages are being redesigned; in the first stage of selection, Primary Sampling Units (PSUs) are stratified and selected with probability proportional to size. These are counties or groups of counties. In the second stage of selection, the methodologies and the frames themselves will change from the current design. Sample will come from three frames: the Unit frame and Group Quarters (GQ) frame, which are created from the Master Address File, and the Coverage Improvement frame, which will provide an area sample of blocks. By far most of the sample will come from the Unit and GQ frames. These are sorted on demographic variables taken from the Decennial Census, and clusters of housing units are systematically selected. The clusters are typically four housing units, although in some cases an interviewer (Field Representative, or FR) will find more housing units than expected assigned to a single cluster. If there are more than 15 housing units when a cluster of four was expected, the FR will subsample, which is the third stage of selection. This event is rare enough that we usually describe our sample design as two-stage, and for simplicity, that is what we will do in this paper.

Any views expressed are those of the author and not necessarily those of the U.S. Census Bureau or the Bureau of Labor Statistics.

The second stage sample is divided into eight replicate panels. Each month, one panel is in sample for the first time and will continue in sample for the following three months. It will come into sample again for four more months after eight months rest. This is referred to as a 4-8-4 rotation design.

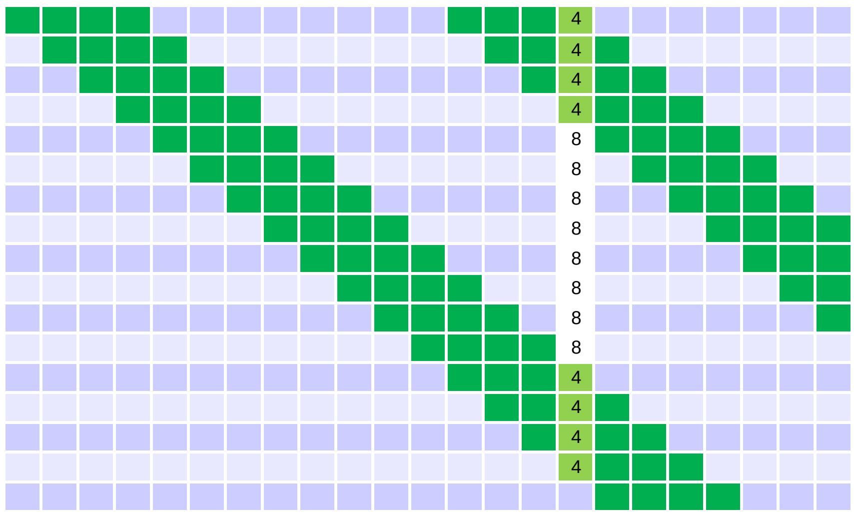

Figure 1: Visualization of the 4-8-4 rotation design. Columns represent panels and rows represent months. Dark green cells identify which panels are in sample for a given month.

In any given month the eight panels are numbered by month-in-sample (MIS), from 1 for the incoming panel to 8 for the outgoing panel, so this would be ordered right-to-left in the green cells of each row of Figure 1. The panels have a complex pattern of correlation within and across months, which we model. The approach to studying different sample allocations involves a variant of the model described in Rottach (2010), and is similar to models described in Bell and Carolan (1998).

2. Design Requirements

The CPS has design requirements at the national and state levels, as well as the District of Columbia and the metropolitan areas of New York City and Los Angeles. In the rest of this paper, the term “state” is used to refer to these 53 areas of interest below the national level.

The official design requirements in the current design are provided in the following list and taken from *CPS Technical Paper 66* (US Census Bureau, 2006):

1. A 0.2% change in the unemployment rate from month-to-month is statistically significant at the 10% level *assuming a 6% unemployment rate*.
2. The maximum state coefficient of variation (*cv*) of an annual average for total unemployed is 8% *assuming a 6% unemployment rate*.

There are also the following unofficial design requirements to consider:

1. Preference for consistency with the current design

2. About 60,000 housing units in sample
3. Reliability for labor force estimates other than unemployment rates and totals
4. A minimum state sample size of about 700

In practice, the assumption of a 6% unemployment rate has been at the state level and not just the national level. Alternatively, states could have had higher or lower assumed unemployment rates, as long as they were consistent with a national rate at 6%.

3. Modeling Variances and Correlations

3.1 Covariance and Correlation Partitions

The inferences made in this paper have as a foundation the Law of Total Covariance. In particular, for a two-stage design, this law implies that a covariance can be partitioned into two components:

$$Cov_{tot}(\hat{Y}_t, \hat{Y}_{t+l}) = Cov_b(\hat{Y}_t, \hat{Y}_{t+l}) + Cov_w(\hat{Y}_t, \hat{Y}_{t+l}) \quad (1)$$

Where the subscripts *tot, b,* and *w* on the covariances refer to total, between-PSU, and within-PSU

From this relationship, a partition of correlation follows, where between-PSU correlation is defined as the between-PSU covariance divided by the between-PSU variance, and the within-PSU correlation is defined similarly. That is:

$$Corr_{tot}(\hat{Y}_t, \hat{Y}_{t+l}) = \alpha Corr_b(\hat{Y}_t, \hat{Y}_{t+l}) + (1-\alpha) Corr_w(\hat{Y}_t, \hat{Y}_{t+l}) \quad (2)$$

Where $\alpha = V_b/V_{tot}$; V and *Corr* represent the variance and correlation, and their subscripts identify the level

3.2 Variance Models

The following graph shows data from August 2005 to October 2010, when the unemployment rate spanned from about 4% to about 10%. This plot motivates the discussion of the indirect relationship between the *cv* and the unemployment rate.

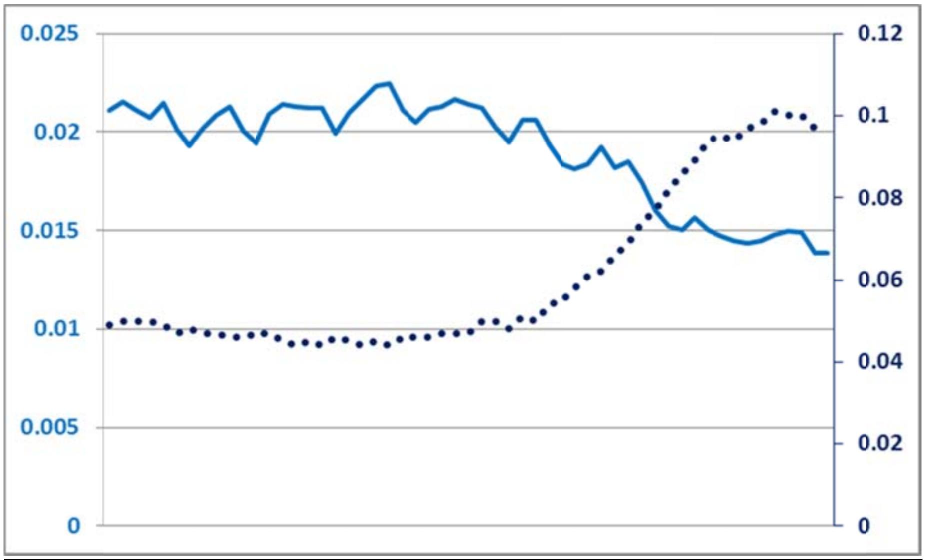

Figure 2: The dark blue dotted line and right axis represent the seasonally adjusted unemployment rate. The light blue solid line and left axis represents the directly estimated total *cv*. August 2005 to October 2010.

3.2.1 Within-PSU variance

For the within-PSU variance of total unemployed, two approaches were considered for modeling. One assumed a constant design effect (equation 3) and the other assumed the variance over the estimate was constant for a fixed sampling interval (equation 4). Both models may work well if $\hat{p}$ is small.

$$deff = \frac{\hat{V}_w(\hat{Y})}{(SI)(\hat{X})(1-\hat{p})} \tag{3}$$

$$altdeff = \frac{\hat{V}_w(\hat{Y})}{(SI)(\hat{X})} \tag{4}$$

Where $\hat{V}_w$ represents the direct within-PSU variance estimator (linearization-based); SI is the sampling interval; $\hat{X}$ is the estimate of Civilian Labor Force level; and $\hat{p}$ is the estimated unemployment rate

These competing models have a foundation in variance stabilizing transformations, in which the first distribution is binomial, and the second is Poisson. Empirically, both models worked similarly well across the changing unemployment rates, so the Poisson model was chosen for simplicity.

3.2.2 Between-PSU variance

The sampling intervals are determined using two iterations, where first the design constraints are met based on an approximation of the between-PSU variance. The results then feed into the PSU stratification process, which will provide better estimates of between-PSU variance when that is finalized. The sampling intervals are then tweaked based on the improved estimates. Prior to PSU stratification, between-PSU variances are estimated based on a Poisson model.

The Poisson model suggests the variance is proportional to the sampling interval. This is approximately true for the between-PSU variance because the number of strata formed, and therefore the number of PSUs selected, increases as the sampling interval decreases.

3.2.3 An illustration of the final variance model

The following graph illustrates how well the modeled *cv*'s follow the directly estimated *cv*'s.

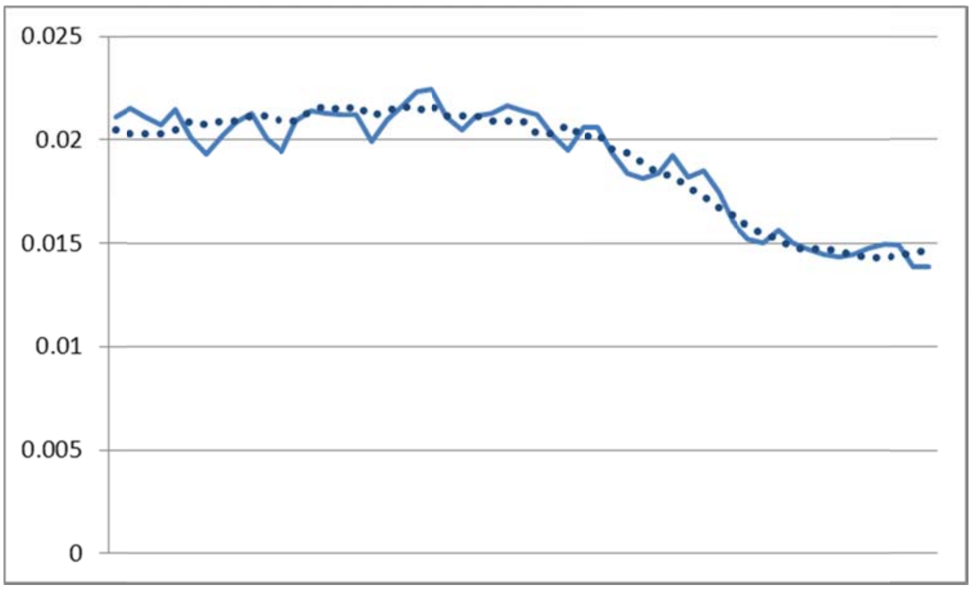

Figure 3: The light blue solid line represents the directly estimated *cv* and the dark blue dotted line represents the *cv* predicted by the seasonally adjusted unemployment rate. August 2005 to October 2010.

3.3 Correlation Models

The following graph shows data from August 2005 to October 2010, and motivates the discussion of the direct relationship between correlations and the unemployment rate.

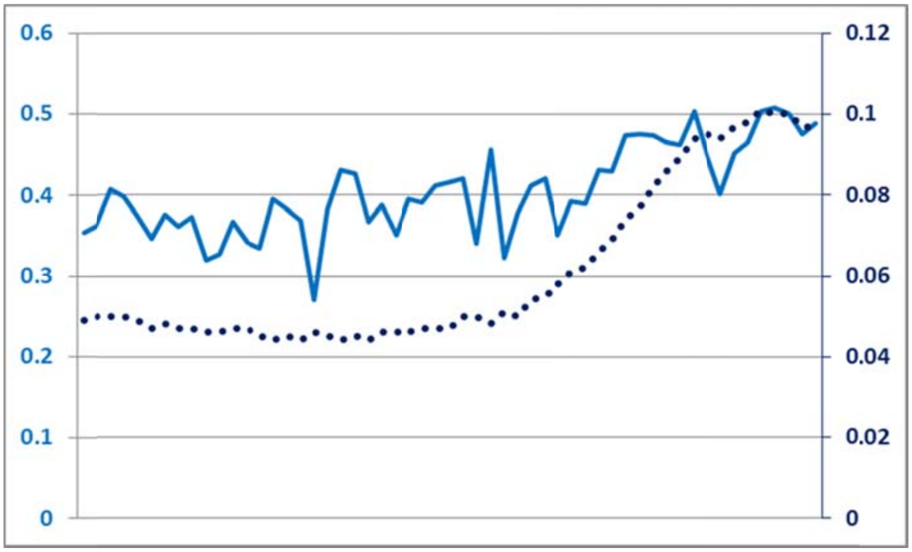

Figure 4: The dark blue dotted line and right axis represent the seasonally adjusted unemployment rate. The light blue solid line and left axis represent the directly estimated one-month-lag total correlation. August 2005 to October 2010.

3.3.1 Within-PSU panel correlations

There is a clear direct relationship between the total correlation and the unemployment rate but we were interested in modeling *panel* correlations and making inferences about overall correlation from these. The non-zero panel correlations were divided into two different types. If two panels had overlapping housing units (that is, they are both dark green in some column of Figure 1), then we expected a positive correlation directly related to the unemployment rate. The second type of correlation was due to weighting adjustments. In each month, the eight panels are divided into four MIS pairs {(1,5), (2,6), (3,7), (4,8)} and weighting adjustments are performed on each of these. This induces a negative correlation between the two panels in a given MIS pair.

The overlapping panel dependence and the weighting adjustment dependence led to a measurable negative correlation of a second order: the MIS pair of an overlapping panel. So, for any lag, two parameters were estimated: a correlation due to overlapping panels and one due to MIS pairs. These parameters were predicted at a 6% unemployment rate using an ordinary least squares regression onto the seasonally adjustment unemployment rate. The parameters were non-zero for lags with overlapping panels; namely {0, 1, 2, 3, 9, 10, 11, 12, 13, 14, 15}.

For the current design, some panels had a correlation due to their selection from the same within-PSU cluster, sometimes referred to informally as neighboring housing units, but in the upcoming design, housing units will not be clustered in the same manner and will lose this component of correlation. These and all other within-PSU panel correlations are assumed to be zero.

Table 1: OLS regression models for panel correlations. The model is $Prediction = \beta_0 + \beta_1 \times UER$, where *UER* is the assumed unemployment rate (6%) times 100.

	Overlapping Panels			*MIS Pair*		
Lag	β_0	β_1	*Prediction*	β_0	β_1	*Prediction*
0	1.000	0.000	1.000	-0.017	-0.011	-0.082
1	0.347	0.032	0.537	-0.010	-0.011	-0.078
2	0.218	0.036	0.432	-0.012	-0.011	-0.077
3	0.156	0.035	0.368	-0.016	-0.010	-0.077
9	0.111	0.020	0.232	-0.018	-0.008	-0.067
10	0.067	0.026	0.223	-0.010	-0.010	-0.067
11	0.045	0.028	0.216	-0.002	-0.009	-0.054
12	0.046	0.030	0.225	0.001	-0.012	-0.072
13	0.020	0.033	0.221	0.009	-0.014	-0.075
14	0.000	0.036	0.217	0.013	-0.015	-0.077
15	-0.007	0.036	0.209	0.016	-0.016	-0.078

3.3.2 Between-PSU panel correlations

When developing sampling intervals for the current design, between-PSU correlations were approximated as 1 across all lags, since they had not been directly estimated. In recent years we have had more stable covariance estimates to work with so were able to produce reasonable direct estimates of between-PSU correlation. They were still volatile enough that a relationship to the unemployment rate was hard to detect, so a straight average was computed across all months of data.

The between-PSU panel correlations were assumed to depend only on the lag, and not whether they were overlapping or in the same MIS pair, for example. This assumption may be reconsidered in future work. The average correlations for the first three lags were {.95, .91, .82}, and these were used to develop an AR(3) model using Yule-Walker equations. From the AR(3) model, the remaining lagged correlations were derived.

3.3.3 The correlation between two monthly estimates

From the panel correlation models, correlations between two monthly estimates can be determined. Assuming that each panel estimate has a constant variance σ^2, the variance of a monthly estimate would be:

$$Var(\mathbf{1}'\boldsymbol{y_0}) = \sigma^2 \mathbf{1}'\boldsymbol{R_0}\mathbf{1} \tag{5}$$

Where **1** is a vector of ones; $\boldsymbol{y_0}$ is a vector of eight panel estimates; $\boldsymbol{R_0}$ is a panel correlation matrix determined by the predictions for lag=0

More generally, consider fifteen months of panel estimates, altogether 120, and a correlation matrix of dimension 120 by 120. The first month's estimate may be written as $\boldsymbol{a}'\boldsymbol{y}$ where $\boldsymbol{y}$ is the vector of 120 panel estimates, and the vector $\boldsymbol{a}$ has ones in the first eight components and zeros in the rest. Similarly, express any other estimate in this time span as a vector product $\boldsymbol{b'y}$, where $\boldsymbol{b}$ is a vector with eight ones corresponding to the month we are interested in and zeros elsewhere. In general, the covariance would be as shown in (6). The variance term cancels out in an expression for the correlation between any two monthly estimates, and will have the form given in (7).

$$Cov(\boldsymbol{a'y}, \boldsymbol{b'y}) = \sigma^2 \boldsymbol{a'Rb} \tag{6}$$

$$Corr(\boldsymbol{a'y}, \boldsymbol{b'y}) = (\boldsymbol{a'Rb})/(\boldsymbol{a'Ra}) \tag{7}$$

Where $\boldsymbol{R}$ is the 120 by 120 panel correlation matrix

These expressions hold for both the within- and between- PSU components and at all lags. The total covariance is a straight sum of its two components, and the total correlation is a weighted average of its components, as shown in equations (1) and (2).

3.3.4 An illustration of the final correlation model

The following graph illustrates how well the modeled lag-1 correlations follow the directly estimated lag-1 correlations.

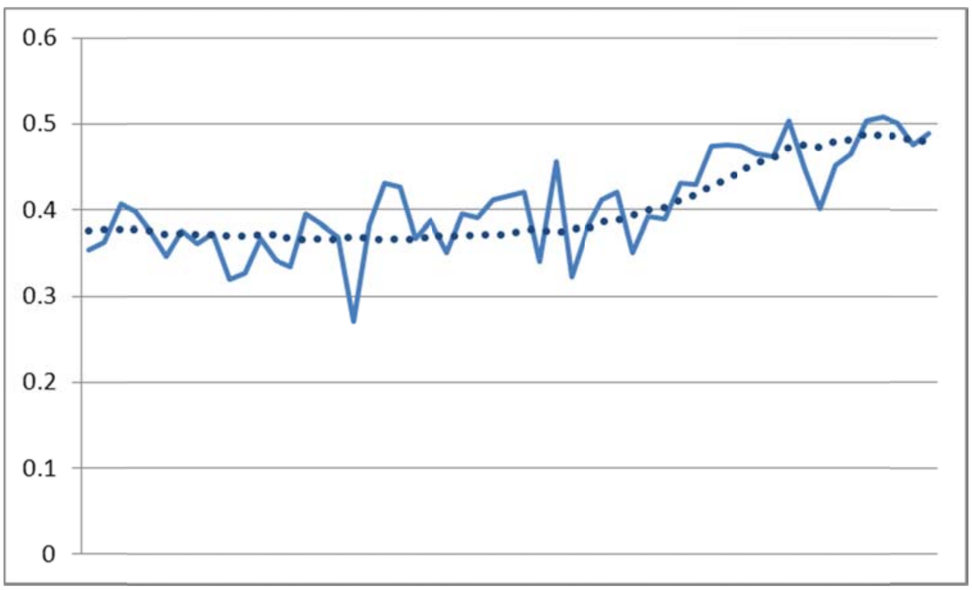

Figure 5: The light blue solid line represents the directly estimated one-month-lag total correlation and the dark blue dotted line represents that correlation as predicted by the seasonally adjusted unemployment rate, using the parameters in Table 1. August 2005 to October 2010.

4. Re-expressing the Design Requirements

4.1 The Squared *cv* of a Monthly Estimate

The design requirements will be written in terms of the squared *cv* of the monthly state estimates, so this is the first relationship needed.

$$cv^2_{stage,s}\left(\hat{Y}_{m,s}\right)=\left(\frac{\hat{V}_{stage,s}}{SI\cdot\hat{Y}_{m,s}}\right)\left(\frac{1}{.06\hat{X}_{m,s}}\right) \tag{8}$$

Where *stage = w or b*, representing either the within- or between-PSU component; *s* represents the state; $\hat{Y}_{m,s}$ represents a monthly state estimate of total unemployed; $\hat{X}_{m,s}$ represents total civilian labor force for the state

The first term on the right hand side is the parameter estimated under the Poisson model, by averaging across all months. The second term leads to a prediction at a 6% unemployment rate. The squared total *cv* is the sum of within and between.

4.2 The *cv* of an Annual Average

The variance of an annual average is:

$$V\left(\frac{1}{12}\sum_{i=1}^{12}\hat{Y}_i\right)=\left(\frac{1}{144}\right)V\left(\hat{Y}_1\right)\sum_{d=0}^{11}\sum_{|i-j|=d}Corr\left(\hat{Y}_i,\hat{Y}_j\right) \tag{9}$$

Where $\hat{Y}_i$, *i*=1, …, 12 are the twelve monthly estimates; this assumes the variance of any of the monthly estimates is the same, in which case it factors out of the summation on the right hand side of the equation

Dividing by the variance of a monthly estimate gives a factor that converts the variance of a monthly estimate to that of an annual average, as in equation (10). Assuming the correlation term in this expression is only a function of the lag (i.e., $Corr\left(\hat{Y}_i,\hat{Y}_j\right)=\rho_{|i-j|}$), then the summation simply counts the number of terms at each of the lags, leading to equation (11). In particular, the estimate of f_{aa} for the within-PSU component is 0.20, and for the between-PSU component is 0.71. The factor that incorporates both is a weighted average of the two, as shown in equation (12).

$$f_{aa,stage}=\left(\frac{1}{144}\right)\sum_{d=0}^{11}\sum_{|i-j|=d}Corr\left(\hat{Y}_i,\hat{Y}_j\right) \tag{10}$$

$$f_{aa,stage}=\left(\frac{1}{144}\right)\{12\rho_0+2(11\rho_1+10\rho_2+9\rho_3\ldots+\rho_{11})\} \tag{11}$$

$$f_{aa,tot,s}=0.71\alpha_s+0.20(1-\alpha_s) \tag{12}$$

Where α_s is the ratio of between to total variance for state *s*

The factors for within and between are assumed to be the same for all states as well as nationally. Differences in the ratio of between to total variance across states lead to differences in the total factor. The maximum monthly squared *cv* for each state is:

$$cv^2_{req1,s}\left(\hat{Y}_{m,s}\right) = \frac{.08^2}{.71\alpha_s + .20(1-\alpha_s)} \tag{13}$$

4.3 The *cv* Associated with a Minimum Sample Size

The relationship we use simply assumes the squared *cv* is inversely proportional to the sample size. Given the estimated monthly *cv* and sample size for the current design, this requirement is expressed as:

$$cv_{req2,s}\left(\hat{Y}_{m,s}\right) = cv_s\left(\hat{Y}_{m,s}\right)\sqrt{n_{current,s}/n_{minimum}} \tag{14}$$

Where $n_{current,s}$ is the number of assigned housing units in the current design and $n_{minimum}$ is the minimum sample size, which will be chosen to be close to 700.

This is an unofficial design constraint, so the minimum sample size may be chosen to maintain a degree of consistency with the current design or to achieve an overall sample size close to 60,000. This constraint helps achieve a basic level of precision for a variety of statistics and not just an annual average unemployment level.

4.4 The *cv* of the Unemployment Rate

The *cv* of a proportion in which the numerator is a subset of the denominator may be approximated using the relationship $cv^2(A/B) \cong cv^2(A) - cv^2(B)$. This result follows from a linearization. The squared *cv* of the estimated unemployment rate is about 99% of that for total unemployed, so the *cv* of the unemployment rate is approximately equal to that of total unemployed. For this work, the *cv* for an unemployment rate was assumed to be equal to that of the unemployment total. Empirically, the state estimates of total unemployed were found to be approximately independent, in which case the squared *cv* of a national unemployment total is:

$$cv_n^2\left(\sum_s \hat{Y}_{m,s}\right) = \sum_s \left(p_s^2 cv_s^2\left(\hat{Y}_{m,s}\right)\right) \tag{15}$$

Where p_s is the ratio of the state civilian labor force total divided by the national total

For a 0.2% difference in the unemployment rate to be significant at the 10% level, the maximum *cv* satisfies the relationship given in equation (16), where the estimated one-month-lag correlation is 0.41. This leads to $cv_n\left(\sum_s \hat{Y}_{m,s}\right) = .0187$. Equation (17) follows.

$$\frac{0.2\%}{cv_n\left(\sum_s \hat{Y}_{m,s}\right)(.06)\sqrt{2(1-.41)}} = 1.645 \tag{16}$$

$$\sum_s \left(p_s^2 cv^2_{req3,s}\left(\hat{Y}_{m,s}\right)\right) = .0187^2 \tag{17}$$

5. Determining the Sampling Intervals

The sampling intervals are determined using a heuristic. Two variables are adjusted to find a solution: a minimum sample size and a maximum sampling interval. The maximum sampling interval is used to limit the variability in sampling intervals from state-to-state, and therefore make the design closer to being nationally self-weighting.

The algorithm runs through the following steps:

- Fix a minimum sample size and a maximum sampling interval with reasonable starting values of 700 and 2500, respectively
- Assign the maximum sampling interval to all states
- Where needed, lower the sampling interval in states to achieve *cv*'s below those given in equations (13) and (14)
- Iterate until equation (17) is satisfied and the overall sample size is close to 60,000

Since there are two inputs to adjust, the solution is not unique, but this flexibility allows us to come close to the expected 60,000 sample size.

6. The AK-Composite Estimator

The CPS produces two types of estimation weights that have different properties and may have led to different results for meeting the design requirements. For the discussion so far in this paper, inferences were made about the calibration estimator, but they could have been made about AK-composite estimator (US Census Bureau, 2006) instead.

The AK-Composite estimator for unemployed is defined recursively in equation (18), which leads to the relationship given in equation (19).

$$Y'_t = (1-K)\hat{Y}_t + K(Y'_{t-1} + \Delta_t) + A\hat{\beta}_t \tag{18}$$

$$Y'_t - K^p Y'_{t-p} = \sum_{i=0}^{p} K^i\{(1-K)\hat{Y}_{t-i} + K\Delta_{t-i} + A\hat{\beta}_{t-i}\} \tag{19}$$

Where:

$\hat{Y}_t = \sum_{i=1}^{8} x_{t,i}$

$\Delta_t = \frac{4}{3}\sum_{i\in s}(x_{t,i} - x_{t-1,i-1})$

$\hat{\beta}_t = \sum_{i\notin s} x_{t,i} - \frac{1}{3}\sum_{i\in s} x_{t,i}$

$i = 1, 2, \dots, 8\ MIS$; $x_{t,i}$= panel calibration estimate of unemployment; s = {2, 3, 4, 6, 7, 8} sample continuing from previous month; K = 0.4; A = 0.3

Since K is less than one, the right hand side of equation (19) may be used to approximate the composite estimator for large enough values of p. For the estimator currently in production with *K=0.4*, using *p=6* will generally lead to a value less than half a percent below the true composite estimator.

The composite estimator is a linear combination of panel estimates, so the results presented in section 3.3.3 on monthly estimates can be generalized to this case. The relationships presented in that section are re-expressed below, but the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ now represent approximations to the composite estimator for two different months. In order to calculate correlations at large time lags, the matrix $\boldsymbol{R}$ may need to be of higher dimension than was used to find correlations of the calibration estimator.

$$Cov(\boldsymbol{a_c}'\boldsymbol{y}, \boldsymbol{b_c}'\boldsymbol{y}) = \sigma^2 \boldsymbol{a_c}'\boldsymbol{R}\boldsymbol{b_c} \tag{20}$$

$$Corr(\boldsymbol{a_c}'\boldsymbol{y}, \boldsymbol{b_c}'\boldsymbol{y}) = (\boldsymbol{a_c}'\boldsymbol{R}\boldsymbol{b_c})/(\boldsymbol{a_c}'\boldsymbol{R}\boldsymbol{a_c}) \tag{21}$$

Where $\boldsymbol{R}$ is a panel correlation matrix

This approach is also described in Rottach (2010), and may be used to find other ratios of covariances, such as the relative sizes of the variance for the composite estimator to the calibration estimator. In a test application, the composite estimator led to results that were different enough from the sampling intervals in the current design, that it is unlikely we will use this estimator for sample allocation for the 2010 redesign.

7. CHIP Sample Allocation

The sample used for CPS estimates includes about 12,000 additional housing units beyond the 60,000 allocated to meet the design requirements. This sample is added to improve estimates that inform legislators on the needs and impact of the Children's Health Insurance Program (CHIP), which was previously referred to as the State Children's Health Insurance Program (SCHIP). The health insurance questionnaire is not part of the monthly survey, but belongs to the Annual Social and Economic (ASEC) Supplement included each March.

The CHIP has complex funding formulas that have changed since the program began in 1997. The funding formulas relate to uninsurance rates among children, as well as uninsurance rates among low-income children. To allocate sample for the CHIP, a minimum number of children in the ASEC sample is set for each state, as well as a minimum number of low-income children (less than 200% of the federal poverty level). Although the ASEC uses the CPS sample, it is a separate survey, and it includes additional sample beyond the 72,000 used for the CPS estimates. Altogether, there are about 99,000 housing units sampled for the ASEC, with the additional 27,000 units sampled from a complex oversampling scheme.

For each state, we determine a sampling interval that would meet the required minimum number of children and low-income children. These sample counts were predicted by taking current sample counts in the ASEC and adjusting them by the ratio of the current sampling interval to a required CHIP sampling interval.

For any state in which the required sampling interval is smaller than the one for the basic CPS, the CHIP sampling interval is:

$$SI_{CHIP} = \frac{1}{1/SI_{required} - 1/SI_{CPS}} \tag{22}$$

The 12,000 CHIP housing units are sampled separately from the basic CPS, so it is possible the CHIP sampling intervals could become unreasonably large to pick up a small number of additional housing units in a state. For this reason, if the expected number of housing units needed for the CHIP sample is close enough to the number in the basic CPS, additional sample will not be added to that state.

There are three parameters that are adjusted to allocate CHIP sample: the minimum number of children in sample, the minimum number of low-income children in sample, and a tolerance on the expected number of housing units. These are adjusted until 12,000 housing units are added and sampling intervals for CHIP are not unreasonably large. Currently, reasonable values are in the neighborhood of 1,250 children, 350 low-income children, and a tolerance of 30 housing units.

8. Future Research

One of the significant changes in the upcoming design is that the second stage sample will be selected annually, rather than every ten years, as has been done in the past. One of the advantages to this is that sampling intervals for future panels may be revised each year, allowing us to maintain a more constant sample size, whereas in the current design, as the number of housing units in the US grows, so does our sample. The current approach is to counteract this by occasional sample cuts that are made across all panels. If sample growth is counteracted by increasing sampling intervals of incoming panels in the new design, the “cuts” will not be made across all panels, but lower-numbered MIS’s will have generally fewer housing units than higher-numbered MIS’s. This suggests the eight panels will not quite be replicates of each other. The consequences of this need further study.

There are many things to study and possibly improve upon with the models used to allocate sample. For example, estimates of correlation are closely related to estimates of gross-flows, and these may have more stability than the correlation estimates we used. This relationship may also offer insight into how to improve the correlation models that currently use ordinary least squares regression.

Furthermore, the estimates of state-level components of correlation were not modeled separately, but instead just used the national estimates. This may be studied more completely.

References

Bell, P.A., and Carolan, A. (1998). “Trend Estimation for Small Areas from a Continuing Survey with Controlled Sample Overlap.” Working Papers in Econometrics and Applied Statistics, Catalogue no. 1351.0, no. 98/1, ABS, Canberra.

Rottach, R. (2010), “Sample Correlations of Current Population Survey Unemployment Statistics,” Proceedings of the Section on Survey Research Methods, American Statistical Association, Washington, DC.

U.S. Census Bureau (2006), “Current Population Survey: Design and Methodology, Technical Paper 66,” October 2006.