Introduction

The Social Security Administration (SSA) produces a popular statistical series, *Income of the Population 55 or Older*, to meet the demand for statistics on the receipt of income from various sources, income distributions, aggregate income, and poverty. This statistical series is based on data from the March Supplement to the Census Bureau’s Current Population Survey (CPS). These statistics are affected by the selection of the CPS as its data source.

This article provides insight into how the data collected in the CPS affect measures describing the economic well-being of the elderly. Because *Income of the Population 55 or Older* is based on a survey, the accuracy of its statistics is dependent on the willingness and ability of CPS respondents to answer survey questions accurately. Different surveys have different strengths and weaknesses, and one method of assessing the differences is to compare one survey’s data with those of other surveys. The article also compares statistics calculated using the CPS with another Census Bureau survey that is particularly strong at measuring income—the Survey of Income and Program Participation (SIPP).

Another method for determining the accuracy of survey data is to compare them with administrative data. The Census Bureau in collaboration with SSA has matched administrative records for the Social Security and Supplemental Security Income (SSI) programs to the SIPP and selected years of the CPS. Although administrative data are not used in *Income of the Population 55 or Older*, they are used in parts of this article to evaluate the accuracy of the reported data for Social Security benefits and Supplemental Security Income in both the CPS and the SIPP.

First, the major features of the CPS and SIPP are outlined, and a description of SSA’s administrative data available to be matched to those two surveys is also given. The article then compares estimates from the SIPP and CPS of the proportion of the elderly receiving income from various sources. Next, administrative data for Social Security benefits and SSI are used to evaluate the accuracy of the estimates derived from the surveys. Finally, the conclusion discusses the tradeoffs involved in selecting a data source.

Data Sources Used in this Analysis

The Social Security Administration has been producing two series of publications on the income of the elderly and near-elderly—*Income of the Population 55 or Older* (IP55), since 1976 and *Income of the Aged Chartbook*, since 1990. Both series are derived from the March Supplement to the Current Population Survey, which is conducted annually by the
U.S. Census Bureau. Another Census survey that asks more detailed questions regarding income is the Survey of Income and Program Participation. This analysis uses income data for 1996 because the SIPP match rate to administrative data declined considerably for the 2001 panel.2

**Survey Data**

Depending on their respective purposes, surveys may differ in many ways: the subjects covered, the questionnaire length and detail, the frequency of interviews, and the sample of the population surveyed are a few of the dimensions along which surveys may vary. The following descriptions of the CPS and SIPP outline some of the broader differences between the surveys that may influence differences in the statistics produced using their data.

**Current Population Survey.** The CPS is a monthly survey conducted by the Census Bureau and sponsored jointly with the Bureau of Labor Statistics. The CPS is a representative sample of the “civilian noninstitutional population.” The survey has been conducted for more than 50 years and is used for official monthly unemployment and labor force statistics. Annual income data have been collected in the March Supplement since 1948, with information gathered concerning income received during the previous calendar year for approximately 35 cash and in-kind sources. Official estimates of income and poverty in the United States are based on the March Supplement (Herz 1996; U.S. Census Bureau 2005 and 2006). Table 1 summarizes major features of the CPS.

This article uses the 1997 March Supplement to the CPS (then called the Annual Demographic Survey), which collected data for annual income received in 1996. Persons are designated as “65 or older” if they were at least age 65 as of the interview in March 1997. In SSA’s two data series (cited earlier), aged units—classified as “65 or older”—are defined as follows:

- nonmarried persons 65 or older (in March 1997),
- or married couples in which either
  - the husband is 65 or older, or
  - the husband is younger than 55 and the wife is 65 or older (in March 1997).

**Survey of Income and Program Participation.** The SIPP is a longitudinal panel survey conducted by the Census Bureau. The survey is designed to collect data on sources and amounts of income to provide improved statistics on the distribution of income in the United States. The sample is also representative of the noninstitutional population. The first SIPP panel began in October 1983; subsequent panels have begun periodically with durations ranging from 2½ to 4 years. Data are collected on approximately 70 cash and in-kind sources of income (U.S. Census Bureau 2005 and 2006). Table 1 summarizes major features of the SIPP.

In this article, analysis of SIPP income data is based on information collected in the 1996 SIPP panel concerning income received during 1996. Persons classified as “65 or older” were at least age 64 as of the interview in March 1996. Aged units classified as “65 or older” were nonmarried persons 64 or older (in March 1996) or married couples in which the husband is 64 or older or the husband was younger than 54 and the wife was 64 or older (in March 1996).3

**Administrative Data**

The Social Security Administration administers two cash benefit programs: (1) Social Security, or Old-Age,
Survivors, and Disability Insurance, and (2) Supplemental Security Income, the separate program for low-income aged and disabled persons. In this study, statistics based on administrative data refer to statistics based on a combination of survey-reported and administrative information. For matched survey records, self-reported data on program benefit receipt were replaced with Social Security’s administrative data. Self-reported data from the survey were used for unmatched survey respondents. A slightly greater proportion of observations in the 1996 SIPP were matched with Social Security administrative records (85 percent of those aged 64 or older present in March 1996) than in the CPS (77 percent of those 65 or older).

Social Security (OASDI). Survey-based Social Security benefit amounts have been replaced with the administrative amount of the Social Security benefit paid to the beneficiary plus the beneficiary’s Medicare Part B premium (when the latter is applicable) in both the SIPP and the CPS.

Supplemental Security Income (SSI). Survey-based SSI payments have also been replaced by administrative payment amounts. The process is somewhat more complicated for the SSI program because there are federal and state components to SSI payments. The SIPP and the CPS treat this differently in their questionnaires: the SIPP asks specifically for federal payments and state payments separately; the CPS asks respondents for a single, combined SSI payment amount.

For states with federally administered state SSI payments, both federal and state SSI amounts were taken directly from administrative payment data files and were used to replace reported SSI payments for matched observations for both the SIPP and the CPS. For states without federally administered state SSI payments, the procedure was different for processing the SIPP and the CPS. For the SIPP, survey-collected federal payments were replaced by administrative data, and survey-based state payments were not changed. For the CPS, the survey-based SSI payment (combined state and federal) amount was replaced by administrative information.

Impact of Survey Selection on Estimates of the Relative Importance of Social Security Benefits

Traditionally, economic resources for the elderly are expected to come from three sources: Social Security, pensions, and savings. Collectively these sources are referred to as the “three-legged stool” of retirement. If one or more of the “legs” are missing, the “stool” may not be able to provide adequate support during retirement, particularly if economic or health shocks occur. Therefore, elderly beneficiaries who receive all of their income from Social Security are considered economically vulnerable. Depending on whether the CPS or the SIPP was used, the proportion of elderly beneficiaries receiving all of their income from Social Security could appear quite different. In the CPS data for 1996, 17.9 percent of aged units 65 or older reported that all of their income came from Social Security benefits compared with only 8.5 percent in the SIPP.

The difference in results from the two surveys may be due in part to two factors: the survey recall period and the structure and detail of the questionnaires. First, respondents in the CPS are asked in March about income from the previous calendar year, while participants in the SIPP are asked about their income in 4-month intervals. Second, respondents in the SIPP are asked more detailed questions about their sources of income. For example, for questions involving asset income, SIPP respondents are asked if they own a particular asset and are then asked if they received income from that asset and how much; CPS respondents are only asked if they received a particular type of asset income and how much. These factors make participants in the SIPP more likely to report small and/or infrequently received amounts of income (U.S. Census Bureau 2006; Hurd, Juster, and Smith 2003).

Table 2 illustrates the differences in the percentages of aged units 65 or older reporting various sources of income. With the exception of earnings, respondents in the SIPP report receipt of income in every major category more often than respondents in the CPS. The gaps between the SIPP and the CPS are particularly large for asset income—especially interest—and private pensions. The large differences in reported receipt of asset income and pensions between the CPS and SIPP may be responsible for the differences in the proportions of beneficiary aged units who report receiving all of their income from Social Security.

According to Income of the Population 55 or Older (based on the CPS), the proportion of aged units 65 or older (referred to here as elderly aged units) reporting that they received all of their income from Social Security benefits rose from a low of 12 percent in 1990 to a high of 20 percent in 2002 (Chart 1). The receipt of asset income appears to be negatively correlated with the percentage of elderly aged units reporting that they received all of their income from Social Security.
Security since 1978, especially from 1990 through 2002 (Chart 2). From 1994 through 2002, receipt of asset income fell from 67 percent to 55 percent. The same observation does not hold for receipt of pension income, which hovered between 41 percent and 43 percent over the same period (Chart 3).

As seen in these charts, restricting the comparison to beneficiary units does not alter the relationship between the estimates of elderly aged units who received asset income and elderly aged units who received Social Security as their sole source of income. This relationship is an indication that the difference in the reporting of asset income may be an important factor contributing to the divergence in measures of the relative importance of Social Security benefits between the CPS and the SIPP.

Since it is more likely that small amounts of income are being overlooked by the CPS rather than by the SIPP, it is useful to consider the other income sources of beneficiary aged units that are receiving most, but not all, of their income from Social Security in the SIPP. Of those beneficiary aged units 65 or older receiving at least 95 percent, but not all, of their income from Social Security, 86.5 percent reported asset income, while only 11.5 percent reported pension income. Similarly, 82.2 percent of those receiving at least 90 percent, but not all, of their income from Social Security reported asset income, and 22.3 percent reported pension income, which again suggests that differences in the reporting of asset income may be responsible for the major portion of the discrepancies in the rates of complete reliance on Social Security between the two surveys.

In addition to the frequency of interviews and detail of the questionnaire, the effect of the order of questions in a survey can be significant. Hurd, Juster, and Smith (2003) examined the Health and Retirement Study (HRS) and Assets and Health Dynamics of the Oldest-Old (AHEAD) and compared waves that asked asset and income questions in separate components to waves that had a “merged” asset and income module. The merged module had questions about particular types of assets followed immediately by questions about income from those assets. They found that asset income was more likely to be reported for the “merged” format that was not able to be explained by an increase in reported assets. They suggested that asset income would be better reported in the CPS if questions about asset ownership preceded the questions about income from those assets.

Other studies have compared the CPS and SIPP with other benchmarks to determine the accuracy of their respective data. Roemer (2000) found that for 1996 the CPS underestimated aggregate income from Social Security and Railroad Retirement by 8 percent, SSI by 16 percent, pensions by 23 percent, and asset income by 29 percent. He also found that for 1996 the SIPP underestimated aggregate income by more than the CPS in these categories, except for pensions, which were underestimated by 14 percent, and SSI, which was overestimated by 1 percent. Koenig (2003) and Fisher (2005), however, found that the SIPP does a better job of classifying Social Security beneficiaries and SSI recipients than does the CPS. Czajka and others (2003) found that asset holdings in the 1996 SIPP were underestimated, which may indicate corresponding underestimates of asset income amounts in the SIPP. Taken together, the work of these authors suggests a difficult situation for analysis: the CPS is more likely than the SIPP to underestimate whether or

| Table 2. Percentage of aged units 65 or older with income from specified source, 1996 |
|-----------------------------------------------|----------------|----------------|
| Source                                      | CPS   | SIPP  |
| Number of aged units (thousands)             | 24,553| 25,671|
| Percentage of aged units reporting—          |       |       |
| Earnings                                     | 20.7  | 21.4  |
| Retirement benefits a                        | 93.1  | 96.1  |
| Social Security a                            | 90.6  | 94.2  |
| Pensions a                                   | 41.2  | 55.7  |
| Public a                                     | 13.6  | 20.3  |
| Private a                                    | 29.9  | 42.9  |
| Asset income a                               | 63.0  | 73.5  |
| Interest a                                   | 60.9  | 71.5  |
| Not interest a                               | 24.7  | 32.0  |
| Public assistance a                          | 5.9   | 9.7   |
| SSI a                                        | 5.6   | 7.9   |
| Other public assistance a                    | 0.4   | 2.5   |
| Veterans’ benefits a                         | 4.6   | 5.9   |
| Unemployment compensation                    | 0.8   | 0.8   |
| Workers’ compensation a                      | 0.6   | 1.0   |


NOTES: Estimates are weighted.


a. Difference between the surveys is significant at the 99 percent level.

b. Difference between the surveys is significant at the 95 percent level.
Chart 1.
Percentage of aged units 65 or older reporting all income from Social Security, 1978–2004

SOURCE: Income of the Population 55 or Older, selected years.

Chart 2.
Percentage of aged units 65 or older reporting asset income, 1978–2004

SOURCE: Income of the Population 55 or Older, selected years.
not a source of income was received, while for those reporting receipt, the SIPP may be more likely than the CPS to underestimate the amount of income received by an individual.

**Impact of Using Administrative Data in Place of Self-Reported Data on Estimates of the Relative Importance of Social Security Benefits**

Several studies have evaluated the effect of substituting Social Security’s administrative data for respondents’ answers for income received from Social Security’s Old-Age, Survivors, and Disability Insurance and SSI programs. Sears and Rupp (2003) and Huynh, Rupp, and Sears (2002) compared administrative data with SIPP-reported data, and Koenig (2003) and Fisher (2005) compared administrative data with reported data for both the CPS and the SIPP. These studies identified several sources of reporting error for both Social Security and SSI income amounts:

- The respondent does not report the income at all.
- The respondent misreports SSI payments as Social Security income or vice versa.
- The respondent reports the current monthly payment amount, which may differ from the payment received in the survey’s retrospective time period.
- The respondent reports the amount received on the Social Security check, which is lower than the actual monthly benefit because it excludes the Medicare Part B premium deducted from the Social Security benefit.12

Koenig (2003) found that Social Security benefit amounts were more likely to be underreported in the 1996 SIPP than the March 1997 CPS because the SIPP was more likely to net out Medicare Part B premiums.13

For the administrative statistics in Tables 3 and 4, self-reported data have been replaced with Social Security’s administrative data for matched observations. Self-reported data continue to be used for unmatched observations. Table 3 illustrates that the CPS reports slightly lower rates of receipt of Social Security benefits and SSI compared with administrative data, while the SIPP reports slightly higher rates of benefit receipt.14 Part of the variation between the two surveys after the substitution of administrative data may be due to the higher administrative match rate in the SIPP.15

The differences in the estimates of the relative importance of Social Security benefits from using administrative data are far smaller than the differences from using a different survey, as shown in Table 4.16 Misreporting SSI as Social Security benefits and vice versa...
versa can affect classification as a beneficiary and as a completely reliant beneficiary, changing estimates of complete reliance on Social Security benefits. Koenig (2003) found that the SIPP is less likely than the CPS to misclassify a person as a beneficiary. For matched elderly persons receiving all of their individual income from Social Security (with either reported or administrative data), Fisher (2005) found that the dominant problem with SIPP was the classification of nonbeneficiaries as Social Security beneficiaries. The CPS, however, was only slightly more likely to misclassify a Social Security beneficiary as a nonbeneficiary than vice versa. SSI payments were much more likely to be omitted or reported as Social Security benefits in the CPS than in the SIPP.  

Conclusion

The purpose of this exercise was to ascertain how the choice of data source affects our understanding of the resources available to the elderly. As demonstrated in this article, the selection of a data source is highly influential on the results; nevertheless, tradeoffs exist when choosing a source.

Several factors must be considered when selecting a survey. The CPS is timely, released within approximately 6 months of data collection, and used by other agencies for official statistics; however, it underreports receipt of income compared with the SIPP. The SIPP may identify recipients missed in the CPS, but panels have not started at regular intervals and the survey may soon be replaced. Other surveys, such as the Survey of Consumer Finances and the Panel Study of Income Dynamics, do not provide enough detailed information on individual income sources, do not have a large enough sample to support detailed distributions of the composition of income by demographic characteristics, or do not provide person-level data. In the cases of *Income of the Population 55 or Older* and *Income of the Aged Chartbook*, availability of timely data at regular intervals and the ability to work with person- and family-level data have led to the selection of the CPS as the underlying source of data.

Although administrative data are more accurate for income receipt, these data alone are not able to fully address important policy analysis questions. SSA’s administrative data on Social Security benefits, SSI, and earnings are limited to information needed to operate its programs with limited demographic information.

Matching administrative data with survey data also results in complicated tradeoffs. For example, the administrative data can be matched to the SIPP and

---

**Table 3.**

<table>
<thead>
<tr>
<th>Income</th>
<th>CPS</th>
<th>SIPP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reported</td>
<td>Administrative</td>
</tr>
<tr>
<td>Retirement benefits</td>
<td>93.1</td>
<td>94.1</td>
</tr>
<tr>
<td>Social Security</td>
<td>90.6</td>
<td>92.0</td>
</tr>
<tr>
<td>Public assistance</td>
<td>5.9</td>
<td>6.9</td>
</tr>
<tr>
<td>Supplemental Security Income</td>
<td>5.6</td>
<td>6.6</td>
</tr>
</tbody>
</table>


NOTES: Estimates are weighted.


---

**Table 4.**

<table>
<thead>
<tr>
<th>Survey/match</th>
<th>90 percent or more of income from Social Security</th>
<th>100 percent of income from Social Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPS reported</td>
<td>30.4</td>
<td>17.9</td>
</tr>
<tr>
<td>CPS administrative</td>
<td>30.0</td>
<td>17.3</td>
</tr>
<tr>
<td>SIPP reported</td>
<td>19.9</td>
<td>8.5</td>
</tr>
<tr>
<td>SIPP administrative</td>
<td>20.5</td>
<td>8.4</td>
</tr>
</tbody>
</table>


NOTES: Estimates are weighted.


Aged units with negative earnings, asset income, or nonpositive total income are excluded.
the CPS only after the data linkages are made available to SSA following a significant processing period. Further, not all persons in the surveys can be matched to their administrative records. Even after the data are matched, the analyst must decide how to accommodate nonmatched respondents—whether nonmatched respondents should be dropped from the analysis or included using reported data. Using only matched respondents would introduce bias if these respondents differed systematically in any way from unmatched respondents. Either choice may lead to false movements in statistics over time if the proportions and/or characteristics of respondents that are matched change over time.

One conclusion that remains clear is that matching administrative data to survey data provides a strong tool in assessing the accuracy of self-reported income data. Based on the results presented here, the percentage of beneficiary aged units receiving 100 percent of their income from Social Security benefits may be overestimated when using the CPS. The statistic published in Income of the Population 55 or Older, 1996—17.9 percent of beneficiary aged units were completely reliant on Social Security in 1996—is much higher than the estimated 8.4 percent using SIPP and administrative data. Because this statistic is often used to make inferences about the well-being of the elderly, it is necessary to keep in mind the issues outlined in this article when interpreting the data.

Notes

1 The March Supplement is currently called the Annual Social and Economic (ASEC) Supplement. Before 2003, the Supplement was called the Annual Demographic Survey (ADS).

2 The more recent 2001 panel of SIPP was able to be matched to only approximately 60 percent of respondents as opposed to over 80 percent of respondents for the 1996 SIPP. Matched records will be discussed further in the Administrative Data subsection.

3 The classification of age 64 as of March 1996 as “65 or older” makes the sample comparable with the CPS sample, which is age 65 as of March 1997. This analysis uses the longitudinal core files for waves 1 through 4 of the 1996 panel of SIPP. For observations that were missing data, either because their first interviews did not collect information on January and/or February 1996 or because of the lack of an interview, the income data for the missing months were replaced with the individual’s average income for the reported months. March 1996 weights were used for three reasons: (1) March 1996 is the first month in which all rotation groups were interviewed; (2) the 1996 calendar year weights would eliminate observations that had not been present the entire year, making attrition bias a concern; and (3) using the March 1996 weights excludes persons who entered the sample because of their association with an original survey respondent.

4 Survey records were matched if they had a valid Social Security number and an age within 5 years of the age on the administrative files.

5 An earlier study by Koenig (2003) also provides analysis of CPS and SIPP data matched to administrative records. The CPS match rate here is higher than that of Koenig’s study because of the discovery in 2004 of additional survey identifiers matched to administrative records.

6 Social Security administrative data come from the monthly benefit credited from the Master Beneficiary Record (MBR), which is usually, but not always, the amount received by the beneficiary. Other studies have used the Payment History Update System (PHUS), which records the actual check amount. Discrepancies may arise between the MBR and the PHUS when payment for retroactive benefits is issued in a single check. Less than 1 percent of elderly observations had discrepancies between the MBR and PHUS. SSI administrative data come from the Supplemental Security Record (SSR).

7 Any state-administered state SSI amounts would be replaced with a value of zero during the substitution of administrative data into the CPS, resulting in the administrative estimate of SSI receipt for the CPS being a lower bound.

8 The statistics presented here were calculated using the most recent methodology from Income of the Population 55 or Older, 2002. Because small changes may have occurred in the data or the methods for calculating the statistics, the statistics presented here for the CPS may not match the statistics published in Income of the Population 55 or Older, 1996.

9 Beneficiary aged units are aged units in which the non-married person or at least one spouse of the married couple is receiving Social Security benefits.

10 Hungerford and others (2002) provide a more complete overview of the trends in income sources and the overall economic status of the elderly.

11 In these studies, SSA’s administrative records are matched to survey reports using Social Security numbers provided during the interview. The administrative amounts of Social Security benefits and SSI payments can then be compared with the respondents’ reported income.

12 Medicare Part B premiums are deducted from Social Security benefits before beneficiaries receive their payments, so the total benefit amount is generally higher than the check (or direct deposit) amount. Not all beneficiaries are enrolled in Medicare Part B, and other beneficiaries may have their premiums paid through other programs. Reporting only the net benefit makes it difficult, if not impossible,
ences in sampling methods. Part of the variation may also be a result of differences may also arise between the surveys because SIPP programs and oversamples lower-income persons to achieve this goal. Part of the variation may also be a result of differences between the surveys for Social Security is no longer statistically significant at the 95 percent level; the differences between the CPS and the SIPP for SSI and the higher-level categories of retirement benefits and public assistance remain statistically significant.

Differences between the CPS and SIPP are significant at the 99 percent level for both reported and administrative estimates. Fisher (2005) also found that the estimates of the elderly in families receiving all of their income from Social Security did not change much after substituting in administrative data, but the changes to estimates of the elderly in beneficiary families receiving at least 90 percent of their income from Social Security were larger, particularly for the SIPP. The most probable explanation for this is that SIPP respondents were more likely to report their net benefits, while CPS respondents were more likely to report their gross benefits (Koenig 2003).

The SIPP is a panel data set that follows respondents for approximately 3 or 4 years. Many respondents drop out of the survey over time, potentially resulting in attrition bias when using later years in the same panel.

The substitution for earnings was not done in this article because administrative earnings data are available annually, not monthly as for Social Security benefits and SSI payments. Although this does not pose a problem with the CPS, it is unclear how to substitute administrative earnings data across months for persons not in the SIPP for the entire year or for persons entering or exiting a marital unit midyear. In addition, while most earnings are available on the administrative file, there may be other earned income that would be reported on a 1040, but not on a W-2 or Schedule SE form, or informal earnings that would not be reported on any administrative record. Additionally, the receipt of earnings does not appear to be misreported as often as the receipt of SSI or Social Security.

References


Committee on Statistical Methodology Research Conference, Washington, DC, November 18.


