



Working Paper No. 618

**Quality of Match for Statistical Matches Used in the 1992 and 2007
LIMEW Estimates for the United States**

by

Thomas Masterson
Levy Economics Institute of Bard College*

September 2010

The Levy Economics Institute Working Paper Collection presents research in progress by Levy Institute scholars and conference participants. The purpose of the series is to disseminate ideas to and elicit comments from academics and professionals.

Levy Economics Institute of Bard College, founded in 1986, is a nonprofit, nonpartisan, independently funded research organization devoted to public service. Through scholarship and economic research it generates viable, effective public policy responses to important economic problems that profoundly affect the quality of life in the United States and abroad.

Levy Economics Institute
P.O. Box 5000
Annandale-on-Hudson, NY 12504-5000
<http://www.levyinstitute.org>

Copyright © Levy Economics Institute 2010 All rights reserved

ABSTRACT

The quality of match of four statistical matches used in the LIMEW estimates for the United States for 1992 and 2007 is described. The first match combines the 1992 Survey of Consumer Finances (SCF) with the 1993 March Supplement to the Current Population Survey, or Annual Demographic Supplement (ADS). The second match combines the 1985 American Use of Time Project survey (AUTP) with the 1993 ADS. The third match combines the 2007 SCF with the 2008 March Supplement to the CPS, now called the Annual Social and Economics Supplement (ASEC). The fourth match combines the 2007 American Time Use Survey with the 2008 ASEC. In each case, the alignment of the two datasets is examined, after which various aspects of the match quality are described. Also in each case, the matches are of high quality, given the nature of the source datasets.

Keywords: Statistical Matching; Wealth Distribution; Time Use; Household Production; United States; LIMEW

JEL Classifications: C14, C40, D31

INTRODUCTION

This paper describes the construction of synthetic datasets created for use in estimation of the LIMEW for the United States for the years 1992 and 2007. This work was carried out for a project supported by the Sloan Foundation to produce international comparisons of economic well-being. Construction of LIMEW estimates requires a variety of information for households. In addition to basic demographics, the estimation process requires information about income, transfers, taxes, time use, and wealth. No single data set has all the required data for the United States. Thus, in order to produce LIMEW estimates, a synthetic data file is created from various source data sets with statistical matching.¹ For the United States, we use the Bureau of Labor Statistics' March Current Population Survey (CPS)² as the base data set, since it contains good information on demographics, income, transfers, and taxes for a regionally representative sample of U.S. households. Wealth data comes from the Survey of Consumer Finances (SCF) carried out by the Federal Reserve. Time use data comes from the American Time Use Survey (ATUS)³ for later years and the Americans' Use of Time Project (AUTP)⁴ for earlier years.

This paper is organized as follows. Each section of the paper details four statistical matches in turn: wealth and time use matches for 1992 and 2007 for the United States. The source datasets are described and their demographic characteristics are compared. Then the quality of the match is reviewed for each.

1992 WEALTH MATCH

Data and Alignment

The matching unit for the wealth match (and the unit of analysis for the LIMEW) is the household. The source data sets for the wealth match for the 1992 LIMEW estimates are the 1993 ADS and the 1992 SCF. The 1993 ADS is used since it has income data for 1992. The 1993 ADS file has records for 155,197 individuals in 58,970 households. After dropping all those living in group quarters (dormitories, nursing homes, etc.), we are left with 58,837

¹ For details of the LIMEW and its construction, see Wolff and Zacharias (2003). See Kum and Masterson (2008) for details of the statistical matching procedure that we use.

² Before 2003, this survey was called the Annual Demographic Supplement (ADS). Subsequently it is called the Annual Social and Economic Supplement (ASEC).

³ Collected by the United States Bureau of Labor Statistics, Department of Labor.

⁴ Collected by Survey Research Center, University of Michigan, Ann Arbor.

households. These records represent 96,391,171 U.S. households after weighting. The 1992 SCF contains multiply imputed information for 3,906 respondents (five implicates for each respondent for a total of 19,530 records). This translates to 95,917,634 households when weighted. In order to perform a successful match, the candidate data sets must be well aligned in the strata variables used in the match procedure.⁵ For the wealth match, strata variables are homeownership, age, race, family type, and household income. Table 1 compares the distribution of households by these five variables in the two data sets. Since both surveys are regionally representative samples carried out a year apart, we can expect them to be well aligned. However, the SCF is drawn from a more complicated sampling frame. Since the SCF is a wealth survey and wealth is highly concentrated, the top of the wealth distribution is oversampled. We expect some misalignment as a result of this important (and necessary for our purposes) difference in sampling frame between the two surveys.

We see that the distribution of family types is slightly different in the two surveys, with married couples more common in the SCF than in the ADS. Also whites are underrepresented in the SCF relative to the ADS. The largest difference is by income category, with those at the lower and higher ends of the household income distribution making up a significantly larger proportion of the SCF sample than of the ADS. These misalignments can make matching a challenge, because it ensures that, for example some households with less than \$20K annual income in the SCF will be matched with households in the middle income categories in the ADS, thereby slightly depressing the wealth profile of the lower middle of the income distribution (corresponding effects can be expected at the upper middle end of the income distribution). Interestingly, experiments with alternative specifications of the income variable (such as using quantiles as category cutoffs rather than absolute amounts) actually decrease the match quality.

Table 2 shows a more detailed breakdown of the alignment of the two surveys, using four of the five strata variables (and replacing more detailed age categories with elder/nonelder indicator variable). Here we can see that the higher prevalence of married couples in the SCF is concentrated among younger renters, both white and nonwhite. The largest single difference is among households with young, white, renter female heads, which are much more prevalent in the

⁵ Statistical matching is done first within subsets of the two data sets defined by key variables, which are referred to as strata variables.

ADS than in the SCF. Based on these observations of the alignment, we can expect that the worst misallocation of wealth variables will be by household income and family type.

Match QC

Turning to the results of the match, we first look to the distribution of matched records by matching round in table 3. Earlier rounds occur in the most detailed cells (round 1 occurs within cells that incorporate all five strata variables). The bulk of the matches occur in the earliest rounds, in fact 91% in the first two rounds alone. This fact means that most of the wealth records will be assigned to records that are similar in age, race, family type, homeownership, and income to their donor records. This bodes well for the quality of the match. Indeed, we can see in figure 1 that the overall distribution of net worth is well carried over into the match file. In fact, it is impossible to see differences at all at this level of detail. Table 4 provides a closer comparison of the distribution of net worth in the SCF and the matched file. The percentile ratios are all quite close, with the exception of p75/p25 and p50/p25. It appears that the bottom tale of the wealth distribution in the matched file is somewhat thinner than in the SCF. The Gini coefficient is quite close, 0.816 in the matched file, compared to 0.810 in the SCF. Table 5 breaks down the mean and median of the five asset and two debt classes that make up net worth in the wealth match.⁶ We can see that for all eight variables the difference in the matched and the source file's mean is small, less than 3% in all cases. For median values, most asset and debt classes are zero. There are larger percentage differences for asset 3 and debt 2 than we saw for average values, but these are small in absolute terms (\$210 and \$80, respectively). The most important asset, asset 1, is precisely matched, and the median net worth is off by 2.5%, but again, this represents a small absolute difference of just over \$1,000.

Examination of the quality of the match within population subgroups shows generally good results. Figure 2 displays ratios of mean net worth between the matched file and the SCF for the five strata variables, as well as alternative specifications of the age and race categories. With some exceptions, the ratios of mean net worth within subcategories of the five strata variables are all within 10% of unity. The middle income group (\$50,000 to \$75,000 in

⁶ The five asset classes are primary residence, other real estate net of debt and business equity, liquid assets, financial and other assets, and retirement assets. The two debt classes are mortgages and equity loans and lines of credit on the primary residence and other debt (exclusive of mortgages on other property, which are subtracted from the value of that property in asset 2).

household income) has 25% higher net worth in the matched file than in the SCF. Table 6 has the actual numbers, and we can see that this represents a substantial difference of \$43,000. However, notice that the median net worth for this group in the matched file is 9% smaller than that of the SCF. The first group in the homeowner panel of figure 2 is renters. We can see that they have 23% smaller net worth in the matched file than in the SCF. We see in table 6 that this translates to \$7,800 less average net worth for renters. The difference in medians is smaller, at only 16%, which translates to a \$19 difference in median net worth. Finally, those households in the youngest age group (those under 35 years old) have 12% higher mean net worth in the matched file than in the SCF. Consulting table 6, we see that this means \$4,400 greater net worth, while their median net worth is 11% lower than in the SCF (a \$290 difference). More detailed race categories are shown in figure 2 than were used in the matching.⁷ We can see that although the nonwhite average net worth in the matched file closely replicates the SCF, differences within the nonwhite group are not carried over as accurately, with blacks and Hispanics appearing wealthier and others less wealthy than in the SCF. For judging the accuracy of the match in preserving the distribution of wealth by subgroups, table 6 displays the ratios of mean and median values for the strata variables' categories. The white/nonwhite ratios of mean and median values are well carried over, while again, the ratios for the groups within nonwhite are not as well. The median net worth of black households is, however, accurately carried over. The rest of the ratios' values in the SCF are reasonably well represented in the match file. The extent to which the match file reproduces the distribution of net worth within matching cells is demonstrated in figure 3.⁸ We can see that, although the tails are attenuated somewhat, the distribution is well preserved in the matching process, even at this level of detail.

Overall, the quality of the match is good. It has its limitations, especially in terms of the race categories (there is, of course, much variation in the “nonwhite” group). But the overall distribution is transferred with remarkable accuracy, and the distributions within even small subgroups, such as young, nonwhite homeowners, is transferred with good precision.

⁷ They are in order, non-Hispanic whites, non-Hispanic blacks, Hispanics, and non-Hispanic others. The latter groups were too small in number to make a statistical match using this categorization feasible.

⁸ Family type is excluded for the sake of clarity of the plot.

1992 TIME USE MATCH

Data and Alignment

The source data sets for the time use match for the 1992 LIMEW estimates are the 1993 ADS and the 1985 AUTP. We use individual records from the 1993 ADS file, excluding those living in group quarters or in the Armed Forces. Since the AUTP covers individuals 18 years old and above, we discard younger individuals from the ADS file. This leaves 112,092 records, which represents 186,837,050 individuals when weighted. The AUTP file includes time use data for 4,024 individuals. The weights provided are probability weights, not frequency weights, so a weighted count of individuals is not possible with the AUTP. For the time use match, the strata variables are sex, parental status, employment status, and marital status. While for the wealth match the matching unit is the household, for the time use match we use individuals. Table 7 compares the distribution of individuals by these variables and household income in the two data sets. Since the two surveys were carried out eight years apart, we can expect them to be somewhat misaligned. We see that the distribution of individuals by sex is different in the two surveys, with females more common in the AUTP than in the ADS. Also the not employed are overrepresented in the AUTP relative to the ADS, reflecting different economic conditions in the two years, as well as the secular trend towards higher labor force participation rates among women. The portion of married individuals is also much higher in the AUTP, perhaps also reflecting trends in household formation and breakup. The closest match is by parental status. The differences by income category are almost as large as by employment category, with those at the lower and higher ends of the household income distribution making up a significantly smaller proportion of the AUTP sample than of the ADS.

Match QC

Turning to the results of the match, we first look to the distribution of matched records by matching round in table 8. The bulk of the matches, 93%, occur in the first round, ensuring as high-quality a match as possible. Table 9 provides a closer comparison of the distribution of weekly hours of household production in the AUTP and the matched file. The percentile ratios are all equivalent. The Gini coefficient is extremely close, 0.4809 in the matched file, compared to 0.4810 in the AUTP. Table 10 breaks down the mean and median of the three classes that

make up total household production in the time use match.⁹ We can see that for all four variables the difference in the matched and the source file's mean and median is very small, less than or equal to 0.1% in all cases.

Examination of the quality of the match within population subgroups shows generally good results. Figure 4 displays ratios of mean weekly hours of household production between the matched file and the SCF for the four strata variables, as well as for household income categories. With some exceptions, the ratios of mean weekly hours of household production within subcategories of the four strata variables are all within 2% of unity. Employed and unmarried individuals have both have 6% higher weekly hours in the matched file than in the AOTP. The largest difference by income group (\$50,000 to \$75,000 in household income) is 4% lower weekly hours of household production in the matched file than in the AOTP for households between \$25,000 and \$35,000 in income. Table 11 has the actual numbers, and we can see that this represents a difference of less than an hour a week. However, notice that the median weekly hours of household production for this group in the matched file is 9% smaller than that of the AOTP, for a difference of 1.75 hours. The larger percentage differences in average weekly hours of household production for unmarried and employed individuals amount to slightly more than one hour per week. The difference in medians for these two groups is smaller, at only 5%, which translates to less than one-hour difference in median weekly hours of household production. For judging the accuracy of the match in preserving the distribution of household production by subgroups, table 11 displays the ratios of mean and median values for the strata variables' and household income categories. The larger deviations in ratios are for the categories already mentioned, but they are small. The rest of the ratios' values in the AOTP are very well represented in the match file. The extent to which the match file reproduces the distribution of weekly hours of household production within matching cells is demonstrated in figure 5.¹⁰ We can see very little difference between the matched file and the AOTP. Thus the distribution of household production is well preserved in the matching process, even at this level of detail.

Overall, the quality of the match is very good. It has its limitations, especially in terms of the marital and employment status categories, but the overall distribution is transferred with

⁹ The three classes are care (child care, education, etc.), procurement (shopping, etc.), and core (cooking, cleaning, laundry, etc.).

¹⁰ Marital status is excluded for the sake of clarity of the plot.

remarkable accuracy, and the distributions within even small subgroups, such as female parent employees, is transferred with good precision.

2007 WEALTH MATCH

Data and Alignment

The source data sets for the wealth match for the 2007 LIMEW estimates are the 2008 ASEC and the 2007 SCF. The 2008 ASEC is used since it has income data for 2007. The 2008 ASEC file contains records for 206,404 individuals in 75,872 households. After dropping those living in group quarters we have records for individuals in 75,813 households. When weighted this gives us data representing 116,783,684 U.S. households. The 2007 SCF has been multiply imputed to replace missing values. There are five replicates for each of the 4,422 original records, making 22,110 household records in the full file. We use all the records. When the weights are appropriately adjusted, the records in the SCF represent 116,122,131 households. As mentioned above, for the wealth match, the strata variables are homeownership, age, race, family type, and household income. Table 12 shows the distribution of households by these five variables in the two data sets. Since both surveys are regionally representative samples carried out a year apart, we can expect them to be well aligned. However, the 2007 SCF is drawn using the same complicated sampling frame as the 1992 SCF. Thus we again expect some misalignment as a result of this important (and necessary for our purposes) difference in sampling frames between the two surveys.

We see that the distribution of family types is slightly different in the two surveys, with married couples once again being more common (3.91%) in the SCF than in the ASEC. Also whites are overrepresented (3.05%) in the SCF relative to the ASEC. This difference is due to underrepresentation of Hispanics and others in the SCF. The differences by income category are smaller than in 1992, with those at the lower end of the household income distribution making up a significantly larger proportion of the SCF sample than of the ASEC, while those at the higher end of the household income scale are a smaller share of the SCF. These misalignments can make matching a challenge, because it ensures that, for example, some households with less than \$20,000 annual income in the SCF will be matched with households in the middle income categories in the ASEC, thereby slightly depressing the wealth profile of the lower middle of the

income distribution (corresponding effects can be expected at the upper middle end of the income distribution).

Table 13 shows a more detailed breakdown of the alignment of the two surveys, using four of the five strata variables (and replacing more detailed age categories with the elder/nonelder indicator variable). Here we can see that the higher prevalence of married couples in the SCF is concentrated among younger renters, both white and nonwhite. Based on these observations of the alignment, we can expect that the worst misallocation of wealth variables will be by race and family type.

Match QC

The match itself, although requiring sixteen rounds of matching to complete, was 89% done after the first round (see table 14). This is a good sign, as so many records were matched within one of 180 very detailed matching cells (formed by combining all of the strata variables). This indicates that the quality of the match should be quite good. Table 15 and figure 6 begin to show that this is in fact the case. The distribution of net worth has been well preserved. There is no discernible difference in the density of log net worth between the SCF and the matched file, and percentile ratios are quite closely carried over. The one exception is the p75/p25 ratio, which is considerably larger in the matched file. This is mostly an example of the denominator problem, however (p25 is \$2,500 in the matched file, compared to \$3,500 in the SCF). The components of net worth are well carried over into the matched file (see table 16). The largest difference is for asset 4, financial assets, which is expected, given the oversampling and consequent difficulty of matching high wealth households, which are more likely to have financial assets.

Figure 7 shows the ratio of mean net worth by strata variable categories. As we can see, net worth has been reproduced in the match file, with generally small variations between the matched file and the SCF. Looking at race, we can see that white wealth has been transferred fairly accurately, more so than nonwhite. The further breakdown shows that the largest problem is among those households in the “other” category (not surprisingly, since this is the smallest category numerically speaking). The comparison by family type looks good for married couples but less so for female-headed and, especially, male-headed households (again the numerically smallest category is the worst). The distribution of wealth by age seems to have been well preserved by the matching, with only small variations between the SCF and the match file. The

transfer within household income categories looks good except that the higher income categories look less wealthy in the match file than in the SCF. This is due to the misalignment between the two files.

Figure 8 shows the distribution of log net worth within collapsed matching cells (again by race, homeownership, and age). The distributions have been carried over very well. The obvious difference is that the upper tails of the distributions haven't been carried over completely. The bulk of the distribution is quite well carried over, however.

Finally, the comparison of mean and median net worth by strata variable categories is found in table 17. The ratios of mean net worth by category are very similar between the SCF and the matched file. The most notable differences is the ratio between "other" and white mean household net worth. While differing considerably in the matched file, the relative position of the "others" vis-à-vis the other race categories is preserved. The ratios of median values are somewhat more concerning, with the same pattern appearing in the race category results and with male-headed households looking slightly worse off than female-headed households in the matched file.

Overall, however, the match has provided us with a fair representation of the original distribution of wealth in the SCF. The differences we observe are small enough not to affect the outcome of the final analysis of the LIMEW greatly.

2007 TIME USE MATCH

Data and Alignment

The source data sets for the time use match for the 2007 LIMEW estimates are the 2008 ASEC and the 2007 ATUS. We use individual records from the 2008 ASEC file, excluding those living in group quarters or in the Armed Forces. Since the ATUS covers individuals 15 years old and above, we discard younger individuals from the ASEC file. This leaves 156,498 records, which represents 237,993,292 individuals when weighted. The ATUS file includes time use data for 12,166 individuals, corresponding to 234,238,921 individuals when weighted. For the time use match, the strata variables are sex, parental status, employment status, and marital status. While for the wealth match the matching unit is the household, for the time use match we use individuals. Table 18 compares the distribution of individuals by these variables and household

income in the two data sets. Since the two surveys were carried out in the same year, we can expect them to be well aligned. We see that the distribution of individuals by sex is only slightly different in the two surveys. Parents are much more prevalent in the ASEC than in the ATUS (4.52%). The not employed are slightly underrepresented in the ATUS relative to the ASEC. The portion of married individuals is lower in the ATUS, by 2.23%. The difference in spouse's labor force status is relatively small (less than 2%). The difference in parental status, reflecting different sampling frames, is the greatest cause for concern in terms of the potential match quality, but the alignment overall is good.

Match QC

Table 19 shows the distribution of matched records by matching round. The fact that 93.9% of records were matched in the first round of matching is a promising sign for the quality of the match. The overall distribution of weekly hours of household production looks perfect, based on the percentile ratios and Gini coefficient displayed in table 20. All but the p75/p25 and p75/p50 are exactly right, while these two ratios are off by very little. The Gini coefficient is off by less than 0.1 Gini points. The mean and median weekly hours of household production and its three components are exactly carried over to the matched file from the ATUS (see table 21). Figure 9 displays ratios of mean weekly hours of household production by the strata variables, as well as education and race. In terms of the strata variables, the match looks good for each one.

Unmarried individuals have 6% greater average weekly hours of household production in the match file, nonparents have 5% greater and parents have 3% fewer household production hours. Of those individuals who are married, those whose spouses are unemployed have 4% greater and those with employed spouses have 3% fewer average weekly hours of household production.

Table 22 gives us a closer look at the numbers behind figure 9, showing the mean and median weekly hours of household production by the strata variables, plus education and race. While the average weekly hours of household production by sex and employment status in the matched file are exactly the same as in the ATUS, the differences in the other strata variables are all one hour, which works out to between 3 and 6%. The ratios by strata variables are correspondingly well reproduced in the matched file. The differences for nonstrata variables are unsurprisingly larger, both in terms of percentage and hours. For example, those without a high school degree have five hours more in the matched file than in the ATUS, amounting to 29%. As

we can see, the ratios of matched to ATUS medians are unity or close to it for all the strata variables. The difference between the matched file and the ATUS for parents, married people, the unemployed, and those with spouses working is one hour per week. The differences for nonstrata variables are again larger, with those with a high school degree registering five hours more per week at the median in the matched file, while those in the black and “other” race categories have four more hours of household production.

REFERENCES

- Kum, Hyunsub, and Thomas Masterson. 2008. "Statistical Matching Using Propensity Scores: Theory and Application to the Levy Institute Measure of Economic Well-Being." Working Paper 535. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College. Available at: http://www.levyinstitute.org/pubs/wp_535.pdf.
- Wolff, Edward N., and Ajit Zacharias. 2003. "The Levy Institute Measure of Economic Well-Being." Working Paper 372. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College. Available at: <http://www.levyinstitute.org/pubs/wp372.pdf>.

TABLES

Table 1. Alignment of Strata Variables for 1992 Wealth Match

	1993 ADS	1992 SCF	Difference
<i># Households</i>	96,391,171	95,917,634	-0.49%
Homeownership			
<i>renter</i>	35.45%	36.07%	0.62%
<i>owner</i>	64.55%	63.93%	-0.62%
Family Type			
<i>MC</i>	55.16%	57.58%	2.42%
<i>FH</i>	28.99%	27.51%	-1.48%
<i>MH</i>	15.85%	14.91%	-0.94%
Age Category			
<i><35</i>	25.69%	25.76%	0.07%
<i>35-44</i>	22.53%	22.76%	0.23%
<i>45-54</i>	17.20%	16.25%	-0.95%
<i>55-64</i>	12.90%	13.20%	0.30%
<i>>=65</i>	21.68%	22.03%	0.35%
Elder			
<i>nonelder</i>	78.32%	77.97%	-0.35%
<i>elder</i>	21.68%	22.03%	0.35%
Race			
<i>white</i>	77.27%	75.29%	-1.98%
<i>black</i>	11.43%	12.67%	1.24%
<i>hispanic</i>	8.34%	7.49%	-0.85%
<i>other</i>	2.96%	4.55%	1.59%
Race Category			
<i>nonwhite</i>	22.73%	24.71%	1.98%
<i>white</i>	77.27%	75.29%	-1.98%
HH Income			
<i><20</i>	32.82%	38.46%	5.64%
<i>20-50</i>	39.78%	35.69%	-4.09%
<i>50-75</i>	16.00%	13.70%	-2.30%
<i>75-100</i>	6.31%	5.47%	-0.84%
<i>>100</i>	5.09%	6.68%	1.59%

Table 2. Matching Cells for 1992 Wealth Match

			Renter			Owner		
			1993 ADS	1992 SCF	Difference	1993 ADS	1992 SCF	Difference
Nonelder	Nonwhite	Married Couple	3,606,835	4,660,214	1,053,379	5,002,567	5,323,171	320,604
		Female Head	4,942,502	4,932,561	(9,941)	1,883,374	2,198,137	314,763
		Male Head	2,349,762	2,466,973	117,211	879,492	786,777	(92,715)
	White	Married Couple	6,771,294	8,034,002	1,262,708	28,483,330	27,997,158	(486,172)
		Female Head	6,338,847	4,792,493	(1,546,354)	5,658,100	5,184,523	(473,577)
		Male Head	5,427,083	5,155,860	(271,223)	4,151,830	3,251,211	(900,619)
Elder	Nonwhite	Married Couple	210,249	265,137	54,888	936,484	782,649	(153,835)
		Female Head	744,509	624,201	(120,308)	832,497	1,178,693	346,196
		Male Head	261,998	224,605	(37,393)	254,761	259,396	4,635
	White	Married Couple	654,611	658,865	4,254	7,505,608	7,503,662	(1,946)
		Female Head	2,265,697	2,244,509	(21,188)	5,274,944	5,234,636	(40,308)
		Male Head	597,718	539,089	(58,629)	1,357,079	1,619,112	262,033

Table 3. Distribution of Matched Records by Matching Round, 1992 Wealth Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	51,999	88.4	88.4
2	1,604	2.7	91.1
3	287	0.5	91.6
4	397	0.7	92.3
5	2,355	4.0	96.3
6	581	1.0	97.3
7	244	0.4	97.7
8	63	0.1	97.8
9	165	0.3	98.1
10	463	0.8	98.9
11	45	0.1	98.9
12	634	1.1	100
Total	58,837	100	

Table 4. Distribution of Net Worth in 1992 Matched File

	p90/p10	p90/p50	p50/p10	p75/p25	p75/p50	p50/p25	gini
SCF 1992	-248.372	8.338	-29.788	75.611	3.335	22.672	0.810
Match	-255.000	8.331	-30.608	104.180	3.351	31.086	0.816

Table 5. Comparison of Mean and Median Wealth Variables in 1992 Matched File to 1992 SCF

	Ave. Asset1	Ave. Asset2	Ave. Asset3	Ave. Asset4	Ave. Asset5	Ave. Debt1	Ave. Debt2	Ave. Networth	Med. Asset1	Med. Asset2	Med. Asset3	Med. Asset4	Med. Asset5	Med. Debt1	Med. Debt2	Med. Networth
SCF 1992	70,209	67,369	21,747	29,790	16,835	22,752	5,616	177,582	40,000	-	3,410	-	-	-	1,080	40,810
Match	69,604	68,145	21,692	30,960	16,405	22,627	5,747	178,432	40,000	-	3,200	-	-	-	1,000	39,790
Ratio	99.14%	101.15%	99.75%	103.93%	97.44%	99.45%	102.32%	100.48%	100.00%		93.84%				92.59%	97.50%

Table 6. Mean and Median Net Worth by Strata Variable, 1992 SCF and Match File

Mean values of Net Worth (in 1992 dollars)			
	SCF1992	Imputed	Ratio
Net Worth	177,582	178,432	100%
House	70,209	69,604	99%
Business Assets	67,369	68,145	101%
Liquid Assets	21,747	21,692	100%
Stocks, bonds	29,790	30,960	104%
Retirement Assets	16,835	16,405	97%
Mortgage debt	22,752	22,627	99%
Other debt	5,616	5,747	102%
Distribution among population subgroups			
Ratio of Mean Values			
Race			
nonwhite	72,117	69,220	96%
white	212,199	210,549	99%
Race			
white	212,199	210,549	99%
black	43,065	59,223	138%
hispanic	49,254	62,469	127%
other	190,603	126,909	67%
Family type			
married couples	240,777	253,565	105%
single males	79,985	75,080	94%
single females	113,655	105,968	93%
Housing Tenure			
renter	33,674	25,875	77%
home owner	258,781	262,215	101%
Age			
nonelderly	157,128	164,105	104%
elderly	249,960	230,193	92%
Age			Ratio to ALL
lt 35	38,131	42,557	112%
35 to 44	123,273	122,073	99%
45 to 54	256,752	276,505	108%
55 to 64	325,172	329,693	101%
ge 65	249,960	230,193	92%
Income group			Ratio to ALL
lt \$20k	47,022	46,063	98%
\$20-50k	100,642	95,770	95%
\$50-75k	169,121	212,118	125%
\$75-100k	337,713	369,906	110%
gt \$100k	1,226,072	1,333,974	109%

Median values of Net Worth (in 1992 dollars)			
	SCF1992	Imputed	Ratio
Net Worth	40,810	39,790	98%
House	40,000	40,000	100%
Business Assets	-	-	-
Liquid Assets	3,410	3,200	94%
Stocks, bonds	-	-	-
Retirement Assets	-	-	-
Mortgage debt	-	-	-
Other debt	1,080	1,000	93%
Distribution among population subgroups			
Ratio of Median Values			
Race			
nonwhite	7,700	6,470	84%
white	60,790	56,600	93%
Race			
white	60,790	56,600	93%
black	5,111	5,030	98%
hispanic	1,175	7,200	613%
other	29,500	15,100	51%
Family type			
married couples	63,680	69,700	109%
single males	18,700	13,300	71%
single females	16,600	14,600	88%
Housing Tenure			
renter	120	101	84%
home owner	88,200	85,870	97%
Age			
nonelderly	30,700	30,700	100%
elderly	85,270	80,250	94%
Age			Ratio to ALL
lt 35	2,590	2,300	89%
35 to 44	33,300	35,830	108%
45 to 54	69,150	62,240	90%
55 to 64	96,400	98,000	102%
ge 65	85,270	80,250	94%
Income group			Ratio to ALL
lt \$20k	9,820	9,700	99%
\$20-50k	39,000	36,800	94%
\$50-75k	85,900	78,350	91%
\$75-100k	166,030	158,200	95%
gt \$100k	443,430	482,480	109%

Table 7. Alignment of Strata Variables for 1992 Time Use Match

	ADS 1993	AUTP 1985	Difference
<i>Individuals</i>	186,837,050	20,120	
HH Income			
<i>15k less</i>	26.31%	20.81%	5.50%
<i>25k less</i>	20.15%	25.26%	-5.11%
<i>35k less</i>	17.36%	23.71%	-6.35%
<i>35k over</i>	36.19%	30.22%	5.97%
Sex			
<i>Female</i>	52.08%	56.41%	-4.33%
<i>Male</i>	47.92%	43.59%	4.33%
Parent			
<i>No</i>	66.71%	68.07%	-1.36%
<i>Yes</i>	33.29%	31.93%	1.36%
Employed			
<i>No</i>	29.99%	36.75%	-6.76%
<i>Yes</i>	70.01%	63.25%	6.76%
Married			
<i>No</i>	42.41%	36.96%	5.45%
<i>Yes</i>	57.59%	63.04%	-5.45%

Table 8. Distribution of Matched Records by Matching Round, 1992 Time Use Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	103,943	92.7	92.7
2	3,365	3.0	95.7
3	115	0.1	95.8
4	1,865	1.7	97.5
5	2,767	2.5	100.0
8	37	0.0	100
Total	112,092	100	

Table 9. Distribution of Weekly Hours of Household Production in 1985 AUTP and Match File

	p90/p10	p90/p50	p10/p50	p75/p25	p75/p50	p25/p50	Gini
AUTP 1985	14.14	1.80	5.00	3.00	1.50	2.00	0.4810
Match	14.14	1.80	5.00	3.00	1.50	2.00	0.4809

Table 10. Comparison of Mean and Median Time Use Variables in 1992 Matched File

	Mean Care	Mean Proc.	Mean Core	Mean HH Prod.	Median Care	Median Proc.	Median Core	Median HH Prod.
AUTP 1985	3.34	5.81	15.02	24.17	0.00	0.00	9.92	18.67
Match	3.34	5.82	15.01	24.16	0.00	0.00	9.92	18.67
Ratio	99.90%	100.10%	99.93%	99.97%			100.00%	100.00%

Table 11. Mean and Median Household Production Weekly Hours, 1985 AOTP and Match

Mean values of HH Production (Weekly Hours)			
	AOTP1985	Imputed	Ratio
HH Production	24.17	24.16	100.0%
Care	3.34	3.34	99.9%
Procurement	5.81	5.82	100.1%
Core	15.02	15.01	99.9%
Distribution among population subgroups			
Ratio of Mean Values			
HH Income			
15k less	25.44	26.01	102.2%
25k less	24.36	24.62	101.1%
35k less	24.29	23.42	96.4%
35k over	23.10	22.92	99.2%
Sex			
Female	30.73	30.73	100.0%
Male	17.04	17.03	99.9%
Parent			
No	21.82	21.48	98.5%
Yes	29.16	29.53	101.3%
Employed			
No	33.17	32.54	98.1%
Yes	19.48	20.58	105.6%
Married			
No	19.51	20.73	106.2%
Yes	26.84	26.69	99.5%

Median values of HH Production (Weekly Hours)			
	AOTP1985	Imputed	Ratio
HH Production	18.67	18.67	100.0%
Care	-	-	
Procurement	-	-	
Core	9.92	9.92	100.0%
Distribution among population subgroups			
Ratio of Median Values			
HH Income			
15k less	21.00	21.00	100.0%
25k less	19.25	19.83	103.0%
35k less	19.25	17.50	90.9%
35k over	17.50	17.50	100.0%
Sex			
Female	27.18	27.18	100.0%
Male	11.08	11.08	100.0%
Parent			
No	16.92	16.68	98.6%
Yes	24.50	24.85	101.4%
Employed			
No	31.50	31.15	98.9%
Yes	14.00	14.70	105.0%
Married			
No	14.82	15.52	104.7%
Yes	21.23	21.12	99.5%

Table 12. Alignment of Strata Variables for 2007 Wealth Match

	ASEC2008	SCF2007	Difference
<i># Households</i>	116,783,684	116,122,131	-0.57%
Homeownership			
<i>renter</i>	32.07%	31.36%	-0.71%
<i>owner</i>	67.93%	68.64%	0.71%
Family Type			
<i>MC</i>	54.89%	58.80%	3.91%
<i>FH</i>	28.03%	27.15%	-0.88%
<i>MH</i>	17.08%	14.04%	-3.04%
Age Category			
<i><35</i>	22.07%	21.66%	-0.41%
<i>35-49</i>	30.08%	30.58%	0.50%
<i>50-65</i>	27.20%	26.63%	-0.57%
<i>>=65</i>	20.65%	21.14%	0.49%
Elder			
<i>nonelder</i>	79.35%	78.86%	-0.49%
<i>elder</i>	20.65%	21.14%	0.49%
Race			
<i>white</i>	70.87%	73.92%	3.05%
<i>black</i>	12.08%	12.58%	0.50%
<i>hispanic</i>	11.42%	9.41%	-2.01%
<i>other</i>	5.62%	4.08%	-1.54%
Race Category			
<i>nonwhite</i>	29.13%	26.08%	-3.05%
<i>white</i>	70.87%	73.92%	3.05%
HH Income			
<i>lt \$20k</i>	18.91%	21.48%	2.57%
<i>\$20-50k</i>	30.35%	32.53%	2.18%
<i>\$50-75k</i>	17.81%	17.25%	-0.56%
<i>\$75-100k</i>	11.93%	10.36%	-1.57%
<i>gt \$100k</i>	21.01%	18.37%	-2.64%

Table 13. Matching Cells for 2007 Wealth Match

White	Nonelder	Married Couple	6,153,278	7,425,911	1,272,633	32,800,000	33,900,000	1,100,000
		Female Head	5,836,162	6,083,420	247,258	7,113,122	8,028,186	915,064
		Male Head	5,265,231	4,776,694	(488,537)	5,932,545	4,791,845	(1,140,700)
	Elder	Married Couple	580,476	568,096	(12,380)	8,171,482	10,000,000	1,828,518
		Female Head	2,080,339	2,216,597	136,258	6,126,145	5,872,957	(253,188)
		Male Head	695,659	673,287	(22,372)	1,970,779	1,949,155	(21,624)
Black	Nonelder	Married Couple	1,463,289	2,181,595	718,306	2,789,680	3,187,606	397,926
		Female Head	3,637,055	3,636,837	(218)	1,764,171	1,812,725	48,554
		Male Head	1,500,803	1,195,575	(305,228)	768,405	786,938	18,533
	Elder	Married Couple	94,325	167,867	73,542	485,445	519,594	34,149
		Female Head	459,034	325,621	(133,413)	708,681	747,557	38,876
		Male Head	212,766	36,066	(176,700)	227,940	90,112	(137,828)
Hispanic	Nonelder	Married Couple	3,037,302	3,062,197	24,895	4,034,777	3,991,619	(43,158)
		Female Head	1,862,413	1,203,244	(659,169)	962,486	493,585	(468,901)
		Male Head	1,407,965	886,464	(521,501)	639,475	363,917	(275,558)
	Elder	Married Couple	141,657	136,705	(4,952)	472,314	325,835	(146,479)
		Female Head	230,583	299,065	68,482	342,797	190,174	(152,623)
		Male Head	106,808		(106,808)	100,168	40,706	(59,462)
Other	Nonelder	Married Couple	1,061,729	838,077	(223,652)	2,343,250	2,083,738	(259,512)
		Female Head	732,970	342,281	(390,689)	535,123	226,277	(308,846)
		Male Head	607,255	357,485	(249,770)	382,962	415,657	32,695
	Elder	Married Couple	91,743	42,122	(49,621)	341,786	194,035	(147,751)
		Female Head	137,023	130,083	(6,940)	204,244	102,379	(101,865)
		Male Head	57,561	37,961	(19,600)	72,838		(72,838)

Table 14. Distribution of Matched Records by Matching Round, 2007 Wealth Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	103,725,998	88.8	88.8
2	1,597,028	1.4	90.2
3	2,344,319	2.0	92.2
4	1,501,386	1.3	93.5
5	375,518	0.3	93.8
6	94,472	0.1	93.9
7	885,159	0.8	94.6
8	2,639,590	2.3	96.9
9	80,495	0.1	97.0
10	505,706	0.4	97.4
11	1,100,900	0.9	98.3
12	104,017	0.1	98.4
13	246,586	0.2	98.6
14	338,367	0.3	98.9
15	244,271	0.2	99.1
16	999,872	0.9	100.0
Total	116,783,684	100.0	

Table 15. Distribution of Net Worth in 2007 SCF and Matched File

	p90/p10	p90/p50	p50/p10	p75/p25	p75/p50	p50/p25	gini
SCF 2007	-154.286	8.533	-0.006	98.316	3.398	0.035	0.835
Match	-154.122	8.643	-0.006	136.627	3.434	0.025	0.831

Table 16. Comparison of Mean and Median Wealth Variables in 2007 Matched File to 2007 SCF

	Ave. Asset1	Ave. Asset2	Ave. Asset3	Ave. Asset4	Ave. Asset5	Ave. Debt1	Ave. Debt2	Ave. Networth	Med. Asset1	Med. Asset2	Med. Asset3	Med. Asset4	Med. Asset5	Med. Debt1	Med. Debt2	Med. Networth
SCF 2007	207,659	194,308	40,541	104,282	70,996	72,449	14,356	530,981	120,000	-	5,200	-	200	-	3,000	101,255
Match	203,446	185,370	40,594	95,436	69,850	71,419	13,956	509,322	120,000	-	5,100	-	70	-	3,000	99,505
Ratio	97.97%	95.40%	100.13%	91.52%	98.39%	98.58%	97.22%	95.92%	100.00%		98.08%		35.00%		100.00%	98.27%

Table 17. Mean and Median Net Worth by Strata Variable, 2007 SCF and Match File

Mean values of Net Worth (in 2007 dollars)			
	SCF 2007	Match	Ratio
Net Worth	530,981	509,322	96%
House	207,659	203,446	98%
Business Assets	194,308	185,370	95%
Liquid Assets	40,541	40,594	100%
Stocks, bonds	104,282	95,436	92%
Retirement Assets	70,996	69,850	98%
Mortgage debt	72,449	71,419	99%
Other debt	14,356	13,956	97%

Distribution among population subgroups				Ratio of Mean Values		
					SCF	Match
Race				Race		
white	645,819	638,287	99%	B/W	0.19	0.19
black	120,640	122,992	102%	H/W	0.26	0.26
hispanic	170,127	166,571	98%	O/W	0.85	0.64
other	547,767	410,324	75%			
Family type				Family type		
married couples	726,413	731,394	101%	SF/MC	0.30	0.29
single females	220,143	214,627	97%	SM/MC	0.43	0.38
single males	313,723	279,274	89%			
Housing Tenure				Housing Tenure		
renter	60,792	61,538	101%	renter/owner	0.08	0.09
home owner	745,799	720,729	97%			
Age				Age		
nonelderly	458,051	454,995	99%	NE/E	0.57	0.63
elderly	803,089	718,115	89%			
Income group				Income group	Ratio to ALL	
lt \$20k	103,192	108,761	105%	lt \$20k	0.19	0.21
\$20-50k	161,648	162,177	100%	\$20-50k	0.30	0.32
\$50-75k	291,666	272,386	93%	\$50-75k	0.55	0.53
\$75-100k	411,726	367,982	89%	\$75-100k	0.78	0.72
gt \$100k	1,976,913	1,652,562	84%	gt \$100k	3.72	3.24

Median values of Net Worth (in 2007 dollars)			
	SCF 2007	Match	Ratio
Net Worth	101,255	99,505	98%
House	120,000	120,000	100%
Business Assets	-	-	
Liquid Assets	5,200	5,100	98%
Stocks, bonds	-	-	
Retirement Assets	200	70	35%
Mortgage debt	-	-	
Other debt	3,000	3,000	100%

Distribution among population subgroups				Ratio of Median Values		
					SCF	Match
Race				Race		
white	142,805	147,570	103%	B/W	0.06	0.05
black	8,000	8,000	100%	H/W	0.06	0.07
hispanic	8,870	10,630	120%	O/W	0.96	0.49
other	137,000	72,850	53%			
Family type				Family type		
married couples	157,000	168,500	107%	SF/MC	0.26	0.24
single females	41,500	39,805	96%	SM/MC	0.27	0.19
single males	43,000	32,650	76%			
Housing Tenure				Housing Tenure		
renter	30	0	0%	renter/owner	0.00	0.00
home owner	212,078	213,100	100%			
Age				Age		
nonelderly	71,700	73,050	102%	NE/E	0.34	0.37
elderly	209,960	196,000	93%			
Income group				Income group	Ratio to ALL	
lt \$20k	4,150	8,071	194%	lt \$20k	0.04	0.08
\$20-50k	47,000	48,300	103%	\$20-50k	0.46	0.49
\$50-75k	113,000	98,005	87%	\$50-75k	1.12	0.98
\$75-100k	212,105	180,300	85%	\$75-100k	2.09	1.81
gt \$100k	560,000	448,001	80%	gt \$100k	5.53	4.50

Table 18. Alignment of Strata Variables for 2007 Time Use Match

	ASEC 2008	ATUS 2007	Difference
<i>Number</i>	237,993,292	234,238,921	-1.58%
Sex			
<i>Female</i>	51.43%	51.62%	-0.19%
<i>Male</i>	48.57%	48.38%	0.19%
Parent			
<i>No</i>	74.40%	69.88%	4.52%
<i>Yes</i>	25.60%	30.12%	-4.52%
Employed			
<i>No</i>	35.00%	34.06%	0.94%
<i>Yes</i>	65.00%	65.94%	-0.94%
Married			
<i>No</i>	44.65%	42.42%	2.23%
<i>Yes</i>	55.35%	57.58%	-2.23%
Spouse's Labor Force Status			
<i>No Spouse</i>	44.65%	42.42%	-2.23%
<i>Not Employed</i>	16.50%	18.27%	1.77%
<i>Employed</i>	38.85%	39.31%	0.46%

Table 19. Distribution of Matched Records by Matching Round, 2007 Time Use Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	223,368,566	93.9	93.9
2	2,273,377	1.0	94.8
3	343,914	0.1	95.0
4	1,118,787	0.5	95.4
5	7,088,604	3.0	98.4
6	3,200,222	1.3	99.7
7	206,122	0.1	99.8
8	84,406	0.0	99.9
9	309,294	0.1	100.0
Total	237,993,292	100.0	

Table 20. Distribution of Weekly Hours of Household Production in 2007 ATUS and Match File

	p90/p10	p90/p50	p10/p50	p75/p25	p75/p50	p25/p50	Gini
ATUS 2007	17.00	2.83	6.00	4.19	1.86	2.25	0.5172
Match	17.00	2.83	6.00	4.16	1.85	2.25	0.5179

Table 21. Comparison of Mean and Median Time Use Variables in 2007 Matched File

	Mean Care	Mean Proc.	Mean Core	Mean HH Prod.	Median Care	Median Proc.	Median Core	Median HH Prod.
ATUS 2007	5.10	4.70	13.00	23.00	0.00	0.00	7.00	17.00
Match	5.10	4.70	13.00	23.00	0.00	0.00	7.00	17.00
Ratio	100.00%	100.00%	100.00%	100.00%			100.00%	100.00%

Table 22. Mean and Median Household Production Weekly Hours, 2007 ATUS and Match

Average HH Production Weekly Hours						
	atus	asec	ratio			
care_wh	5.10	5.10	100%			
proc_wh	4.70	4.70	100%			
core_wh	13.00	13.00	100%			
hhprod_wh	23.00	23.00	100%			
				Ratios		
					atus	asec
Female	28.00	28.00	100%	fem/male	1.556	1.556
Male	18.00	18.00	100%			
Unmarried	18.00	19.00	106%	sing/marr	0.667	0.704
Married	27.00	27.00	100%			
no kid	20.00	21.00	105%	no kid/kid	0.645	0.700
kid	31.00	30.00	97%			
unemp	28.00	28.00	100%	unemp/emp	1.333	1.333
emp	21.00	21.00	100%			
no spouse	18.00	19.00	106%	spun/nosp	1.333	1.316
spouse unemp	24.00	25.00	104%	spemp/nosp	1.611	1.474
spouse emp	29.00	28.00	97%			
less high school	17.00	22.00	129%	ltHS/coll	0.708	0.917
high school grad	25.00	23.00	92%	HS/Coll	1.042	0.958
some college	24.00	23.00	96%	ltColl/coll	1.000	0.958
college grad	24.00	24.00	100%			
white	24.00	23.00	96%			
black	19.00	22.00	116%	blk/wh	0.792	0.957
other	23.00	24.00	104%	oth/wh	0.958	1.043
hispanic	23.00	23.00	100%	hisp/wh	0.958	1.000

Median HH Production Weekly Hours						
	atus	asec	ratio			
care_wh	0.00	0.00	#DIV/0!			
proc_wh	0.00	0.00	#DIV/0!			
core_wh	7.00	7.00	100%			
hhprod_wh	17.00	17.00	100%			
				Ratios		
					atus	asec
Female	23.00	23.00	100%	fem/male	1.917	1.917
Male	12.00	12.00	100%			
Unmarried	12.00	12.00	100%	sing/marr	0.545	0.571
Married	22.00	21.00	95%			
no kid	14.00	14.00	100%	no kid/kid	0.538	0.560
kid	26.00	25.00	96%			
unemp	23.00	22.00	96%	unemp/emp	1.533	1.467
emp	15.00	15.00	100%			
no spouse	12.00	12.00	100%	spun/nosp	#DIV/0!	#DIV/0!
spouse unemp	20.00	20.00	100%	spemp/nosp	#DIV/0!	#DIV/0!
spouse emp	23.00	22.00	96%			
less high school	11.00	16.00	145%	ltHS/coll	0.579	0.889
high school grad	19.00	18.00	95%	HS/Coll	1.000	1.000
some college	18.00	17.00	94%	ltColl/coll	0.947	0.944
college grad	19.00	18.00	95%			
white	18.00	17.00	94%			
black	12.00	16.00	133%	blk/wh	0.667	0.941
other	14.00	18.00	129%	oth/wh	0.778	1.059
hispanic	16.00	17.00	106%	hisp/wh	0.889	1.000

FIGURES

Figure 1. Distribution of Log Net Worth, 1992 SCF and Match File

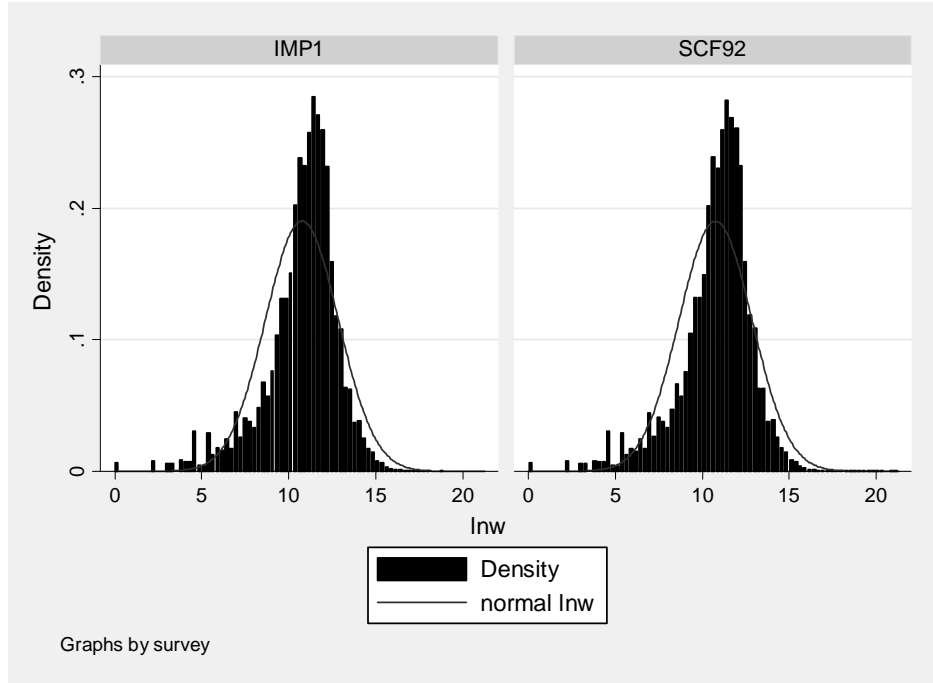


Figure 2. Ratio of Mean Net Worth by Category (Match/SCF 1992)

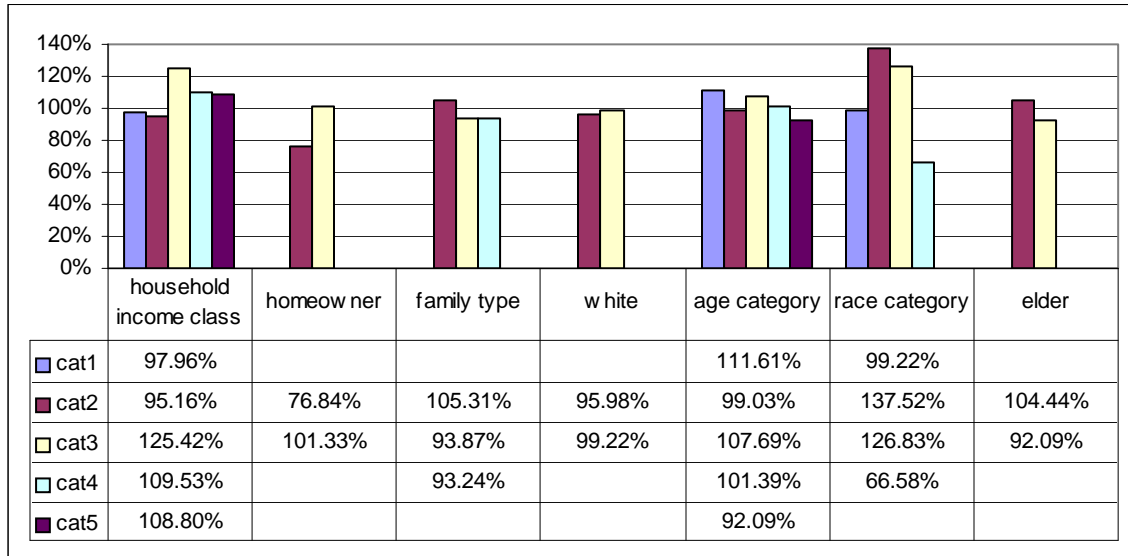


Figure 3. Net Worth by Matching Cells, 1992 SCF and Match File

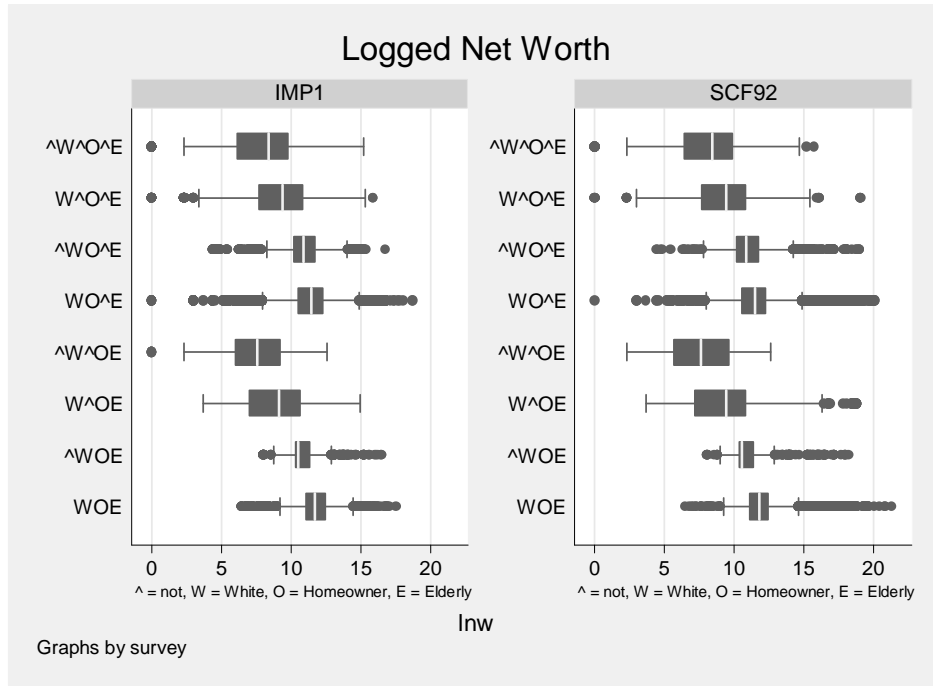


Figure 4. Ratio of Mean HH Production by Category (Match/AUTP 1985)

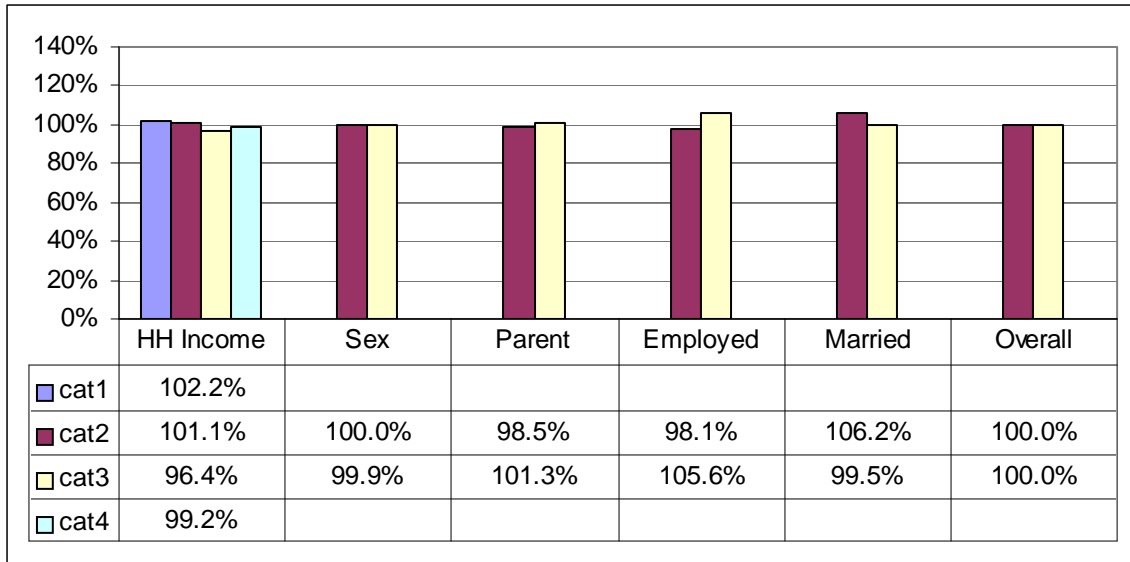


Figure 5. Household Production by Matching Cells, 1985 AUP and Match File

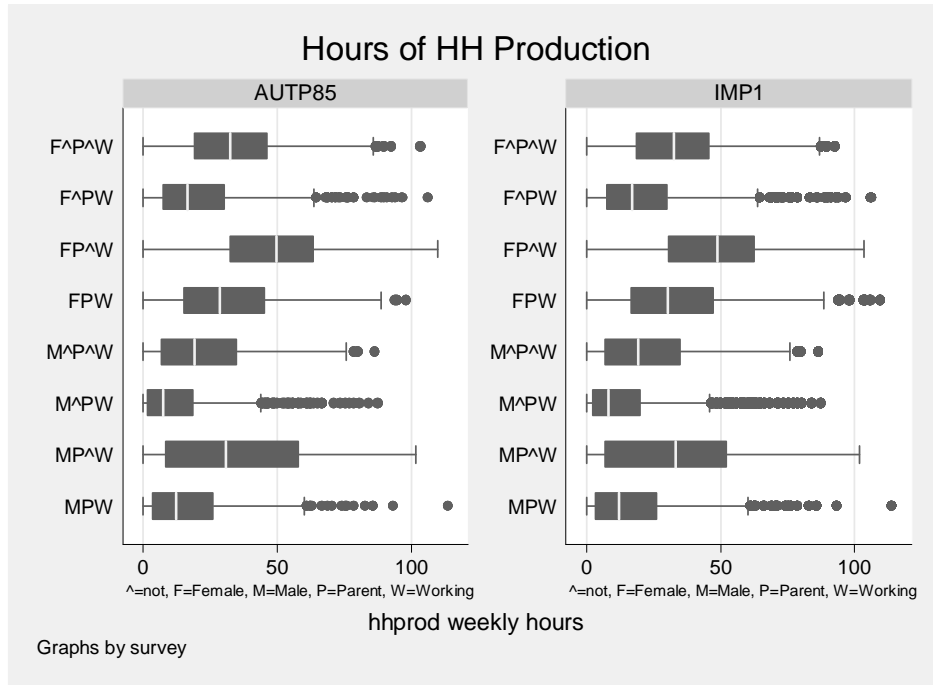


Figure 6. Distribution of Log Net Worth, 2007 SCF and Match File

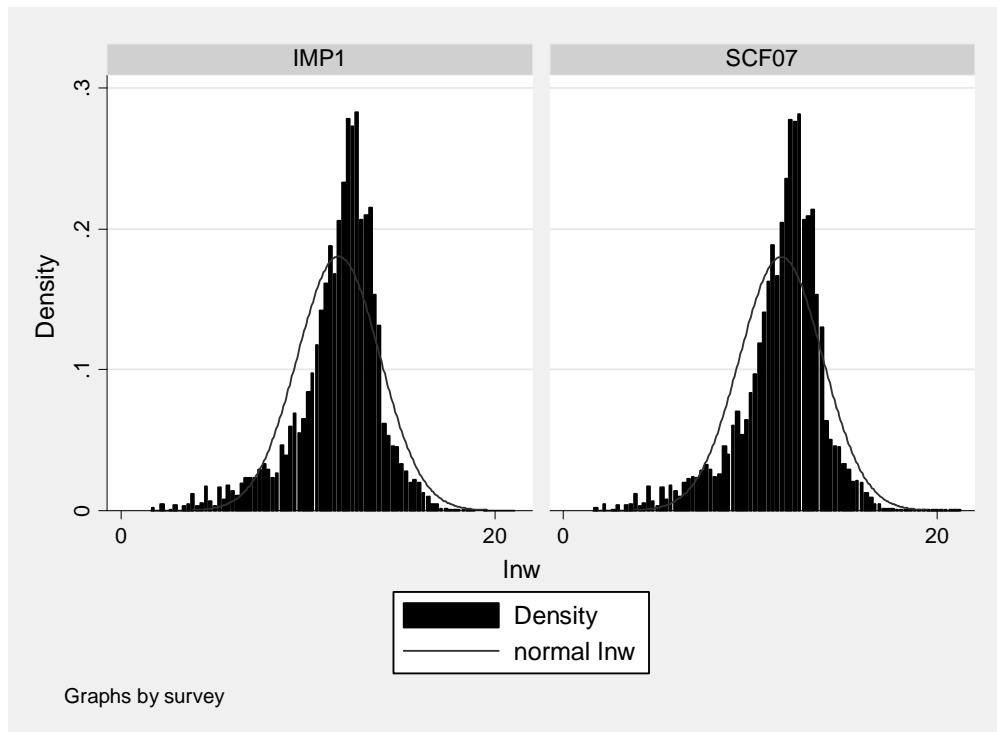


Figure 7. Ratio of Mean Net Worth by Category (Match/SCF 1992)

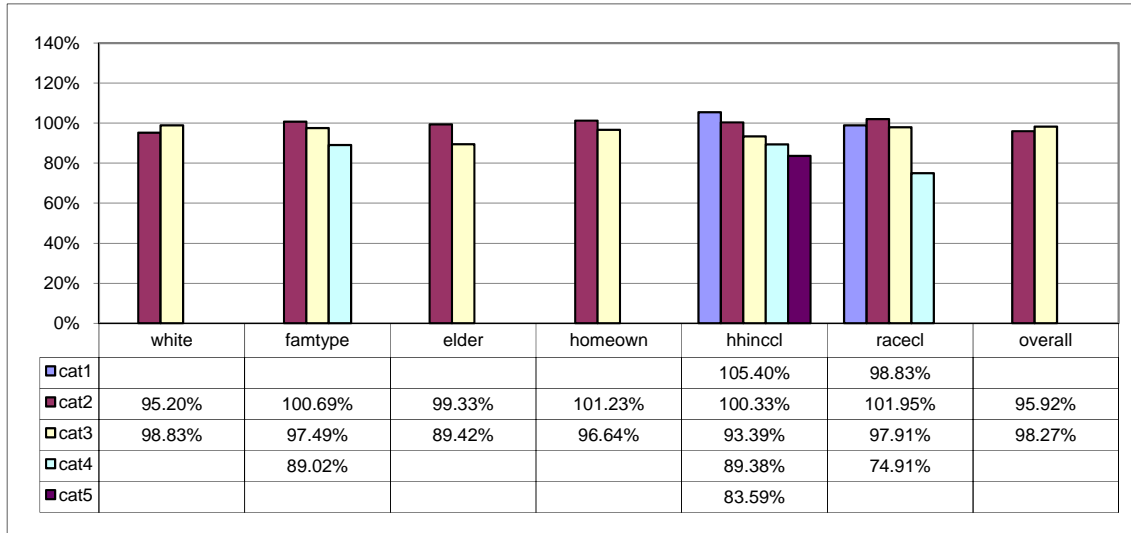


Figure 8. Net Worth by Matching Cells, 2007 SCF and Match File

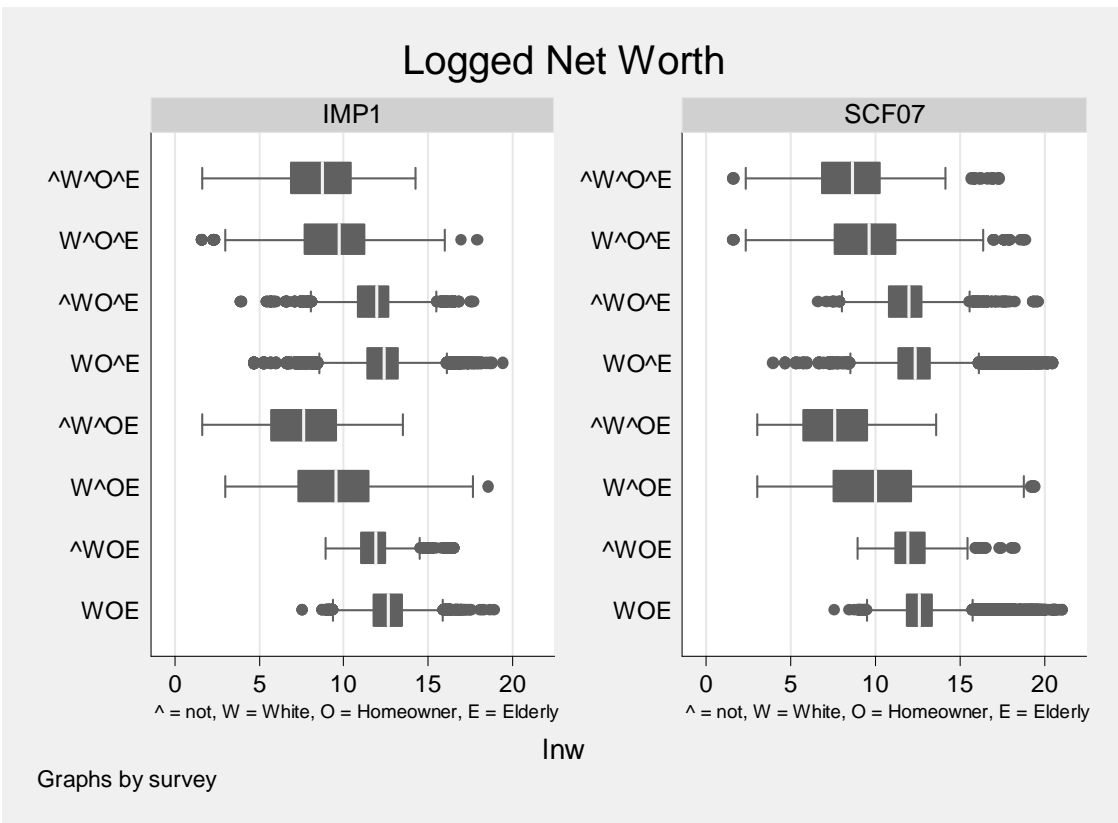


Figure 9. Ratio of Mean Household Production by Category (Match/ATUS 2007)

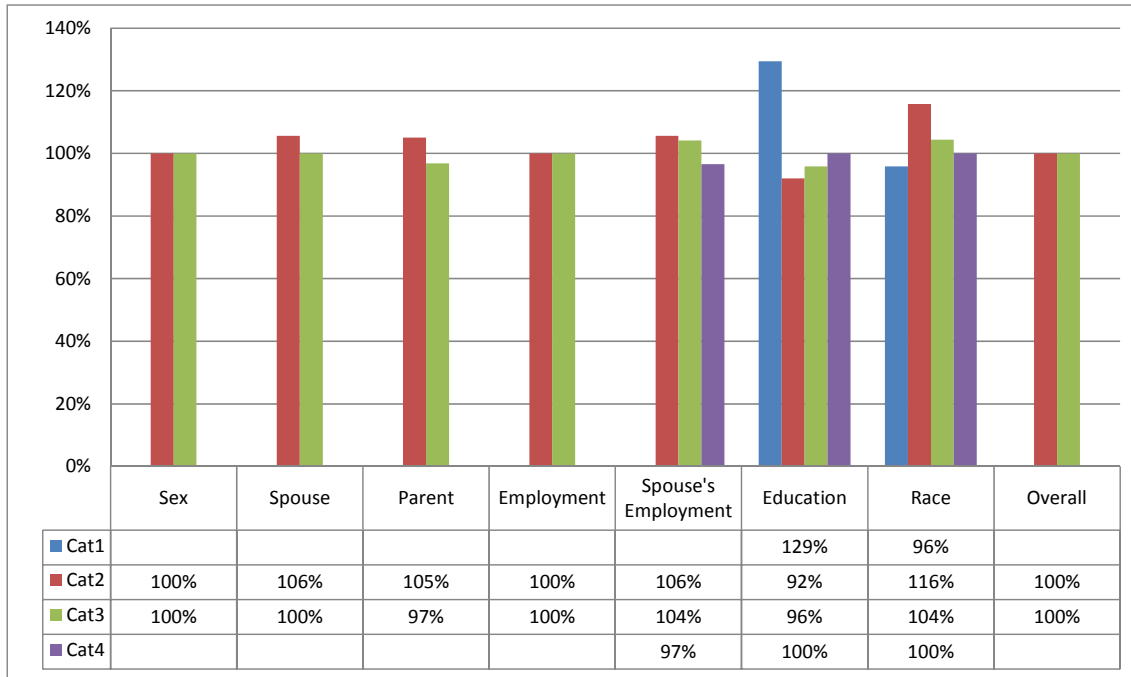


Figure 10. Household Production by Matching Cells, 2007 ATUS and Match File

