Differences in Income Estimates Derived from Survey and Tax Data

by Barry W. Johnson, Internal Revenue Service and Kevin Moore, Baord of Govenors, Federal Reserve System

The Statistics of Income Division of the United States Internal Revenue Service collects statistical data from samples of most major federal tax and information returns. Among these are annual studies of *U.S. Individual Income Tax Returns* (Form 1040). These data are used by both the U.S. Congress and the Executive Branch of the Government to evaluate and develop tax and economic policy, and by other government agencies and the general public for a variety of different purposes.

Form 1040 is filed annually by individuals or married couples to report income, including wages, interest, dividends, capital gains, and some types of business income. Also reported are data on deductions, expenses, and tax credits. The SOI sample of these returns is stratified based on: (1) the larger in absolute value of positive income or negative income; (2) the size of business and farm receipts; (3) the presence or absence of specific forms or schedules; and (4) the usefulness of returns for tax policy modeling purposes (see Internal Revenue Service, 2005).

The Survey of Consumer Finances is a survey of household balance sheets conducted by the Board of Governors of the Federal Reserve System in cooperation with the SOI. Beginning with 1983, the survey has been conducted triennially, with data collected by the Survey Research Center at the University of Michigan in 1983, 1986, and 1989, and by NORC, a national organization for social science and survey research at the University of Chicago, from 1992 forward. In addition to collecting information on assets and liabilities, the SCF collects information on household demographics, income, relationships with financial institutions, attitudes toward risk and credit, current and past employment, and pensions (for more details on the SCF, see Bucks, Kennickell, and Moore, 2006).

The SCF uses a dual-frame sample design to provide adequate representation of the financial behavior of all households in the United States. One part of the sample is a standard multistage national area probability sample (Tourangeau et al., 1993), while the list sample uses the SOI individual income tax data file to oversample wealthy households (Kennickell, 2001). This dual-frame design provides the SCF with efficient representation of both assets widely held in the population, such as cars or houses, and assets more narrowly held by wealthy families, such as private businesses and bonds. Wealth data from the SCF are widely regarded as the most comprehensive survey data available for the United States.

Sample weights constructed for the SCF allow aggregation of estimates to the U.S. household population level in a given survey year (Kennickell and Woodburn, 1999; Kennickell, 1999). Missing values in the 1989-2004 SCF were imputed using a multiple imputation technique (Kennickell, 1991, 1998b).

► Income Data

Both the SCF and the SOI file are important sources of data on the different types of income received by households and tax filers. There are a number of differences between the two sources, including the population covered, unit of observation, available data, and the motivations people face in providing data. It is also worth noting the difference in the sample size. The 2004 SOI file is a sample of approximately 200,000 tax records out of a population of about 130 million, while the sample size for the 2004 SCF is much smaller, about 4,500 households. Although the SCF has a smaller sample, the detail and scope of the data allow for a broader range of research than is possible with the tax data.

This paper was originally presented at the American Statistical Association's 2008 Joint Statistical Meetings in Denver, CO.

The population of Federal income tax filers includes only those U.S. citizens and resident aliens whose gross income, a concept defined by statute, was above legislatively prescribed thresholds. Nonresident aliens are subject to different filing requirements, based on income earned in the U.S. Income tax filers represent roughly 61 percent of the U.S. individual population (see Sailer and Weber, 1999). In addition, recent income tax filing gap estimates for Tax Year 2000 suggest that as many as 11 million taxpayers, or about 9 percent of the potential income tax filing population, either file returns late or not at all (see Brown and Mazur, 2003). In contrast, the SCF sample design ensures coverage of the entire U.S. population.

The unit of observation in the case of federal income taxes can vary according to current filing regulations. Married couples may file returns jointly, but they are also allowed to file separately when marginal tax rates favor treating the two incomes separately. Dependent children and others living in a home may also be required to file separate returns to report both earned and unearned income. Differences in the economic unit reported on income tax returns limit the data's usefulness for some types of research.

In the SCF, the area-probability sample uses a sampling frame in which the household is the unit of observation, but, for the list sample, the unit of observation is the tax-filing unit. Often the tax-filing unit is analogous to the household, but, for certain households, such as households where a married couple files separately and those with multiple subhouseholds located within a household, there are differences. While there is the possibility of frame errors in the list sample, adjustments are made during the construction of the frame and during the sampling stage to limit these distortions (see Kennickell and McManus, 1993; Frankel and Kennickell, 1995; Kennickell, 1998a; and Kennickell, 2001).

Because income tax reporting requirements are established by legislation, data concepts and definitions may not necessarily coincide with those required for economic analysis. For example, income is combined for couples who file a joint income tax return, however, for some research purposes, it would be useful to know the amounts earned by each individual. Another consideration is that, while a precise geographic location is often useful for analytical purposes, mailing addresses present on tax records may not always be the appropriate location, as when a post office box number is supplied rather than a street address. Addresses on tax returns might also be those of paid preparers rather than the filers. In some instances, a filer who owns multiple residences may even file from the address that provides the best tax advantages, rather than the address that he or she would consider 'home.'

An important aspect of data content is continuity over time, both in the items included and in the data definitions. SOI goes to great lengths to ensure both in its annual data files. However, coverage and content are subject to discontinuities resulting from changes to laws, regulations, administrative practices, and program scope. For example, income tax law revisions in 1981, 1986, 1990, and 1993 all made significant changes, both to the components of income subject to taxation and to the allowable deductions from income, that had significant impact on the statistical uses of tax return data (see Petska and Strudler, 1999).

Since surveys have more flexibility than administrative systems to specify a conceptual framework, many issues related directly to the definition and scope of the data are less pressing. However, content and valuation issues of a different sort are present in survey data. Unit and item nonresponse are two important sources of nonsampling error in surveys, though there are methods to help deal with both these issues, such as sample weight adjustments and imputation. For respondents who agree to participate and answer all the survey questions, measurement error is still a concern in survey data. Respondents may "guestimate" answers to questions; even if respondents' guesses overall are not biased, such approximation reduces the estimation efficiency of the data. Respondents may also have difficulty recalling past events. Other typical measurement errors include rounding dollar amounts, misunderstanding questions, and altering responses due to stigma or prestige attached to certain behaviors or a desire to protect privacy. A large volume of research exists on measurement error and its effects on survey

data (see Lessler and Kalsbeek, 1992 and the references therein).

While it is true that, for administrative data, unit and item non-response are usually not a problem on core items, it is not clear that administrative data are always more accurate than survey data. For example, some individuals may intentionally misreport values on tax returns to reduce their tax liabilities–-it is estimated that underreporting may have resulted in underpayment of as much as $120 billion in income taxes for Tax Year 1998 (Brown and Mazur, 2003). Those same individuals may report the true value in response to a survey question since there is no benefit to misreporting in the survey.

The income questions in the SCF are structured to allow respondents to reference their tax forms when answering the income questions. Figure 1 shows the correspondence between the income questions in the SCF and the line number on IRS Form 1040. The SCF income questions were designed to cover most forms of income that a household reports on its tax form. The figure shows that there is much overlap between the two data sources, although there are some differences. Since the SCF is interested in all sources of household income and not just income subject to taxation, the questions on pensions, IRA/401(k) distributions, annuities, and Social Security payments refer to the total amounts. The SCF also asks about any income received from government transfer programs (such TANF, SSI, and food stamps). Households are not questioned about any adjustments to total income, but households are questioned about their Adjusted Gross Income (AGI). All income amounts reported in the SCF are for the year prior to the survey year.

Even with the close correspondence between the income questions in the SCF and IRS Form 1040, accurate classification and reporting of income amounts are still a potential problem in the SCF. To improve comparability, respondents are encouraged to reference documents, including tax forms, during the interview. Figure 2 shows that, for the 2004 SCF, almost 21 percent of all households referenced their tax forms. This represents a significant increase over earlier surveys. Higher income respondents were more likely to use tax returns during their interviews. Almost 25 percent of those reporting at least $50,000 in adjusted gross income referenced tax forms in answering the income module of the SCF in 2004, compared to fewer than 18 percent of those with lower incomes.

▶ Comparisons Between SCF and SOI Estimates

Figure 3 provides a comparison of SCF and SOI estimates for Tax Years 1988, 1991, 1994, 1997, 2000, and 2003 and highlights the difference in the unit of observation between the two data sources. In the SCF, the unit of observation is the household, which can sometimes contain more than one tax unit. The SCF asks the filing status of the core individual or couple in a household, thus allowing married or partnered households filing separately to be counted as two returns. The SCF consistently underestimates the number of returns in the tax filing population, no doubt in large part because the SCF does not ask about the filing status of other individuals within the household. These individuals include dependents who may also file a return and other members of the household who are not financially dependent on the household head or the core couple. Estimates of the income tax filing population produced using the SCF have improved over time and differed from the actual total by less than 23 percent for Tax Year 2003. Despite significant differences in filing population estimates, the SCF and SOI estimates of total income differ by no more than approximately 10 percent in each Tax Year shown, with the SCF estimate larger in each case.

SCF estimates of wages and salaries, unemployment and alimony payments, and other income are consistently larger than those produced by SOI. The difference between the estimates of alimony income is due to definitional differences; the SCF question on alimony income instructs the respondent to include child support payments. Since child support payments are nontaxable, such payments should not be included in the SOI estimate. The differences between the SCF and SOI estimates of "other income" are difficult to pinpoint, given the wide range of types of income

Figure 1

Form **1040** Department of the Treasury—Internal Revenue Service **U.S. Individual Income Tax Return** **2000** (99) IRS Use Only—Do not write or staple in this space.

For the year Jan. 1–Dec. 31, 2000, or other tax year beginning , 2000, ending , 20 OMB No. 1545-0074

Label (See instructions on page 19.) **Use the IRS label.** Otherwise, please print or type. LABEL HERE

Your first name and initial | Last name | **Your social security number**

If a joint return, spouse's first name and initial | Last name | **Spouse's social security number**

Home address (number and street). If you have a P.O. box, see page 19. | Apt. no.

City, town or post office, state, and ZIP code. If you have a foreign address, see page 19.

▲ **Important!** ▲ You **must** enter your SSN(s) above.

Presidential Election Campaign (See page 19.) **Note.** Checking "Yes" will not change your tax or reduce your refund. Do you, or your spouse if filing a joint return, want $3 to go to this fund? ► ☐ Yes ☐ No ☐ Yes ☐ No (You / Spouse)

Filing Status Check only one box.

1 ☐ Single
2 ☐ Married filing joint return (even if only one had income)
3 ☐ Married filing separate return. Enter spouse's social security no. above and full name here. ►
4 ☐ Head of household (with qualifying person). (See page 19.) If the qualifying person is a child but not your dependent, enter this child's name here. ►
5 ☐ Qualifying widow(er) with dependent child (year spouse died ►). (See page 19.)

Exemptions

6a ☐ **Yourself.** If your parent (or someone else) can claim you as a dependent on his or her tax return, **do not** check box 6a
b ☐ **Spouse**

No. of boxes checked on 6a and 6b

c **Dependents:**

(1) First name Last name	(2) Dependent's social security number	(3) Dependent's relationship to you	(4) ✓ if qualifying child for child tax credit (see page 20)
			☐
			☐
			☐
			☐
			☐
			☐

If more than six dependents, see page 20.

No. of your children on 6c who:
- **lived with you**
- **did not live with you due to divorce or separation (see page 20)**

Dependents on 6c not entered above

Add numbers entered on lines above ►

d Total number of exemptions claimed

Income

Attach Forms W-2 and W-2G here. Also attach Form(s) 1099-R if tax was withheld.

If you did not get a W-2, see page 21.

Enclose, but do not attach, any payment. Also, please use **Form 1040-V.**

7	Wages, salaries, tips, etc. Attach Form(s) W-2	7	X5702
8a	**Taxable** interest. Attach Schedule B if required	8a	X5708
b	**Tax-exempt** interest. **Do not** include on line 8a — 8b X5706		
9	Ordinary dividends. Attach Schedule B if required	9	X5710
10	Taxable refunds, credits, or offsets of state and local income taxes (see page 22)	10	
11	Alimony received	11	X5718
12	Business income or (loss). Attach Schedule C or C-EZ	12	X5704
13	Capital gain or (loss). Attach Schedule D if required. If not required, check here ► ☐	13	X5712
14	Other gains or (losses). Attach Form 4797	14	X5712
15a	Total IRA distributions — 15a X5724 — b Taxable amount (see page 23)	15b	
16a	Total pensions and annuities — 16a X5722 — b Taxable amount (see page 23)	16b	
17	Rental real estate, royalties, partnerships, S corporations, trusts, etc. Attach Schedule E	17	X5714
18	Farm income or (loss). Attach Schedule F	18	X5704
19	Unemployment compensation	19	X5716
20a	Social security benefits — 20a X5722 — b Taxable amount (see page 25)	20b	
21	Other income. List type and amount (see page 25)	21	X5724
22	Add the amounts in the far right column for lines 7 through 21. This is your **total income** ►	22	

Adjusted Gross Income

23	IRA deduction (see page 27)	23	
24	Student loan interest deduction (see page 27)	24	
25	Medical savings account deduction. Attach Form 8853	25	
26	Moving expenses. Attach Form 3903	26	
27	One-half of self-employment tax. Attach Schedule SE	27	
28	Self-employed health insurance deduction (see page 29)	28	
29	Self-employed SEP, SIMPLE, and qualified plans	29	
30	Penalty on early withdrawal of savings	30	
31a	Alimony paid b Recipient's SSN ►	31a	
32	Add lines 23 through 31a	32	
33	Subtract line 32 from line 22. This is your **adjusted gross income** ►	33	X5751.X7651.X7652

For Disclosure, Privacy Act, and Paperwork Reduction Act Notice, see page 56. Cat. No. 11320B Form **1040** (2000)

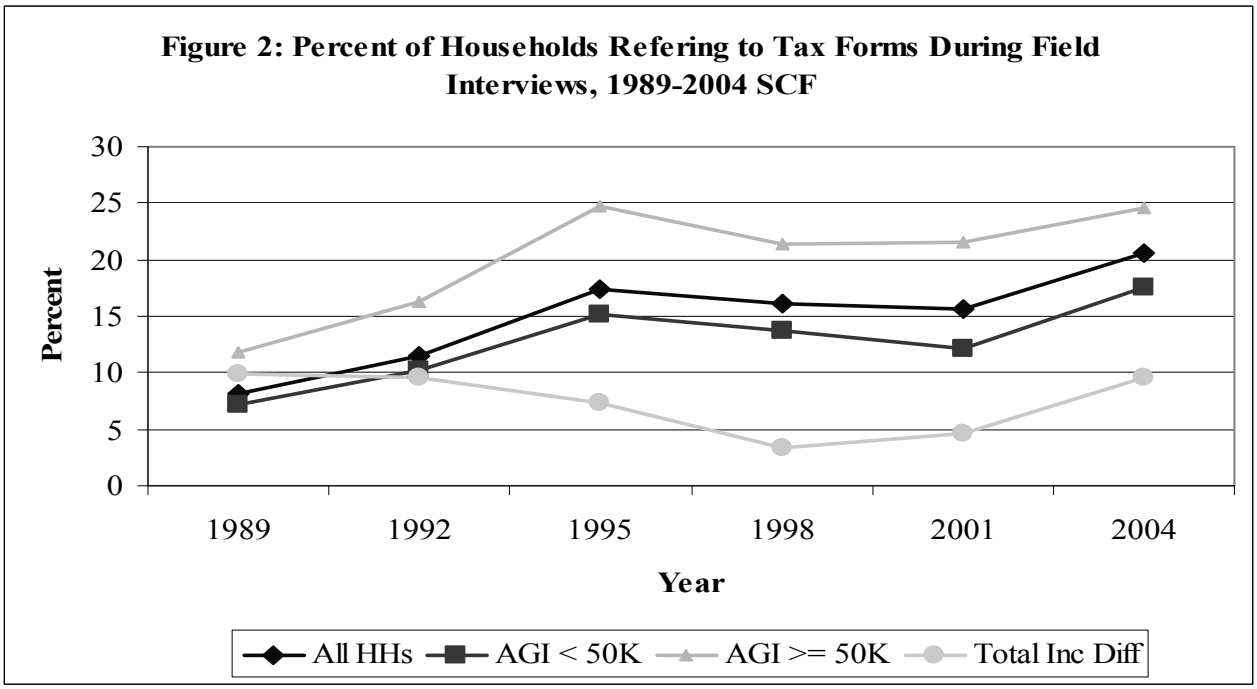


Figure 2: Percent of Households Refering to Tax Forms During Field Interviews, 1989-2004 SCF
30
25
20
15
10
5
0
Percent
1989
1992
1995
1998
2001
2004
Year
All HHs
AGI < 50K
AGI >= 50K
Total Inc Diff

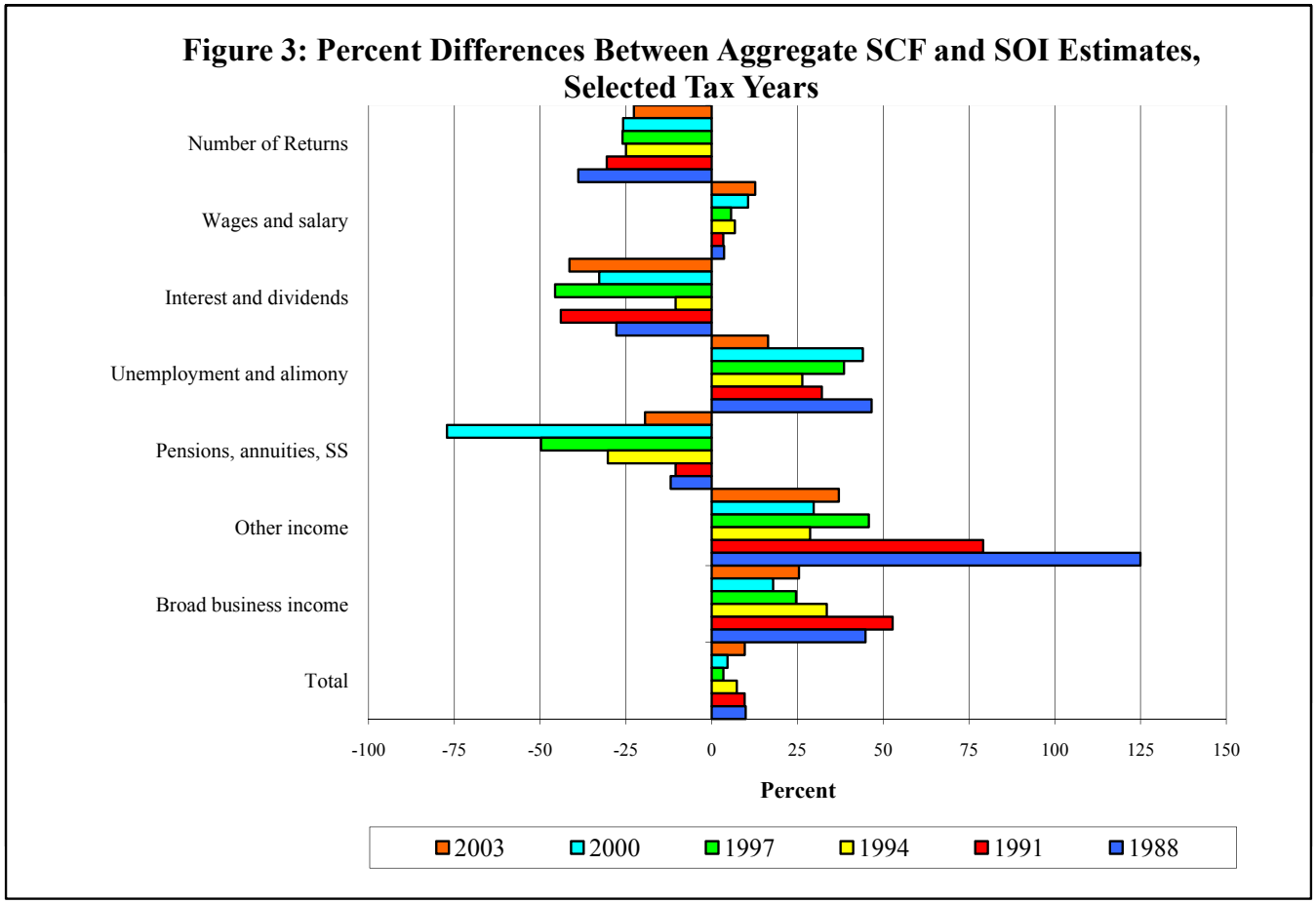


Figure 3: Percent Differences Between Aggregate SCF and SOI Estimates, Selected Tax Years
Number of Returns
Wages and salary
Interest and dividends
Unemployment and alimony
Pensions, annuities, SS
Other income
Broad business income
Total
-100
-75
-50
-25
0
25
50
75
100
125
150
Percent
2003
2000
1997
1994
1991
1988

potentially included in that category. Of the income categories shown, estimates for wages and salaries derived from the two data sources are relatively small, increasing from just 3.6 percent for Tax Year 1988 to 12.7 percent for Tax Year 2003.

The SCF estimates of broad business income are also consistently larger than the SOI estimates. Broad business income combines sole proprietorship and farm income, capital gains, and rent, royalties, and subchapter S corporation income. These components are combined because households in the SCF may misclassify capital gains or rent, royalties, and subchapter S corporation income as sole proprietor income. This could be partially due to the order of the income questions in the SCF, since the sole proprietor and farm business income questions are asked early in the income sequence, while the capital gains and rent, royalties, and subchapter S corporation income questions are asked later in the sequence. Constructing a broader measure of business income eliminates some of these classification issues and reduces the differences substantially, especially for the three most recent tax years shown.

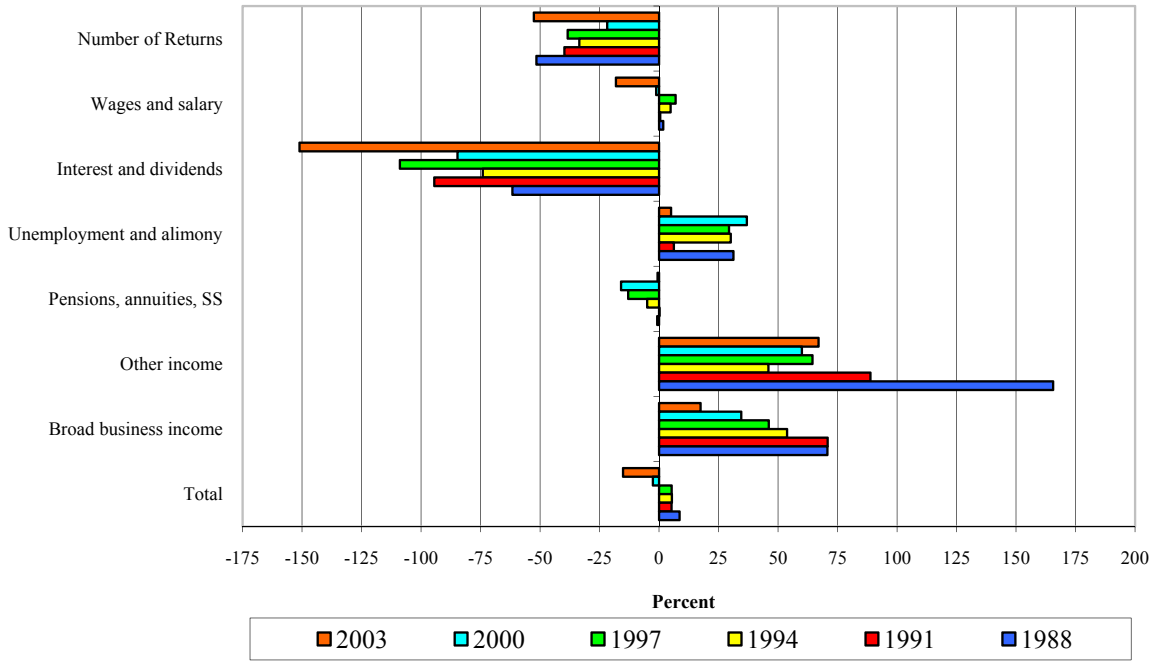
The SCF consistently underestimates the amount of interest (taxable and nontaxable) and dividends, as well as income from pensions, annuities, and Social Security. Differences between the SOI and SCF estimates of interest and dividends range from -10.5 percent to as much as -45.6 percent. One possible reason for these lower estimates is that households that receive only small amounts of taxable interest or dividend income may forget to report these amounts in the SCF questionnaire. Another possible reason is that households may not think they have "received" this income, particularly in the case of interest earned on bank accounts and money market funds. Even households with relatively large dividend and interest incomes may underestimate these values, due to the inherent variability of annual earnings, especially if they are not in a phase of life where such income is an important source of disposable income. The SCF understates the total of pension, annuity, and Social Security incomes by -10.5 percent to as much -77.1 percent, depending on the year. Using information reported in other sections of the SCF, it is possible to compute alternative estimates of pension, annuity and Social Security income. This computation reveals that (1) information in other sections of the survey corresponds closely with information provided in the income module of the SCF and (2) the SCF estimates of Social Security income are consistently similar to, but larger than the SOI estimates, while the SCF estimates of pension and annuity income are substantially less than the SOI estimates.

As noted previously, households in the SCF with at least $50,000 in AGI were much more likely to have referenced tax forms during the interview than lower income households. This suggests that households in the SCF with higher AGI should do a better job of reporting and classifying income. Data for respondents in these two AGI classes are shown in Figures 4 and 5.

For respondents in the less than $50,000 AGI group, estimates derived from SCF and SOI data for wages and salaries, unemployment and alimony, pensions, annuities and Social Security, and total income are all reasonably close. In contrast, estimates for interest and dividends are substantially different between the two sources. Again, this may be due to a large number of households neglecting to report relatively small amounts of interest income on the SCF. For example, in the 2004 SCF, only about a quarter of households with less than $50,000 in AGI that owned interest-bearing assets reported any interest income The median amount of interest-bearing assets for these households was only $1,200, suggesting that unreported interest would have been very small.

Figure 4 also shows that there is a sizeable difference in the estimate of broad business income for the less than $50,000 AGI group, although the difference has declined over time. Much of this difference is due to much larger estimates of rent, royalties, and subchapter S corporation income in the SCF and may be partly due to the treatment of losses in the survey. Although the SCF allows households to record negative amounts for certain income questions, households often report zero instead of the actual loss. Given the potentially favorable tax treatment of losses, actual losses are more likely to be reported to the IRS.

Figure 4: Percent Differences Between Aggregate SCF and SOI Estimates, AGI < $50,000, Selected Tax Years

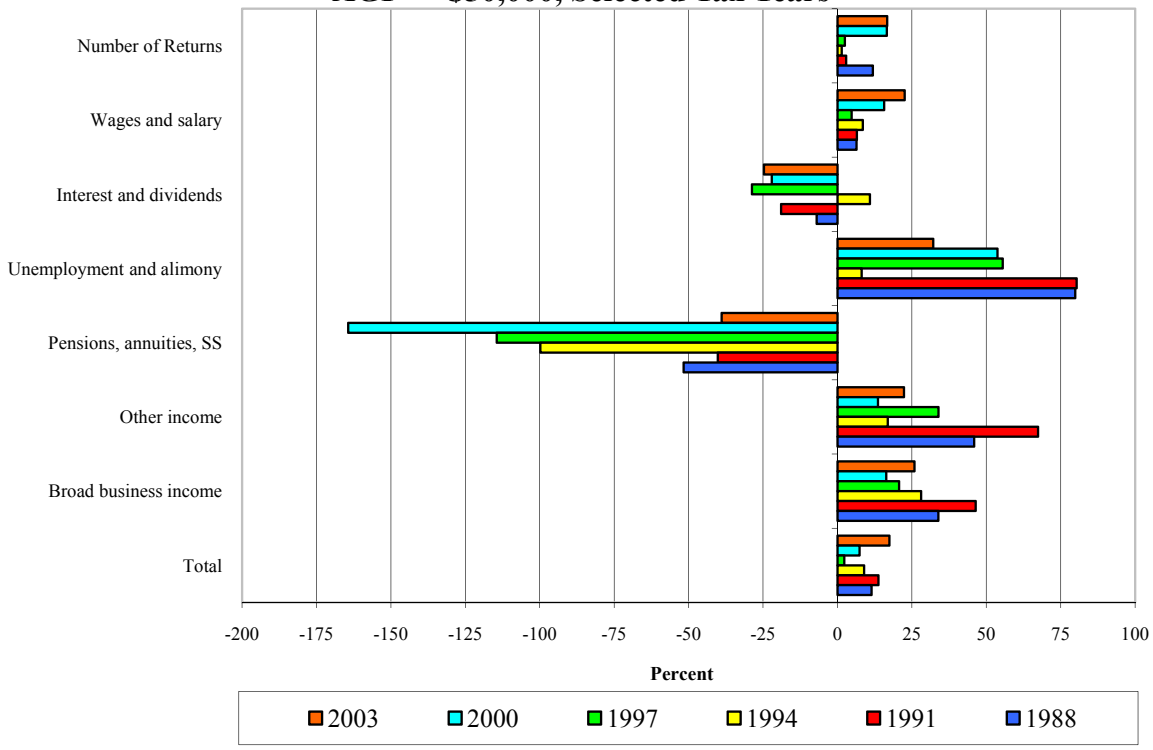


Turning to households with $50,000 or more in AGI, there is some evidence that the increased use of tax forms as references by members of this group improves the comparability between SCF and SOI estimates (see Figure 5). Estimates from these two data sources for the number of tax returns filed, as well as total income, wages and salaries, and interest and dividends differ by less than 30 percent. Also, the percentage differences in the broad business income estimates are smaller for households with $50,000 or more in AGI than for the lower income group. The SCF estimate for interest and dividends is less than the SOI amount in all but one year. Here again, only about 44 percent of households with $50,000 or more in AGI that owned interest-bearing assets reported any interest income, suggesting that even these respondents may neglect to report relatively small amounts. The median value of interest-bearing assets for these nonreporting households was about $6,000.

Most striking for the $50,000 or more AGI group are differences between the SCF and SOI estimates of pension, annuity, and Social Security income for all tax years shown. As with the estimates for all households, the summation of the alternative SCF estimates of pension, annuity, and Social Security incomes are very similar to the SCF estimate derived directly from the income questions. Also, the SCF estimates of Social Security income are typically fairly close to the SOI estimates. Thus, the bulk of the difference between the SCF and SOI estimates is due to pension and annuity income.

One possible reason for this discrepancy is the treatment of rollovers from one tax-deferred retirement account to another tax-deferred retirement account. For example, if a household transfers the balance of one IRA account to another IRA account, the transfer is not taxable, but the transfer amount should ap-

Figure 5: Percent Differences Between Aggregate SCF and SOI Estimates, AGI >= $50,000, Selected Tax Years



pear on line 16a of Form 1040 (see Figure 1). Often, households neglect to report these rollovers on their tax forms since there are no tax implications. However, the SOI estimate will include these rollovers, even if the household does not include them on its tax form.[1] Since households in the $50,000 or more AGI group are about twice as likely to have some sort of tax-deferred retirement account, these households are likely to have more rollovers. In published SOI estimates, a rough measure of the amount of rollovers is the difference between total and taxable pension and annuity income. For filers with $50,000 or more in AGI, about 60 percent of pension and annuity income is taxable, compared to about 80 percent for filers with less than $50,000 in AGI. If households in the SCF are not reporting their rollovers in the pension income question, this could explain most of the difference between these SCF and SOI estimates.

▶ Conclusion

In summary, the Survey of Consumer Finances contains an income module that is designed to capture information comparable to that reported on IRS Form 1040 for the tax year prior to the year in which the survey is conducted. Estimates produced from these data should closely match those produced by the Statistics of Income Division of the IRS. Indeed, taking into account differences in the reporting unit between the two data sources and sample variance, estimates of total income for each AGI group and tax year examined are remarkably close. Disaggregating total income into more detailed categories, however, reveals important differences.

Differences between estimates produced using SOI and SCF data are due in part to the idiosyncrasies of

[1] A rollover transaction generates a Form 1099-R that SOI matches to Form 1040. If a filer neglects to report the rollover on his or her tax form, the value from Form 1099-R is added to the filer's Form 1040.

the Tax Code. Some income items, including a portion of Social Security income, certain components of payments from a divorced spouse, and interest earned on some investments are exempt from taxation and are therefore excluded from SOI estimates. However, for the purpose of studying a household's economic condition, these items are necessarily included in estimates produced by the SCF. Other items, such as the allocation of depreciation on rental properties or the carry-forward (or even backward) of business losses, are an important part of good tax planning, but are not easily captured within the structure of a household survey. The relative consistency of differences between SCF and SOI estimates over time, as shown in Figures 3, 4, and 5, suggests that they may be attributed primarily to these types of inherent disparities.

Figures 3, 4, and 5 do show significant improvements in the comparability of SCF and SOI estimates over time, which suggests that households sometimes classify income items differently in their survey responses than on tax returns. Some of this improvement is due to changes in the structure of the SCF over time. Cognitive testing and experience have led to some changes in both question design and the order in which questions are asked. An important change was the transition from a paper survey instrument to computer-assisted personal interviewing (CAPI) after the 1992 SCF. The CAPI instrument helps improve the quality of data collected by performing real-time tests intended to ensure that all dollar values are entered as reported by the respondent. CAPI also facilitates online tools, such as definitions and code lists, which improve the quality of data collected in the field. The research presented here also suggests that encouraging households to reference their tax forms is critical for improving the comparability of data between the SCF and SOI. Where classification differences persist, it is it is often possible to use information from other sections of the survey to make adjustments in order to better align the SCF and SOI income definitions. Ultimately, these classification differences may highlight the challenges some taxpayers face in classifying their incomes according to IRS reporting requirements. It is clear that, for some taxpayers, IRS distinctions between certain forms of income are blurred.

The goal of the research presented in this paper has not been to declare either the SCF or SOI data superior. Instead, we have attempted to document important similarities and differences between the two data sources. The detail and scope of the data collected in the SCF allow for a broader range of research than in the SOI tax data. The large sample size and administrative nature of SOI tax data make them appealing for certain types of research, such as studying some aspects of tax policy. The key, then, is that both data sources have strengths and weaknesses that need to be understood and carefully considered before attempting to use them to answer any set of research questions.

▶ Acknowledgements

The authors gratefully acknowledge the insightful comments, editing suggestions and encouragement provided by Arthur Kennickell, Martha Gangi, and James Dalton.

▶ References

Bucks, Brian , Kennickell, Arthur B., and Moore, Kevin B. (2006), "Recent Changes in U.S. Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances," *Federal Reserve Bulletin*, vol. 92, A1-A38.

Brown, Robert E and Mark J. Mazur, (June 2003) "IRS' Comprehensive Approach to Compliance Measurement," 2003 National Tax Association Spring Symposium http://www.irs.gov/pub/irs-soi/mazur.pdf.

Frankel, Martin and Arthur B. Kennickell (1995) "Toward the Development of an Optimal Stratification Paradigm for the Survey of Consumer Finances," *Proceedings of the Section on Survey Research Methods,* American Statistical Association.

Kennickell, Arthur B., (2001), "Modeling Wealth with Multiple Observations of Income: Redesign of the Sample for the 2001 Survey of Consumer Finances," working paper, Board of Governors of the Federal Reserve System.

Kennickell, Arthur B., (1999), "Revisions to the SCF Weighting Methodology: Accounting for Race/ Ethnicity and Homeownership," working paper, Board of Governors of the Federal Reserve System.

Kennickell, Arthur B. (1998a) "List Sample Design for the 1998 Survey of Consumer Finances," working paper, Board of Governors of the Federal Reserve System.

Kennickell, Arthur B., (1998b), "Multiple Imputation in the Survey of Consumer Finances," *Proceedings of the Section on Business and Economic Statistics*, American Statistical Association.

Kennickell, Arthur B. (1991), "Imputation of the 1989 Survey of Consumer Finances: Stochastic Relaxation and Multiple Imputation," *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Kennickell, Arthur B. and Douglas McManus (1993) "Sampling for Household Financial Characteristics Using Frame Information on Past Income," working paper, Board of Governors of the Federal Reserve System.

Kennickell, Arthur B. and Woodburn, R. Louise, (1999), "Consistent Weight Design for the 1989, 1992, and 1995 SCFs, and the Distribution of Wealth," *Review of Income and Wealth*, Series 45, 2, 193-215.

Lessler, Judith T. and William D. Kalsbeek (1992) *Nonsampling Error in Surveys*, New York: John Wiley & Sons.

Petska, Tom and Mike Strudler, (1999) "The Distribution of Individual Income Taxes: A New Look at an Old Issue," *Turning Administrative Systems Into Information Systems, Internal Revenue Service*, pp. 7-22.

Sailer, Peter and Michael Weber, (1999) "The IRS Population Count: An Update," Proceedings, Section on Survey Research Methods, American Statistical Association.

Statistics of Income—2003 Individual Income Tax Returns, Internal Revenue Service, Washington, DC, 2005.

Tourangeau, Roger, Johnson, Robert A., Qian, Jiahe, Shin, Hee-Choon, and Frankel, Martin R., (1993), "Selection of NORC's 1990 National Sample," working paper, National Opinion Research Center at the University of Chicago.