THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Augmenting Data Reported in the Survey of Income and Program Participation with Administrative Record Data - A Brief Discussion

No. 22

Douglas K. Sater

U.S. Census Bureau

September 1986

U.S. Department of Commerce U.S. CENSUS BUREAU

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

AUGMENTING DATA REPORTED IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION WITH ADMINISTRATIVE RECORD DATA-- A BREIF DISCUSSION

by
Douglas K. Sater 22

No. 8614

SIPP Working Paper #22

September, 1986

TABLE OF CONTENTS

		Page No.
A.	Introduction	1
B.	Matching Keys	4
	1. SSNs	4
	2. Person Names	5
	3. Employer Names	6
C.	SIPP-Program Benefits Evaluation	6
D.	SIPP-Economic Censuses Exact Match Study	9
E.	SIPP-Summary Earnings Records Match Study	12
F.	SIPP-Master Beneficiary Records Match Study	13
G.	Other Administrative Record Data Sources	15
H.	Problems	15
I.	Summary	17
J.	Footnotes and References	17

AUGMENTING DATA REPORTED IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION WITH ADMINISTRATIVE RECORD DATA -- A BRIEF DISCUSSION

Douglas K. Sater
Nov. 15, 1985

A. INTRODUCTION

This paper briefly outlines the potential to augment the Survey of Income and Program Participation (SIPP) with data from various administrative record sources and points to possible research and evaluation activities using the augmented data set. The research initiatives listed below are at various stages of development, ranging from ideas or project proposals to operational studies. This outline is not intended as a treatise on the subject, but merely as a general introduction to provide a flavor of the potential of such record linkages. The statements made herein are the responsibility of the author and do not necessarily reflect official Census Bureau position.

The SIPP is a new Census Bureau survey designed to collect a wealth of information on the social, demographic, and economic situation of the Nation's individuals and families. The SIPP is a longitudinal personal interview survey which began in October 1983 with three interviews, or waves, per year. The SIPP collects data on cash and noncash sources of income, participation in selected benefit programs, taxes, assets, liabilities, numerous demographic characteristics, data relating to the work experience, characteristics of the current and previous job, and health and disability. Data from the survey will provide a multiyear perspective on changes in income and their relationship to participation in government programs, changes in household composition, and so forth. 1/

The SIPP started in October 1983 as an ongoing survey program of the Bu the Census with one sample panel of approximately 21,000 households in sample units (PSUs) selected to represent the noninstitutional populati United States.

Each household is interviewed once every 4 months for 2 1/2 years to pr sufficient data for longitudinal analyses while providing a relatively recall period for reporting monthly income. The reference period for t survey items is the 4 months preceding the interview. For example, in reference period is June through September; when the household is inter again in February, it is October through January. This interviewing pl result in eight interviews per household.

The data in the SIPP will, in themselves, be extremely valuable for num research efforts, both static (cross-sectional) and dynamic (longitudin However, for some studies, the usefulness of the SIPP may be somewhat l For example, there are little data on the characteristics of the employ the sample person worked. Another example is the earnings history info It is limited to nine interviews covering a span of 2 1/2 years. 2/ La analysis and human capital estimation would be greatly enhanced by the employer characteristics or of more earnings history information. Anot for data augmentation is the addition of data similar to that collected SIPP, but from an alternative source for which the data are (or are pr fairly accurate. This would permit an evaluation of the SIPP data on a basis to determine response error and response bias in the SIPP. This especially important for the SIPP because many of the data items in th new to household collection efforts.

There are numerous techniques by which additional data could be prepared such as:

(1) a follow-up survey,
(2) modeling and imputation using statistical techniques (such as regression analysis),
(3) statistical matching, and
(4) exact matching.

Of course, there is no "best" technique, but the technique is dependent on a number of factors: 3/

(1) availability and structure of auxiliary data (e.g., tabulations, nonidentifiable microdata or identifiable microdata);
(2) overlap between the survey and the auxiliary data set;
(3) for microdata, presence of common and unique matching identifers;
(4) requirements imposed by the statistical use (such as data accuracy); and
(5) cost and timing considerations for the project under the alternative techniques.

This paper will focus on projects which use the exact matching technique to enhance data in the SIPP with data from other sources, primarily administrative records. The use of exact matching for evaluation or survey data enhancement is well recognized as an established and valuable approach. Examples of such matches would be the March 1973 CPS-IRS-SSA Exact Match file, 4/ the February 1978 CPS-IRS Exact Match file, the April 1980 CPS-SSA Exact Match file, and the post-1980 annual March CPS-SSA and March CPS-IRS Exact Matches. The advantages to the exact match approach are fairly well known and will not be elucidated here, except to note that this is the best approach to the projects delineated below in terms of the availability of identifiable microdata, matching ease, data utility, project requirements, data reliability, and of course, cost.

In fact, "From the survey's inception, the SIPP data system has always been planned as a combination of data from administrative records and household surveys.

This reduces respondent burden by looking to other data sources for di
to obtain information. Household interview responses will be supplem
information from program files such as the earnings and benefit record
Social Security Administration. This allows, for example, analysis o
term impact of various social security benefit formulas, which are a
postearnings histories and retirement decisions of currently ineligibl
rather than characteristics of current recipients. To make these lin
[and efficient], social security numbers are required of the individua

B. MATCHING KEYS

For the exact match projects outlined below, there are three basic ma
keys: the social security number, the person's name, and the employe
address/employer identification number reported in the SIPP.

1. SSNs

A social security number (SSN) is requested of all household mem
in the SIPP, including children under 16 years of age. Potential
responses accepted in the SIPP are (a) the SSN, (b) a missing SS
(c) don't have an SSN, or (d) refused. For all SSNs reported in
SIPP a computer validation is performed by comparing the SIPP su
(first six digits), age, sex, and race to the SSA's SS-5 data it
(name, age, sex, and race) and preparing a validity code (0 to 1
showing the combinations of agreement/disagreement. Valid SSNs
accepted; invalid SSNs are treated as SSN nonresponse. The secor
step in the process is to attempt to obtain an SSN for those cas
with a missing or invalid SSN. The SSN "refusals" and "don't ha
SSN" are, of course, excluded from this process. This is a manu
search operation using the reported name, age, race, and sex, a

the SSA's SOUNDEX [illegible] search operation. New SSNs located in this operation are then appended. If the SSN is still unavailable, then a field followup is conducted to try to obtain the SSN from the respondent and to obtain any missing data needed to successfully conduct the SOUNDEX search. These cases are then recycled through the operation. At that point, SSNs still missing are accepted. An additional complexity is the longitudinal nature of the SIPP. This process is done on a rotation-by-rotation, wave-by-wave flow basis, and new survey entrants 1/ ("200 + level" persons) are also added into the process. Of persons who have an SSN, we expect about 80 percent to have reported and valid codes, 10 percent to be located via SOUNDEX searching, with about 2 percent refusals and 8 percent nonresponses. Note that these figures are based on the Current Population Survey and the 1979 Income Survey Development Program. Since more effort is expended in the SIPP to collect complete and valid SSNs, these rates may not accurately reflect the final SIPP rates. 7/ One final note, although reliability of the SSN is quite high, it is not without its own problems and uniqueness 8/ which may lead to potential mismatches and potential data problems. These potential problems are, in general, quite minor.

2. Person Names

The use of name (other than in conjunction with the SSN operation) is not expected to be used in the projects outlined below. However, recent advances in computer technology and name-matching algorithms 9/ make such matching operations a process to be considered for future work.

3. Employer Names

The project on augmenting the SIPP with employer characteristics from the Standard Statistical Establishment List (SSEL) 10/ and the economic censuses uses a computer-assisted manual search process on employer names, addresses, and EIN (where reported). This process matches the name to the SSEL to pick up a unique identification code (Census File Number (CFN)) which is subsequen used to match to the economic censuses. For a description of the process, problems and solutions, and advantages and disadvantages which is too long to include here, see Sater (1985). 11/

The remainder of the paper outlines several research initiatives, plans, and potential activities. Again, this is not intended to be exhaustive, but merely to provide a flavor of the types of activities involving the exact matching of the SIPP to auxiliary data sources.

C. SIPP-PROGRAM BENEFITS EVALUATION STUDY

Since much of the benefits data is new to a household survey, they nee evaluated. More specifically, the ultimate goal of this project is t the quality of reporting for selected questions asked in the SIPP. Re this project are expected to provide greater understanding of the acc degree of estimation bias in the data, provide an empirical basis for on questionnaire revision, and to serve as a benchmarking device for statistical adjustment. To perform this evaluation, the Census Burea administrative records from a variety of benefit programs, Federal and

individuals in the SIPP. The benefit data covered in this evaluation may include several of the following programs:

(1) Aid to Families with Dependent Children (AFDC) 12/
(2) Unemployment Compensation 13/
(3) Food Stamps 14/
(4) Workers Compensation 15/
(5) Social Security 16/
(6) Supplemental Security Income (SSI) 17/
(7) Veterans Compensation and Benefits 18/
(8) PELL Grants 19/
(9) Civil Service Retirement 20/

This requires separate arrangements with each state and, within state, each state agency. To make this a manageable task, the evaluation will be conducted only for SIPP respondents in selected states. Further confidentiality laws and restrictions 21/ cause problems in some states. The following chart shows the agency/program, government involved, and current (October 1985) status of data acquisition:

Program	Federal Government	FLA.	N.Y. City	N.Y. State	PA.	WI.
AFDC.....	-	2	5	6	2	3
Unemployment Comp.....	-	3	-	5	3	3
Food Stamps.....	-	2	5	6	2	3
Workers Compensation....	-	3	-	2	2	3
Social Security.....	3	-	-	-	-	-
SSI.....	2	-	-	-	-	-
Veterans Compensation...	3,4	-	-	-	-	-
Pell Grants.....	1	-	-	-	-	-
Civil Service Ret.....	2	-	-	-	-	-

- Means program/agency not included in project.
1 Data received, but may not be usable because of a lack of recipient address, no N.Y. recipients.
2 Data delivery expected by December 1985.
3 Data received.
4 Data exclude direct deposits and payments to institutions.
5 Still negotiating legal/administrative/cost details.
6 Data delivery expected by March 1986.

As evidenced by this table, the acquisition of the administrative rec various stages of completion. We expect this process to be completed spring, the actual matching work to be conducted in 1986, and analysi

The primary match key will be the SSN. Even though we have as accura complete set of SSNs as possible, there will still be missing SSNs wh true nonmatches or may be false nonmatches. Since the bias introduce nonmatches is unknown, we want to eliminate these as much as possible. An alternate is to select the file of beneficiaries based on individu received benefits in the reference month and who lived in the state d reference month. There are two problems with this alternative: pers address in the administrative record source was not in the state, and have moved out of the state subsequent to the reference month and have their address on the administrative record source. This will also in nonmatches.

Because both matching techniques introduce bias due to false nonmatche want to reduce this universe to as great an extent as possible. We e two-stage blend of these two techniques. Briefly stated, the match p be as follows:

(1) Eliminate all individuals not in the study universe; that is, persons not in Florida or Wisconsin and persons with SSN refusals.

(2) Provide the SSA with a finder file of reported SSNs, obtain an extract and match to the study file.

(3) Produce a residual nonmatch file.

(4) Obtain from the SSA a file of all program beneficiaries in the reference month(s) in the two states.

(5) Perform a computer/clerical match of the "nonmatch" file with the SSA extract.

The design of the SIPP file to be used in the analysis is currently being prepared and a detailed study plan for the linked data is planned for early 1986.

As with other evaluation studies, all matching work will be done at the Bureau of the Census and the linked sets restricted to individuals at the Census Bureau working on this project. Any work will be conducted so as to ensure the confidentiality of the respondents as required under Title 13 and the various Federal and state statutes for the administrative data. The only data released from this study will be in the form of summary statistics which will be subject to suppression procedures to ensure the confidentiality of the individuals involved in the study.

D. SIPP-ECONOMIC CENSUSES EXACT MATCH STUDY

The Census Bureau is currently involved in a development project to match the persons in the SIPP to the employers for which the sample person worked and then to the employer characteristics in the economic censuses. The primary goals of this pilot study are to perform the link to show it is possible, identify problems and devise solutions, evaluate the quality and completeness of the linked data set, and demonstrate the reliability and usefulness of the augmented data set. We have, so far, conducted the basic research on matching methodology, prepared the plan for the matching process, prepared the SIPP extract file, and developed numerous study plans incorporating the linked data set into the analysis. The next step is to begin the matching activity. I should stress that this is a pilot project and may be limited in scope as dictated by project funding.

The data in the SIPP will be, in themselves, valuable to labor market analysis; however, inclusion of characteristics of the employer for which the sample persons worked will greatly enhance the analysis. The potential uses suggest the need to augment the SIPP data with data on the employer.

The addition of economic data to the demographic data for individuals will enable researchers to obtain improved estimates of the impact of institutional forces which have been intensively studied but are only understood or measured. 22/ It will also enable researchers to exami of labor market outcomes and production processes that have, heretofo cult to study. Some areas in which the matched file can yield new i the relationship between capital and wage rates, labor mobility betwe high-wage employers, differences in individual earnings in low- and h firms due to the characteristics of the workers and the degree of cap of the employer, the effect of minimum wage legislation, structural u characteristics of high-tech workers in high-tech firms, implications transition from a goods to a service economy, the effect of unions on market, the substitutability of capital and labor, and productivity a

The matched data will also be useful in evaluating the reporting of by respondents in the SIPP.

Such data can be found in the Census Bureau's numerous economic censu veys, such as the Census of Manufactures (CM). For example, data fo ments from the CM include: number of employees, payroll, sales, rece worked by production employees, cost of materials and services used, capital expenditures, value of shipments; and, on a sample basis, su labor costs, value of depreciable assets, value of retired depreciabl depreciation, rental payments, and value of shipments to other plant company. These data, of course, cannot be collected in a household s

must be obtained through other methods. There are two methods: a follow-on survey of employers that were reported in the SIPP or an exact match of the employer to the data already collected in the SIPP. The former option is clearly an expensive undertaking, especially because the data are needed for establishments, are technically complex, involve issues of respondent burden/duplication and the Reduction of Paper Work Act, and would result in unknown nonresponse problems. The exact match option, however, is feasible, fairly cheap (relative to a survey effort), has known response rates (at an acceptable level), and satisfies the provisions of the Reduction in Paper Work Act. The first study initiated in this area is a pilot study.

In general, there are two approaches to this match process. The first is to match the SIPP to the economic censuses directly. Due to the content (matching identifiers), the match would have to be based solely on the employer name, address, and industry code, and would require numerous independent matches. The second process is to match the SIPP to the Standard Statistical Establishment List (SSEL) to pickup the Census File Number (CFN) and use this CFN to match to the economic censuses (using the name and address as validity checks). A potential drawback to this approach relates to the file reference periods and organizational structure changes of the employers. The economic censuses represent corporate structures as of 1982, while the SSEL represents corporate structures as of 1984. If such changes affect the CFN--that is, an establishment or legal entity changes from one enterprise to another, then there may be a problem matching the SSEL CFN to the economic censuses. For this project, the second process was chosen-the use of the SSEL as an intermediary in the record linkage process. We have conducted

preliminary work on the matching process and have shown that the proce implemented; it is feasible and should give fairly good match rates fo lishments and enterprises. More detail can be found in the recent pap later. 23/

E. SIPP-SUMMARY EARNINGS RECORDS MATCH STUDY

Another potential for improvement of the SIPP data for labor market an lates to the 2 1/2-year span covered by the SIPP and the need for more earnings history data. This can, to some extent, be remedied by match Summary Earnings Records (SER) to the SIPP. As mentioned previously, of this nature is not new; for example, the 1973 CPS-IRS-SSA exact mat the CPS with, among other things, SSA's earnings history data. The SE covered earnings information annually since 1957, with an aggregate an all years prior to 1957. "Covered earnings" include wage and salary nonfarm self-employment income for all persons covered by the Social S system. There are numerous studies for which this information would b A few examples would be

(1) Evaluation of lifetime earnings estimates from cross-sectio data (e.g., CPS).

(2) Improvement in income distribution estimation, and developme of nonsurvey based income distribution estimates.

(3) Estimation of "human capital," estimation of the impact of employment breaks on earnings, and incorporation of earn history into economic and labor market analysis.

(4) Estimation of future Social Security Program costs assessme of effects of Social Security policy changes.

(5) To evaluate the efficacy of SSA procedures for estimating c wages above the taxable maximum.

(6) To augment SSA's longitudinal earnings information with variables such as noncovered wages, education, and family status; this approach may lead to a better understanding of the lifetime covered employment patterns of persons who have contributed to the Social Security Program.

(7) How does the level of assets and income of the retired vary with their earnings history?

(8) To what extent does the relative economic status of people remain the same throughout their lives?

(9) Are different cohorts of nonaged men (and/or women) with similar wage histories accumulating assets at a different rate? How do the income and assets of men and women with the same wage history differ? At different ages with different earnings histories?

F. SIPP-MASTER BENEFICIARY RECORDS MATCH STUDY

Another potential for the improvement of the SIPP data would be the augmentation of the SIPP data with data from the SSA's Master Beneficiary Records (MBR). The MBR 24/ is a complete file of all individuals who have been awarded Title II benefits and includes wage earners and/or dependents of wage earners who have been awarded old age, survivors, and disability insurance under Title II of the Social Security Act. The MBR also contains cross-references to applicants filing since the period of 1972 to 1975 for Railroad Retirement, Supplemental Security Income, health insurance (Medicare), and black lung. There are some nuances to this universe: terminations prior to 1962 are excluded, and, since the period of 1975 to 1979, all individuals who have filed an application regardless of award status are included. The file contains about 90,000,000 individuals (as of March 1982) with 36,000,000 in current pay status; 6,000,000 disallowed or denied; 40,000,000 terminated; and the remainder in other statuses 24/.
The file includes the following data items, if applicable to the individual: name, mailing address, SSN, age, race, sex, health status (for disability cases

only , benefit status and type, benefit entitlement amount (on a histor and the amount of the current month's check. Note that entitlement am actually equal the amount of the check due to status changes and adjust previous over/under payments.

The data from the applications for benefits (Forms 1-19) are transcrib SSA Form 450 and then data keyed. This initiates the generation of an record and checks for eligibility. The data are collected and changed and then the master file is updated monthly. Corrections are similarl Because these are administrative data, we generally assume it is of fai quality and errors (as to the extent they affect award status/amount) quickly detected and corrected. Apparently, results of periodic evalua be found in "Report on Policies and Procedures for Establishing Initia ment to RSDI Benefits" from 1964 to 1979.

There are numerous studies for which the SIPP-MBR linked data set woul invaluable. A few are

(1) Preparation of characteristics of program participants and t testing of social science theories with regard to the Social Security Program.

(2) An analysis of participation and benefit amount stability.

(3) Assessment of the degree of dependency on the benefit, economic status of beneficiaries, and how this changes with respect to duration of entitlement.

(4) Analysis of the changes in beneficiary characteristics (old vs. new beneficiaries).

These studies could be undertaken with a SIPP-MBR Exact Match and the listed in Section E could be undertaken with a SIPP-SER Exact Match; ho

interaction between the MBR and the SER data would also enhance these studies. Thus, if a SIPP-MBR or SIPP-SER is to be undertaken, it would seem logical that both should be; i.e., a SIPP-SER-MBR match.

G. OTHER ADMINISTRATIVE RECORD DATA SOURCES

The projects noted in Sections E and F, the SIPP-SER and SIPP-MBR, are potential studies and in the early discussion stages. There are a few other data sources that may be worth investigating to assess the data availability, data content, matching feasibility, and potential uses of the data. I should stress that these are only ideas that may be worth investigating.

(1) Medicare enrollment and usage data
(2) IRS income information
(3) Asset and wealth information from IRS sources
(4) Unemployment insurance information
(5) National Death Index--for respondents who have died during the panel collection period

H. PROBLEMS

Clearly, the potential for analysis using SIPP data is greatly enhanced by augmentation from other data sources. However, it is not without its own problems. There will be legal and administrative access and use problems. For data such as the SER or MBR, the access problems in the development of a SIPP-SER or SIPP-MBR Exact Match are not major stumbling blocks 21/ and there are precedents, such as the 1973 CPS-IRS-SSA Exact Match studies. 4/ For other data sources, there may or may not be major problems. However, there may be usage problems. The data in the SIPP are protected under the confidentiality provisions of Title 13, and each administrative record source has its own confidentiality provisions. These provisions can be met by the Census Bureau and the linked file can be made available

to Census Bureau staff for analysis. Publications or data releases fr analysis can be made in such a fashion to satisfy all the applicable tiality of the respondents. With the addition of administrative recor containing sufficient detail and accuracy for analysis, the maintenan confidentiality cannot be assured. We are working on solutions to th but, in the interim, we cannot release a microdata tape of linked SIPP Administrative Record data in public-use form.

Although a substantial amount of work is done to verify and improve t key (i.e., SSN), there are still missing or refused SSNs as well as er SSNs.6/ Thus, the data must be adjusted for matching problems to acco false matches, false nonmatches, and mismatches.8/ 9/ The identificat these is not at all straightforward, and adjustment technique may be (such as reweighting). Further, the technique may be different depen the analysis needed.

The SER data are only for earnings up to the social security maximum wage base. This now covers over 90 percent of all wages, but has dip 80 percent in the 1970's. Depending upon the analysis, adjustments ma approximate total covered earnings. The SER data have no data on oth income or assets and, therefore, at best, can only be a proxy for pas status. (Again, social security covered earnings are the major sourc the United States; the proportion of total income it represents varie with income.)

Since about 1976, the SER is based on an annual reporting system. The quarters of coverage are estimates based on the total annual covered earnings. That is, many breaks in employment will not be captured because the total covered earnings qualifies for more quarters of coverage than was actually incurred.

The SER data like the SIPP data are "before tax" income; therefore, "after tax" income would have to be estimated to determine savings behavior. If this is required for the analysis on a historical basis, it could be difficult indeed. Also, the SER and MBR data are in nominal dollars. Thus, analysis using the data may require adjustment to real income.

I. SUMMARY

This paper has outlined, in general terms, the potential to augment the SIPP data with data from alternate sources such as the economic censuses and selected administrative records, and pointed to the potential uses of the linked data set. While the data augment process will be extremely valuable for the SIPP, it is not without some problems which need to be addressed. This outline lists several projects which range from just ideas to operational projects. Clearly, more work in these areas is needed.

J. FOOTNOTES AND REFERENCES

1/ For more information on the SIPP, see:

a. Frankel, Delma T., "A Summary of the Content of the SIPP," SIPP Working Paper Series, No. 8504, 1985.

b. Herriot, Roger A., and Daniel Kasprzyk, "The Survey of Income and Program Participation," SIPP Working Paper Series, No. 8405, 1984.

2/ There are, of course, valid reasons why such data cannot be collec
household survey such as the SIPP. One reason is that the inform
may not be known by the respondent (such as the employer characte
Another is a question of reliability--collection of earnings histo
formation for a long time span will have substantial recall error
may be survey cost and respondent burden. Collection of earnings
for a 20- or 30-year career would be a substantial undertaking.

3/ For more information on the statistical uses of administrative re
on matching techniques see:

a. U.S. Department of Commerce, Statistical Policy Working Paper 5--Report on Exact and Statistical Matching Techniques, 1980.

b. U.S. Department of Commerce, Statistical Policy Working Paper 6--Report on Statistical Uses of Administrative Records.

4/ The 1973 CPS-IRS-SSA Exact Match Study, a joint undertaking betwe
Bureau and the Social Security Administration, linked survey reco
in the March 1973 Current Population Survey to earnings history an
information from the SSA as well as selected items from the 1972
income tax returns. Quite a plethora of papers were prepared on
project. For more information see:

a. Kilss, Beth, and Frederick J. Scheuren, "The 1973 CPS-IR
Exact Match Study," Social Security Bulletin, October 197

b. U.S. Department of Health and Human Services, Studies fr
Interagency Data Linkages, Reports 1 through 11.

c. Various authors in the American Statistical Association Proceedings, Social Statistics Section, annual reports, 1973 through 1985.

Also of interest is the following paper:

Butz, William P., "The Future of Administrative Records in the Census Bureau's Demographic Activities," American Statis Association Proceedings, Social Statistics Section, 1984.

5/ Kasprzyk, Daniel, and Charles A. Lininger, "The 1982 Survey of In Program Participation," American Statistical Association Proceedi Social Statistics Section, 1980.

6/ Kasprzyk, Daniel, "Social Security Number Reporting, the Use of A strative Records, and the Multiple Frame Design in the Income Sur Development Program," Technical, Conceptual and Administrative Les in the Income Survey Development Program (ISDP), 1983.

7/ A report on the SIPP process for validating and searching for missing SSNs is currently being prepared by Nora Bea Parker and is expected to be released in the near future. In the interim, the following documents may be of value for further information.

a. Cobleign, Cynthia, and Wendy Alvey, "Validating Reported Social Security Numbers," American Statistical Association Proceedings, Social Statistics Section, 1974.

b. Kilss, Beth, and Barbara Tyler, "Searching for Missing Social Security Numbers."

c. Remington-Rand, "Soundex-Foolproof Filing System For Finding Any Name in the File," Brochure LBV809 (undated).

8/ See Jabine, Thomas B., "Properties of the SSN Relevent to Its Use in Record Linkage," Record Linkage Techniques--1985. This is a collection of papers presented at or resulting from the Workshop on Exact Matching Methodology that was held May 10, 1985. The report is expected to be released in the very near future.

9/ See selected papers in Record Linkage Techniques--1985, a collection of papers from the Workshop on Exact Matching Methodology that was held May 10, 1985. The report is expected to be released in the near future.

10/ U.S. Department of Commerce, The Standard Statistical Establishment List Program, Technical Paper 44, 1979.

11/ For more details on the project outline, matching methodology, problems, and solutions, see Sater, Douglas K., "Enhancing Data from the Survey of Income and Program Participation with Data from Economic Censuses and Surveys," SIPP Working Paper Series, No. 8505.

12/ AFDC--Aid to Families with Dependent Children (AFDC) or Aid for Dependent Children (ADC) is a program that is state or locally administered. Eligibility for AFDC (or ADC) is based upon (1) the incapacity or death of a parent; (2) a parent's continued absence from the home; or (3) the unemployment of a parent.

13/ Money received from state government unemployment insurance funds to help replace wages lost during times of layoff or unemployment.

14/ Food stamps are coupons that can be used to purchase food. The food stamps program is a joint Federal-state program which is administered by state and local governments.

15/ Worker's Compensation payments are generally made under state laws set up to compensate workers for loss of pay because of accidents or illnesses caused by or happening on the job. Employees of the Federal Government, longshoremen, and harbor workers are also covered under special Federal programs.

Payments may be one-time compensation awards or may be made for periods in situations where the disabling condition persists fo In cases of permanent and total disability, about three out of make payments for life. The great majority of cases involve cas in situations where a worker is unable to work for a temporary which the employee will be able to return to work. Payment amo calculated as a fraction of the worker's usual wage -- most comm two-thirds. One-third of the states provide additional benefit dependents.

Administration of a worker's compensation payment is not uniform states or even within each state. Payments may come directly f employer, from an insurance company retained by the employer, o a state worker's compensation fund.

16/ Social security payments are received by persons who have worke enough in "covered" employment (that is, employment in which soc deductions were made from a person's salary) to be entitled to Social security payments may also be made to the spouse or depe of covered workers.

The Social Security Administration also pays benefits to studen (under age 22) of eligible social security annuitants who are d deceased. Benefits are paid by the Federal Government directly student.

Note that other programs administered by the SSA, such as black fits, are not included.

17/ SSI is a Federally-administered program that pays monthly benef disabled, and blind people who have limited income and assets. can also be paid to disabled and/or blind persons regardless of Further, a person may be eligible for SSI payments even if they worked.

Even though the Social Security Administration administers the program, income from SSI is NOT the same as social security. T for these programs comes from two distinct sources and, if eligi can get SSI in addition to social security.

The state SSI program is administered by the Social Security Adm Each state may add to the Federal payment from its own funds. additional money may be included in the Federal payment or it m ceived as a separate check. If it is combined with the Federal the words "STATE PAYMENT INCLUDED" will appear on the Federal c A few states make SSI payments to individuals who do not receiv payment.

18/ There are two types of veterans disability pensions a person can receive.

a. Compensation for service-connected disability--These payments are made to veterans with a disability that was caused or aggravated by military service. Payment amounts are related to extent of disability which may range from 0 to 100 percent. Veterans with ratings of 50 percent or more may receive additional compensation for dependent wives, children under age 18 or under 22 and attending school full-time, and dependent parents.

b. Nonservice-connected disability pensions--The second type are VA PENSIONS made to veterans whose disability is not service connected, but who served on active duty in wartime and are in financial need. Veterans who are under age 65 must be permanently and totally disabled. Veterans who are 65 and over and not working are also eligible for those pensions.

Payments vary depending on such factors as income, number of dependents, need for the regular aid and attendance of another person, whether the veteran is permanently housebound, and whether time is spent in a VA medical facility.

When the pension is originally awarded and in January during each year that the pension is received, the veteran receives a letter from the Veterans Administration indicating the amount of the monthly payment and which person in his immediate family besides himself is included in it. In certain instances, the dollar amount for each person is also indicated in the letter.

19/ PELL Grants--These are direct payments to students to permit access to post-secondary education for all students. Persons eligible are undergraduate students enrolled in institutions of higher education on at least half-time basis who are making satisfactory progress. The institution in which the student attends generally determines the amount of the award using a payment schedule based on the cost of attendance, the student's financial need, and the level of appropriation. Generally, the institutions also act as dispersing agents, but the program is administered by the Department of Education.

20/ Federal employee retirement pensions are payments by the Federal Government to its retired and/or disabled civilian employees. Civilian Federal Government employees usually worked in the U.S. Civil Service but may also have been postal workers or former employees of other civilian branches of the Federal Government, such as the Federal Reserve Bank System, Foreign Service, Tennessee Valley Authority, and Public Health Service.

Payments are made by green-colored U.S. Treasury checks or via deposit in the person's bank account. At the present time, mos payments are still made by check. They are made monthly and us arrive in the first few days of the month.

This does not include U.S. military retirement pay.

21/ Issues relating to privacy, confidentiality, and access to administrative records, are discussed in:

a. Robbin, Alice, and Linda Jozefacki, Compendium of Stat Legislation on Privacy and Access, 1983.

b. Sater, Douglas K. et al, "Administrative Records Subcommittee on Privacy, Confidentiality, and Public Perception," Draft Report, 1983.

c. Butz, William P., "The Future of Administrative Records in the Census Bureau's Demographic Activities, American Statistical Association Proceedings, Social Statistics Section, 1984.

d. Alexander, Lois A., "Statistical Progeny of Administrative Records: Some Legal Issues," American Statistica Association Proceedings, Social Statistics Section, 19

e. Alexander, Lois A., and Thomas B. Jabine, "Access to Social Security Microdata Files for Research and Statistical Purposes," Social Security Bulletin, August 1978.

22/ Extracted from:

a. Haber, Sheldon et al, "Matching Economic Data to the Survey of Income and Program Participation: A Pilot S American Statistical Association Proceedings, Social Statistics Section, 1984.

b. Haber, Sheldon, "Applications of a Matched File Linkin the Bureau of the Census Survey of Income and Program Participation and Economic Data," SIPP Working Paper Series, No. 8502, 1985.

23/ Sater, Douglas K., "Enhancing Data from the Survey of Income an Program Participation with Data from Economic Censuses and Sur SIPP Working Paper Series, No. 8505, 1985.

24/ Data Linkage Working Group, Subcommittee on Statistical Uses of Administrative Records, Federal Committee on Statistical Methodology, "Record System Description of the Master Beneficiary Record File," Statistical Uses of Administrative Records: Recent Research and Present Prospects, Vol. 1, 1984.