

IV. OVERVIEW OF THE IMPLEMENTATION PLAN

To develop an implementation plan, we focus on a pilot linkage between one monthly CPS extract (**March 2000**) and one state administrative extract (**Florida**). The March 2000 CPS has several advantages for this pilot because it includes recent information on income source; is collected around the same period as other data with strong potential for other data linking projects (e.g., ACS and Decennial Census); and the Census has worked on "verifying" the SSNs that will be used in the match.⁸ While the choice of any given state FSP administrative file is somewhat arbitrary, Florida has several comparable advantages for the match, including a large population of FSP participants; full history file of all FSP participants since 1993; use in previous research projects; and potential links to several other programs (e.g., TANF).

A. Sample Size

Based on our calculations, there should be a sufficient sample size of FSP recipients in the CPS for the linkage, though it could be limited for specific subgroups. In the March 2000 CPS, 145 (unweighted) Florida CPS households (313 recipients) reported receipt of FSP benefits. However, this estimate significantly undercounts the number of potential linked FSP households for two reasons. First, the CPS survey significantly undercounts FSP participation. Consequently, we suspect that some CPS respondents who do not self-report FSP participation will be in the FSP administrative records. Second, and more importantly, the Florida state administrative file includes *all recipients since 1993*. Hence, there should be a large sample of linked records, particularly for participants who were in the FSP just before the passage of welfare reform. For example, the 1996 CPS included 245 Florida FSP households.

B. Potential Implementation Barriers

The two primary implementation barriers, which are common with any data linking project, are legal issues associated in obtaining, processing, and storing confidential data and technical associated in generating a reliable match across separate files. The restricted CPS research files and the state administrative data are both governed by confidentiality agreements that may inhibit the use of the data in one form or another. While many of the confidentiality restrictions across the CPS and state data are likely be similar, special regulations for the administrative or survey data could affect the final linkage. For example, restricted CPS files can only be access at the Census RDCs, which will likely limit the broad usage of these files. The technical issues primarily involve missing, inconsistent, or purged data from either the CPS or state administrative file. If these issues are prevalent in either file, the reliability of the match will be significantly reduced.

⁸ Census has already assessed the reliability of the March 2000 files by cross-checking the CPS data with the Social Security Administration (SSA) files from the Alphadent and the Numident records. The Alphadent files include an alphabetical listing of a person's name and SSN. The Numident files include a list sorted by SSN of names. Both files are used in checking the reliability of the name and address information on the CPS. Presumably, these checks could be in place for other data sources, such as the 2001 March CPS.

C. Assumptions for Cost Estimates

To generate cost estimates, we make assumptions for the activities undertaken by each agency. For ease of illustration, we assume that the costs will be borne by the agency closest to the activity. For example, because the Census houses restricted CPS data, we assume that all costs associated in manipulating the data will be absorbed by the Census.

We assume that three agencies will be involved in the following activities:

- **USDA:** The USDA will coordinate all activities across agencies, including applying for access to use restricted files from the March 2000 CPS and Florida FSP files. The USDA will also assist in developing protocols for future use.
- **Census:** The Census will coordinate the details of the data linkage, which includes obtaining and cleaning the data and performing the statistical match. Because the data matching process is labor intensive, the bulk of the costs fall to the Census.
- **Florida's Department of Children and Families (DCF):** The DCF will provide technical assistance and the data from the state administrative records.

It is important to note that these costs are an estimate. The actual costs of implementation for each agency will depend on the funding available and further agreements across agencies. For example, while the majority of costs described in the next section fall to the Census, these costs could be offset if the USDA provided funds to either the Census or a private contractor to complete the tasks below.

To be consistent with Bell (2001) and Kenyon, et al. (2001), we assume that each agency has three labor categories: Senior Manager, Senior Analyst, and Research Assistant. The Senior Manager category includes experts who have at least ten years of research and/or program experience. The Senior Analyst category includes researchers and programmers who have between three to nine years of research and/or program experience. The Research Assistant category includes individuals with very limited experience. We generate cost estimates for these categories using a list of hourly wage rates from industry averages.⁹ *Appendix A* provides a detailed summary of these wage rates, along with a brief description of the workplan for each task outlined in the next section.

⁹ We developed these industry averages from our original proposal to ERS for this task order.