NHATS Technical Paper #42

National Health and Aging Trends Study
Five Hot Deck Income Imputations:

Rounds 1 - 11

April 2024

Suggested citation: Hu, Mengyao, Vicki A. Freedman & Rui Jiao. 2024. National Health and Aging Trends Study Five Hot Deck Income Imputations: Rounds 1-11. NHATS Technical Paper #42. Baltimore: Johns Hopkins Bloomberg School of Public Health. Available at www.nhats.org. This technical paper was prepared with funding from the National Institute on Aging (U01AG032947).

Contents

Overview

The National Health and Aging Trends Study (NHATS) asks participants to report their total income for the prior calendar year in Rounds 1, 3, 5, 7, 9 and 11 forward. For those who do not report an exact value for total annual income, information is collected in broad categories. For single Sample Persons, individual income is reported; for those with a spouse or partner, couple income is reported.

In order to facilitate the use of these income data by the research community, we have developed multiple total income variables that include imputed values for those with missing income information. Through Round 11, NHATS included 5 income imputations on the Sample Person (SP) release files. These initial imputations were prepared using a hot deck approach. Beginning in Round 12, and for all prior rounds, NHATS expanded the number of imputations provided to 20 to improve standard error estimation. These expanded imputations are provided in a set of auxiliary files and were prepared using an interval regression methodology.

This technical paper describes the initial income imputation methodology using a hot deck approach. It draws upon previously released round-specific technical papers (Montaquila et al. 2012, 2015; DeMatteis et al. 2016, 2018, 2020; Jiao et al. 2022) now available in the document archive section of the NHATS website. A separate technical paper describes the methodology for the more expanded set of imputations using interval regression (Hu & Freedman 2024).

Although this technical paper demonstrates that percentile distributions for the two sets of estimates are close, means and standard errors are lower for estimates based upon the updated methodology. We therefore recommend that researchers use the auxiliary set of 20 imputations.

Income Imputation Variables on the SP Files

Table 1 summarizes income imputation variables on the SP Files through Round 11, including the original (reported) values of total income, 5 imputed values (actual and imputed incorporated into each variable, except in Rounds 3 and 5), and a flag variable indicating which cases have imputed (vs. reported) information. In addition, actual and imputed categorical variables are provided for respondents with and without a spouse/partner, along with a flag indicating which cases have imputed values.

Table 1. Total Income Imputation Variables included in the NHATS SP Files (through Round 11)

Variable name	Label	Description
Total Income Value		
Ia#totinc	R# IA50 TOTAL INCOME	Actual reported amounts
ia#toincimf	R# F IMPUTED TOTAL INC FLG	Flag indicating imputation
ia#toincim1-5	R# IA50 IMPUTED TOTAL INC1-INC5	Imputed values 1-5 for missing amounts and reported bracket amounts[1]
Total Income Range – Respondents who have spouse/partner		
ia#toincesjt	R# IA51A JOINT EST TOT INCOME	Actual reported amounts

ia#eincimjf	R# F IMPTD JOINT EST TOT INC FLG	Flag indicating imputation
ia#eincimj1-5	R# IA51A IMP EST JOINT TOT INC1-INC5	Imputed values 1-5 for missing amounts and reported bracket amounts[1]
Total Income Range – Respondents who are single		
ia#toincessg	R# IA51B SNGLE EST TOT INC	Actual reported amounts
ia#eincimsf	R# F IMPUTED SGL EST TOT INC FLG	Flag indicating imputation
ia#eincims1-5	R# IA51B IMP EST SP SGL TOT INC1-INC5	Imputed values 1-5 for missing amounts and reported bracket amounts[1]

[1]Actual reported and imputed amounts are incorporated into the imputation variable in all Rounds, except 3 and 5. For Rounds 3 and 5, researchers must combine actual reported amounts with the imputed values.

Using Imputed Income in Analyses

We recommend that researchers use the auxiliary set of 20 imputations, since they yield more reliable estimates with narrower standard errors than the 5 hot deck imputations previously provided.

Extent of Missing Data for Total Income

Table 2 shows the percentage of participants that require imputation and the reason through Round 11. The percentage of respondents who did not require imputation (i.e., provided an exact value of total income) increased across rounds from 55.8% in Round 1 to 68.8% in Round 11. A categorical response (requiring imputation within a bracket) was reported by 13.6%-21.7%, depending on the round and 13.1%-30.6% were missing both an exact valued and a categorical response (requiring full imputation).

Table 2. Percentage of NHATS Participants Requiring Income Imputation and Reason

	n*	% Not imputed: exact value reported	% Imputed: bracket response only	% Imputed: missing exact value and bracket response
Round 11	3,388	68.8	17.2	14.0
Round 9	4,460	69.0	17.9	13.1
Round 7	5,566	65.1	19.5	15.4
Round 5	7,576	55.9	19.4	24.7
Round 3	4,884	59.2	21.7	19.2
Round 1	7,609	55.8	13.6	30.6
*Includes those who completed a living SP interview				

Imputation Approach

Westat's AutoImpute software was used to impute five values of the total income items. AutoImpute uses a cyclical n-partition hot deck (an approach analogous to the Gibbs sampler but using the hot deck to generate the imputations; see Judkins 1997; Judkins et al. 2007; Judkins, Piesse, and Krenzke 2008; Krenzke and Judkins 2008). This software is designed to facilitate preservation of multivariate distributions while also ensuring that imputations

maintain skip patterns and adhere to constraints. In this application an example of a constraint is ensuring imputations for specific amounts fall within reported (or imputed) bracket ranges.

The cyclical n-partition hot deck procedure initially imputes all target variables in sequence (i.e., covariates used in the imputation models and items requiring imputation) using a simple hot deck that uses specified auxiliary variables and skip controllers. Using the initial imputed variables, a model is fit for each target variable using simple forward stepwise regression selection in an iterative process until the convergence criterion is met. The predicted values of the target variable from the final model are used to generate imputed values by using predictive mean matching for ordinal (ordered categorical) target variables and clustering for unordered categorical target variables. Predictive mean matching uses a hot deck with the skip controllers as hard boundaries and the predicted values from the final model as soft boundaries. For unordered categorical target variables, a k-means clustering algorithm is used on the vector of predicted values for each level, and then a hot deck is used to impute the target variable with the skip controllers as hard boundaries and the cluster membership indicators as soft boundaries. For more details on the procedure see Judkins et al. (2007). Total income value was also constrained to fall within the reported/imputed bracket amount. The whole procedure is performed 5 times independently. In this technical paper, we refer this approach as the Hot Deck approach.

Imputation Groups

The hot deck approach is applied uniformly to all cases with missing total income.

Variables Used in Imputation

Here we provide an overview of covariates included in the imputation models. See the Appendix Table for details by round.

For Rounds 1 and 5, which included a more detailed IA section than the other rounds, three classes of variables were used in the imputation:

1. Source variables that indicate (yes/no) whether the respondent (and his/her spouse/partner, if applicable) has the particular source of income (referred to below as "fencepost" variables). Sources include Social Security and/or Railroad Retirement payments; Supplemental Security Income; Veteran's Administration payments; and whether they worked for pay last week or absent from work last week, and whether SP and spouse worked for pay last month; pension plans that were a job-related or union benefit; other retirement plans such as 401K plans, 403B plans or IRAs; mutual funds or stocks; any bonds; checking account; savings account; certificates of deposit (CDs); and whether they own a business, a farm, or any other real estate besides their home.
2. Source amount variables that indicate the amount of income from the particular source. Amounts include most recent monthly Social Security or Railroad Retirement payment, Supplemental Security Income payment, Veteran's Administration payments, and job-related pension plans payments; current amounts in retirement accounts; amounts

received or withdrawn from all retirement accounts last month and last year; amounts for mutual funds or stocks and government, corporate, or other bonds that are not part of retirement accounts; total amounts for checking accounts, savings or money market accounts, and certificates of deposit or CDs; total interest and dividend income from mutual funds or stocks, government, corporate, or other bonds, and bank accounts or CDs last year; worth value of owned business, farm, or real estate; income from these businesses or property before tax; amount from any work for SPs and how much earned by SP's spouse/partner (if applicable) last month. All amounts are collected for SP and spouse/partner of SP if applicable.

3. Auxiliary variables that included respondent characteristics (e.g., age, race/ethnicity, gender, educational attainment, marital status, home ownership, ever work for pay, veteran's status, labor force status, spouse/partner's labor force status, whether live in facility, whether live with spouse/partner, occupation for most of life and whether still doing this work, interviewer observations about the home condition), whether more than one job last week and hours worked last week, payment schedule and amount from main job, payment method and month and year payment started for Social Security retirement and/or Railroad Retirement payments in the last month, whether replenishment sample for Round 5, and skip pattern controllers.

For Rounds 3, 7, 9, and 11, which included a briefer version of the IA section, the following three classes of variables were used in the imputation:

1. Source variables that indicate (yes/no) whether the respondent (and his/her spouse/partner, if applicable) has the particular source of income (referred to below as "fencepost" variables); Sources include Social Security and/or Railroad Retirement payments; Supplemental Security Income; Veteran's Administration payments; and whether they worked for pay last week or absent from work last week, and whether SP and spouse worked for pay last month; pension plans that were a job-related or union benefit; other retirement plans such as 401K plans, 403B plans or IRAs; mutual funds or stocks; any bonds; checking account; savings account; certificates of deposit (CDs); and whether they own a business, a farm, or any other real estate besides their home.
2. Auxiliary variables that included respondent characteristics (e.g., age, race/ethnicity, gender, educational attainment, marital status, home ownership, ever work for pay, veteran's status, labor force status, spouse/partner's labor force status, whether live in facility, whether live with spouse/partner, interviewer observations about the home condition), whether more than one job last week and hours worked last week, payment schedule from main job, and skip pattern controllers; and
3. Prior round total income and source variables. Sources include Social Security and/or Railroad Retirement payments; Supplemental Security Income; Veteran's Administration payments; and whether they worked for pay last week or absent from work last week, and whether SP and spouse worked for pay last month; pension plans that were a job-related or union benefit; other retirement plans such as 401K plans, 403B plans or IRAs; mutual funds or stocks; any bonds; checking account; savings account; certificates of deposit (CDs); whether they own a business, a farm, or any other real estate besides their

home; and total income. These variables are from the last round in which income was collected.

Income Imputation Procedure

In order to preserve the joint distribution of the full set of income variables, all missing fencepost, source amount (Round 1 and 5 only), auxiliary, and prior round total income and source variables were imputed. For respondents with a spouse/partner, separate respondent and spouse/partner variables as well as the combined variable were imputed for each source.

Descriptive Statistics for Total Income

Table 3 shows weighted descriptive statistics through Round 11 using 5 imputed values generated through the hot deck approach. The weighted means are based on averages of the 5 imputations and the standard error calculations take into account the extra variation introduced by the uncertainty of the imputation.

Table 3. Total income based on 5 hot deck imputations, NHATS Round 1-11 (in current dollars)

	Reference year	Ages	n	Mean	Standard error	25th percentile	Median	75th percentile
Round 11	2020	71+	3,388	67,979	3,541	24,000	46,000	80,000
Round 9	2018	69+	4,460	62,740	2,501	22,000	44,000	80,000
Round 7	2016	67+	5,566	90,636	21,811	21,272	41,800	76,600
Round 5	2014	65+	7,576	66,548	4,493	20,000	40,000	75,000
Round 3	2012	67+	4,884	69,428	7,447	18,600	35,600	70,000
Round 1	2010	65+	7,609	58,050	3,222	16,668	32,200	60,000
Note. Weighted means and standard errors were computed using "mi estimate: svy" command. Median and other percentile measures were computed using the "_pctile" command for the average of the 20 imputed variables. Final analytic weights were used in these analyses. Total income is for individuals and their spouse/partners (if applicable).								

Comparison with Revised Interval Regression Imputation Approach

NHATS now provides 20 imputations prepared using an interval regression approach for all Rounds (for details see Hu & Freedman, 2024). Table 4 summarizes major differences between the two imputation approaches.

Table 4. Differences Between Hot Deck and Interval Regression Imputation Approaches

Hot Deck (5 imputations)	Interval Regression (20 imputations)
Imputation preparation: Editing before multiple imputation	
• Reports of zero income were accepted as valid and eligible to be used as donors for imputation	• Reports of less than $200 were treated as missing an exact value and imputed into the lowest income category
• Reports of very high incomes were accepted as valid and eligible to be used as donors for imputation	• Reports above $900,000 were treated as missing an exact value and imputed

Imputation	
• Cyclical n-partition hot deck procedure	• Interval regression
• Models do not distinguish different imputation groups; most but not all variables used in imputation are the same across rounds	• Models estimated for separate groups; most but not all variables used in imputation are the same across rounds
• Total Income Range variables for respondents who have spouse/partner and respondents who are single are imputed first; Total Income is then imputed within the income range.	• Total Income Range variables are not imputed as an interim step
Delivery	
• 5 imputed variables (ia#toincim1-5) and a flag variable (ia#toincimf) indicating which cases were imputed • Total income range for respondents who have spouse/partner (ia#eincimj1-5) and a flag variable (ia#eincimjf) • Total income range for respondents who are single (ia#eincims1-5) and a flag variable (ia#eincimsf)	• 20 imputed variables (ia#dtoincimi1-ia#dtoincimi20), a flag variable (ia#toincimif) indicating which cases were imputed, and a derived variable (ia#dtoincimreas) indicating reason for imputation
• Included in SP data file	• Included in an auxiliary file zipped with the NHATS data files

Table 5 compares descriptive statistics for the hot-deck and interval regression approaches through Round 11. The percentile distributions for the two sets of estimates are close. However, but means are higher in the hot deck approach, likely because outliers were included as possible donors in the hot deck approach. Standard errors are also higher in the hot deck approach, likely because outliers were included as possible donors and fewer imputations were generated (5 vs. 20).

Table 5. NHATS imputed total income through Round 11: hot deck vs. interval regression approaches

	n	Mean	Standard error	25th percentile	Median	75th percentile
Round 11						
Hot deck	3,388	67,979	3,541	24,000	46,000	80,000
Interval regression	3,388	64,321	2,656	24,637	47,000	80,000
Round 9						
Hot deck	4,460	62,740	2,501	22,000	44,000	80,000
Interval regression	4,460	61,986	2,058	22,889	44,048	80,000
Round 7						
Hot deck	5,566	90,636[a]	21,811[a]	21,272	41,800	76,600
Interval regression	5,566	60,411	2,135	22,000	41,741	76,602
Round 5						
Hot deck	7,576	66,548	4,493	20,000	40,000	75,000
Interval regression	7,576	57,744	1,841	20,400	40,000	74,000
Round 3						
Hot deck	4,884	69,428	7,447	18,600	35,600	70,000
Interval regression	4,884	52,774	1,711	18,980	35,270	68,985
Round 1						

Hot deck	7,609	58,050	3,222	16,668	32,200	60,000
Interval regression	7,609	46,540	1,335	17,320	32,000	59,610

Note: Weighted means and standard errors were computed using "mi estimate: svy" command. Median and percentiles were computed using the "_pctile" command for the average imputed variables. Final analytic weights were used in these analyses.

[a] In Round 7, there were several reported (n=12) and imputed (n=2-9, depending on the imputation) values above $900,000, including reported and imputed values of $99,999,999 and $9,999,999.

References

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural equation modeling: a multidisciplinary journal*, 15(4), 651-675.

DeMatteis J. M., Freedman, V. A., & Kasper, J. D. (2016). National Health and Aging Trends Study Round 5 Income Imputation. NHATS Technical Paper #15. Johns Hopkins University School of Public Health. Available at www.NHATS.org.

DeMatteis J. M., Freedman, V. A., & Kasper, J. D. (2018). National Health and Aging Trends Study Round 7 Income Imputation. NHATS Technical Paper #21. Johns Hopkins University School of Public Health. Available at www.NHATS.org.

DeMatteis J. M., Freedman, V. A., Kasper, J. D., & Jiao, R. (2020). National Health and Aging Trends Study Round 9 Income Imputation. NHATS Technical Paper #27. Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention science*, 8, 206-213.

Hu, M. & Freedman, V. A. (2024). National Health and Aging Trends Study Twenty Interval Regression Income Imputations: Rounds 1-12. NHATS Technical Paper #43. Baltimore: Johns Hopkins Bloomberg School of Public Health. Available at www.nhats.org.

Jiao, R., Freedman, V. A., Kasper, J. D., & Schrack, J. (2022). National Health and Aging Trends Study Round 11 Income Imputation. NHATS Technical Paper #31. Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Judkins, D. (1997). Imputing for Swiss cheese patterns of missing data. Proceedings of Statistics Canada Symposium '97, 143-148.

Judkins, D., Krenzke, T., Piesse, A., Fan, Z., and Haung, W.C. (2007). Preservation of skip patterns and covariate structure through semi-parametric whole questionnaire imputation. Proceedings of the Section on Survey Research Methods of the American Statistical Association, 3211 3218.

Judkins, D., Piesse, A., and Krenzke, T. (2008). Multiple semi-parametric imputation. Proceedings of the Section on Survey Research Methods of the American Statistical Association.

Krenzke, T. and Judkins, D. (2008). Filling in blanks: Some guesses are better than others – Illustrating the impact of covariate selection when imputing complex survey items. Chance, 21(3), 7-13.

Marker, D. A., Judkins, D. R., and Winglee, M. (2001). Large-scale imputation for complex surveys. In Survey Nonresponse, Eds. R. M. Groves, D. A. Dillman, E. L. Eltinge, and R. J. A. Little. New York: John Wiley & Sons.

Montaquila, J., Freedman, V. A., & Kasper, J. D. (2012). National Health and Aging Trends Study Round 1 Income Imputation. Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Montaquila, J., Freedman, V. A., & Kasper, J. D. (2015). National Health and Aging Trends Study Round 3 Income Imputation. NHATS Technical Paper #10. Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Appendix

Lists of Variables Used in NHATS Rounds 1 and 5 Income Imputation

Appendix Table 1. Source ("Fencepost") Variables for Rounds 1 and 5

#	Variable name	Label	Round 1	Round 5
1	ia#recsspa1	R# IA1 SP REC SOCIAL SECURITY	Y	Y
2	ia#recsspa2	R# IA1 SPOUSE PART REC SOC SEC	Y	Y
3	ia#recsspa3	R# IA1 NO SOC SECURTY PAYMNT REC	Y	Y
4	ia#recssils1	R# IA4 SP RECEIVD SSI LAST MONTH	Y	Y
5	ia#recssils2	R# IA4 SPOUSE PRT REC SSI LST MO	Y	Y
6	ia#recssils3	R# IA4 NO SSI RECEIVD LAST MONTH	Y	Y
7	ia#rvapayls1	R# IA5 SP REC PAY FRM VA LAST MO	Y	Y
8	ia#rvapayls2	R# IA5 SPOUS PA REC VA PAY LSTMO	Y	Y
9	ia#rvapayls3	R# IA5 NO VA PAY REC LAST MONTH	Y	Y
10	ia#penjobou1	R# IA6 SP HAS PENSION PLAN	Y	Y
11	ia#penjobou2	R# IA6 SPOUSE HAS PENSION PLAN	Y	Y
12	ia#penjobou3	R# IA6 NO PENSION PLAN	Y	Y
13	ia#iraothac1	R# IA7 SP HAS IRA OTH RETIRE ACC	Y	Y
14	ia#iraothac2	R# IA7 SPOUSE HAS IRA OTHR ACC	Y	Y
15	ia#iraothac3	R# IA7 NO IRA OTHR RETIRE ACCT	Y	Y
16	ia#mutfdstk1	R# IA8 SP OWNS MUTUAL FUND STOCK	Y	Y
17	ia#mutfdstk2	R# IA8 SPOUSE OWNS FUNDS STOCK	Y	Y
18	ia#mutfdstk3	R# IA8 SP SPOUSE OWN FUNDS STOCK	Y	Y
19	ia#mutfdstk4	R# IA8 NO FUNDS OR STOCK OWNED	Y	Y
20	ia#ownbond1	R# IA9 SP OWNS BONDS	Y	Y
21	ia#ownbond2	R# IA9 SPOUSE OWNS BONDS	Y	Y
22	ia#ownbond3	R# IA9 SP SPOUSE OWN BONDS	Y	Y
23	ia#ownbond4	R# IA9 NO BONDS OWNED	Y	Y
24	ia#bnkacccd1	R# IA10 SP OWNS CHECK ACCT	Y	Y
25	ia#bnkacccd2	R# IA10 SPOUSE OWNS CHECK ACCT	Y	Y
26	ia#bnkacccd3	R# IA10 SP SPOUSE OWN CHECK ACCT	Y	Y
27	ia#bnkacccd4	R# IA10 NO CHECK ACCT OWNED	Y	Y
28	ia#bnkacccd5	R# IA10 SP OWNS SAVINGS ACCT	Y	Y
29	ia#bnkacccd6	R# IA10 SPOUSE OWNS SAVING ACCT	Y	Y
30	ia#bnkacccd7	R# IA10 SP SPOUSE OWN SAVNG ACT	Y	Y
31	ia#bnkacccd8	R# IA10 NO SAVINGS ACCT OWNED	Y	Y
32	ia#bnkacccd9	R# IA10 SP OWNS CDS	Y	Y
33	ia#bnkaccc10 for Round 1, ia#bnkacccd10 for Round 5	R# IA10 SPOUSE OWNS CDS	Y	Y
34	ia#bnkaccc11 for Round 1, ia#bnkacccd11 for Round 5	R# IA10 SP SPOUSE OWN CDS	Y	Y
35	ia#bnkaccc12 for Round 1, ia#bnkacccd12 for Round 5	R# IA10 NO CDS OWNED	Y	Y
36	ia#realestt1	R# IA13 SP OWNS REAL ESTATE	Y	Y
37	ia#realestt2	R# IA13 SPOUSE OWNS REAL ESTATE	Y	Y
38	ia#realestt3	R# IA13 SP SPOUSE OWN REAL ESTTE	Y	Y
39	ia#realestt4	R# IA13 NO REAL ESTATE OWNED	Y	Y
40	lf#workfpay	R# LF1 WORKED FOR PAY RECENTLY	Y	Y
41	lf#abstlstwk	R# LF2 ABSENT FRM JOB LAST WEEK	Y	Y
42	lf#wrkplstmn	R# LF3 WORK FOR PAY IN LST MONTH	Y	Y
43	lf#huswifwrk	R# LF13 HUSB/WIFE/PARTN PAY WORK	Y	Y

Appendix Table 2. Source Amount Variables for Rounds 1 and 5

	Variable name	Label	Round 1	Round 5
1	ia#ssrrpymnt	R# IA14 RECENT MTHLY SS RR PYMNT	Y	Y
2	ia#ssrrjtamt, ia#ssrrspamt, ia#ssrrptamt	IA14A-IA16A AMOUNT OF SOCIAL SECURITY/RAILROAD RETIREMNT	Y	Y
3	ia#ssrrjtest, ia#ssrrspest, ia#ssrrptest	IA14B-IA16B RANGE OF SOCIAL SECURITY/RAILROAD RETIREMNT	Y	Y
4	ia#ssipymnt	R# IA17 RECENT MTHLY SSI PAYMENT	Y	Y
5	ia#ssijtamt, ia#ssispamt, ia#ssiptamt	IA17A-IA19A AMOUNT OF SUPPLEMENTAL SECURITY INCOME	Y	Y
6	ia#ssijtest, ia#ssispest, ia#ssiptest	IA17B-IA19B RANGE OF SUPPLEMENTAL SECURITY INCOME	Y	Y
7	ia#vapymnt	R# IA20 RECENT MNTHLY VA PAYMENT	Y	Y
8	ia#vajtamt, ia#vaspamt, ia#vaptamt	IA20A-IA22A AMOUNT OF VETERANS ADMINISTRATION	Y	Y
9	ia#vajtest, ia#vaspest, ia#vaptest	IA20B-IA22B RANGE OF VETERANS ADMINISTRATION	Y	Y
10	ia#penpymt	R# IA23 RCNT MTH JOBPENSION PYMT	Y	Y
11	ia#penjtamt, ia#penspamt, ia#penptamt	IA23A-IA25A AMOUNT OF PENSION PLAN	Y	Y
12	ia#penjtest, ia#penspest, ia#penptest	IA23B-IA25B RANGE OF PENSION PLAN	Y	Y
13	ia#retworth	R# IA26 RETIREMENT ACCOUNT WORTH	Y	Y
14	ia#retjtwrt, ia#retspwrt, ia#retptwrt	IA26A-IA28A AMOUNT OF ANY RETIREMENT ACCOUNTS	Y	Y
15	ia#retjtest, ia#retspest, ia#retptest	IA26B-IA28B RANGE OF ANY RETIREMENT ACCOUNTS	Y	Y
16	ia#rtlmwdrw	R# IA29 RETRMNT WDRW AMT LST MTH	Y	Y
17	ia#rtlmjtwdr, ia#rtlmspwdr, ia#rtlmptwdr	IA29A-IA31A AMOUNT OF RETIREMENT ACCTS WITHDRAW LST MTH	Y	Y
18	ia#rtlmjtest, ia#rtlmspest, ia#rtlmptest	IA29B-IA31B RANGE OF RETIREMENT ACCTS WITHDRAW LST MTH	Y	Y
19	ia#rtyrwdrw	R# IA32 RETRMNT WDRWL AMT LST YR	Y	Y
20	ia#rtyrjtamt, ia#rtyrspamt, ia#rtyrptamt	IA32A-IA34A AMOUNT OF PARTNERS ACCTS WITHDRAW LST YR	Y	Y
21	ia#rtyrjtest, ia#rtyrspest, ia#rtyrptest	IA32B-IA34B RANGE OF PARTNERS ACCTS WITHDRAW LST YR	Y	Y
22	ia#skbdwrth	R# IA35 NONRETR STKS BNDS WRTH	Y	Y
23	ia#skbdjtwrt, ia#skbdspwrt, ia#skbdptwrt	IA35A-IA37A AMOUNT OF WORTH OF FUNDS & STOCKS	Y	Y
24	ia#skbdjtest, ia#bndjtest, ia#skbdspest, ia#bndspest, ia#skbdptest, ia#bndptest	IA35B-IA37C RANGE OF WORTH OF FUNDS & STOCKS	Y	Y
25	ia#bkcdwrth	R# IA38 BANK ACCT CD WORTH	Y	Y
26	ia#bkcdjtwrt, ia#bkcdspwrt, ia#bkcdptwrt	IA38A-IA40A AMOUNT OF WORTH OF BANK ACCDS & CDS	Y	Y
27	ia#bkcdjtest, ia#bnkjtest, ia#bkcdspest, ia#bnkspest, ia#bkcdptest, ia#bnkptest	IA38B-IA40C RANGE OF WORTH OF BANK ACCDS & CDS	Y	Y
28	ia#itdvinc	R# IA41 AMT INT DIV INCOME LS YR	Y	Y
29	ia#itdvjtamt, ia#itdvspamt, ia#itdvptamt	IA41A-IA43A AMOUNT OF TOT WRTH COMBNED INTERST	Y	Y
30	ia#itdvjtest, ia#itdvspest, ia#itdvptest	IA41B-IA43B RANGE OF TOT WRTH COMBNED INTERST	Y	Y
31	ia#brewrt	R# IA44 BUSINESS REALESTATE WRTH	Y	Y
32	ia#brejtwrt, ia#brespwrt, ia#breptwrt	IA44A-IA46A AMOUNT OF TOT WRTH ALL REAL ESTATE	Y	Y
33	ia#brejtest, ia#brespest, ia#breptest	IA44B-IA46B RANGE OF TOT WRTH ALL REAL ESTATE	Y	Y
34	ia#breiinc	R# IA47 BUS REALESTATE INC LS YR	Y	Y
35	ia#breijtamt, ia#breispamt, ia#breiptamt	IA47A-IA49A AMOUNT OF COMB INCOME FROM REL ESTAT	Y	Y
36	ia#breijtest, ia#breispest, ia#breiptest	IA47B-IA49B RANGE OF COMB INCOME FROM REL ESTAT	Y	Y
37	lf#ernfrmwrk	R# LF10 AMT EARN FRM WRK LST MTH	Y	Y
38	lf#huwpaearn	R# LF14 HUS/WIFE/PAR PAY LST MTH	Y	Y

Appendix Table 3. Auxiliary Variables for Rounds 1 and 5

#	Variable name	Label	Round 1	Round 5
1	hh#martlstat for Round 1, hh5dmarstat for Round 5	R# HH1 MARITAL STATUS	Y	Y
2	pa#workfrpay	R# PA17 EVER WORK FOR PAY	Y	Y
3	el#higstschl for Round 1, el#dhigstsch[1] for Round 5	R# EL10 HGHST DGREE SCOOL COMPLD	Y	Y
4	rl#hisplatno for Round 1, rl#dracehisp for Round 5	R# RL3 CNSDR YRSF HSPAN OR LATNO	Y	Y
5	rl#yourrace1	R# RL1 RACE OF SP WHITE	Y	N
6	rl#yourrace2	R# RL1 RACE OF SP AFRICN AMERICN	Y	N
7	rl#yourrace3	R# RL1 RACE OF SP AMERICN INDIAN	Y	N
8	rl#yourrace4	R# RL1 RACE OF SP ALASKA NATIVE	Y	N
9	rl#yourrace5	R# RL1 RACE OF SP ASIAN	Y	N
10	rl#yourrace6	R# RL1 RACE OF SP NATIVE HWAIIAN	Y	N
11	rl#yourrace7	R# RL1 RACE OF SP PACIFIC ISLNDR	Y	N
12	rl#yourrace8	R# RL1 RACE OF SP OTHER SPECIFY	Y	N
13	va#serarmfor	R# VA1 SERVED IN ARMED FORCES	Y	Y
14	va#memnatgrd	R# VA3 MEMBER OF NATIONAL GUARD	Y	Y
15	lf#mrthnonjb	R# LF4 MOR THN ONE JOB LAST WEEK	Y	Y
16	lf#hrswkwork	R# LF5 HRS PR WEEK WORK MAIN JOB	Y	Y
17	lf#hrwrkltwk	R# LF6 HOURS WORK LAST WEEK	Y	Y
18	lf#hrwrklstw	R# LF7 HOW MNY HOURS DID YOU WRK	Y	Y
19	lf#oftpaid	R# LF8 HOW OFTN PAID ON MAIN JOB	Y	Y
20	lf#lstpaychk	R# LF9 HOW MUCH LAST PAYCHECK	Y	Y
21	hp#ownrentot	R# HP1 OWN RENT OR OTHER	Y	Y
22	ia#howrecssp	R# IA2 HOW REC SOC SEC PAY	Y	Y
23	ia#msrtrecss	R# IA3A MTH STRD REC SOC SEC PAY	Y	Y
24	ia#yrstressp	R# IA3B YR STRTD REC SOC SEC PAY	Y	Y
25	ir#areacond1	R# IR15 LITTER GLASS ON SDWLK ST	Y	Y
26	ir#areacond2	R# IR15 GRAFFITI ON BUILDG WALLS	Y	Y
27	ir#areacond3	R# IR15 VACANT HOUSES OR STORES	Y	Y
28	ir#condhome1	R# IR16 BROKEN WINDOWS IN HOME	Y	Y
29	ir#condhome2	R# IR16 CRUMBLNG FOUNDTN IN HOME	Y	Y
30	ir#condhome3	R# IR16 MISSNG BRCKS SIDNG IN HM	Y	Y
31	ir#condhome4	R# IR16 ROOF PROBLEM IN HOME	Y	Y
32	ir#condhome5	R# IR16 BROKEN STEPS TO HOME	Y	Y
33	fl#facility	R# F SP IN FACILITY	Y	Y
34	sex	HISKEW GENDER	Y	Y
35	agecat for Round 1, agecat_r5 for Round 5	HISKEW AGE CATEGORY	Y	Y
36	rtirace	HISKEW RACEETH, 3-CATEGORY	Y	N
37	per_cap_inc_5yr	PER CAPITA INCOME [WT1; ACS]	Y	Y
38	hh#livwthspo	R# HH11 LIVE WITH SPOUSE PARTNER	Y	N
39	hh#placekind	R# HH12 KIND OF PLACE LIVE IN	Y	N
25	ir#condhome6	R# IR16 CONTINUOUS SIDEWALKS	N	Y
30	lf#occupaton	R# LF11 OCCUPATION MOST OF LIFE	N	Y
31	lf#diffwrknw	R# LF12 DO THS WRK SOMETHNG DIFF	N	Y
32	smptype	O=Original smp, R=Replenishment smp	N	Y

[1] Equal to el1higstsch for original sample cases and el5higstsch for replenishment sample cases.

Lists of Variables Used in NHATS Rounds 3, 7, 9 and 11 Income Imputation

Appendix Table 4. Source ("Fencepost") Variables for Rounds 3, 7, 9 and 11

#	Variable name	Label	Round 3	Round 7	Round 9	Round 11
1	ia#recsspa1	R# IA1 SP REC SOCIAL SECURITY	Y	Y	Y	Y
2	ia#recsspa2	R# IA1 SPOUSE PART REC SOC SEC	Y	Y	Y	Y
3	ia#recsspa3	R# IA1 NO SOC SECURTY PAYMNT REC	Y	Y	Y	Y
4	ia#recssils1	R# IA4 SP RECEIVD SSI LAST MONTH	Y	Y	Y	Y
5	ia#recssils2	R# IA4 SPOUSE PRT REC SSI LST MO	Y	Y	Y	Y
6	ia#recssils3	R# IA4 NO SSI RECEIVD LAST MONTH	Y	Y	Y	Y
7	ia#rvapayls1	R# IA5 SP REC PAY FRM VA LAST MO	Y	Y	Y	Y
8	ia#rvapayls2	R# IA5 SPOUS PA REC VA PAY LSTMO	Y	Y	Y	Y
9	ia#rvapayls3	R# IA5 NO VA PAY REC LAST MONTH	Y	Y	Y	Y
10	ia#penjobou1	R# IA6 SP HAS PENSION PLAN	Y	Y	Y	Y
11	ia#penjobou2	R# IA6 SPOUSE HAS PENSION PLAN	Y	Y	Y	Y
12	ia#penjobou3	R# IA6 NO PENSION PLAN	Y	Y	Y	Y
13	ia#iraothac1	R# IA7 SP HAS IRA OTH RETIRE ACC	Y	Y	Y	Y
14	ia#iraothac2	R# IA7 SPOUSE HAS IRA OTHR ACC	Y	Y	Y	Y
15	ia#iraothac3	R# IA7 NO IRA OTHR RETIRE ACCT	Y	Y	Y	Y
16	ia#mutfdstk1	R# IA8 SP OWNS MUTUAL FUND STOCK	Y	Y	Y	Y
17	ia#mutfdstk2	R# IA8 SPOUSE OWNS FUNDS STOCK	Y	Y	Y	Y
18	ia#mutfdstk3	R# IA8 SP SPOUSE OWN FUNDS STOCK	Y	Y	Y	Y
19	ia#mutfdstk4	R# IA8 NO FUNDS OR STOCK OWNED	Y	Y	Y	Y
20	ia#ownbond1	R# IA9 SP OWNS BONDS	Y	Y	Y	Y
21	ia#ownbond2	R# IA9 SPOUSE OWNS BONDS	Y	Y	Y	Y
22	ia#ownbond3	R# IA9 SP SPOUSE OWN BONDS	Y	Y	Y	Y
23	ia#ownbond4	R# IA9 NO BONDS OWNED	Y	Y	Y	Y
24	ia#bnkacccd1	R# IA10 SP OWNS CHECK ACCT	Y	Y	Y	Y
25	ia#bnkacccd2	R# IA10 SPOUSE OWNS CHECK ACCT	Y	Y	Y	Y
26	ia#bnkacccd3	R# IA10 SP SPOUSE OWN CHECK ACCT	Y	Y	Y	Y
27	ia#bnkacccd4	R# IA10 NO CHECK ACCT OWNED	Y	Y	Y	Y
28	ia#bnkacccd5	R# IA10 SP OWNS SAVINGS ACCT	Y	Y	Y	Y
29	ia#bnkacccd6	R# IA10 SPOUSE OWNS SAVING ACCT	Y	Y	Y	Y
30	ia#bnkacccd7	R# IA10 SP SPOUSE OWN SAVNG ACT	Y	Y	Y	Y
31	ia#bnkacccd8	R# IA10 NO SAVINGS ACCT OWNED	Y	Y	Y	Y
32	ia#bnkacccd9	R# IA10 SP OWNS CDS	Y	Y	Y	Y
33	ia#bnkacccd10 for Rounds 3, 7, & 9, ia#bnkaccc10 for Round 11	R# IA10 SPOUSE OWNS CDS	Y	Y	Y	Y
34	ia#bnkacccd11 for Rounds 3, 7, & 9, ia#bnkaccc11 for Round 11	R# IA10 SP SPOUSE OWN CDS	Y	Y	Y	Y
35	ia#bnkacccd12 for Rounds 3, 7, & 9, ia#bnkaccc12 for Round 11	R# IA10 NO CDS OWNED	Y	Y	Y	Y
36	ia#realestt1	R# IA13 SP OWNS REAL ESTATE	Y	Y	Y	Y
37	ia#realestt2	R# IA13 SPOUSE OWNS REAL ESTATE	Y	Y	Y	Y
38	ia#realestt3	R# IA13 SP SPOUSE OWN REAL ESTTE	Y	Y	Y	Y
39	ia#realestt4	R# IA13 NO REAL ESTATE OWNED	Y	Y	Y	Y
40	lf#workfpay	R# LF1 WORKED FOR PAY RECENTLY	Y	Y	Y	Y
41	lf#abstlstwk	R# LF2 ABSENT FRM JOB LAST WEEK	Y	Y	Y	Y
42	lf#wrkplstmn	R# LF3 WORK FOR PAY IN LST MONTH	Y	Y	Y	Y
43	lf#huswifwrk	R# LF13 HUSB WIFE PARTN PAY WORK	Y	Y	Y	Y

Appendix Table 5. Auxiliary Variables for Rounds 3, 7, 9 and 11

#	Variable name	Label	Round 3	Round 7	Round 9	Round 11
1	Sex	HISKEW GENDER	Y	Y	Y	Y
2	agecat	HISKEW AGE CATEGORY	Y	Y	Y	Y
3	rtirace	HISKEW RACEETH, 3-CATEGORY	Y	Y	Y	Y
4	per_cap_inc_5yr	PER CAPITA INCOME [WT1; ACS]	Y	Y	Y	Y
5	el#higstschl	R1 (/5) EL10 HGHST DGREE SCOOL COMPLD	Y	Y	Y	Y
6	rl#hisplatno	R1 RL3 CNSDR YRSF HSPAN OR LATNO	Y	Y	Y	Y
7	rl#yourrace1	R1 RL1 RACE OF SP WHITE	Y	Y	Y	Y
8	rl#yourrace2	R1 RL1 RACE OF SP AFRICN AMERICN	Y	N	N	N
9	rl#yourrace3	R1 RL1 RACE OF SP AMERICN INDIAN	Y	N	N	N
10	rl#yourrace4	R1 RL1 RACE OF SP ALASKA NATIVE	Y	N	N	N
11	rl#yourrace5	R1 RL1 RACE OF SP ASIAN	Y	N	N	N
12	rl#yourrace6	R1 RL1 RACE OF SP NATIVE HWAIIAN	Y	N	N	N
13	rl#yourrace7	R1 RL1 RACE OF SP PACIFIC ISLNDR	Y	N	N	N
14	rl#yourrace8	R1 RL1 RACE OF SP OTHER SPECIFY	Y	Y	Y	Y
15	rl5dracehisp	R5 D RACE AND HISPANIC ETHNICITY	N	Y	Y	Y
16	va1serarmfor for Round 3, va5serarmfor for other rounds	R1 (/5) VA1 SERVED IN ARMED FORCES	Y	Y	Y	Y
17	va1memnatgrd for Round 3, va5memnatgrd for other rounds	R1 (/5) VA3 MEMBER OF NATIONAL GUARD	Y	Y	Y	Y
18	fl#facility	R# F ROUTING FLAG FROM RE4f HT3 5 6 7	Y	Y	Y	Y
19	ir#areacond1	R# IR15 LITTER GLASS ON SDWLK ST	Y	Y	Y	Y
20	ir#areacond2	R# IR15 GRAFFITI ON BUILDG WALLS	Y	Y	Y	Y
21	ir#areacond3	R# IR15 VACANT HOUSES OR STORES	Y	Y	Y	Y
22	ir#condhome1	R# IR16 BROKEN WINDOWS IN HOME	Y	Y	Y	Y
23	ir#condhome2	R# IR16 CRUMBLNG FOUNDTN IN HOME	Y	Y	Y	Y
24	ir#condhome3	R# IR16 MISSNG BRCKS SIDNG IN HM	Y	Y	Y	Y
25	ir#condhome4	R# IR16 ROOF PROBLEM IN HOME	Y	Y	Y	Y
26	ir#condhome5	R# IR16 BROKEN STEPS TO HOME	Y	Y	Y	Y
27	ir#condhome6	R# IR16 CONTINUOUS SIDEWALKS	Y	Y	Y	Y
28	hh#dmarstat	R# D MARITAL STATUS AT R3	Y	Y	Y	Y
29	hh#livwthspo	R# HH11 LIVE WITH SPOUSE PARTNER	Y	Y	Y	Y
30	hh#placekind	R# HH12 KIND OF PLACE LIVE IN	Y	Y	Y	Y
31	pa#workfrpay	R# PA17 EVER WORK FOR PAY	Y	Y	Y	Y
32	lf#mrthnonjb	R# LF4 MOR THN ONE JOB LAST WEEK	Y	Y	Y	Y
33	lf#hrswkwork	R# LF5 HRS PR WEEK WORK MAIN JOB	Y	Y	Y	Y
34	lf#hrwrkltwk	R# LF6 HOURS WORK LAST WEEK	Y	Y	Y	Y
35	lf#hrwrklstw	R# LF7 HOW MNY HOURS DID YOU WRK	Y	Y	Y	Y
36	lf#oftpaid	R# LF8 HOW OFTN PAID ON MAIN JOB	Y	Y	Y	Y
37	hp#ownrentot	R# HP1 OWN RENT OR OTHER	Y	Y	Y	Y

Appendix Table 6. Source ("Fencepost") Variables from Last Round Income was Collected (&=current round #-2) for Rounds 3, 7, 9 and 11[2]

#	Variable name	Label	Round 3	Round 7	Round 9	Round 11
1	ia&recsspa1	R& IA1 SP REC SOCIAL SECURITY	Y	Y	Y	Y
2	ia&recsspa2	R& IA1 SPOUSE PART REC SOC SEC	Y	Y	Y	Y
3	ia&recsspa3	R& IA1 NO SOC SECURTY PAYMNT REC	Y	Y	Y	Y
4	ia&recssils1	R& IA4 SP RECEIVD SSI LAST MONTH	Y	Y	Y	Y
5	ia&recssils2	R& IA4 SPOUSE PRT REC SSI LST MO	Y	Y	Y	Y
6	ia&recssils3	R& IA4 NO SSI RECEIVD LAST MONTH	Y	Y	Y	Y
7	ia&rvapayls1	R& IA5 SP REC PAY FRM VA LAST MO	Y	Y	Y	Y
8	ia&rvapayls2	R& IA5 SPOUS PA REC VA PAY LSTMO	Y	Y	Y	Y
9	ia&rvapayls3	R& IA5 NO VA PAY REC LAST MONTH	Y	Y	Y	Y
10	ia&penjobou1	R& IA6 SP HAS PENSION PLAN	Y	Y	Y	Y
11	ia&penjobou2	R& IA6 SPOUSE HAS PENSION PLAN	Y	Y	Y	Y
12	ia&penjobou3	R& IA6 NO PENSION PLAN	Y	Y	Y	Y
13	ia&iraothac1	R& IA7 SP HAS IRA OTH RETIRE ACC	Y	Y	Y	Y
14	ia&iraothac2	R& IA7 SPOUSE HAS IRA OTHR ACC	Y	Y	Y	Y
15	ia&iraothac3	R& IA7 NO IRA OTHR RETIRE ACCT	Y	Y	Y	Y
16	ia&mutfdstk1	R& IA8 SP OWNS MUTUAL FUND STOCK	Y	Y	Y	Y
17	ia&mutfdstk2	R& IA8 SPOUSE OWNS FUNDS STOCK	Y	Y	Y	Y
18	ia&mutfdstk3	R& IA8 SP SPOUSE OWN FUNDS STOCK	Y	Y	Y	Y
19	ia&mutfdstk4	R& IA8 NO FUNDS OR STOCK OWNED	Y	Y	Y	Y
20	ia&ownbond1	R& IA9 SP OWNS BONDS	Y	Y	Y	Y
21	ia&ownbond2	R& IA9 SPOUSE OWNS BONDS	Y	Y	Y	Y
22	ia&ownbond3	R& IA9 SP SPOUSE OWN BONDS	Y	Y	Y	Y
23	ia&ownbond4	R& IA9 NO BONDS OWNED	Y	Y	Y	Y
24	ia&bnkacccd1	R& IA10 SP OWNS CHECK ACCT	Y	Y	Y	Y
25	ia&bnkacccd2	R& IA10 SPOUSE OWNS CHECK ACCT	Y	Y	Y	Y
26	ia&bnkacccd3	R& IA10 SP SPOUSE OWN CHECK ACCT	Y	Y	Y	Y
27	ia&bnkacccd4	R& IA10 NO CHECK ACCT OWNED	Y	Y	Y	Y
28	ia&bnkacccd5	R& IA10 SP OWNS SAVINGS ACCT	Y	Y	Y	Y
29	ia&bnkacccd6	R& IA10 SPOUSE OWNS SAVING ACCT	Y	Y	Y	Y
30	ia&bnkacccd7	R& IA10 SP SPOUSE OWN SAVNG ACT	Y	Y	Y	Y
31	ia&bnkacccd8	R& IA10 NO SAVINGS ACCT OWNED	Y	Y	Y	Y
32	ia&bnkacccd9	R& IA10 SP OWNS CDS	Y	Y	Y	Y
33	ia&bnkaccc10	R& IA10 SPOUSE OWNS CDS	Y	Y	Y	Y
34	ia&bnkaccc11	R& IA10 SP SPOUSE OWN CDS	Y	Y	Y	Y
35	ia&bnkaccc12	R& IA10 NO CDS OWNED	Y	Y	Y	Y
36	ia&realestt1	R& IA13 SP OWNS REAL ESTATE	Y	Y	Y	Y
37	ia&realestt2	R& IA13 SPOUSE OWNS REAL ESTATE	Y	Y	Y	Y
38	ia&realestt3	R& IA13 SP SPOUSE OWN REAL ESTTE	Y	Y	Y	Y
39	ia&realestt4	R& IA13 NO REAL ESTATE OWNED	Y	Y	Y	Y
40	lf#workfpay	R& LF1 WORKED FOR PAY RECENTLY	Y	Y	Y	Y
41	lf#abstlstwk	R& LF2 ABSENT FRM JOB LAST WEEK	Y	Y	Y	Y
42	lf#wrkplstmn	R& LF3 WORK FOR PAY IN LST MONTH	Y	Y	Y	Y
43	lf#huswifwrk	R& LF13 HUSB/WIFE/PARTN PAY WORK	Y	Y	Y	Y
44	ia&totinc	R& IA50 TOTAL INCOME	Y	Y	Y	Y
45	ia&toincesjt	R& IA51A JOINT EST TOT INCOME	Y	Y	N	N
46	ia&toincessg	R& IA51B SNGLE EST TOT INC	Y	Y	N	N

[2] For Round 11, Round 9 source variables are not available for 3 Round 9 respondents who had completed facility questionnaires only, so their imputation process only used Table 1 and Table 2 variables. For Round 9 imputation, Round 7 source variables are not available for 16 Round 7 respondents who had completed facility questionnaires only, so their imputation process only used Table 1 and Table 2 variables.