

III. INITIATIVE SUMMARY

Restricted files of the CPS can be linked to state FSP administrative extracts using unique identifiers that exist on both records, including Social Security Numbers (SSNs), name, and date of birth information. The CPS gathers this information during each monthly interview, but only makes it available in restricted research files. Several state FSP databases include SSN and/or name information to administer benefits and establish links to other programs. Researchers must satisfy specific confidentiality restrictions before using restricted CPS or state FSP administrative files. Because the match will use individual identifiers, information on FSP participation will need to be imputed to the household level. Researchers who use survey data often have to make similar imputations.⁵

Why use the CPS rather than another survey? The CPS has several advantages over other surveys for creating a data linkage to state FSP records. First, most surveys do not collect information on unique identifiers, such as SSNs, that are critical for the linkage (e.g., Panel Survey of Income Dynamics). Second, the CPS includes a large sample that can be used in state level analyses, which is particularly important given that micro-level FSP administrative data are only available at the state level. Third, researchers have used the CPS extensively in past research. Consequently, improving these data should enhance options for future FSP research. Finally, the CPS is an on-going survey planned each month, which provides some flexibility for future data linking projects.

Why use state FSP administrative data? State administrative data are necessary to provide program participation histories on FSP participants. In general, administrative data provide a low cost mechanism for obtaining program histories for FSP participants. Many states track FSP over multiple periods in their databases.

A linked CPS-FSP file would include CPS survey information on demographic, employment, and income characteristics that are linked to lifetime history records on FSP participation from a particular state.⁶ This file would be accessible from a restricted Research Data Center (RDC) (which would most likely be located in Washington, DC) by sworn special agents of the Census and USDA. Researchers would become sworn special agents by filing a proposal to the Census and USDA that ensures the confidentiality of the data.

⁵ This may pose a bigger problem for observing longer term changes in FSP patterns for other household member, particularly if researchers expect major changes in household composition. Unfortunately, we cannot observe household changes in the CPS. Nonetheless, researchers will be able to follow FSP histories for the heads of households (as recorded on the FSP file), but will need to make assumptions on the FSP patterns of other members.

⁶ The file may also include other FSP administrative information, such as benefit amount and income. The general rule in using administrative files is that the variables in these files are only as reliable as their importance for administering the program. For example, past research has found that certain variables, such as education, are very unreliable in many administrative systems because they are not necessary in administering the program

A. Advantages

The combination of survey and FSP administrative records would significantly expand research opportunities beyond that provided by survey or administrative data files alone, particularly in the area of FSP dynamics. The survey information would provide detailed background information on demographic, income, health, and other program characteristics of FSP participants and non-participants. This additional information would address a major limitation of administrative records, which contain very limited background information on individuals. The administrative records would supplement this information by providing lifetime FSP histories for each CPS respondent. These records would address a major limitation of the CPS, which provides a one period "snapshot" of FSP outcomes.

Researchers could use these data to observe detailed transitions of CPS respondents before, during, and after their interviews. While transitions onto FSP can be observed using FSP administrative data alone (e.g., the FSP Quality Control data), the linked survey data allow for the construction of detailed profiles of family characteristics at these transition points on FSP and non-FSP participants. Consequently, the linked data would provide detailed contextual information related to program dynamics unavailable elsewhere.

Researchers could use the linked files to address several questions related to the dynamics of FSP participation, including:

- What are the income and program characteristics of participants who cycle on and off FSP programs in specific states?
- How many eligible FSP non-participants have ever participated in the FSP? If so, are they more likely to participate in recent periods (e.g., past two years)?
- How do long-term FSP participation patterns vary across specific subgroups (e.g., low-income) of participants and non-participants?
- How do FSP participation patterns of CPS respondents vary before, during, and after their interviews?

Researchers could also use these files to assess the reliability of reported program participation in various states, which could be valuable to the Census data collection efforts for the CPS. For example,

- How do self-reported records for FSP differ from administrative records? Are there potential explanations for under-reporting of FSP participation in the CPS?
- Does the CPS adequately sample a representative sample of FSP households based on the information from the administrative records?
- Can the administrative records be used to improve current CPS imputation processes?

Finally, and potentially most importantly, the creation of a matched CPS-FSP file could generate other linked data initiatives, which, by virtue of having access of state

administrative records, would become significantly cheaper. Some natural extensions include:

- *Creation of other Matched Census-FSP data sources:* Other Census data sources could also be linked to state FSP administrative records, including the SIPP, American Community Survey, and the Decennial Census. Each of these data sources also contain the same unique identifiers as the CPS. Presumably, the same agreements developed to match the CPS to FSP records could also applied to these other data sources. These linkages could generate a set of unique research opportunities for both the Census and USDA.⁷
- *Creation of Longitudinal CPS-FSP File:* As mentioned above, researchers could link the publicly available CPS files from several months with the linked data. This linkage would provide multiple months of survey data on labor force attachment (from the survey) and FSP participation (from the administrative records). For example, researchers could use this information to assess questions such as "How do changes in employment status influence FSP participation?"
- *Matched CPS- FSP-SSA Administrative File:* Certain CPS files have already been linked to SSA administrative records on earnings and program participation. In theory, a matched CPS-FSP file could be linked with the existing linked CPS-SSA file. If SSA files become available, it would be a relatively low cost exercise to combine the information from the linked surveys. Unfortunately, obtaining the matched SSA restricted files is problematic because of significant data restrictions on SSA earnings files. Consequently, it is not clear, given the current rules, if this is a viable option.

B. Limitations

There are three important limitations of the linked data. First, detailed characteristics from CPS interviews are only available for a very limited period. Therefore, it is not possible to observe whether certain factors (found in the CPS), such as changes in household composition, influence FSP participation in other periods. Second, the linked administrative files only track program participants while they are participating in the state programs. In one sense, this is not a problem since non-participation is an important outcome in its own right signaled by the lack of inclusion in program files. However, researchers will not be able to distinguish between participants who left the program and those who moved from the state—a problem faced by all prior uses of administrative data and not considered crippling by most researchers. Finally, because the analysis is based on a state match, the sample sizes for specific state respondents could be somewhat limited, especially for small subgroups (e.g., Able-bodied adults). We discuss specific sample size issues in more detail in the next section.

⁷ For example, researchers could use a linkage to the longitudinal survey information from the SIPP records to address questions such as "How do the short-term program patterns observed in the SIPP relate to long-term patterns from the administrative records?"