



Working Paper No. 676

Quality of Match for Statistical Matches Used in the 1989 and 2000 LIMEW Estimates for France*

by

Thomas Masterson

Levy Economics Institute of Bard College

July 2011

* The assistance of Ramzi Hadji, Centre de Recherche en Économie de Paris Nord (CEPN), was invaluable in preparing the datasets for the statistical matches presented here. Correspondence: masterso@levy.org.

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>

ABSTRACT

The quality of match for each of four statistical matches used in the LIMEW estimates for France for 1989 and 2000 is described. The first match combines the 1992 Enquête sur les Actifs Financiers with the 1989–90 Enquête Budget de Famille (BDF). The second match combines the 1998 General Social Survey (EDT) with the 1989–90 BDF. The third match combines the 2003–04 Enquête Patrimoine with the 2000–01 BDF. The fourth match combines the 1999 EDT with the 2000 BDF. In each case, the alignment of the two datasets is examined, after which various aspects of the match quality are described. 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; France; LIMEW

JEL Classifications: C14, C40, D31

INTRODUCTION

This paper describes the construction of synthetic datasets created for use in estimation of the LIMEW for France for the years 1989 and 2000. 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 France. Thus, in order to produce LIMEW estimates, a synthetic data file is created from various source data sets with statistical matching.¹ We use the Institute National de la Statistique et des Études Économiques' (INSEE) Enquête Budget de Famille (BDF) as the base data set, since it contains good information on demographics, income, transfers, and taxes for a regionally representative sample of French households. Wealth data for 1989 comes from the 1992 Enquête sur les Actifs Financiers (EAF), and for 2000 from the 2004 Enquête Patrimoine (PAT), both carried out by INSEE. Time use data comes from the Enquête Emploi du Temps (EDT) also carried out by INSEE.

This paper is organized as follows. Each section of the paper details four statistical matches in turn: wealth and time use matches for 1989 and 2000 for France. The source datasets are described and their demographic characteristics are compared. Then the quality of the match is reviewed for each.

1989 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 1989 French LIMEW estimates are the 1989-90 BDF and the 1992 EAF. The 1989-90 BDF is used since it has income data for 1989. The 1989-90 BDF file has records for 24,595 individuals in 9,038 households. These records represent 54,658,197 individuals in 21,201,890 French households after weighting. The 1992 EAF contains 9,530 household records. Many of the wealth and income variables were

¹ For details of the LIMEW and its construction, see Wolff and Zacharias (2003). See Kum and Masterson (2010,) for details of the statistical matching procedure that we use.

categorical. In these cases, we replaced those above the median category with a random draw from a Pareto distribution within the record's category range. We dealt with the missing values² in the data with the method of multiple imputation with chained equations. We created five imputates for each record for a total of 47,650 records. This translates to 22,145,405 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.³ For the 1989 French wealth match, strata variables are homeownership, age of the household head, educational achievement of the household head, 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 three years apart, we can expect them to be reasonably well aligned.

The largest differences between the two surveys are in terms of income category, with those at the lower and higher ends of the household income distribution making up a smaller proportion of the EAF sample than of the BDF, while those in the middle income categories make up a larger proportion. These misalignments can make matching a challenge, because it ensures that, for example some households with less than 50,000 Francs⁴ annual income in the BDF will be matched with households in the middle income categories in the EAF, thereby slightly exaggerating the wealth profile of the lower end of the income distribution (corresponding effects can be expected at the upper end of the income distribution). The other strata variables are better aligned, with home ownership and family type having one percent or less difference between the surveys. The former is especially significant for the wealth match, of course, since home ownership constitutes a major proportion of most households' assets.

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/non-elder indicator variable). Here we can see that the higher prevalence of young homeowners in the BDF is concentrated among married couples, with by far the largest absolute differences, especially

² Variables with missing values were: home ownership, dwelling type, household income class, home value, and most of the asset value variables.

³ Statistical matching is done first within subsets of the two data sets defined by key variables, which are referred to as strata variables.

⁴ All monetary values are in nominal French Francs for 1989 and in nominal Euros for 2000.

among the less educated. We can see that the quality of the match will likely be worst according to educational achievement.

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 majority of the matches usually happen in the earliest rounds, but generally a much greater percentage than in this case. Only 92% of the records are matched in the first five rounds. This demonstrates the effect of the misalignment noted above. This fact means that although most of the wealth records will be assigned to records that are similar in age, education, family type, home ownership and income to their donor records, a great many will be mismatched in one or more of these dimensions. In all, twenty-two rounds of matching were required to match every donor record. The final round includes all those recipient records for which no match could be found. In the latter case, each recipient record was assigned the average value from the corresponding subcell in the donor data set for each variable. 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 EAF and the matched file. The p75/p50 and p90/p50 ratios are quite close, but the others are not as good. It appears that the bottom tail of the wealth distribution in the matched file is somewhat thinner than in the EAF. For example, p10 for net worth in the matched file is 285F, while it is 1,304F in the EAF. The Gini coefficient is quite close, 0.681 in the matched file, compared to 0.677 in the EAF. 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.⁵ 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, only assets 1 and 3 are non-zero. Asset 3 is almost four percent smaller in the matched file, but this amounts to less than 2,000F. The most important asset, asset 1, is precisely matched, and the median net worth is off by 2.8%, but again, this represents a small absolute difference of just 8,000F.

⁵ 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).

Examination of the quality of the match within population sub-groups shows generally good results. Figure 2 displays ratios of mean net worth between the matched file and the EAF for the five strata variables. With one exception, the ratios of mean net worth within sub-categories of the five strata variables are all within 10% of unity. The fourth income group (from 100,000 to 130,000 Francs in household income) has 15% lower net worth in the matched file than in the EAF. Table 6 has the actual numbers, and we can see that this represents a substantial difference of 79,000F. The median net worth for this group in the matched file is 18% smaller than that of the EAF, though this difference is less than 63,000F. The second group in the homeowner panel of Figure 2 is homeowners. We can see that they have 3.2% smaller net worth in the matched file than in the EAF. We see in Table 6 that this translates to 30,000F less average net worth for homeowners in the matched file. The corresponding difference in medians is 8,000F. Those households with elderly heads have 6% lower mean net worth in the matched file than in the EAF. Consulting Table 6, we see that this means 40,000F smaller net worth, while their median net worth is 9.5% lower than in the EAF (a 39,000F difference). For judging the accuracy of the match in preserving the distribution of wealth by sub-groups, Table 6 displays the ratios of mean and median values for the strata variables' categories. The renter-owner ratios of mean and median values are well-carried over, while the ratios for the elder/non-elder ratio are as well. The ratios by household income group are surprisingly well reproduced in the match file, considering the misalignment in this variable. The rest of the ratios' values in the EAF 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 3.⁶ 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 household income. But the overall distribution is transferred with remarkable accuracy, and the distribution within even small sub-groups is transferred with good precision.

⁶ Household income and educational achievement are excluded for the sake of clarity of the plot.

1989 TIME USE MATCH

Data and Alignment

The source data sets for the time use match for the 1989 LIMEW estimates are the 1989-90 BDF and the 1985 EDT. We use individual records from the 1989-90 BDF file, excluding those living in group quarters or in the Armed Forces. Since the EDT covers individuals 15 years old and above, we discard younger individuals from the BDF file. This leaves 19,293 records, which represents 43,496,343 individuals when weighted. The EDT file includes time use data for 16,047 individuals, representing 43,183,035 individuals when weighted. For the time use match, the strata variables are sex, parental status, employment status, marital status, and spouse's employment 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 in the two data sets. We see that the distribution of individuals by sex is very closely aligned in the two surveys. The next closest match is by labor force status, with more employed persons in the EDT. Parental status is also well-aligned. However, the portion of married individuals is much higher in the BDF. Spouse's labor force status, on the other hand, is relatively close (among those with spouses). Clearly marital status is the most troubling in terms of alignment and we can expect there to be some discrepancy between the matched file and the EDT in this variable.

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, 92%, 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 EDT and the matched file. The percentile ratios are almost all equivalent. P75 is slightly off between the matched file (35.93 hours) and the EDT (35.58 hours), a very small difference. The Gini coefficient is extremely close, 0.4875 in the matched file, compared to 0.4866 in the EDT. Table 10 breaks down the mean and median of the

three classes that make up total household production in the time use match.⁷ We can see that for all four variables the matched and the source file's mean and median are equal.

Examination of the quality of the match within population sub-groups shows generally good results. Figure 4 displays ratios of mean weekly hours of household production between the matched file and the EDT for the five strata variables. When not equal, the ratios of mean weekly hours of household production within sub-categories of the strata variables are all within 5% of unity. Unmarried individuals and those individuals whose spouse is not working have weekly hours that are 5% lower and higher, respectively, in the matched file than in the EDT. Table 11 has the actual numbers, and we can see that these differences amount to one hour a week in each case. However, notice that the median weekly hours of household production for unmarried individuals in the matched file is two hours lower than that of the EDT, for a difference of 13%. The median weekly hours for those not working is one hour lower in the matched file, a difference of 4%. All other means and medians in the matched file perfectly mirror the EDT. For judging the accuracy of the match in preserving the distribution of household production by sub-groups, 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 still small. The rest of the ratios' values in the EDT are perfectly represented in the match file. The extent to which the match file reproduces the distribution of weekly hours of household production within collapsed matching cells is demonstrated in Figure 5.⁸ We can see very little difference between the matched file and the EDT. 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. The overall distribution is transferred with remarkable accuracy, and the distributions within sub-groups, such as female non-parent employees, are transferred with good precision. Even in the case of marital status, the transfer of weekly hours of household production is quite precise.

⁷ The three classes are care (child care, education, etc.), procurement (shopping, etc.), and core (cooking, cleaning, laundry, etc.).

⁸ Marital status and spouse's employment status are excluded for the sake of clarity of the plot.

2000 WEALTH MATCH

Data and Alignment

The source data sets for the wealth match for the 2000 LIMEW estimates for France are the 2000 BDF and the 2004 PAT. The 2000 BDF is used since it has income and demographic data for 2000. The 2000 BDF file contains records for 25,803 individuals in 10,305 households. These records represent 59,450,271 individuals in 24,525,505 French households after weighting. Missing values have been replaced using the method of multiple imputation with chained equations.⁹ This resulted in five replicates for each original observation for a total of 129,015 individual records and 51,525 household records. The 2004 PAT contains 9,692 household records. When the weights are appropriately adjusted, the records in the PAT represent 24,737,820 households. As for the EAF 1992, many of the asset and income values were categorical and so were transformed using the Pareto distribution in the manner described above. Again, missing values were replaced using the method of multiple imputation with chained equations.¹⁰ This process produced five implicates for each original record, resulting in a total of 48,460 records. The strata variables for this wealth match are homeownership, age, family type, household income, and education. Table 12 shows the distribution of households by these five variables in the two data sets. Both surveys are regionally representative samples carried four years apart, we can expect them to be reasonably well aligned.

We see that as with the 1989 wealth match, the distribution of household income is fairly poorly aligned. In this case, however, the upper and lower income categories are over-represented in the PAT, while the middle income categories are under-represented, with respect to the BDF. The distribution of the other strata variables is very close in the two surveys, within one percent in all cases but family type. In the latter case, married couples are 1.9% more prevalent in the BDF than the PAT, while male-headed households are 1.3% less prevalent in the PAT. These misalignments carry the cautions mentioned above in terms of what we can expect from the match quality along these dimensions, at least.

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/non-

⁹ Variables with missing values were educational attainment and occupational category.

¹⁰ Variables with missing values were occupational category, dwelling type, and nearly all of the financial variables.

elder indicator variable). Here we can see that the higher prevalence of homeownership in the BDF is concentrated among younger households, especially single male-headed. Based on these observations of the alignment, we can expect that the worst misallocation of wealth variables will be by homeownership and household income.

Match QC

The match itself required twenty rounds of matching to complete and was 85 percent done after the first round (see Table 14), within one of 162 very detailed matching cells (formed by combinations of all five strata variables). After five rounds over 95% of the records were matched. These characteristics of the matching process indicate that the quality of the match should be good. Table 15 and Figure 6 begin to show that this is in fact the case. The distribution of net worth has been fairly well-preserved. There are very small discernible differences in the density of log net worth between the PAT and the matched file (Figure 6). Percentile ratios are closely carried over (Table 15). The differences in the ratios between the matched file and the PAT, are due to the lower half of the distribution in the matched file having larger values than the PAT and vice versa for the upper half of the distribution. For example, the p10 value for net worth in the matched file is €375, as opposed to €354 in the PAT file, while the p90 is €348,645 and €349,089 in the match file and the Pat, respectively. The Gini coefficients are, nonetheless, almost identical. The components of net worth are well carried over into the matched file (see Table 16). The largest difference in means is for debt 1, home debt, which is 10% (€1,000) lower in the matched file. The rest are within 2% of the PAT. The largest difference in the medians is for asset 1 which is 14.5% (€5,500) lower in the matched file.

Figure 7 shows the ratio of mean net worth in the matched file to the PAT by strata variable categories. As we can see, average values of net worth for various demographic groups has been fairly well reproduced in the match file, with generally small variations between the matched file and the PAT. In most cases the differences are within 5%. Exceptions include male-headed households, with 7.2% lower net worth in the matched file, elders with 7% greater net worth, renters with 9.7% greater net worth in the matched file, and household heads with less than a baccalaureate, with 5.8% greater net worth. The greatest differences are by household income category. Households with between €10,000 and €20,000 in household income per year have 7.4% lower net worth in the matched file, while those with between €30,000 and €60,000

and greater than €60,000 in household income per year have 10.3% and 17.8% greater net worth in the matched file, respectively. These relatively large differences are due to the misalignment in household income categories between the two files noted above.

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 PAT and the matched file. The most notable difference is the ratio between non-elder and elder mean household net worth. While the means in the matched file differ considerably from the PAT, the relative position of the non-elders vis-à-vis elders is preserved. The matched file to PAT ratios in median values are somewhat more concerning. Non-elders have 13% lower median net worth in the matched file (a €8,400 difference), while households between €10,000 and €20,000 in household income per year have 27% lower median net worth (€8,800). However, the ratios of non-elder to elder median net worth are close enough and the ratios of the individual income categories to the highest category are well reproduced in the matched file.

Finally, Figure 8 shows the distribution of log net worth within collapsed matching cells (by family type, homeownership, and age). The distributions have been carried over very well. The most obvious difference is that the lower tails of the distributions have not been carried over completely in some of the larger cells (for example, non-elder renter married couples). The bulk of the distribution is quite well carried over, however.

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

2000 TIME USE MATCH

Data and Alignment

The source data sets for the time use match for the 2000 LIMEW estimates are the 2000 BDF and the 1999 EDT. We use individual records from the 2000 BDF file, excluding those living in group quarters or in the Armed Forces. Since the EDT covers individuals 15 years old and above, we discard younger individuals from the BDF file. This leaves 103,320 records, which represents 47,659,195 individuals when weighted. The EDT file includes time use data for 15,466

individuals, corresponding to 47,302,220 individuals when weighted. Due to missing values,¹¹ we used multiple imputation with chained equations on the 1999 EDT. For the time use match, the strata variables are sex, parental status, employment status, marital status, and spouse's employment 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 in the two data sets. Since the two surveys were carried out just one year apart, 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 less prevalent in the BDF than in the EDT (by 7.5%). The employed are slightly under-represented by 2.1%, in the EDT relative to the BDF. The portion of married individuals is lower in the EDT, by 1.8%. The difference in spouse's labor force status is quite small (0.4%). The difference in parental status, possibly 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 only seven rounds were required to complete the match is a promising sign for the quality of the match. Indeed, 90.8 percent of records were matched in the first round of matching. The overall distribution of weekly hours of household production in the matched file is very close to that in the EDT, based on the percentile ratios and Gini coefficients displayed in Table 20. Only the p90/p50 ratio is off, by very little. The Gini coefficient is off by only 0.01 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 EDT (see Table 21). Figure 9 displays ratios of mean weekly hours of household production by the strata variables, as well as household income and education. In terms of the strata variables, the match looks very good for each one. With one exception the matched file exactly reproduces the EDT. Non-parents have 6% greater average weekly hours of household production in the match file. In terms of household income and education, the differences are greater, but still mostly within 10%. The lowest household income category is the farthest off, 18% lower in the matched file than in the EDT, while the highest

¹¹ The one variable with missing values was household income.

income category and those with greater than baccalaureates had 13% and 12% greater weekly hours of household production, respectively, in the matched file than in the EDT.

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 household income. Here we can see that the 6% difference in mean weekly hours for non-parents translates to one hour per week, as do the differences by education and income for the most part. The exceptions are for those households with less than €10,000 (four hours less) and more than €50,000 and greater than baccalaureate (two hours more). The ratios by strata variables are correspondingly well reproduced in the matched file. As we can see, the ratios of matched to EDT medians are unity for all the strata variable categories except non-parents. For the latter the difference is 7%, but still only a one hour difference. The differences for non-strata variables are again larger, with those with less than a baccalaureate registering two hours less per week and those with greater than a baccalaureate one more at the median in the matched file, while those in households with less than €10,000 incomes have six fewer, those in households with €10,000 to €20,000 two fewer, and those with €20,000 to €30,000 and greater than €50,000 two more hours of household production. The ratios of household income categories to the highest category are thus not well-retained in the matched file.

Finally, Figure 10 displays the distributions of household production weekly hours in collapsed matching cells (by sex, parent, and employment status). There are few noticeable differences between the EDT and the matched file, indicating that even within cells, there has been good transference of the distributions of household production. In many of the cells the upper tail has not been well-transferred.

In summary the reproduction of the weekly hours of household production in the EDT in the matched file is very good. The remaining differences are small, and will not greatly impact the final LIMEW estimates for France.

REFERENCES

- Kum, Hyunsub, and Thomas Neal Masterson. 2010. "Statistical matching using propensity scores: Theory and application to the analysis of the distribution of income and wealth." *Journal of Economic and Social Measurement* 35(3): 177–196.
- 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. <http://www.levyinstitute.org/pubs/wp372.pdf>.

Tables

Table 1 Alignment of Strata Variables for 1989 Wealth Match

	BDF 1989	EAF 1992	Diff (%)
<i>Households</i>	21,201,890	22,145,485	4.45%
HH Income Category			
<i><50,000F</i>	16.35%	15.17%	-1.18%
<i>50,000-75,000 F</i>	13.24%	16.99%	3.75%
<i>75,000-100,000 F</i>	12.93%	15.86%	2.93%
<i>100,000-130,000 F</i>	15.06%	16.53%	1.47%
<i>130,000-200,000 F</i>	24.21%	20.63%	-3.58%
<i>>= 200,000 F</i>	18.21%	14.81%	-3.40%
Home ownership			
<i>Renter</i>	44.52%	45.57%	1.05%
<i>Owner</i>	55.48%	54.43%	-1.05%
Family Type			
<i>Married Couple</i>	65.54%	65.74%	0.20%
<i>Female Head</i>	23.71%	22.98%	-0.73%
<i>Male Head</i>	10.74%	11.28%	0.54%
Age Category			
<i>Nonelder</i>	75.41%	73.64%	-1.77%
<i>Elder</i>	24.59%	26.36%	1.77%
Age Category			
<i>Less than 35</i>	22.50%	20.34%	-2.16%
<i>35 to 44</i>	21.08%	21.44%	0.36%
<i>45 to 54</i>	15.22%	16.35%	1.13%
<i>55 to 64</i>	16.61%	15.51%	-1.10%
<i>65 and older</i>	24.59%	26.36%	1.77%
Educational Attainment			
<i>Less than BAC</i>	26.54%	24.09%	-2.45%
<i>BAC</i>	59.73%	61.69%	1.96%
<i>More than BAC</i>	13.73%	14.23%	0.50%

Table 2 Matching Cells for 1989 Wealth Match

			Less than BAC			BAC			More than BAC		
			BDF 1989-90	EAF 1992	Difference	BDF 1989-90	EAF 1992	Difference	BDF 1989-90	EAF 1992	Difference
Renter	Nonelder	Married Couple	1,208,486	1,186,345	(22,141)	2,647,512	2,801,121	153,609	709,276	675,370	(33,906)
		Female Head	421,506	412,231	(9,275)	1,004,986	978,935	(26,051)	372,410	365,860	(6,550)
		Male Head	230,523	238,420	7,897	653,365	742,785	89,420	264,251	267,200	2,949
	Elder	Married Couple	237,893	228,660	(9,233)	294,314	372,250	77,936	38,885	29,230	(9,655)
		Female Head	640,240	600,865	(39,375)	461,154	498,475	37,321	30,149	15,715	(14,434)
		Male Head	100,439	72,915	(27,524)	107,608	157,275	49,667	15,567	17,730	2,163
Owner	Nonelder	Married Couple	1,267,229	1,008,395	(258,834)	4,779,300	4,484,474	(294,826)	949,119	1,087,415	138,296
		Female Head	205,788	118,739	(87,049)	557,058	498,120	(58,938)	164,055	186,845	22,790
		Male Head	143,126	123,835	(19,291)	327,214	337,025	9,811	82,661	100,075	17,414
	Elder	Married Couple	532,164	565,575	33,411	1,038,371	1,322,410	284,039	193,820	176,300	(17,520)
		Female Head	470,669	404,955	(65,714)	638,088	727,565	89,477	61,877	64,780	2,903
		Male Head	169,013	145,650	(23,363)	155,430	158,515	3,085	28,344	29,605	1,261

Table 3 Distribution of Matched Records by Matching Round, 1989 Wealth Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	17,739,636	83.7	83.7
2	477,146	2.3	85.9
3	268400	1.3	87.2
4	321913	1.5	88.7
5	624,181	2.9	91.6
6	128956	0.6	92.2
7	78142	0.4	92.6
8	61835	0.3	92.9
9	234760	1.1	94.0
10	12051	0.1	94.1
11	648176	3.1	97.1
12	3982	0.0	97.2
13	62,811	0.3	97.5
14	51,217	0.24	97.7
15	11,124	0.05	97.7
16	82,827	0.39	98.1
17	4,356	0.02	98.2
18	10,184	0.05	98.2
19	4,140	0.02	98.2
20	19,976	0.09	98.3
21	3,947	0.02	98.3
22	356,077	1.68	100.0
Total	21,205,837	100.0	

Table 4 Distribution of Net Worth in 1989 Matched File

	p90/p10	p90/p50	p50/p10	p75/p25	p75/p50	p50/p25	Gini
Match	4979.45	4.88	1021.31	25.97	2.56	10.16	0.681
EAF	1108.89	4.83	229.60	22.40	2.53	8.87	0.677

Table 5 Comparison of Mean and Median Wealth Variables in 1989 Matched File to 1992 EAF

		Asset1	Asset2	Asset3	Asset4	Asset5	Debt1	Debt2	Networth
Mean	Match	336,387	165,774	101,147	21,090	14,108	48,446	8,399	581,660
	EAF92	338,982	169,960	103,525	21,646	14,489	49,860	8,326	590,417
	ratio	99.23%	97.54%	97.70%	97.43%	97.37%	97.16%	100.88%	98.52%
Median	Match	175,000	-	42,637	-	-	-	-	291,074
	EAF92	175,000	-	44,332	-	-	-	-	299,392
	ratio	100.00%		96.18%					97.22%

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

Average Net Worth

	EAF1992	Match	Ratio		EAF1992	Match
Asset1	338,982	336,387	99.23%			
Asset2	169,960	165,774	97.54%			
Asset3	103,525	101,147	97.70%			
Asset4	21,646	21,090	97.43%			
Asset5	14,489	14,108	97.37%			
Debt1	49,860	48,446	97.16%			
Debt2	8,326	8,399	100.88%			
Networth	590,417	581,660	98.52%			
renter	160,263	159,278	99.39%	ren/own	0.169	0.173
homeowner	950,537	920,567	96.85%			
non-elder	561,165	565,304	100.74%	non/eld	0.835	0.895
elder	672,142	631,813	94.00%			
MC	706,390	694,513	98.32%			
FH	362,526	363,688	100.32%	fh/mc	0.513	0.524
MH	378,856	374,287	98.79%	mh/mc	0.536	0.539
LT BAC	348,102	373,980	107.43%	ltBAC/gtBAC	0.335	0.375
BAC	581,608	578,480	99.46%	BAC/gtBAC	0.560	0.580
GT BAC	1,038,878	997,030	95.97%			
Less than 50K	235,667	256,977	109.04%	lt 50k	0.156	0.197
50K to 75K	297,237	323,337	108.78%	50-75k	0.197	0.248
75K to 100K	395,661	383,635	96.96%	75-100k	0.262	0.294
100K to 130K	521,959	442,518	84.78%	100-130k	0.345	0.340
130K to 200K	636,620	592,127	93.01%	130-200k	0.421	0.454
200K or more	1,511,029	1,302,947	86.23%			

Median Net Worth

	EAF1992	Match	Ratio		EAF1992	Match
Asset1	175,000	175,000	100.00%			
Asset2	0	0				
Asset3	44,332	42,637	96.18%			
Asset4	0	0				
Asset5	0	0				
Debt1	0	0				
Debt2	0	0				
Networth	299,392	291,074	97.22%			
renter	30,936	30,660	99.11%	ren/own	0.049	0.050
homeowner	633,302	612,812	96.76%			
non-elder	262,321	271,824	103.62%	non/eld	0.646	0.740
elder	405,827	367,279	90.50%			
MC	413,935	407,236	98.38%			
FH	112,214	106,244	94.68%	fh/mc	0.271	0.261
MH	121,317	132,088	108.88%	mh/mc	0.293	0.324
LT BAC	139,508	150,893	108.16%	ltBAC/gtBAC	0.288	0.406
BAC	350,718	358,847	102.32%	BAC/gtBAC	0.724	0.966
GT BAC	484,654	371,615	76.68%			
Less than 50K	44,009	44,640	101.43%	lt 50k	0.046	0.057
50K to 75K	109,691	140,000	127.63%	50-75k	0.114	0.179
75K to 100K	223,971	205,644	91.82%	75-100k	0.233	0.262
100K to 130K	348,414	285,243	81.87%	100-130k	0.363	0.364
130K to 200K	444,889	403,634	90.73%	130-200k	0.464	0.515
200K or more	959,759	783,938	81.68%			

Table 7 Alignment of Strata Variables for 1989 Time Use Match

	BDF1989	EDT1985	Diff (%)
<i>Individuals</i>	43,511,114	43,183,035	-0.75%
Sex			
<i>Female</i>	52.58%	52.15%	-0.43%
<i>Male</i>	47.42%	47.85%	0.43%
Parental Status			
<i>No</i>	72.43%	71.61%	-0.82%
<i>Yes</i>	27.57%	28.39%	0.82%
Labor Force Status			
<i>Not employed</i>	50.89%	50.36%	-0.53%
<i>Employed</i>	49.11%	49.64%	0.53%
Spouse			
<i>No</i>	36.58%	44.43%	7.85%
<i>Yes</i>	63.42%	55.57%	-7.85%
Spouse's Labor Force Status			
<i>Spouse not employed</i>	41.19%	40.35%	-0.84%
<i>Spouse employed</i>	58.81%	59.65%	0.84%

Table 8 Distribution of Matched Records by Matching Round, 1989 Time Use Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	40,060,981	92.2	92.2
2	19,518	0.0	92.2
3	2821934	6.5	98.7
4	113,828	0.3	99.0
5	116,908	0.3	99.3
6	229801	0.5	99.8
7	53704	0.1	99.9
8	37472	0.1	100.0
Total	43,454,146	100.0	

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

	p90/p10	p90/p50	p50/p10	p75/p25	p75/p50	p50/p25	Gini
EDT	17.600	2.667	6.600	4.692	1.848	2.538	0.4866
IMP	17.600	2.667	6.600	4.738	1.867	2.538	0.4875

**Table 10 Comparison of Mean and Median Time Use Variables in 1989
Matched File (weekly hours)**

	Mean Total	Mean Care	Mean Procure ment	Mean Core	Median Total	Median Care	Median Procure ment	Median Core
EdT	22.0	16.0	2.8	2.9	17.0	11.0	0.6	0.0
IMP	22.0	16.0	2.8	2.9	17.0	11.0	0.6	0.0
RATIO	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	

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

Mean Values of Household Production

	EdT85	Match	Ratio		EdT85	Match
Care	16.0	16.0	100.0%			
Procurement	2.8	2.8	100.0%			
Core	2.9	2.9	100.0%			
Total	22.0	22.0	100.0%			
Female	32.0	32.0	100.0%	F/M	2.909	2.909
Male	11.0	11.0	100.0%			
Unmarried	19.0	18.0	94.7%	S/M	0.826	0.750
Married	23.0	24.0	104.3%			
Non-parent	20.0	20.0	100.0%	NP/P	0.741	0.741
Parent	27.0	27.0	100.0%			
Not Working	27.0	27.0	100.0%	NW/W	1.588	1.588
Working	17.0	17.0	100.0%			
No Spouse	19.0	18.0	94.7%	NoSp/SpW	0.950	0.857
Spouse Not Workin	20.0	21.0	105.0%	NoSp/SpNW	0.731	0.692
Spouse Working	26.0	26.0	100.0%			

Median Values of Household Production

	EdT85	Match	Ratio		EdT85	Match
Care	11.0	11.0	100.0%			
Procurement	0.6	0.6	100.0%			
Core	0.0	0.0				
Total	17.0	17.0	100.0%			
Female	30.0	30.0	100.0%	F/M	4.286	4.286
Male	7.0	7.0	100.0%			
Unmarried	15.0	13.0	86.7%	S/M	0.833	0.684
Married	18.0	19.0	105.6%			
Non-parent	15.0	15.0	100.0%	NP/P	0.682	0.682
Parent	22.0	22.0	100.0%			
Not Working	24.0	23.0	95.8%	NW/W	2.000	1.917
Working	12.0	12.0	100.0%			
No Spouse	15.0	13.0	86.7%	NoSp/SpW	1.000	0.867
Spouse Not Workin	15.0	15.0	100.0%	NoSp/SpNW	0.714	0.619
Spouse Working	21.0	21.0	100.0%			

Table 12 Alignment of Strata Variables for 2000 Wealth Match

	BDF 2001	PAT 2004	Diff (%)
<i>Households</i>	24,525,505	24,737,820	0.87%
HH Income Category			
<i>Less than 10K</i>	10.58%	16.41%	5.83%
<i>10K to 20K</i>	30.54%	30.84%	0.30%
<i>20K to 30K</i>	25.64%	22.47%	-3.17%
<i>30K to 60K</i>	27.87%	20.03%	-7.84%
<i>60K or more</i>	5.37%	10.25%	4.88%
Home ownership			
<i>Renter</i>	45.19%	44.28%	-0.91%
<i>Owner</i>	54.81%	55.72%	0.91%
Family Type			
<i>Married Couple</i>	63.41%	61.55%	-1.86%
<i>Female Head</i>	24.29%	24.85%	0.56%
<i>Male Head</i>	12.30%	13.60%	1.30%
Age Category			
<i>Nonelder</i>	74.12%	73.11%	-1.01%
<i>Elder</i>	25.88%	26.89%	1.01%
Age Category			
<i>Less than 35</i>	19.95%	19.13%	-0.82%
<i>35 to 44</i>	20.35%	19.65%	-0.70%
<i>45 to 54</i>	20.11%	19.53%	-0.58%
<i>55 to 64</i>	13.71%	14.81%	1.10%
<i>65 and older</i>	25.88%	26.89%	1.01%
Educational Attainment			
<i>Less than BAC</i>	21.09%	20.82%	-0.27%
<i>BAC</i>	69.88%	70.17%	0.29%
<i>More than BAC</i>	9.03%	9.01%	-0.02%

Table 13 Matching Cells for 2000 Wealth Match

			Less than BAC			BAC			More than BAC		
			BDF 2000-1	Pat 2004	Difference	BDF 2000-1	Pat 2004	Difference	BDF 2000-1	Pat 2004	Difference
Renter	Nonelder	Married Couple	1,094,640	983,430	(111,210)	3,472,945	2,939,875	(533,070)	477,560	426,505	(51,055)
		Female Head	446,076	384,470	(61,606)	1,917,617	1,724,795	(192,822)	162,782	206,685	43,903
		Male Head	290,273	321,650	31,377	1,113,012	1,278,320	165,308	159,625	189,080	29,455
	Elder	Married Couple	217,901	285,150	67,249	419,714	462,175	42,461	26,680	33,385	6,705
		Female Head	351,385	546,070	194,685	607,024	742,465	135,441	48,446	57,190	8,744
		Male Head	107,318	103,325	(3,993)	161,934	156,365	(5,569)	9,183	18,265	9,082
Owner	Nonelder	Married Couple	1,023,024	1,024,670	1,646	5,441,576	5,740,505	298,929	796,465	852,610	56,145
		Female Head	144,321	138,350	(5,971)	834,083	824,800	(9,283)	102,211	120,215	18,004
		Male Head	106,300	126,580	20,280	510,745	570,265	59,520	85,060	75,140	(9,920)
	Elder	Married Couple	789,826	625,775	(164,051)	1,584,789	1,575,125	(9,664)	206,955	144,655	(62,300)
		Female Head	438,215	426,790	(11,425)	794,429	869,695	75,266	110,231	51,090	(59,141)
		Male Head	163,639	138,320	(25,319)	279,425	324,125	44,700	30,096	34,515	4,419

Table 14 Distribution of Matched Records by Matching Round, 2000 Wealth Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	20,775,605	84.7	84.7
2	305,639	1.3	86.0
3	157,082	0.6	86.6
4	150,318	0.6	87.2
5	2,012,298	8.2	95.4
6	2,020	0.0	95.4
7	80,449	0.3	95.8
8	291,313	1.2	96.9
9	123,398	0.5	97.4
10	136,370	0.6	98.0
11	33,933	0.1	98.1
12	26,201	0.1	98.2
13	10,617	0.0	98.3
14	116,646	0.5	98.8
15	40,731	0.2	98.9
16	14,570	0.1	99.0
17	44,778	0.2	99.2
18	20,072	0.1	99.3
19	21,901	0.1	99.3
20	161,564	0.7	100.0
Total	24,525,505	100.0	

Table 15 Distribution of Net Worth in 2004 Pat and Matched File

	p90/p10	p90/p50	p50/p10	p75/p25	p75/p50	p50/p25	Gini
Match	929.72	4.63	200.87	28.32	2.42	11.71	0.6786
PAT	986.13	4.56	216.45	29.12	2.40	12.14	0.6788

Table 16 Comparison of Mean and Median Wealth Variables in 2000 Matched File to 2004 Pat

		Asset1	Asset2	Asset3	Asset4	Debt1	Debt2	Networth
Mean	Match	76,899	46,840	13,456	23,657	8,729	1,531	150,592
	Pat04	78,008	47,240	13,615	23,937	9,693	1,536	151,570
	Ratio	98.58%	99.15%	98.83%	98.83%	90.05%	99.67%	99.35%
Median	Match	32,471	-	4,890	1,962	-	-	75,327
	Pat04	37,984	-	4,815	1,888	-	-	76,623
	Ratio	85.49%		101.56%	103.92%			98.31%

Table 17 Mean and Median Net Worth by Strata Variable, 2004 Pat and Match File

Average Net Worth

	PAT2004	Match	Ratio		PAT2004	Match
Asset1	78,008	76,899	98.58%			
Asset2	47,240	46,840	99.15%			
Asset3	13,615	13,456	98.83%			
Asset4	23,937	23,657	98.83%			
Debt1	9,693	8,729	90.05%			
Debt2	1,536	1,531	99.67%			
Networth	151,570	150,592	99.35%			
renter	37,265	40,886	109.72%	ren/own	0.154	0.170
homeowner	242,418	241,059	99.44%			
non-elder	147,887	142,819	96.57%	non/eld	0.915	0.826
elder	161,583	172,854	106.98%			
MC	193,378	189,387	97.94%			
FH	81,190	82,815	102.00%	fh/mc	0.420	0.437
MH	90,944	84,426	92.83%	mh/mc	0.470	0.446
LT BAC	88,565	93,699	105.80%	ltBAC/gtBAC	0.522	0.533
BAC	167,921	164,502	97.96%	BAC/gtBAC	0.989	0.936
GT BAC	169,789	175,842	103.57%			
Less than 10K	52,678	52,720	100.08%	lt 10k	0.104	0.089
10K to 20K	79,192	73,352	92.63%	10-20k	0.157	0.123
20K to 30K	117,851	113,046	95.92%	20-30k	0.233	0.190
30K to 60K	200,613	221,258	110.29%	30-60k	0.397	0.371
60K or more	505,600	595,656	117.81%	60k or more	1.000	1.000

Median Net Worth

	PAT2004	Match	Ratio		PAT2004	Match
Asset1	37,984	32,471	85.49%			
Asset2	0	0				
Asset3	4,815	4,890	101.56%			
Asset4	1,888	1,962	103.92%			
Debt1	0	0				
Debt2	0	0				
Networth	76,623	75,327	98.31%			
renter	4,612	5,140	111.45%	ren/own	0.030	0.033
homeowner	155,229	155,244	100.01%			
non-elder	66,642	58,228	87.37%	non/eld	0.649	0.514
elder	102,677	113,385	110.43%			
MC	114,906	109,644	95.42%			
FH	23,547	26,378	112.02%	fh/mc	0.205	0.241
MH	20,084	19,381	96.50%	mh/mc	0.175	0.177
LT BAC	33,825	47,569	140.63%	ltBAC/gtBAC	0.380	0.511
BAC	88,071	82,880	94.11%	BAC/gtBAC	0.988	0.890
GT BAC	89,108	93,111	104.49%			
Less than 10K	6,485	5,778	89.10%	lt 10k	0.022	0.018
10K to 20K	31,974	23,197	72.55%	10-20k	0.108	0.074
20K to 30K	82,774	73,195	88.43%	20-30k	0.280	0.234
30K to 60K	141,141	147,248	104.33%	30-60k	0.477	0.471
60K or more	295,876	312,454	105.60%	60k or more	1.000	1.000

Table 18 Alignment of Strata Variables for 2000 Time Use Match

	BDF 2001	EdT 1999	diff (%)
<i>Individuals</i>	47,659,195	47,302,220	-0.75%
Sex			
<i>Female</i>	52.29%	51.90%	-0.39%
<i>Male</i>	47.71%	48.10%	0.39%
Parental Status			
<i>No</i>	69.23%	61.73%	-7.50%
<i>Yes</i>	30.77%	38.27%	7.50%
Labor Force Status			
<i>Not employed</i>	50.10%	52.23%	2.13%
<i>Employed</i>	49.90%	47.77%	-2.13%
Spouse			
<i>No</i>	36.04%	37.89%	1.85%
<i>Yes</i>	63.96%	62.11%	-1.85%
Spouse's Labor Force Status			
<i>Spouse not employed</i>	43.04%	42.63%	-0.41%
<i>Spouse employed</i>	56.96%	57.37%	0.41%

Table 19 Distribution of Matched Records by Matching Round, 2000 Time Use Match

Matching Round	Records Matched	Percent	Cumulative Percent
1	43,273,229	90.8	90.8
2	808,810	1.7	92.5
3	2,504,740	5.3	97.8
4	52,069	0.1	97.9
5	705,783	1.5	99.4
6	233,510	0.5	99.9
7	67,233	0.1	100.0
Total	47,645,374	100.0	

Table 20 Distribution of Weekly Hours of Household Production in 1999 EDT and Match File

	p90/p10	p90/p50	p50/p10	p75/p25	p75/p50	p50/p25	Gini
EDT1999	.	3.08	.	8.67	2.00	4.33	0.5084
MATCH	.	3.00	.	8.67	2.00	4.33	0.5085

Table 21 Comparison of Mean and Median Time Use Variables in 2000 Matched File (weekly hours)

	Mean HH Prod.	Mean Care	Mean Proc.	Mean Core	Median HH Prod.	Median Care	Median Proc.	Median Core
EDT1999	20.00	14.00	3.60	2.20	15.00	9.30	0.00	0.00
MATCH	20.00	14.00	3.60	2.20	15.00	9.30	0.00	0.00
ratio	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%		

Table 22 Mean and Median Household Production Weekly Hours, 1999 EDT and Match

Mean Values of Household Production

	EDT1999	Match	Ratio		EDT1999	Match
Care	14.0	14.0	100.0%			
Procurement	3.6	3.6	100.0%			
Core	2.2	2.2	100.0%			
Total	20.0	20.0	100.0%			
Not married	15.0	15.0	100.0%	S/M	0.682	0.682
Married	22.0	22.0	100.0%			
Non-parent	17.0	18.0	105.9%	NP/P	0.708	0.750
Parent	24.0	24.0	100.0%			
Female	28.0	28.0	100.0%	F/M	2.545	2.545
Male	11.0	11.0	100.0%			
Not Working	23.0	23.0	100.0%	NW/W	1.438	1.438
Working	16.0	16.0	100.0%			
No Spouse	15.0	15.0	100.0%	NoSp/SpW	0.714	0.714
Not Working	21.0	21.0	100.0%	NoSp/SpNW	0.652	0.652
Working	23.0	23.0	100.0%			
LT BAC	21.0	20.0	95.2%	LT BAC/GTB	1.235	1.053
BAC	20.0	19.0	95.0%	BAC/GTB	1.176	1.000
GT BAC	17.0	19.0	111.8%			
<€10,000	22.0	18.0	81.8%	lt€10k/ge€50k	1.375	1.000
€10,000-19,999	21.0	20.0	95.2%	€10-20k/ge€50k	1.313	1.111
€20,000-29,999	20.0	20.0	100.0%	€20-30k/ge€50k	1.250	1.111
€30,000-49,999	18.0	19.0	105.6%	€30-50k/ge€50k	1.125	1.056
>=€50,000	16.0	18.0	112.5%			

Median Values of Household Production

	EDT1999	Match	Ratio		EDT1999	Match
Care	9.3	9.3	100.0%			
Procurement	0.0	0.0				
Core	0.0	0.0				
Total	15.0	15.0	100.0%			
Not married	11.0	11.0	100.0%	S/M	0.611	0.611
Married	18.0	18.0	100.0%			
Non-parent	13.0	14.0	107.7%	NP/P	0.650	0.700
Parent	20.0	20.0	100.0%			
Female	26.0	26.0	100.0%	F/M	3.714	3.714
Male	7.0	7.0	100.0%			
Not Working	20.0	20.0	100.0%	NW/W	1.667	1.667
Working	12.0	12.0	100.0%			
No Spouse	11.0	11.0	100.0%	NoSp/SpW	0.611	0.611
Not Working	18.0	18.0	100.0%	NoSp/SpNW	0.611	0.611
Working	18.0	18.0	100.0%			
LT BAC	18.0	16.0	88.9%	LT BAC/GTB	1.385	1.143
BAC	15.0	15.0	100.0%	BAC/GTB	1.154	1.071
GT BAC	13.0	14.0	107.7%			
<€10,000	20.0	14.0	70.0%	lt€10k/ge€50k	1.250	0.778
€10,000-19,999	18.0	16.0	88.9%	€10-20k/ge€50k	1.125	0.889
€20,000-29,999	14.0	16.0	114.3%	€20-30k/ge€50k	0.875	0.889
€30,000-49,999	14.0	14.0	100.0%	€30-50k/ge€50k	0.875	0.778
>=€50,000	11.0	13.0	118.2%			

Figures

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

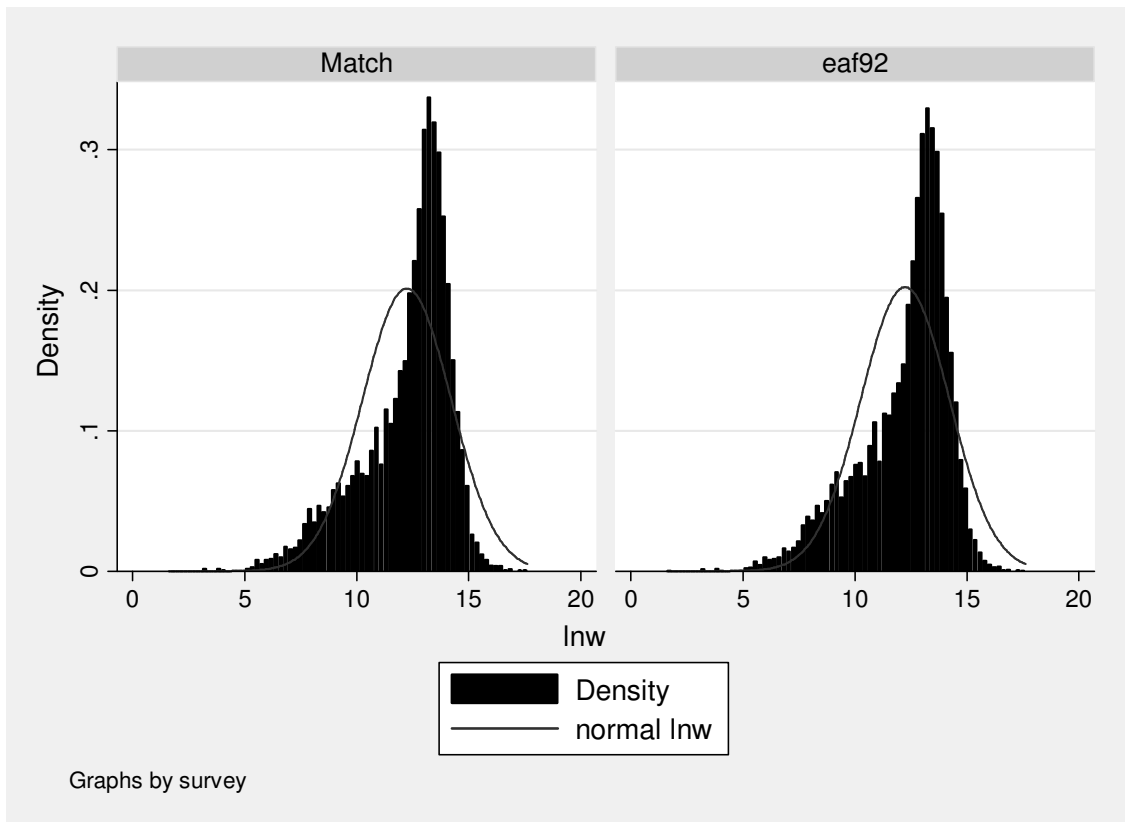


Figure 2 Ratio of Mean Net Worth by Category (Match/EAF 1992)

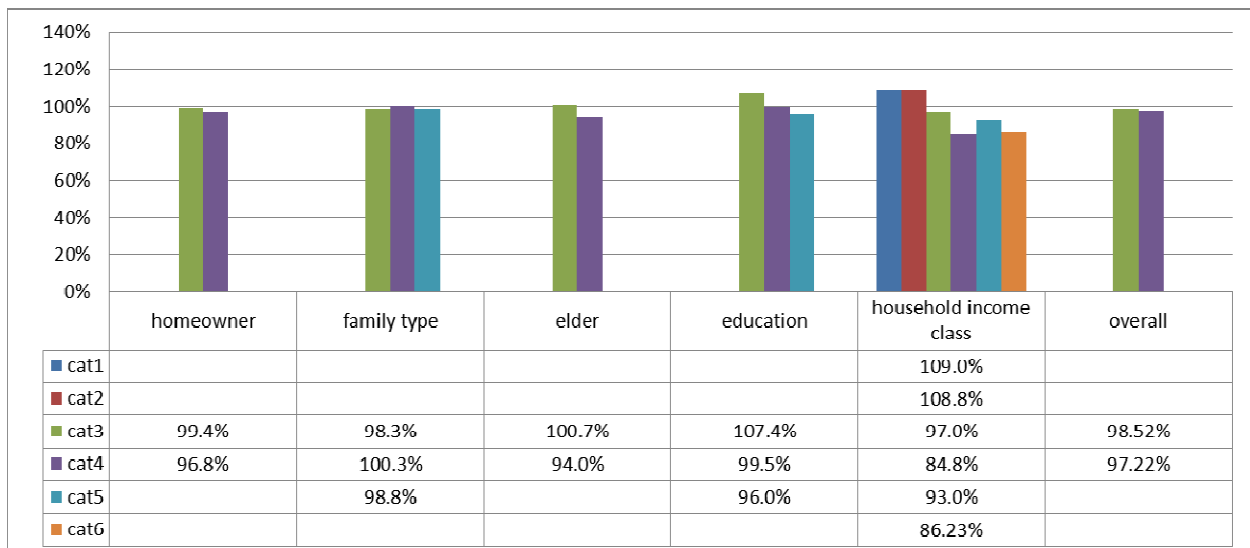


Figure 3 Net Worth by Matching Cells, 1992 EAF and Match File

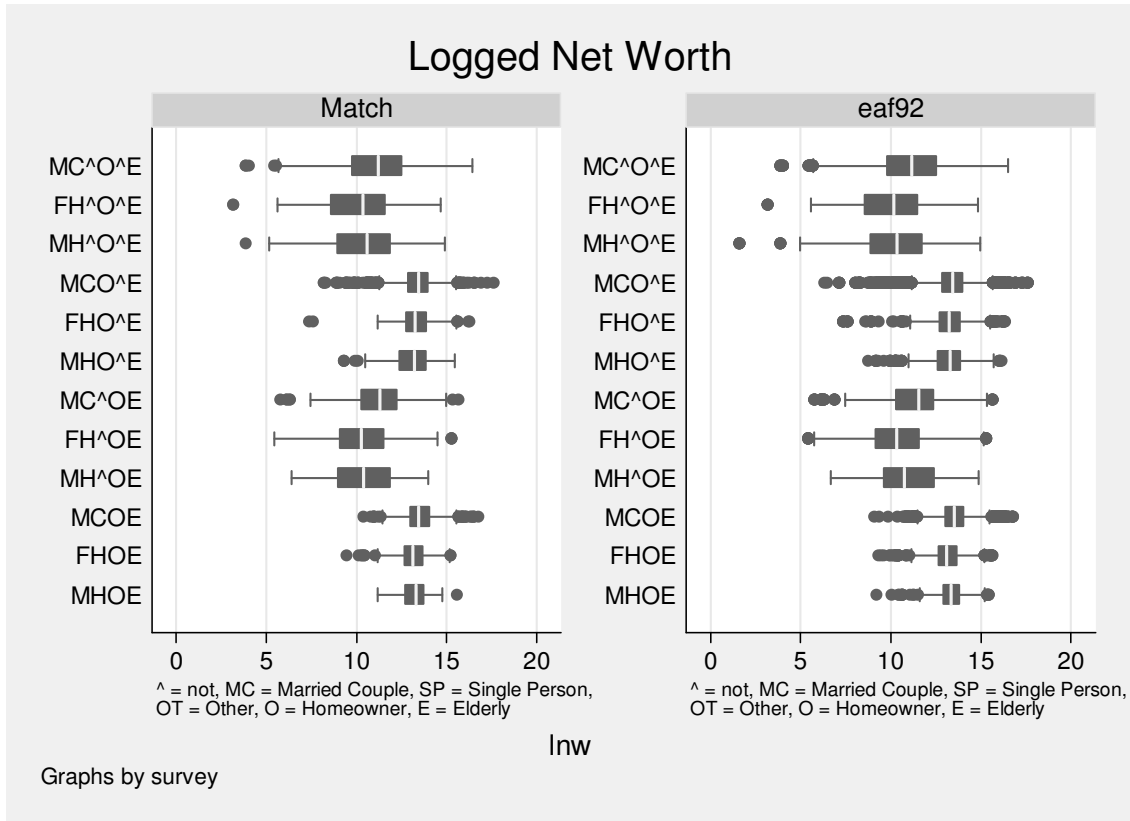


Figure 4 Ratio of Mean Hours of HH Production by Category (Match/EDT 1985)

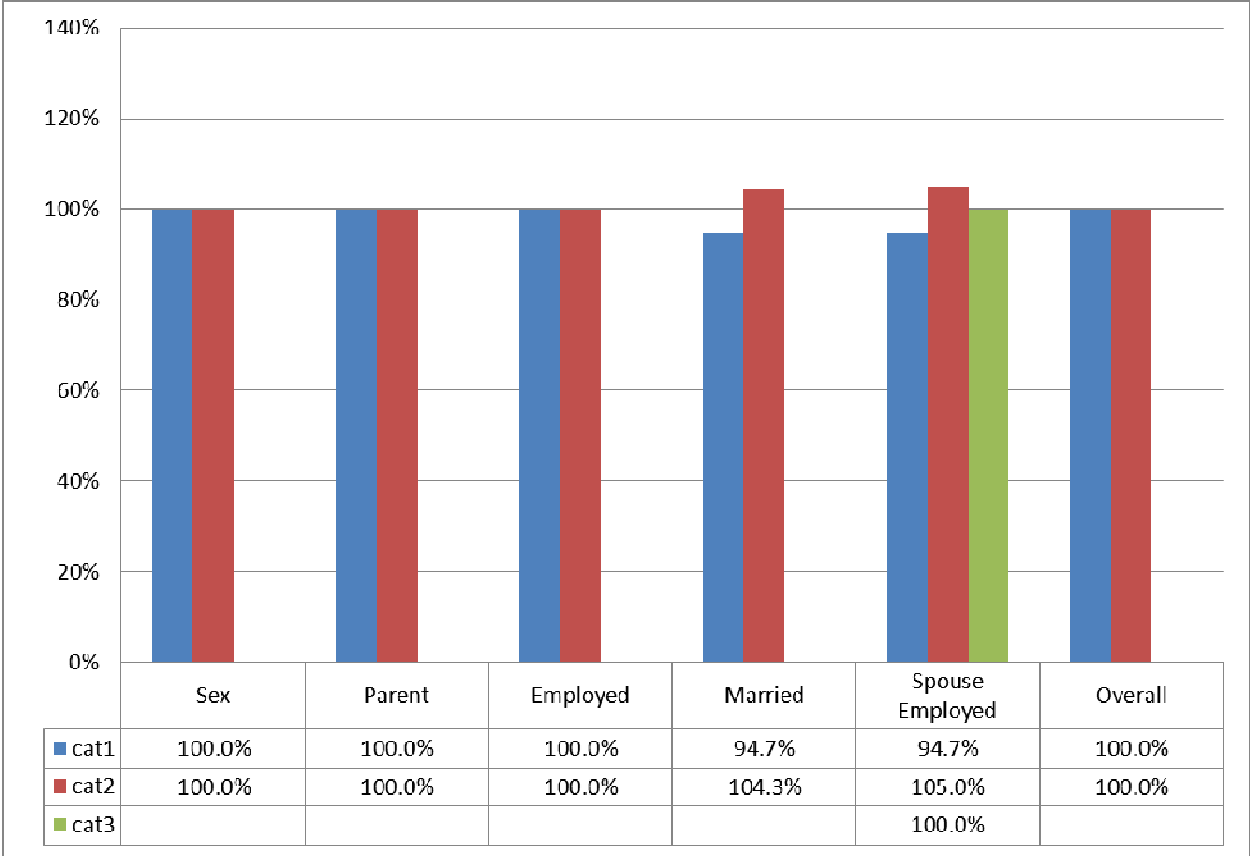


Figure 5 Hours of Household Production by Matching Cells, 1985 EDT and Match File

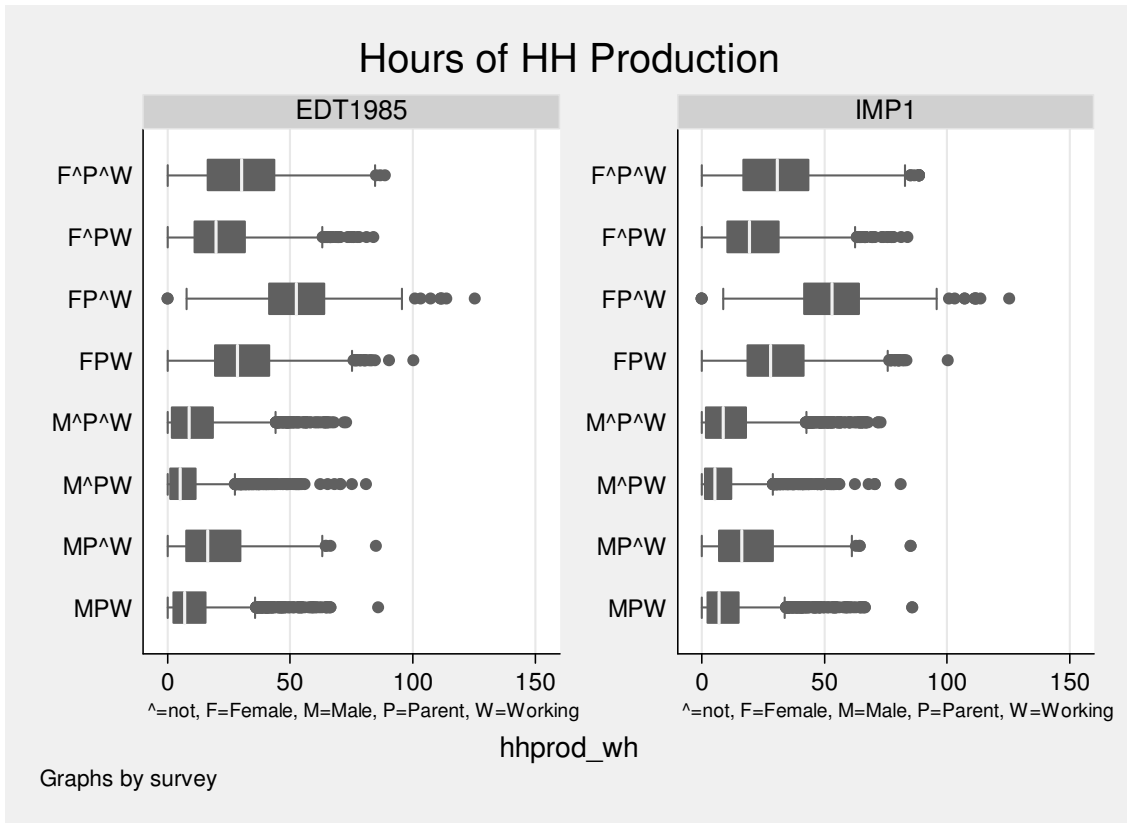


Figure 6 Distribution of Log Net Worth, 2004 Pat and Match File

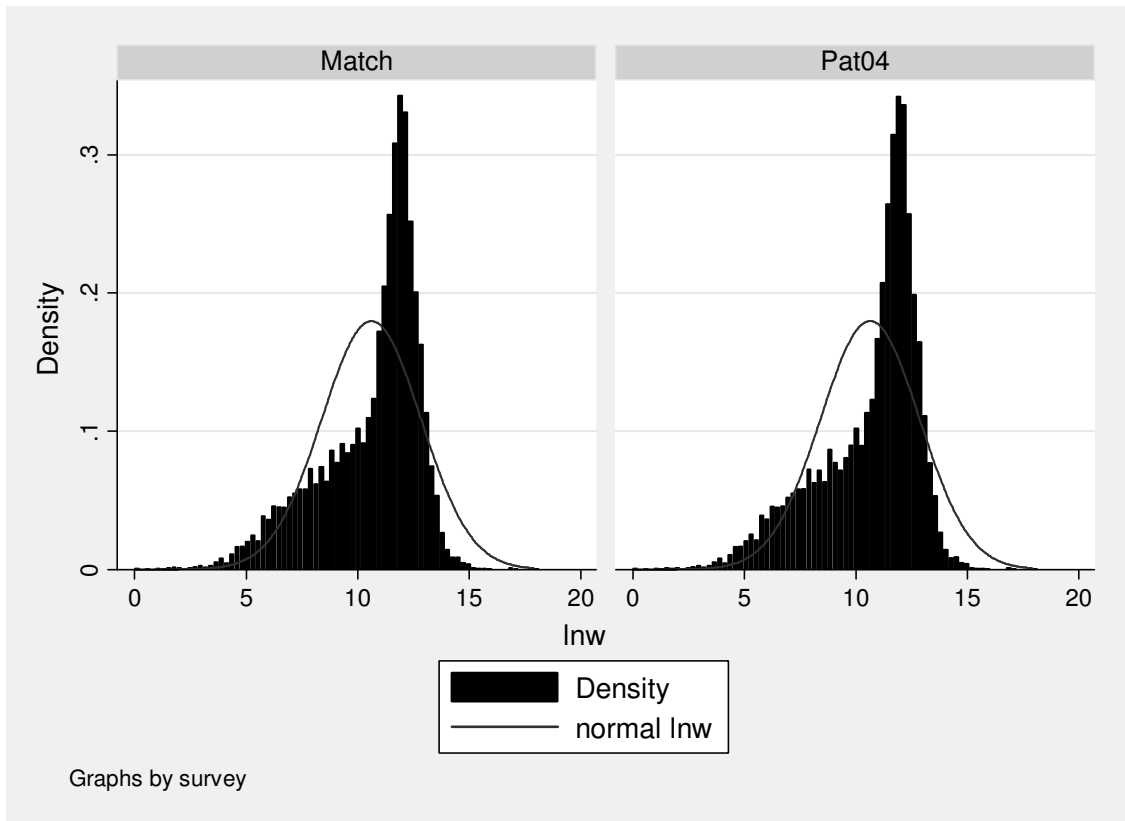


Figure 7 Ratio of Mean Net Worth by Category (Match/PAT 2004)

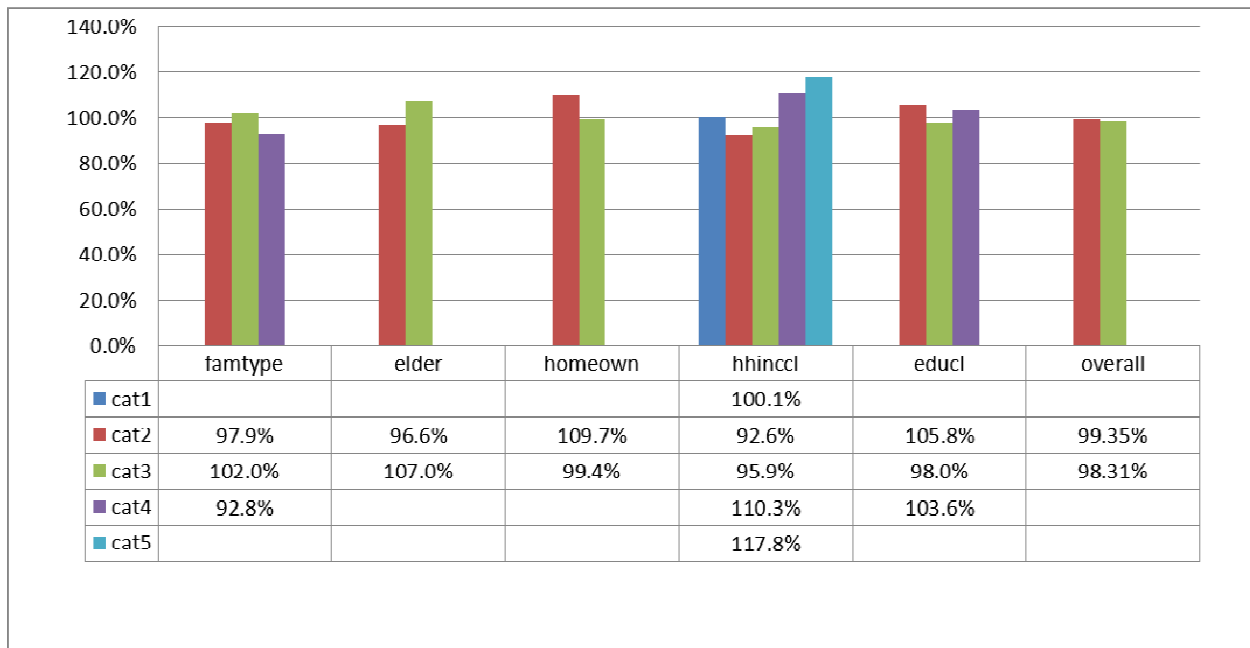


Figure 8 Net Worth by Matching Cells, 2004 Pat and Match File

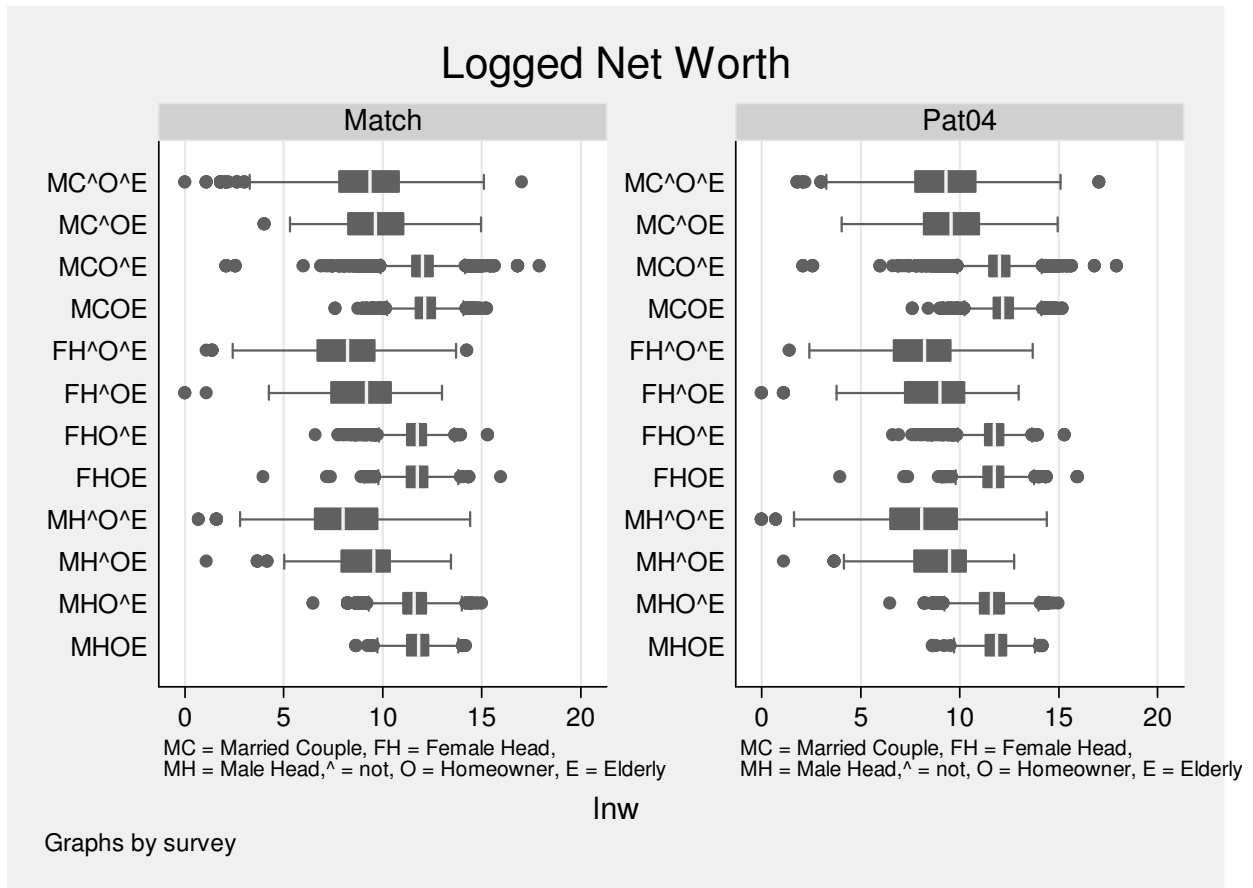


Figure 9 Ratio of Mean Hours of HH Production by Category (Match/EDT 1999)

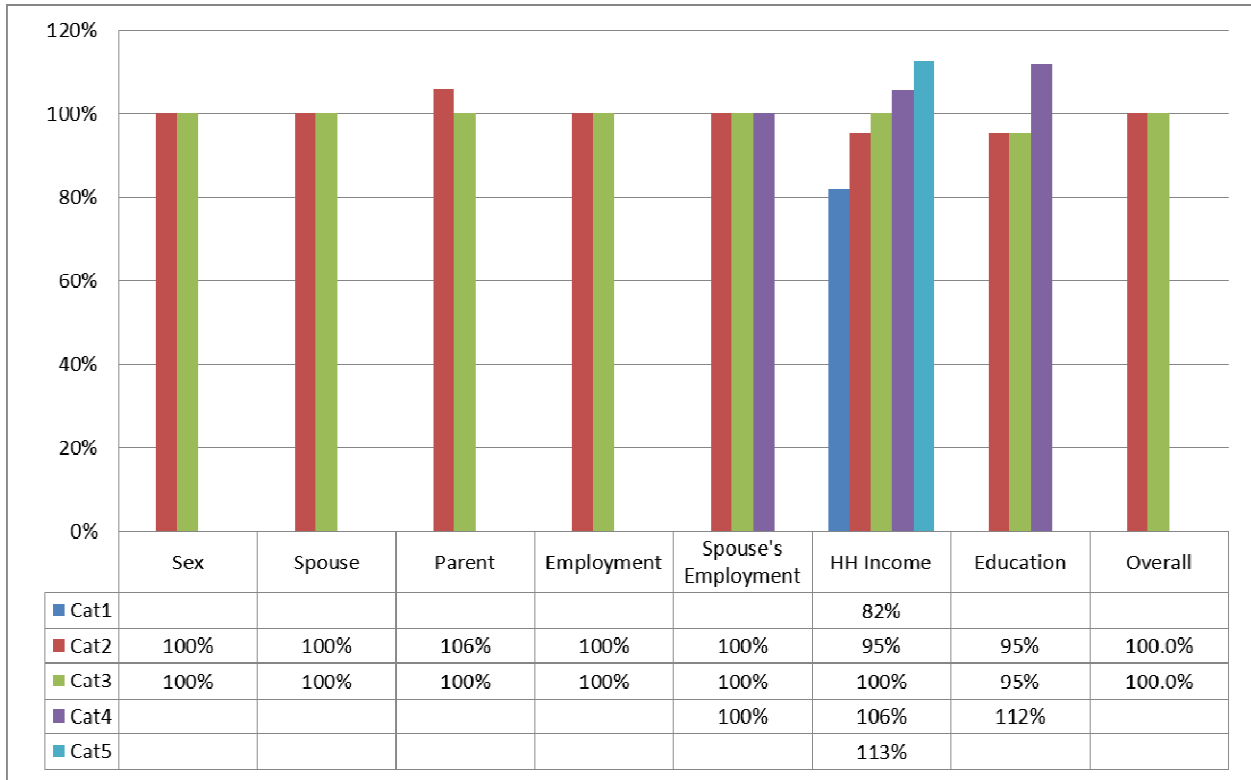


Figure 10 Hours of Household Production by Matching Cells, 1999 EDT and Match File

