USING THE AMERICAN COMMUNITY SURVEY TO ENHANCE THE CURRENT POPULATION SURVEY: OPPORTUNITIES AND ISSUES
Shail Butani1, Charles Alexander2, James Esposito1
Shail Butani and James Esposito of U.S. Bureau of Labor Statistics
Charles Alexander of U.S. Bureau of the Census

Key Words: Multi-stage sampling, stratified systematic sample, coefficient of variation, ratio adjustment, time-series models, small area estimation, rolling sample

I. Introduction

The Current Population Survey (CPS) provides labor force estimates for various demographic groups at the national and state levels. It is a monthly survey of about 50,000 housing units that is conducted by the U.S. Bureau of Census (BOC) for the U.S. Bureau of Labor Statistics (BLS). It is also the vehicle for supplements collecting national information on income, poverty, and other topics. The CPS sampling frame is based on the list of addresses for the decennial census, updated for new construction. Additionally, data from the census are used in CPS stratification and estimation.

The American Community Survey (ACS) is a new program that is meant to collect census “long form” type data giving basic population characteristics continuously throughout the decade. Starting in 2003, the ACS will use a rolling sample of about 250,000 different housing units per month, spread evenly throughout the country, based on a continuously updated address list. Both the regular availability of “census” type data and the updated address list provide opportunities and additional flexibility for the CPS design and estimates.

In this paper, we discuss some of the most promising opportunities. The two major opportunities include uses of ACS data to improve CPS design in terms of selection of housing units from an updated list, and adjusting various stages of weighting. In addition, ACS provides the ability to target CPS sample to households with specific demographic characteristics, and greater flexibility to expand the CPS sample as needed. The greatest potential benefit to BLS from the ACS, because of its large sample size of 3,000,000 addresses per year, lies in enhancements of the models for labor force estimates at the state and sub-state levels. The demand for timely and good quality local data is ever increasing, especially in light of the passage of the Workforce Investment Act. Issues and concerns such as coverage of the updated lists and measurement error are also discussed.

We start by giving a brief description of the CPS and ACS sample designs, respectively, in Sections II and III. In Section IV, we discuss the use of an updated Master Address File (MAF) for ACS. In Section V, we discuss use of ACS to improve the various stages of weighting in CPS. Next in Section VI, we compare and contrast the labor force questions and data collection procedures, including reference period, of the two surveys. In Section VII, we state possible

---

1 Any opinions expressed in this paper are those of the authors and do not constitute policy of the U.S. Bureau of the Labor Statistics.
2 This paper reports the results of research and analysis conducted by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.
methodologies to reconcile or calibrate ACS to agree with CPS at the national level for civilian labor force statistics, including unemployment rates. In Section VIII, we mention some methods to improve state and sub-state labor force estimates. Finally in Section IX, we conclude with some thoughts for future areas of research.

II. CPS Sample Design

The CPS is conducted monthly using a probability sample of about 50,000 housing units. Each housing unit is in the sample for 4 consecutive months, out of the sample for 8 consecutive months, and then back in sample for the next 4 months; this is known as 4-8-4 rotation. This rotation yields a 75 percent month-to-month sample overlap and 50 percent overlap with the same month a year ago. Generally, the data for the first and fifth months-in-sample are collected by personal interviews while for other months interviews are over the telephone; computerized instruments are used to collect data; and all interviewers go through extensive training. The reference period for the survey is the week containing the 12th day of the month and data collection occurs during the week of the month that includes the 19th; generally, the data are published the first Friday of the month following the reference period. That is, CPS provides timely estimates with measurable reliability of employment, unemployment, and other characteristics of the labor force, as well as information on persons not in the labor force; this information is provided for demographic groups (age, sex, race, ethnicity, etc.).

After each decennial census, the sample is redesigned and a new sample of housing units is selected for the coming decade. A brief overview of the sample design is given below.

The CPS has a state-based design that is augmented to meet national reliability requirements; the design assumes a 6 percent unemployment rate. The first design requirement specifies that coefficient of variation (c.v.) is 8 percent or lower on annual average unemployment rates for each of the 50 States, District of Columbia, and several sub-state areas. The sub-state areas are: the Los Angeles-Long Beach metropolitan area, the balance of California; New York City (5 boroughs), and the balance of New York State. After this requirement is satisfied, then the remaining sample is distributed in a manner that would minimize c.v. on the national monthly unemployment rate to about 1.9 percent. This translates into a 0.2 percentage point change in the monthly unemployment rate being significant at a 90 percent confidence level. For most states, a difference of 0.8 percentage point in the annual average unemployment rate is significant at the 90 percent confidence level, while for month-to-month change, in the unemployment rate, a change of 1.6 percentage points is significant.

The CPS is a stratified systematic sample of housing unit (HU) clusters. In the first stage of sampling, large geographic areas called primary sampling units (PSU) are selected in each state. The most populous PSUs are selected with certainty, while smaller or non-self-representing PSUs are stratified and sampled with a probability proportional to size. Within sampled PSUs, decennial census blocks are sorted. There are two types of blocks. About three-fourths are address list blocks where direct use is made of decennial census address lists to create and sample ultimate sampling units (USU) defined as a cluster of four contiguous (or nearby) HUs. The other blocks require area-based sampling techniques. To keep the CPS sample up-to-date for either type of block, new construction HUs are located and sampled, usually using lists of HUs obtained from permit offices. In the second stage, USUs are systematically sampled within each PSU at intervals that vary by PSU. At present, the entire sample of USUs or HUs is selected to span a decade.
The basic form of the CPS estimator is: \( \sum w_i x_i \); where \( w_i \) is the final weight that consists of several factors associated with various stages of weight adjustments in the CPS estimation process. The final weight at time \( t \) is defined as: 
\[
    w_i = w_{psu} x w_{hu} x w_s x w_{ni} x w_{1st} x w_{2nd},
\]
where, 

- \( w_{psu} \) = weight associated with the selection of PSU (it is 1.000 for self representing PSUs and the inverse of the probability of selection for non-self representing PSUs).
- \( w_{hu} \) = weight of a household or person within a PSU.
- \( w_s \) = special weight for the person’s USU or HU, to account for subsampling in the field.
- \( w_{ni} \) = noninterview adjustment weight for the HU.
- \( w_{1st} \) = first-stage ratio adjustment for the person.
- \( w_{2nd} \) = second-stage ratio adjustment for the person with built-in composite weights.
- \( x_i \) = person possessing characteristic of interest (1 or 0).

For item nonresponse, a hot deck imputation procedure is used (see CPS Technical Paper 63).

**III. ACS Sample Design**

Starting in 2003, the ACS will mail to about 250,000 addresses at the start of each month, selected from an updated MAF (Master Address File). After two mailings during the month, an attempt is made to conduct the interview by telephone during the next month. In the third month, one-third of the remaining nonrespondents are contacted in person. The reference period is defined as the week preceding the date on which the respondent completed the form. Thus, the March data consist of mail returns from the March mailout, telephone follow-up (and some late mail returns) from the February mailout, and personal visit follow-up from the January mailout. This pattern is repeated each month. For occupied units, we expect that about 60 percent will be interviewed by mail, 10 percent by telephone, and one-third of the remaining 30 percent in person, with each of the last group receiving a tripled weight.

Each address is in sample one time, and not eligible again until five years later. Thus the sample cumulates to 3,000,000 addresses each year, and to about 15,000,000 over a five year period. The survey will be in all counties each month. The data are weighted to make annual average estimates. The population totals by age, race, sex, and Hispanic origin are controlled to be consistent with the official intercensal population estimates. Details on the ACS weighting are given in Alexander, Dahl, and Weidman (1997).

The ACS sample size is sufficient to make useful annual estimates for areas down to about 65,000 population. For smaller areas, several years of data may be cumulated. For very small areas, such as census tracts, which average about 4,000 population, cumulating five years of data is typically required. This gives precision close to that of the census long form sample.

The questionnaire content will be similar to that of the decennial census long form. A description of the content, and other aspects of the ACS data design, and testing plans are given on the Census website [www.census.gov](http://www.census.gov). One important feature is that the survey uses a “current residence” rule: each person is counted as a resident where he/she is currently living or staying, with the exception of short stays (two months or less) away from a “usual residence.”
IV. Master Address File

The Master Address File, used in Census 2000, linked to the TIGER geographic information system, will be kept up-to-date after the census. For areas containing about 80 percent of the addresses, the main source of updates will be the Postal Service Delivery Sequence File; new addresses will be added twice a year using a computer match, supplemented by a clerical operation.

In the remaining (“non-city-style”) areas, where the mail delivery addresses do not correspond to a house number and street name, most new addresses will come from a new Community Address Updating System (CAUS).

Field representatives using an automated listing and mapping instrument will verify and locate addresses obtained from local governments. This will be supplemented by “targeted listing” based on a combination of field observations, centralized administrative records counts, and local sources. The targeted listing methods will also be used to enhance the coverage of the Delivery Sequence File in the “city-style” areas.

The CAUS is currently being tested, and will go into full operation in 2002. There will be an ongoing area sample to measure the coverage of the MAF, and supply a coverage improvement sample of any omitted addresses. Our evidence, from ACS estimates using similar updating methods, and from Ott, et al (1997), suggests that the coverage of CAUS will be fairly complete. The main exception is expected to be several months lag in updating the most recent new construction. The Delivery Sequence File, and some of the sources for targeting listing, may only pick up units once they have been built and occupied. The CPS sample is currently updated using building permits, which can be sampled before the unit is even built. A study is underway to match the MAF and the CPS building permit sample to measure this lag.

Partly because of the concern about the coverage of recently built units, the main CPS sample will continue to be updated using building permits rather than the updates from the Delivery Sequence File and CAUS.

V. ACS to Improve CPS Design

Using ACS or any other data source to update the PSUs more than once a decade is impractical both in terms of costs and data quality; that is, it is disruptive in terms of hiring/firing interviewers, training new interviewers, etc. The ACS and MAF, however, have the potential to considerably improve sampling of USUs or HUs within PSUs (see Dippo, 1998). As mentioned earlier, the CPS sample for the entire decade is selected at one time and later supplemented with new construction. Direct use is made of decennial census address lists to create and sample about 75 percent of the USUs. This list is anywhere from 4 to 14 years old for a particular month of CPS interviewing. To the extent that PSUs are growing or declining in population at different rates, the non-self-representing PSUs weights (\(w_{psu}\)) that are based on census populations are inefficient. The ACS data can remove the inefficiencies of relying on census population numbers that are anywhere from 4 to 14 years old. The weights (\(w_{hu}\)) assigned to USUs within PSUs can also become inefficient, due mostly to the increasing number of housing units on the census lists that are demolished through the decade; an up-to-date address file from the ACS can greatly reduce the problem.
The use of ACS data can be beneficial in the first-stage ($w_{1st}$) and second-stage ($w_{2nd}$) ratio adjustments. The first-stage ratio adjustment is applied to persons for black/non-black in 20 selected states; it forces the weighted 16+ decennial census data from sampled non-self-representing PSUs to equal actual data for all non-self-representing PSUs in the state. To the extent that the demographic composition of the non-self representing PSUs is changing over the decade, this ratio based on the decennial census is inefficient. Having a more up-to-date “count” from ACS would improve this weight adjustment. The second-stage ratio adjustment is applied to persons within each of the eight rotation groups and is raked six times to independent population controls for 50 States and District of Columbia, and various national demographic groups. The derivation of independent population controls for various demographic groups is extremely difficult and complex, and is an inexact science. This process can be greatly enhanced and validated by the ACS data.

The ACS interviews can be used to target housing units containing demographic groups so that CPS can oversample these groups within the CPS sample PSUs to provide more precise estimates. The MAF could also be used to give greater flexibility to expand the CPS sample. For example, the possibility of a telephone only component across a state can substantially reduce between-PSU variance.

Before using the ACS and MAF for these purposes, however, many issues such as coverage, address quality, and timeliness need to be studied and evaluated. In particular, if there is a lag in picking up new construction on the MAF, it may be necessary to give higher weight to the most recent new construction in the basic CPS building permit frame to make up for the absence of these units in the supplemental MAF frame. The determination of “frame status” for this dual frame estimation problem could be done in several ways. The CPS permit sample might be matched against the new MAF address each time it is updated, or information about when the unit was first occupied could be collected on the CPS interview.

**VI. Labor Force Questions and Data Collection Procedures—ACS vs. CPS**

In the following subsections, we summarize some of the more important content and procedural differences between the CPS and the ACS.

**Labor Force Questions.** On the CPS, there are at least 16 questionnaire items (see Table 1) that are used to assign the target person to one of seven labor force (LF) categories. On the ACS, while there are four numbered questions (items 22, 28, 29, and 30), there are actually seven labor force questions—item 28 has three parts and the response options to Q30 can be viewed as being equivalent to one of the CPS LF items (see Table 2). While it will **not** be possible to assign the target person to one of the seven CPS labor force categories using the seven LF questions on the ACS, it will be possible to assign persons to one of the three major LF categories (employed, unemployed, not in labor force). Given the different sets of questions asked (and ignoring for the time being other important considerations, like mode effects), at issue here is whether the two surveys are capable of producing equivalent estimates with respect to these three major LF categories. In our view, considering question content **alone**, the two sets of LF questions (ACS vs. CPS) are capable of producing fairly similar estimates of major labor force categories. One key estimate, the unemployment rate, will **not** be equivalent—but the discrepancy may not be that large. In the CPS, most (but certainly not all) persons are classified as unemployed on the basis of answers provided to a sequence of four questions (LK, LKM1, LKAVL, and LKAVR).
The ACS asks the equivalent of three of these questions. The only question not specifically asked in the ACS is LKM1. [Note: Field-testing would need to be conducted in order to determine actual CPS-versus-ACS differences for each major labor force category (employed, unemployed, not in labor force).]

Mode of Administration. All CPS interviews are conducted by Census Bureau interviewers using computers; approximately one-quarter of these interviews are scheduled to be in person with the remainder being conducted over the telephone. Over 60 percent of the ACS cases will involve a self-administered interview. All of the remaining cases (40 percent) will involve computer-assisted interviews conducted by Census Bureau interviewers; approximately one-quarter of that 40 percent will be telephone interviews and the remainder will be personal visits. Given that interviewers will not be on hand to answer questions respondents might have when completing the ACS, one might presume that data quality for the self-administered ACS cases will be inferior to that collected by interviewers for both the ACS and the CPS. The following example illustrates a potential data-quality problem. During the quality assessment phase of the redesign of the CPS, interviewers consistently reported that some respondents were experiencing confusion on how to respond to CPS item WK: “LAST WEEK, did you do ANY work for (either) pay (or profit)?” A common response to this question was: “Just my job.” Apparently this happens fairly frequently, but it is not a serious data quality issue for the CPS because interviewers know to code such answers as a “yes” response and move on. It is not known how respondents will deal with this confusion in the self-administered context of the ACS. Almost certainly, some percentage of respondents will answer “no” to Q22 (assuming that this question is asking about work other than that associated with their jobs) and, as a consequence, will be classified as either unemployed or not in labor force by the ACS. This is a potentially serious data-quality problem.

Reference Week. The CPS has a fixed reference week (the week containing the 12th of the month), whereas the ACS has a floating reference week (the respondent reports for the calendar week preceding the date on which the survey, or interview, was completed). As noted in ACS documentation under Limitation of the Data: “The reference week for the employment data is not the same for all persons. Since persons can change their employment status from one week to another, the lack of a uniform reference week may mean that the employment data do not reflect the reality of the employment situation in any given week.”

Household Composition. The CPS has a complex set of rules for determining who is to be included as a member of the household (see CPS interviewers manual, p. C3-7). These rules are not captured in the instructions provided to ACS respondents (see left-hand margin on p. 2 of the ACS form). Since the two surveys do not define household membership in equivalent ways, there will be differences in labor force estimates, and it is important to study and quantify them.

Other Issues. There are a variety of other issues that potentially impact the comparability of labor force estimates generated by the ACS vis-à-vis the CPS, such as:

- Unit nonresponse and imputation procedures differ for the two surveys.
- Differences in data-collection phases affect the accuracy of reported data. For example, missing data from the self-administered phase of the ACS will be collected by phone approximately one month after questionnaires are returned, whereas CPS interviewers are trained to minimize missing data problems.
- Processing paper-and-pencil data (ACS), relative to electronic processing (CPS), may impact data accuracy.
For the CPS, there is approximately a 75 percent overlap for CPS-interviewed households from one month to the next—that is, about 75 percent of the households have been interviewed in the previous month. The expected value of unemployment rate for months-in-sample one and five (incoming group) is higher than for the other six months. This is documented in the paper “The Effects of Rotation Group Bias on Estimates from Panel Surveys,” in Bailar (1975). There is no common sample component from month to month or from year to year in the ACS.

Table 1: CPS Items Used in the Determination of Labor Force Status

<table>
<thead>
<tr>
<th>Label</th>
<th>Wording</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUS/Q19A</td>
<td>Does anyone in this household have a business or a farm?</td>
</tr>
<tr>
<td>WK/Q20</td>
<td>LAST WEEK, did you do ANY work for (either) pay (or profit)?  [Note: The interviewer reads the words in parentheses if the respondent answers “yes” to BUS/ Q19A.]</td>
</tr>
<tr>
<td>BUS1/Q20-1</td>
<td>LAST WEEK, did you do any unpaid work in the family business or farm?  [Note: One also needs data on actual number of hours worked for “yes” responses to this question.]</td>
</tr>
<tr>
<td>RET1/Q20-num</td>
<td>Do you currently want a job, either full or part time?</td>
</tr>
<tr>
<td>ABS/Q20B-a</td>
<td>LAST WEEK, did you have a job either full or part time? Include any job from which you were temporarily absent.</td>
</tr>
<tr>
<td>LAY/Q20B-b</td>
<td>LAST WEEK, were you on layoff from a job?</td>
</tr>
<tr>
<td>ABSSRN/Q20B-1</td>
<td>What was the main reason you were absent from work LAST WEEK?</td>
</tr>
<tr>
<td>ABSQD/Q20B-2</td>
<td>Are you being paid by your employer for any of the time off last week?  [Note: A person who has a job, but who did not work during the reference week (e.g., vacation, illness), will still be classified as employed even if she/he was not paid for that missed work.]</td>
</tr>
<tr>
<td>LAYDT/Q21</td>
<td>Has your employer given you a date to return to work?  [Note: This question is read if the respondent answers “yes” to LAY/Q20B-b.]</td>
</tr>
<tr>
<td>LAY6M/Q21A</td>
<td>Have you been given any indication that you will be recalled to work within the next 6 months?  [Note: This question is read if the respondent answers “no” to LAYDT/Q21.]</td>
</tr>
<tr>
<td>LAYAVL/Q21A-1</td>
<td>Could you have returned to work LAST WEEK if you had been recalled?</td>
</tr>
<tr>
<td>LAYAVR/Q21A-2</td>
<td>Why is that?  [Note: This question is read if the respondent answers “no” to LAYAVL/Q21A-1. Persons who could not return to work due to a temporary illness are classified as “unemployed—on layoff”.]</td>
</tr>
<tr>
<td>LK/Q22</td>
<td>Have you been doing anything to find work during the last 4 weeks?</td>
</tr>
<tr>
<td>LKM1/Q22A</td>
<td>What are all the things you have done to find work during the last 4 weeks?</td>
</tr>
<tr>
<td>LKAVL/Q22B</td>
<td>LAST WEEK, could you have started a job if one had been offered?  [Note: This question is read if the respondent answers “no” to LKAVL/Q22B. Persons who could not have started a job either because they were waiting for a new job to begin or because of temporary illness are classified as “unemployed—looking”.]</td>
</tr>
</tbody>
</table>
Table 2: ACS Items Used in the Determination of Labor Force Status

<table>
<thead>
<tr>
<th>Label</th>
<th>Wording</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q22</td>
<td>LAST WEEK, did this person do ANY work for either pay or profit? Mark (X) in the “yes” box even if the person only worked 1 hour, or helped without pay in a family business or farm for 15 hours or more, or was on active duty in the Armed Forces.</td>
</tr>
</tbody>
</table>
| Q28   | Part A: LAST WEEK, was this person on layoff from a job?  
Part B: LAST WEEK, was this person TEMPORARILY absent from a job or business?  
Part C: Has this person been informed that he or she will be recalled to work within the next 6 months OR given a date to return to work? |
| Q29   | Has this person been looking for work during the last 4 weeks? |
| Q30   | LAST WEEK, could this person have started a job if offered one, or returned to work if recalled? [Response options: (1) YES, could have gone to work; (2) NO, because of own temporary illness; (3) NO, because of all other reasons (e.g., in school, etc.)] |

VII. Methodologies to Reconcile ACS to CPS National Data

At present, it is not possible to produce national or any state level estimates from ACS, because the survey is being conducted only in 31 sites for comparison to Census 2000. However, once the ACS is implemented nationwide, there will be annual average estimates from both CPS and ACS both at the national and state level, for the three major labor force categories. Given that CPS is specifically designed to measure labor force status of individuals and is the internationally recognized standard for such measurements, the national labor force estimates from the ACS will be inferior to those produced from the CPS with respect to reliability and timeliness (see Brown, 1997). The plans are to adjust the ACS annual average labor force estimates to agree with the CPS measures of employment and unemployment at the national level. This adjustment is being done at the national level because CPS estimates for even the largest State, California, have large sampling errors.

Several methods can be considered to calibrate ACS data to agree with CPS labor force estimates. Various constrained estimation techniques, such as “calibration estimation”, can modify the ACS weights, so that the ACS weighted annual average equals selected labor force estimates. However, these weights may not be suitable for ACS characteristics other than labor force characteristics. Prediction models, using the CPS as the dependent variables and ACS as the independent variables, could produce ACS-based estimates which are generally in agreement with the CPS estimates, but again this only applies to the variables included in the model. A more flexible approach, allowing the calibrated variables to be cross-tabulated with the full range of ACS data, would be imputation of “adjusted” responses for sample individuals, based on measurement error models whose parameters are estimated from the differences between CPS and ACS. The idea is to estimate what proportion of the ACS respondents must have given the wrong answers to produce the observed differences, and then to impute the necessary proportion of different answers to bring agreement. (Chand and Alexander, 1997)
VIII. Improvements to State and sub-State Estimates

While the CPS sample is designed to achieve a high degree of reliability for monthly unemployment estimates for the Nation as a whole, for each of the 50 States and the District of Columbia this sample is too thinly spread geographically to support “useful” sample-based monthly estimates. Most states have monthly estimates that have standard errors of 1.0 percentage point. In other words, there is a 90 percent chance that the estimate of the unemployment rate will be as high as 7.6 percent or as low as 4.4 percent when the true rate is 6.0 percent. At present, statewide employment and unemployment estimates are produced using time series methodology developed by BLS staff. The model represents an observed state CPS series as the sum of “signal plus noise”: \[ \text{CPS}_t = \text{Signal}_t + \text{Noise}_t \]

The signal represents the true value of the labor force characteristic, either unemployment or employment, which would be produced if a complete census of the population were taken. The noise represents mainly error due to sampling a portion of the population. Changes in CPS data represent both signal and noise. The state and four area models use two types of information to estimate the signal: 1) characteristics of the CPS error to estimate the noise; and 2) historical state CPS (employment and unemployment), UI claims, and payroll employment data to estimate a time series model of the signal. An important feature of the CPS design that the model accounts for is the positive autocorrelation structure induced by sample overlap from the 4-8-4 rotation. While the models remove noise from the monthly estimates on a current basis, some of the survey error is included back in the estimates during the annual benchmarking process that forces the annual average of the model estimates to equal to the CPS annual average. As mentioned earlier, for most states, a considerable amount of sampling error is associated even with annual average estimates from CPS.

The ACS data can be of help in developing alternative benchmarking methods. For example, we can compare and evaluate the annual average percentage changes in unemployment rates and employment-to-population ratios between the ACS, CPS, and model estimates across states and time to determine whether it is even necessary to benchmark on a priori basis. Perhaps, benchmarking should occur only when certain conditions are not met (i.e., model failure). Alternatively, data may indicate that it is sufficient to benchmark the sum of the state estimates from time series models to the CPS national estimates on a monthly basis. Another immediate use of the ACS data can be made in the diagnostic analysis of additive outliers, temporary level shifts that are generally induced by 4-8-4 sample rotation, and permanent level shifts. It would, however, take time to build up an ACS data time series for direct use in the state models.

With the exception of the New York City, the balance of New York, the Los Angeles metropolitan area, and the balance of California, the labor force estimates for the sub-state areas are produced using a building block approach. Monthly estimates are produced for over 6,000 areas—metropolitan and small labor market areas, counties, cities of 25,000 population or more, and cities and towns of all sizes in New England. This method uses very little direct information from the CPS because in many of these areas the sample sizes are too small or nonexistent. Data from a variety of sources are used and some input, particularly items utilized from the decennial census, can be out-of-date. It is in the area of sub-state level estimates that the data from ACS when combined with CPS has the greatest potential benefit. At these levels, ACS data are expected to have smaller variances than CPS and CPS to have smaller bias than ACS. Models that utilize both sets of data would, on average, yield smaller total error as measured by mean
squared error (variance + bias²). These models can be built at higher geographic levels (e.g., state, large metropolitan areas) that have sufficiently large CPS sample. Using these relationships, labor force estimates for smaller geographic entities can be produced that are, on average, more reliable than the ones produced using current methodology.

IX. Summary

The data from the ACS will provide information similar to that currently being collected on the census long form on a continuous basis rather than once a decade. Additionally, it has the potential to vastly improve state and sub-state estimates that are currently produced from other household surveys. In particular, we feel this rich data source will greatly complement but not replace CPS data to improve labor force estimates at the state and sub-state levels. Before new procedures are adopted and implemented, however, we believe much research remains to be done in evaluating the ACS and in constructing and testing labor force models.

X. References


