Technical Report
Construction and Evaluation of the 2013 Longitudinal Individual and Family Weights

May 26, 2015

Patricia A. Berglund, Steven G. Heeringa

University of Michigan, Ann Arbor, MI

This document describes the construction of the 2013 Core/Immigrant individual and family longitudinal sample weights for the Panel Study of Income Dynamics (PSID). This technical report is organized in four sections. Section I provides an overview of the PSID sample, defines PSID sample and non-sample persons and explains the following rules. Section II outlines the methodology for computing the 2013 longitudinal individual and family weights. The report concludes in Section III with a descriptive analysis of the weights, including comparisons of distributions of U.S. socioeconomic characteristics using weighted estimates from the CPS and PSID.

I. The PSID Sample and Following Strategy in 2013

The 2013 PSID panel is based on the dynamic, longitudinal follow-up of individuals and their families originally identified in a combination of three probability samples of U.S. households: the Survey Research Center 1960 National Sample (SRC), a subsample of families interviewed in 1967 by the Bureau of the Census for the Office of Economic Opportunity (SEO) (McGonagle and Schoeni, 2006) and the 1997 PSID Immigrant Supplement (Heeringa and Connor, 1998). Sample persons and their descendents identified in the baseline SRC and SEO samples (termed the PSID "Core" in many publications) have been interviewed since 1968. In 1997 and 1999 the baseline sample of the post-1968 immigrants was added and these new immigrant sample persons have been followed continuously since the late 90s. More detailed information on the PSID 1968 and 1997/1999 immigrant samples is available from the PSID website, (psid.org).

Under the "dynamic" sample follow-up design, PSID interviewed 9,063 families in 2013. Included in these families are 24,952 individuals: 17,785 PSID "sample persons" (see Table 1) and 7,167 "non-sample" spouses and family members.

PSID traditionally categorizes persons into one of two groups: sample persons and non-sample persons. The definition of these categories has changed slightly over the years. From 1968 to 1993, a sample person was defined as someone who was either an original sample person; i.e., resident of a PSID sample family in 1968, or an offspring born to or adopted by a sample individual who was actively participating in the study at the time. A newborn child had to appear in the study at the wave immediately following their birth to be considered a sample person. In 1994, the definition of a sample person was expanded to include children born to or adopted by a sample person when the sample person was not participating in the study; i.e., the child need not be residing with a responding panel family at birth or adoption. The same current PSID definition of sample persons (implemented in 1994) applies to the immigrant sample.

All of the 9,063 PSID families interviewed in 2013 are members of the Core or the Immigrant samples. In 2013, the rules for following sample persons and interviewing their existing or newly formed families were the same as in the prior 2011 wave[1]. Specifically, sample persons who participated in the previous wave survey were followed. In addition to following sample persons who were respondents in 2011, the PSID attempted to obtain an interview with sample individuals who did not respond in the prior wave (2011 survey year), but responded in the 2009 survey year.

Each sample person successfully interviewed for 2013 receives a positive value for their 2013 longitudinal individual weight. Nonsample persons receive a 2013 PSID individual longitudinal weight equal to zero (0).

II. Methodological Approach to the 2013 PSID Longitudinal Weight Construction

The methodology for the calculation of PSID longitudinal weights follows a four year (two wave) cycle. At the beginning of each cycle, the calculation of weights incorporates an explicit adjustment for panel attrition due to nonresponse that has occurred over the past four years. The current cycle began in 2011 and a full nonresponse adjustment was incorporated in the 2011 longitudinal weights for individuals and families (see Gouskova, et al., 2008 for a description of the longitudinal weight nonresponse adjustment methodology). At the second wave of each four year weight development cycle, a simpler procedure is used to carry forward the individuals' weights from the previous wave and to update the weights for new births, sample panel members who "reappear" and are interviewed again after one or more waves of nonresponse. Family weights are also updated to reflect changes in family composition due to marriage, divorce, death, and other additions of new members. No explicit nonresponse adjustment is incorporated in the weight computations for the second wave of each of the 4-year weight updating cycles.

The 2013 weights are "carry-over" weights. The last attrition adjustment of the PSID longitudinal individual weights was done in 2011 and thus, the construction of the 2013 individual weights starts with the 2011 longitudinal weight as the basis. For sample persons who were interviewed in both 2011 and 2013, the 2013 individual weights were assigned by carrying forward the 2011 longitudinal weight. For sample persons who were interviewed in 2013 but not in 2011, the most recent non-zero individual weight or "reference weight" for the case was carried forward as the 2013 individual weight. All "nonsample" individuals in the panel receive a zero (0) value for their longitudinal weight. The PSID provides an optional cross-sectional weight that is

---

[1] For more detail on the following rules in 1993-2007 survey years see Table 1 in Gouskova et al. (2008), (http://psidonline.isr.umich.edu/)

designed for single wave analysis of all cases in the PSID individual data. The PSID cross-sectional weight is a positive weight for all sample and nonsample members of interviewed PSID families (see Heeringa, et al., 2011).

For sample newborns under 2 years of age in 2013, the 2013 individual longitudinal weight was calculated as the average of head's and spouse's individual weight in 2013. If a PSID sample person moved into a PSID family during the period between 2011 and 2013 and they had no existing reference weight, that sample individual was assigned a new individual weight equal to the average of all positive 2013 individual weights in the family unit.

Once individual longitudinal weights had been constructed for each sample person interviewed in 2013, the 2013 longitudinal family weight was computed as the average of the positive individual weights for sample persons and the zero-value weights for the nonsample persons in the family. For example, consider a 2013 PSID family that consisted of a young married couple in which the female spouse was a PSID sample person and had an individual longitudinal weight of 60. Her new spouse was PSID "nonsample" and therefore is assigned a "0" value for his longitudinal individual weight. The 2013 family weight for this two-person family is (60+0)/2=30. Figure 1 is a simple schematic that illustrates the dynamic process of family level weighting for four waves of data collection. At the baseline wave, families "A" and "B" are chosen to the sample. Families "C" and "D" were eligible for probability sample selection at baseline but were not chosen. Over the next three waves, there are "split-offs" from sample families, marriage or new family formation by members of original sample and non-sample family members. At each wave's change in family composition the family weights, $W_{t, FAM}$, are recomputed as the average of the current individual weights for the sample and nonsample persons that comprise the family unit. Note from the final column that the sum of all family weights constructed in this fashion remains consistent with the total number of all family units in the hypothetical dynamic population.

Unlike the 2013 PSID individual weight which is available in both the longitudinal form (sample persons only) and cross-sectional analysis form (sample and nonsample persons have non-zero weights), there is only one version of the 2013 PSID family weight. The longitudinal family weight can be used for cross-sectional analysis of PSID family data.

III. 2013 PSID Individual Respondents: Transition from 2011 Status

As indicated above, the 2013 PSID completed interviews with 24,952 individuals. The columns of Table 1 show the 2013 status of each individual respondent by the sample/nonsample status and selected PSID special classifications for persons in these two major subpopulations. In this table, all respondents from 2013 are cross-classified against 2011 status using 6 sample status categories. The top row of the body of the table contains information about those that responded in 2013 but were classified as "non-sample persons, not part of the elderly group" in 2011. Of these, the key sample status categories are highlighted in dark shading and represent 2013 sample person members: original sample from 1968, born-in sample person consisting of newborns born in 2011, 2012, or 2013 and others born into a sample family, and those that moved into a sample family during this two year period.

The "born in" sample person group (751 people) consists mainly of newborns with 746 of the 751 assigned a 2013 weight as the average of the head and wife's weight. The remaining individuals in this group were assigned the average of all positive 2013 individual weights within the family unit. Of the "move in" sample group, 27 of the 105 were newborns and 78 were non-newborn individuals who received the average of the family unit 2013 weights. All other sample persons who were interviewed in both years, (highlighted in light shading) were assigned the "carry-forward" value of their 2011 individual longitudinal weight.

Family longitudinal weights were constructed as the average of all non-zero individual weights in the family unit during 2013. Note that the family units do change from year to year. See the PSID family level data set documentation and codebooks for more information. See also Duncan and Hill (1985) for a discussion of the issues involved in longitudinal analysis of family units.

IV. Descriptive Statistics for the 2013 PSID Longitudinal Weights

Tables 2 through 6 provide descriptive information on the 2013 PSID Longitudinal weights. To enable comparison of the longitudinal weights across years, the same set of descriptors is reported for the longitudinal weights from the six prior waves (2001-2011).

Tables 2 and 3 summarize the total number of cases with positive, zero, and missing values for individual and family weights and the total numbers of sample and non-sample individuals (families with and without sample members). For individual weights, the number of weights with a positive value is equal to the number of sample persons, and the number of the zero-valued individual weights is the same as the number of non-sample

5

persons (Table 2). As with the 2011 survey, in 2013 all families had at least one sample member (Table 3). As a result all PSID families in 2013 carry a non-zero, positive longitudinal family weight.

Tables 4 and 5 report summary statistics for the longitudinal individual and family weights.   Based on the summary statistics, the distributions of the 2013 longitudinal weights are similar to those in the six most recent survey waves. Across years, the measures of dispersion indicate that there is an increasing trend in variability of the distribution in the individual and family weights. This year over year increase in the variability of the PSID longitudinal weights can be attributed to the periodic nonresponse adjustment (every four years) and for family weights, the reweighting that is required to reflect changes in family composition (e.g. new family formations).

Table 6 provides a key to the PSID variables names for longitudinal individual and family weight variables.

 IV. Evaluation of the PSID Cross-sectional Weights: Comparisons with the CPS.

 Tables 7 through 9 compare PSID and CPS weighted estimates for selected demographic statistics based on characteristics including age, gender and race of household head. Each table reports the unweighted PSID estimates, PSID estimates weighted (as applicable) by the PSID family or individual longitudinal weight and the CPS weighted estimate.   For age (Table 7) and race (Table 9), the first panel of the table compares weighted estimates for family (household) head and the second panel of the table provides estimates of mean or percent values for individuals. The statistics in the right most columns of each table are simple ratios of the weighted PSID and CPS estimates.   These tables are useful for examining three features of the PSID data: consistency of un-weighted and weighted estimates across years, the effect of the longitudinal weights on the distributions of estimates of family and individual population characteristics, and, finally, the consistency of the PSID weighted estimates with those obtained from the CPS data[2].

The tables show that consistency across years of the weighted distributions is comparable to the consistency of the unweighted distributions. Comparison of the unweighted and weighted PSID distributions with the CPS distributions reveals that in a majority of cases the weighted estimates are closer to CPS estimates than are the

---

[2] Note, that some characteristics are not strictly comparable between the two surveys. For example, in the PSID, race is not asked of all individuals while in the CPS data all individuals are asked to provide detailed race information. To calculate proportions of black and non-black individuals in the PSID data, individual race was approximated using the race of the family head.

estimates obtained without weights. This is to be expected since due to the 1968 SEO oversample, the baseline inclusion probabilities for African American and lower income PSID families and individuals were substantially greater than for other domains of the U.S. household population.

With some noticeable difference in the weighted distribution by race of household head and race of individuals the weighted PSID and CPS estimates align fairly closely for these three demographic characteristics. However, caution is advised in placing too much emphasis on minor differences between the PSID and CPS weighted distribution.   Analysts should keep in mind that for any given wave, the simple comparison of weighted demographic distributions does not explicitly take into account PSID non-coverage of immigrant populations after 1997. Immigrants arriving after 1997 when the immigrant sample was added to the PSID are not fully represented in the PSID. Another limitation of this comparison is that the CPS does not cover the institutionalized population while PSID due to the dynamic nature of the sample may include institutionalized persons. There are differences in the definitions that PSID and CPS use to code household composition and disaggregate households into family and non-family units.   Finally, the PSID longitudinal weights for families and individual do not include any recent adjustment to external population controls (e.g. 2000 Census or annual CPS or American Community Survey (ACS) population totals).   The question of whether to introduce explicit post-stratification controls to the PSID weights is a topic of research for the PSID weight development program.

V. References

Duncan, G.J. and Hill, M.S. (1985). "Conceptions of longitudinal households: Fertile or futile?", *Journal of Social and Economic Measurement, 13,* 361-375.

Ernst, L.R. (1989). "Weighting issues for longitudinal household and family estimates." In *Panel Surveys* (Eds. D. Kaspryzk, G. Duncan, G. Kalton and M.P. Singh). New York: John Wiley, 139-159.

Gouskova E., S. Heeringa, K. McGonagle, R. Schoeni, and F. Stafford, 2008. "Panel Study of Income Dynamics Revised Longitudinal Weights 2007", PSID Technical Report #08-05, ISR, University of Michigan. http://psidonline.isr.umich.edu/Publications/Papers/tsp/2008-05_PSID_Revised_Longitudinal_Weights_1993-2005%20.pdf

Heeringa, S.G. and Connor, J.H. (1997). "Technical documentation for the 1997 PSID Sample". Panel Study of Income Dynamics Technical Report. Survey Research Center, University of Michigan, Ann Arbor.

Heeringa, S.G. and Connor J.H. (1998). "Technical documentation for the 1997 PSID Immigrant Supplement". Panel Study of Income Dynamics Technical Report. Survey Research Center, University of Michigan, Ann Arbor.

Heeringa, S.G., Berglund, P.A., McGonagle, K., Schoeni, R., (2013) Panel Study of Income Dynamics Construction and Evaluation of the Longitudinal Sample Weights 2011. Survey Research Center, University of Michigan, Ann Arbor.

Hill, M.S. (1992). *The Panel Study of Income Dynamics: A User's Guide.* Newbury Park, CA: Sage Publications.

Huang, H. (1984). "Obtaining cross-sectional estimates from a longitudinal survey: Experiences of the Income Survey Development Program." , *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 670-675.

Kalton, G. (1987). "Including nonsample persons in PSID analyses". Panel Study of Income Dynamics Working Paper, Ann Arbor: University of Michigan.

Kalton, G. and Brick, J.M. (1995). "Weighting Schemes for Household Panel Surveys", *Survey Methodology, Vol 21, No. 1,* pp. 33-34, Statistics Canada.

Kasprzyk, D. (1988). *The Survey of Income and Program Participation: An Overview and Discussion of Research Issues.* SIPP Working Paper No. 8830. Washington, D.C.: U.S. Bureau of the Census.

Lavallee, P. (1995). "Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method. *Survey Methodology*, 21 25-32.

Lavallee, P., Michaud, S. and Webber, M. (1993). "The Survey of Labour and Income Dynamics, design issues for a new longitudinal survey in Canada. *Bulletin of the International Statistical Institute,* 49th Session, Contributed Papers, Book 2, 99-100.

Little, R.J.A. (1989). "Sampling weights in the PSID: Issues and comments. Panel Study of Income Dynamics Working Paper, Ann Arbor: University of Michigan.

Lynn, P., Buck, N. Burton, J., Laurie, H, Uhrig, S.C.N. (2006). *Quality Profile: British Household Panel Survey, Version 2: Waves 1-13:1991-2003. Essex: University of Essex, Institute for Social and Economic Research.*

McGonagle, K. and Schoeni, R. (2006). "The Panel Study of Income Dynamics: Overview and Summary of Scientific Contributions After Nearly 40 Years." Panel Study of Income Dynamics Technical Paper Series. Available at: http://psidonline.isr.umich.edu/Publications/Papers/tsp/2006-01_PSID_Overview_and_summary_40_years.pdf.

Table 1. Table of 2011 Sample Status for 2013 PSID Individual Respondents

Sample status counts in 2011, among 2013 responders	Sample Status Counts in 2013						
	"Non-sample persons", not part of elderly group	Original "sample persons"	Born in "sample persons"	Move in "sample persons"	Followable "non-sample parents"	"Non-sample persons", part of elderly group	Total
"Non-sample persons", not part of elderly group	5182	0	751	105	0	0	6038
Original "sample persons"	0	5350	0	0	0	0	5350
Born in "sample persons"	2	0	10366	2	0	0	10370
Move in "sample persons"	2	0	1	1210	0	0	1213
Followable "non-sample parents"	0	0	0	0	1961	0	1961
"Non-sample persons", part of elderly group	0	0	0	0	0	20	20
Total	5186	5350	11118	1317	1961	20	24952

Table 2. PSID Longitudinal Individual Weights, 2001-2013

Year	Core sample (SRC, SEO) and Immigrant sample					
	Total number of individuals in the study	Total number of "sample persons"	Total number of "non-sample persons"	Number of cases with positive individual weight	Number of cases with zero individual weight	Number of cases with missing individual weight
2001	21400	15646	5754	15646	5754	0
2003	22290	16012	6278	16012	6278	0
2005	22918	16620	6298	16620	6298	0
2007	23508	16906	6602	16906	6602	0
2009	24385	17471	6814	17471	6814	0
2011	24661	17643	7018	17643	7018	0
2013	24952	17785	7167	17785	7167	0

Table 3. PSID Longitudinal Family Weights, 2001-2013

Year	Core sample (SRC, SEO) and Immigrant sample				
	Total number of families in the study	Number of families with no "sample person"	Number of families with positive weight	Number of families with zero weight	Number of families with missing weight
2001	7406	211	7195	211	0
2003	7822	257	7565	257	0
2005	8002	0	8002	0	0
2007	8289	0	8289	0	0
2009	8690	0	8690	0	0
2011	8907	0	8907	0	0
2013	9063	0	9063	0	0

Table 4. Summary Statistics of the PSID Longitudinal Individual Weights, 2001-2013 (Sample Persons Only)

Year	N	Mean	Standard Deviation	Min	Max	Coefficient of Variation
2001	15646	25.07	18.97	0.25	167.68	0.76
2003	16012	25.62	19.54	0.25	173.56	0.76
2005	16620	24.81	19.33	0.23	173.56	0.78
2007	16906	25.38	20.09	0.20	181.45	0.79
2009	17471	24.57	19.90	0.23	181.45	0.81
2011	17643	25.65	21.47	0.25	196.44	0.84
2013	17785	24.75	21.11	0.25	196.44	0.85

Table 5. Summary Statistics for the PSID Longitudinal Family Weights, 2001-2013 (With 2001 and 2003 Based on Families with Positive Weights Only)

Year	N	Mean	Standard Deviation	Min	Max	Coefficient of Variation
2001	7195	22.03	16.74	0.06	167.68	0.76
2003	7565	22.06	17.06	0.12	132.64	0.77
2005	8002	21.04	16.82	0.12	136.03	0.80
2007	8289	21.32	17.40	0.10	139.34	0.82
2009	8690	20.66	17.28	0.10	139.34	0.84
2011	9807	21.71	18.75	0.12	150.89	0.87
2013	9063	20.85	18.44	0.08	150.89	0.89

Table 6. Names of the PSID Longitudinal Weight Variables, 1993-2013

Year	Core Longitudinal Weight	
	Individual	Family
1993	ER30864	V23361
1994	ER33119	ER4160
1995	ER33275	ER7000
1996	ER33318	ER9251
	Core/Immigrant Longitudinal Weight	
	Individual	Family
1997	ER33430	ER12084
1999	ER33546	ER16518
2001	ER33637	ER20394
2003	ER33740	ER24179
2005	ER33848	ER28078
2007	ER33950	ER41069
2009	ER34045	ER47014
2011	ER34154	ER52436
2013	ER34268	ER58257

Table 7. Comparison of PSID and CPS Weighted Estimates of Mean and Median Age, 2001-2013

A. Family Level Data (age of head)								
	PSID unweighted		PSID weighted		CPS weighted		Ratio	
Year	Mean years [1]	Median years [2]	Mean years [3]	Median years [4]	Mean years [5]	Median years [6]	Mean [3]/[5]	Median [4]/[6]
2001	44.91	43.00	49.39	47.00	48.72	46.00	1.01	1.02
2003	44.98	43.00	49.60	48.00	48.69	47.00	1.02	1.02
2005	45.08	44.00	49.96	48.00	49.04	47.00	1.02	1.02
2007	45.04	44.00	50.13	49.00	49.30	48.00	1.02	1.02
2009	45.79	44.00	49.82	49.00	47.60	47.00	1.05	1.04
2011	45.21	43.00	50.60	50.00	48.11	47.00	1.05	1.06
2013	45.68	43.00	51.21	51.00	48.56	48.00	1.05	1.06

B. Individual Level Data								
	PSID unweighted		PSID weighted		CPS weighted		Ratio	
Year	Mean years [1]	Median years [2]	Mean years [3]	Median years [4]	Mean years [5]	Median years [6]	Mean [3]/[5]	Median [4]/[6]
2001	30.86	29.00	36.30	36.00	35.65	35.00	1.02	1.03
2003	31.25	29.00	36.53	36.00	35.82	35.00	1.02	1.03
2005	31.41	29.00	36.93	36.00	36.17	36.00	1.02	1.00
2007	31.61	29.00	37.35	37.00	36.44	36.00	1.02	1.03
2009	32.30	29.00	37.90	37.00	36.80	36.00	1.03	1.03
2011	31.95	29.00	38.75	38.00	37.00	36.00	1.05	1.06
2013	32.91	30.00	39.27	38.00	37.64	37.00	1.04	1.03

Table 8. Comparison of PSID and CPS Weighted Estimates of % Population by Gender, 2001-2013
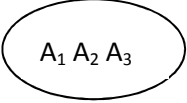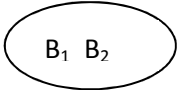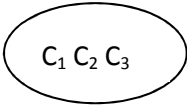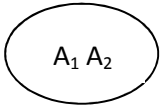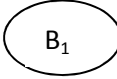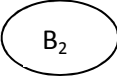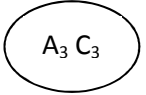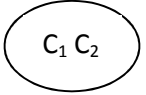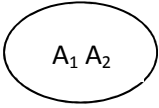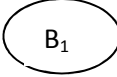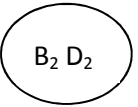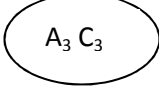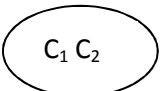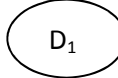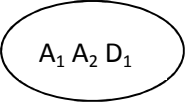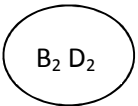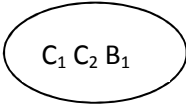
Year	PSID unweighted		PSID weighted		CPS weighted		Ratio	
	Male [1]	Female [2]	Male [3]	Female [4]	Male [5]	Female [6]	Male [3]/[5]	Female [4]/[6]
2001	47.93	52.07	48.08	51.92	48.86	51.14	0.98	1.02
2003	47.98	52.02	48.17	51.83	48.92	51.08	0.98	1.01
2005	47.88	52.12	48.23	51.77	49.03	50.97	0.98	1.02
2007	47.88	52.12	48.58	51.42	49.08	50.92	0.99	1.01
2009	47.48	52.52	48.40	51.60	49.10	50.90	0.99	1.01
2011	47.87	52.13	48.74	51.26	49.21	50.79	0.99	1.01
2013	47.69	52.31	48.83	51.17	48.96	51.04	1.00	1.00

Table 9. Comparison of PSID and CPS Weighted Estimates of %Population by Race, 2001-2013

A. Family Level Data (race of head)

Year	PSID unweighted		PSID weighted		CPS weighted		Ratio	
	Non-black [1]	Black [2]	Non-black [3]	Black [4]	Non-black [5]	Black [6]	Non-black [3]/[5]	Black [4]/[6]
2001	69.60	30.40	87.40	12.60	87.80	12.20	1.00	1.03
2003	68.40	31.60	87.20	12.80	87.90	12.10	0.99	1.06
2005	66.70	33.30	86.10	13.90	87.80	12.20	0.98	1.14
2007	65.70	34.30	85.90	14.10	87.60	12.40	0.98	1.14
2009	64.60	35.40	84.40	15.60	87.50	12.50	0.96	1.25
2011	62.93	37.07	85.18	14.82	87.35	12.65	0.98	1.17
2013	61.84	38.16	83.54	16.46	86.97	13.03	0.96	1.26

B. Individual Level Data (individual race is proxied by the race of head in PSID data)

Year	PSID unweighted		PSID weighted		CPS weighted		Ratio	
	Non-black [1]	Black [2]	Non-black [3]	Black [4]	Non-black [5]	Black [6]	Non-black [3]/[5]	Black [4]/[6]
2001	67.00	33.00	86.90	13.10	87.30	12.70	1.00	1.03
2003	66.10	33.90	86.60	13.40	87.50	12.50	0.99	1.07
2005	64.60	35.40	86.00	14.00	87.40	12.60	0.98	1.11
2007	64.20	35.80	85.90	14.10	87.40	12.60	0.98	1.12
2009	63.70	36.30	85.20	14.80	86.70	13.30	0.98	1.11
2011	63.35	36.65	84.19	15.81	86.43	13.57	0.97	1.17
2013	61.88	38.12	84.79	15.21	85.95	14.05	0.99	1.08

Figure 1: Illustration of Dynamic Weighting for PSID Families

Wave		Total
t_0	$A_1 A_2 A_3$ B_1 B_2 $C_1 C_2 C_3$ $D_1 D_2$	
$W_{0, IND}$	2, 2, 2 2, 2 0, 0, 0 0, 0	10
$W_{0,FAM}$	2 2 0 0	4
t_1	$A_1 A_2$ B_1 B_2 $A_3 C_3$ $C_1 C_2$ D_1 D_2	
$W_{1,FAM}$	2 2 2 1 0 0 0	7
t_2	$A_1 A_2$ B_1 $B_2 D_2$ $A_3 C_3$ $C_1 C_2$ D_1	
$W_{2,FAM}$	2 2 1 1 0 0	6
t_3	$A_1 A_2 D_1$ $B_2 D_2$ $A_3 C_3$ $C_1 C_2 B_1$	
$W_{3,FAM}$	1.33 1 1 0.67	4

15