National Study of Caregiving IV

Twenty Interval Regression Income Imputations:

Rounds 11 - 12

April 2024

Contents

Overview

The National Study of Caregiving (NSOC) IV asks caregivers to report their total income for the prior calendar year. For those who do not report an exact value for total annual income, information is collected in broad categories. For single caregivers individual income is reported; for those with a spouse or partner, couple income is reported. Caregivers who are spouses/partners of NHATS participants are asked income questions in NSOC IV and also have information collected or imputed in NHATS (Freedman et al. 2023; Hu & Freedman 2024a).

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. For NHATS and NSOC through Round 11, we initially released 5 sets of imputed values. For NSOC these values were based on interval regression models and were released in a set of enhanced files. Starting in Round 12, to ensure more reliable standard error estimation and consistency with NHATS, we now provide 20 imputed values for all rounds of NSOC. The 20 imputations are provided in a set of auxiliary files and were prepared using an interval regression methodology.

This technical paper describes the updated income imputation methodology using interval regression in NSOC IV (Rounds 11 and 12). Comparisons with the initial release of five sets of imputed values are provided for Round 11. A separate technical paper describes the methodology for NSOC I-III (Rounds 1, 5 and 7) (Hu & Freedman 2024b).

Although this technical paper demonstrates that percentile distributions for the two sets of estimates in Round 11 are close, means and standard deviations are lower for estimates based upon the updated methodology. We therefore recommend that researchers use the auxiliary set of 20 imputations.

Auxiliary Imputed Income Files and Variables

The NSOC IV Imputed Income file includes one observation for each NSOC participant in the given round. In both rounds, all caregivers have reported or imputed values of total income. The income imputation file for each round of NSOC IV includes 20 values of total income. In addition, a flag variable indicating which cases have imputed (vs. reported) values in NSOC or NHATS is included along with a derived variable indicating the reason for imputation.

Table 1. Imputed variables included in NSOC IV – 20 imputations.

Variable name	Label	Description
chi#toincimif	C# F IMPUTED TOTAL INC FLG INTERVAL	Flag indicating imputation: 1. Reported in NSOC 2. Reported in NHATS (Spouse/partner of SP) 3. Imputed in NSOC 4. Imputed in NHATS (Spouse/partner of SP)

chi#dtoincimi1 - chi#dtoincimi20	C# D HI10 IMPUTED TOTAL INC1-20 INTERVAL	Includes reported and imputed amounts
chi#dtoincimreas	C# D IMPUTED TOTAL INC REASON	Derived variable indicating reasons for imputation: 1. Not imputed: exact value reported 2. Imputed: bracket response only 3. Imputed: missing exact value and bracket response 4. Imputed: other

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 imputations previously provided.

To adjust coefficients and standard errors for the variability between imputations following the combination rules by Rubin (1987), users can use the "mi estimate" command in Stata to run estimations on the imputed dataset. To use this procedure, researchers must first generate an income variable that only includes non-imputed income values (that is, they must replace outliers that were imputed with a missing value. We have included code below for merging the auxiliary file, generating a non-imputed version of total income, and estimating weighted mean total income.

Below is an example of Stata code that may be used to estimate unweighted and weighted mean income using NSOC Rounds 11 and 12. The # stands for round number.

```
*merge income imputation file with SP file
use "[path\]NSOC_R#_Interval_Inc_Imp_File.dta", clear
merge 1:1 spid opid using "[path\]NSOC_R#_Crss_File.dta"
drop _merge

*create a weight for CGs to both living and deceased SPs
gen w#cg_lmlwgt0 = w#cgfinwgt0 if w#cgfinwgt0 > 0
replace w#cg_lmlwgt0 = w#cglmlfinwgt0 if w#cgfinwgt0 == 0

*generate a total income variable that includes only non-imputed income values
gen chi#income2  = chi#dtoincimi1
replace chi#income2  = . if chi#dtoincimreas > 1

*save it to a new dataset
save newdataset, replace

*use this new dataset for analysis
use newdataset, clear
```

*use mi import to import data that include reported income and additional imputed income variables
mi import wide, imputed (chi#income2 = chi#dtoincimi1 chi#dtoincimi2 chi#dtoincimi3 chi#dtoincimi4 chi#dtoincimi5 chi#dtoincimi6 chi#dtoincimi7 chi#dtoincimi8 chi#dtoincimi9 chi#dtoincimi10 chi#dtoincimi11 chi#dtoincimi# chi#dtoincimi13 chi#dtoincimi14 chi#dtoincimi15 chi#dtoincimi16 chi#dtoincimi17 chi#dtoincimi18 chi#dtoincimi19 chi#dtoincimi20) drop
mi set wide

*estimate unweighted mean income of all caregivers
mi estimate: mean chi#income2

*estimate weighted mean income

mi svyset c#varunit [pweight=w#cg_lmlwgt0], strata(c#varstrat) singleunit(centered)
mi estimate: svy: mean chi#income2

Extent of Missing Data for Total Income

Table 2 shows the percentage of participants that require imputation and the reason for NSOC Rounds 11 and 12. The percentage of respondents who did not require imputation (i.e., they provided an exact value of total income in either NSOC or their spouse provided the amount in NHATS) ranged from 71%-76%. An additional 2% had values imputed in NHATS and brought over to NSOC. A categorical response (requiring imputation within a bracket) was reported by 9%-14% and 10%-11% were missing both an exact valued and a categorical response (requiring a full imputation). Another 2%-3% provided values that were considered outliers (extremely low or high values) and were also imputed.

Table 2. Percentage of NSOC IV Respondents Requiring Income Imputation and Reason, Rounds 11 and 12

	Round 11		Round 12	
Reason for Imputation (%)	n	%	n	%
Not imputed: exact value reported[a]	1,382	71.3	1,844	75.9
Reported in NSOC	1,347	69.5	1,788	73.6
Reported in NHATS (Spouse / partner of SP)	35	1.8	56	2.3
Imputed in NHATS	45	2.3	55	2.3
Imputed in NSOC: bracket response only	265	13.7	219	9.0
Imputed in NSOC: missing exact value and bracket response	210	10.8	247	10.2
Imputed in NSOC: other	36	1.9	66	2.7
Total	1,938	100.0	2,431	100.0

[a]When spouse/partner caregivers with only categorical income information in NSOC gave answers consistent with income information in NHATS, we used the exact value from NHATS (n=35 in Rd 11 and 56 in Rd 12).

Imputation Approach

For purposes of imputation, we transformed income reports to log income (after adding $1). For respondents missing an exact value (or reporting an outlier value suspected to be misreporting[1]), we used multiple imputation by chained equations (MICE) to impute missing covariates and missing income. We used interval regression in Stata to impute income, which fills in the missing values of partially observed (censored) income (that is, within a range).

To implement the interval regression approach, we specified a lower and upper limit for all cases. For cases with an exact value, we added $1 to the reported value, and specified the lower and upper limits to be the log-transform of that amount (essentially assigning the reported value to each imputation). For cases with only categorical information, we used the log-transformed upper and lower amounts of the bracket (plus $1) as limits. Cases with only an upper limit received $1 as their lower limit. Cases with only a lower received a missing value for their upper limit, which can take any positive value above the lower limit. For low-end outliers (<$200), we assumed that the exact value was in the lowest category. For cases missing both sources of income (exact value and categorical information) and for high-end outliers (>$900,000), the lower limit was specified as $1 and upper limit set to missing. Before releasing the data, we transformed reported and imputed values back to the original (non-logged) dollar scale.

Imputation Groups

Imputation groups were formed based on available covariates. For NSOC IV (Rounds 11 and 12), four groups were formed.

1. Non-spouse/partner caregivers to living SPs who completed the entire interview (non-breakoff);
2. Spouse/partner caregivers to living SPs;
3. Non-spouse/partner caregivers to living SPs who did not complete the interview (breakoff);
4. Caregivers to deceased SPs (i.e., Last month of life, LML)

Note that Group 2 has two potential sources of information: (1) reported information from NSOC and (2) reported or imputed information from NHATS. We developed the following approach for determining which source to use for Spouse/partner caregivers to living SPs:

- If the caregiver reported a valid exact value ($200 or above and $900,000 or less) in NSOC, we used the information from NSOC.
- If the caregiver reported a bracketed amount in NSOC, we compared the bracket to the amount reported in NHATS (ia#totinc) or the average imputed amount in NHATS (ia#dtoincimi1-ia#dtoincimi20); and:

---

[1] Cases with reported income <$200 and >$900,000 were confirmed to be substantially different from reports in prior rounds and therefore imputed.

o If the value from NHATS fell within $10,000 of the NSOC bracket range, we used the NHATS amount; otherwise,

o we imputed income within the NSOC bracket.

- If the caregiver was missing a valid exact value and missing bracket information in NSOC, or reported an outlier (<$200 or more than $900,000), we used the reported or imputed income from NHATS.

To ensure adequate sample sizes in the imputation models, when imputing Groups 3 and 4, we also included Group 1 and 2 cases (either their reported or the average of their 20 imputed values) in the models.

Variables Used in Imputation

Here we provide an overview of covariates included in the imputation models. See Appendix Table 2 for details by imputation group and round.

- <u>Caregiver characteristics</u> include age, gender, race/ethnicity, number of other people in the household, education and spouse/partner's education, relationship of caregiver to sample person and (final) survey mode (web/telephone).
- <u>NHATS Sample Person characteristics</u> include census division and metro/nonmetro residence, income, and whether the caregiver lives with the SP in the current round.
- <u>Caregiver economic resources</u> include insurance coverage (private insurance, Medicaid enrollment), assets (e.g., whether the CG and their spouse/partner own a home, have a checking account, savings account, certificate of deposit, retirement plan, and stocks or mutual funds); business/farm ownership, and whether worked for pay.

Income Imputation Procedure

For each imputation group, we estimated multiple imputations by chained equations (MICE) first to fill in missing values for covariates and logged income (starting with the least missing). Interval regression was used to impute 20 values of logged income. We chose 20 to be consistent with the NHATS imputation methodology. We also tested various numbers of imputations, ranging from 5 to 50. We found that 20 imputations consistently resulted in stable standard error estimation. The final number of imputations also aligns with recommendations from the literature when up to 30% of respondents have missing values (Graham et al., 2007; Bodener, 2008). For each imputation group, we specified 20 burn-in iterations (i.e., the number of iterations before the first set of imputed values is drawn). Final imputed variables were transformed back to a non-logged dollar scale.

We incorporated the NSOC survey design into the imputation procedure by using NSOC weights, which build upon NHATS weights, in the imputation model (using a pweight statement).[2]

---

[2]Although we did include publicly available geographic indicators (SP's Census division and metro/non-metro residence) as controls, we could not directly account for strata and cluster variables in the imputation strategy. We

A summary of imputation model results can be found in the Appendix.

Descriptive Statistics for Total Income

Table 3 presents weighted descriptive statistics for total income using reported and 20 imputed values generated through the interval regression approach. The weighted means are based on averages over the 20 imputations and the standard error calculations take into account the extra variation introduced by the uncertainty of the imputation.

Table 3. Weighted descriptive statistics, Total Income, NSOC Rounds 1-12 (in dollars)

	Reference Year	n	Mean	Standard error	25th percentile	Median	75th percentile
Round 12	2021	2,431	71,075	3,459	22,393	50,000	90,000
Round 11	2020	1,938	74,670	2,877	25,349	55,000	100,000
Round 7[a]							
Crosssectional	2016	2,612	66,251	2,246	24,000	50,000	90,000
Longitudinal	2016	1,392	56,070	2,053	21,000	41,000	75,000
Round 5	2014	2,204	59,050	2,246	21,000	43,953	83,436
Round 1	2010	2,007	53,652	2,106	19,039	40,000	75,000

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. For Rounds 7-12 cross-sectional files, final analytic weights were constructed using both caregiver final weight (w#cgfinwgt0) and caregiver final LML weight (w#cglmlfinwgt0).
[a]Caregivers who did not help last month or in the last month of life are excluded.

Comparison with Previous Interval Regression Imputation Approach

Previously, NSOC provided on the enhanced caregiver files 5 imputations prepared using an interval regression methodology. Table 4 summarizes the main differences between the current interval regression imputation (yielding 20 imputations) and the previous approach (yielding 5 imputations).

Table 4. Differences between previous vs. current interval regression imputation approaches

Previous imputation (5 imputations)	Current imputation (20 imputations)
Imputation preparation: Editing before multiple imputation	

---

considered other approaches–e.g., adding a variable for each strata by cluster combination or imputing separately by each unique cluster (Heeringa, West & Berglund, 2017; Kalpourtzi et al. 2024; Reiter 2006). However, the study has too many cluster/strata combinations to implement these approaches. For NHATS imputations, we also explored controlling for the full set of replicate weights, which captures the survey's complex design, but found that models did not consistently converge across imputation groups.

| | Reports of zero income were treated as missing and were imputed | | Reports of less than $200 were treated as missing exact value and imputed into the lowest income category |

- Reports of zero income were treated as missing and were imputed
- Reports of very high incomes were accepted as valid reports
- Income brought over from NHATS for caregivers who were spouse/partner of SP was based on 5 imputed hot deck values

Imputation

- For Groups 1, 3 and 4, SP's income from NHATS was reported or based on 5 imputed hot deck values
- After bringing over income from NHATS, Group 2 cases with missing income (values of zero) were imputed
- For Groups 3 and 4, Group 1 cases are included in the imputation model

Delivery

- 5 imputed variables (chi#incim1-chi#incim5) and a flag variable (chi#incimf) indicating which cases were imputed

- Included in enhanced NSOC I-III data file

- Reports of less than $200 were treated as missing exact value and imputed into the lowest income category
- Reports above $900,000 were treated as missing an exact value and imputed
- Income brought over from NHATS for caregivers who were spouse/partner of SP was based on 20 interval regression imputations

- For Groups 1, 3 and 4, SP's income from NHATS was reported or based on 20 interval regression imputations
- After bringing over income from NHATS, Group 2 cases did not have missing values

- For Groups 3 and 4, both Groups 1 and cases are included in the imputation model

- 20 imputed variables (chi#dtoincimi1 - chi#dtoincimi20), a flag variable (chi#toincimf) indicating which cases were imputed, and a derived variable indicating reason for imputation (chi#dtoincimreas)
- Included in an auxiliary file zipped with the NSOC data files

Table 5 compares imputed income for NSOC IV (Round 11) based on the two sets of imputations. The percentile distributions for the two sets of estimates are close, but means and standard errors are lower in the 20 interval regression approach, likely because outliers were imputed.

Table 5. NSOC IV Imputed Income Round 11: 5 vs. 20 Interval Regression Imputations

	n	Mean	Standard error	25th percentile	Median	75th percentile
Round 11 (5 imputations)	1,938	82,727	5,938	25,000	55,108	98,953
Round 11 (20 imputations)	1,938	74,670	2,877	25,349	55,000	100,000

References

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

Freedman, V. A., Hu, M., and Wolff, J. (2023). National Study of Caregiving IV User Guide: Rounds 11-12 Final Release. Baltimore: Johns Hopkins Bloomberg 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.

Heeringa, S. G., West, B. T., & Berglund, P. A. (2017). Applied survey data analysis 2nd edition. CRC press.

Hu, M. & Freedman, V. A. (2024a). 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.

Hu, M. & Freedman, V. A. (2024b). National Study of Caregiving I-III Twenty Interval Regression Income Imputations: Rounds 1, 5 and 7. NHATS Technical Paper #44. Baltimore: Johns Hopkins Bloomberg School of Public Health. Available at www.nhats.org.

Hu, M., Freedman, V. A. & Jiao, R. (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.

Kalpourtzi, N., Carpenter, J. R., & Touloumi, G. (2024). Handling missing values in surveys with complex study design: A simulation study. *Journal of Survey Statistics and Methodology*, 12(1), 105-129.

Reiter, J. P., Raghunathan, T. E., & Kinney, S. K. (2006). The importance of modeling the sampling design in multiple imputation for missing data. *Survey Methodology*, 32(2), 143.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley

Rubin, D.B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 91(434), 473-489.

Appendix

Appendix Table 1. Sample Sizes for Income Imputation Groups: NSOC IV Rounds 11-12

	1 Non-Spouse/ Partner	2 Spouse/ Partner	3 Breakoff	4 LML
Round 11				
Number with Reported Income	930	247	0	205
Number with Imputed Income	378	59[e]	28	91
Total Model Sample Size	1,308	306	1,642[f]	1,910[f]
Round 12				
Number with Reported Income	1,249	429	6	160
Number with Imputed Income	423	68[e]	49	47
Total Model Sample Size	1,672	497	2,224[f]	2,376[f]
Weight used	w#cgfinwgt0	w#cgfinwgt0	w#cgfinwgt0	w#cgfinwgt0 (for CGs to living SPs) and w#cglmlfinwgt0 (for LML CGs)

[e] Includes cases imputed income from NHATS and brought over to NSOC.
[f] Group 1 & 2 cases (reported or average of 20 imputed values) are also included in model estimation.

Appendix Table 2. Covariates used to impute income in NSOC IV for each subgroup

Covariates	Imputation Group			
	1 Non-Spouse/ Partner	2 Spouse/ Partner	3 Breakoff	4 LML
Age	Y	Y	Y	Y
Race / ethnicity	Y	Y	Y	Y
Gender	Y	Y	Y	Y
Number of other people live in the household	Y	Y	Y	Y
CG education	Y	Y	Y	Y
CG spouse/partner's education[a]	Y	Y	Y	Y
SP census division	Y	Y	Y	Y
SP Metro / non-metro residence	Y	Y	Y	Y
Relationship to SP	Y		Y	Y
SP's income from NHATS (log-transformed)	Y		Y	Y [b]
CG in household with SP	Y		Y	Y
CG has private insurance coverage	Y	Y		Y
CG has Medicaid coverage	Y	Y		Y
CG owns home	Y	Y		Y
CG has checking acct	Y	Y		Y
CG has savings account	Y	Y		Y
CG has cert of deposit	Y	Y		Y
CG has retirement plan	Y	Y		Y
CG has stocks mutual funds	Y	Y		Y
Work for pay last week or in the last month of SP's life	Y	Y		Y [c]
Own farm business	Y	Y		
Final survey mode[d]	Y	Y	Y	Y

[a] For Group 2, CG spouse's education was imputed using ordinal logistic model. For the other three groups, CG spouse's education was imputed using multinomial logistic model. In Round 11, for group 1 imputation, due to convergency issues, missing for CG spouse's education is treated as a separate category.
[b] For Round 11 LML group, SP income from Round 9 were used for everyone in the model. For Round 12, when SP's income from NHATS Round 12 is missing (for deceased SPs), we filled in SP's income from Round 11. We did not use Round 11 income for everyone in the model for the LML group in Round 12 is because Round 11 SP income was missing for the replenished sample.
[c] Constructed using cel#wrk4pay and cec#wrk4pay for Rounds 11 and 12. Not included in prior rounds.
[d] Final survey mode is used because the income questions are close to the end of the instrument.

Appendix Table 3. Imputation model results for NSOC IV Round 11 by imputation groups

	1 Non-Spouse/ Partner		2 Spouse/ Partner		3 Breakoff		4 LML	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
CG Age	0.01	0.00	0.01	0.01	0.01	0.00	-0.01	0.00
CG Race / ethnicity								
1 White, non-Hispanic								
2 Black, non-Hispanic	0.14	0.08	-0.21	0.20	0.00	0.08	-0.41	0.05
3 Others/Missing	0.26	0.11	-0.44	0.32	0.09	0.14	-0.36	0.10
4 Hispanic	0.38	0.13	-0.11	0.17	0.22	0.11	-0.35	0.09
CG Gender								
1 Male								
2 Female	-0.06	0.06	0.07	0.10	-0.02	0.06	-0.05	0.04
Number of other people live in CG household	-0.01	0.03	-0.05	0.09	-0.04	0.03	-0.03	0.02
CG education								
1 <HS								
2 HS	0.54	0.24	0.15	0.16	0.58	0.16	0.05	0.08
3 >HS, <College	0.68	0.24	0.08	0.16	0.81	0.16	0.20	0.08
4 College+	0.93	0.25	0.11	0.22	1.17	0.17	0.39	0.10
CG spouse/partner's education								
0 No spouse								
1 <HS	0.85	0.17			0.88	0.14	-0.28	0.08
2 HS	0.57	0.09	0.07	0.15	0.87	0.08	-0.12	0.06
3 >HS, <College	0.52	0.08	0.07	0.20	0.77	0.08	-0.05	0.06
4 College+	0.78	0.10	0.46	0.17	1.12	0.09	0.09	0.07
9 Missing	0.95	0.21			1.23	0.18	0.10	0.19
SP census division								
1 Northeast Region: New England Division								
2 Northeast Region: Middle Atlantic Division	-0.21	0.13	-0.57	0.30	-0.18	0.12	0.07	0.10
3 Midwest Region: East North Central Division	-0.26	0.10	-0.56	0.28	-0.18	0.09	0.11	0.10
4 Midwest Region: West North Central Division	-0.21	0.11	-0.34	0.27	-0.11	0.10	0.02	0.10
5 South Region: South Atlantic Division	-0.18	0.11	-0.39	0.23	-0.05	0.09	0.16	0.09
6 South Region: East South Central Division	-0.43	0.12	-0.55	0.27	-0.33	0.11	0.06	0.11
7 South Region: West South Central Division	-0.38	0.13	-0.66	0.30	-0.39	0.11	0.02	0.10
8 West Region: Mountain Division	-0.25	0.17	-0.28	0.28	-0.18	0.14	0.25	0.14
9 West Region: Pacific Division	-0.14	0.11	-0.47	0.29	-0.07	0.11	0.24	0.10
SP metro / non-metro residence								
1 Metropolitan								

2 Non-metropolitan	-0.06	0.07	0.08	0.12	0.02	0.07	-0.08	0.05
CG relationship to SP								
2 Spouse/partner of SP								
3 Children					0.50	0.13	-0.71	0.07
4 Other relatives	0.00	0.07			0.38	0.14	-0.82	0.09
5 Non-relatives	-0.09	0.07			0.16	0.14	-0.64	0.10
CG in household with SP	0.01	0.04			0.15	0.05		
SP's income from NHATS (log-transformed)	0.01	0.04					0.04	0.04
CG in household with SP								
1 Yes								
2 No	0.19	0.08			0.37	0.08	0.07	0.06
CG has private insurance coverage								
No								
Yes	0.25	0.07	0.13	0.14			-0.01	0.05
CG has Medicaid coverage								
No								
Yes	-0.10	0.12	-0.49	0.20			-0.08	0.07
CG owns home								
No								
Yes	0.17	0.07	0.22	0.13			0.00	0.05
CG has checking acct								
No								
Yes	0.57	0.16	0.02	0.23			-0.11	0.09
CG has savings account								
No								
Yes	0.06	0.09	0.43	0.16			0.22	0.06
CG has certificate of deposit								
No								
Yes	-0.02	0.08	-0.06	0.14			0.01	0.06
CG has retirement plan								
No								
Yes	0.44	0.08	0.06	0.11			0.12	0.05
CG has stocks or mutual funds								
No								
Yes	0.21	0.06	0.23	0.14			0.09	0.05
CG worked for pay last week								
Yes								
No	-0.35	0.08	-0.23	0.22			0.04	0.05
Retired	-0.32	0.08	-0.15	0.24			0.04	0.06
CG owns business or farm								
No								
Yes	0.15	0.07	-0.46	0.53				
Survey mode								
1 Phone								
2 Web	-0.18	0.07	-0.27	0.18	-0.16	0.07	-0.08	0.05
Constant	7.97	0.67	9.65	0.70	6.88	0.59	10.53	0.37

Log of sigma	-0.36	0.05	-0.25	0.16	-0.22	0.05	-0.42	0.02
Sigma	0.70	0.04	0.78	0.13	0.80	0.04	0.66	0.01

Appendix Table 4. Imputation model results for NSOC IV Round 12 by imputation groups

	1 Non-Spouse/ Partner		2 Spouse/ Partner		3 Breakoff		4 LML	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
CG Age	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
CG Race / ethnicity								
1 White, non-Hispanic								
2 Black, non-Hispanic	-0.11	0.10	-0.13	0.13	-0.11	0.08	-0.35	0.05
3 Others/Missing	0.06	0.16	-0.06	0.11	0.02	0.10	-0.06	0.07
4 Hispanic	0.08	0.11	-0.15	0.12	0.06	0.09	-0.48	0.07
CG Gender								
1 Male								
2 Female	0.05	0.07	-0.08	0.07	0.04	0.06	-0.05	0.04
Number of other people live in CG household	0.06	0.03	0.02	0.03	0.03	0.02	-0.02	0.02
CG education								
1 <HS								
2 HS	0.09	0.22	0.05	0.13	0.38	0.14	-0.01	0.08
3 >HS, <College	0.25	0.21	0.03	0.11	0.62	0.13	0.00	0.08
4 College+	0.58	0.22	0.33	0.13	1.09	0.14	0.16	0.09
CG spouse/partner's education								
0 No spouse								
1 <HS	0.42	0.20			0.73	0.15	-0.34	0.09
2 HS	0.46	0.16	-0.05	0.14	0.88	0.12	-0.30	0.06
3 >HS, <College	0.39	0.09	0.09	0.14	0.83	0.09	-0.21	0.07
4 College+	0.56	0.12	0.40	0.14	1.04	0.11	-0.12	0.07
9 Missing								
SP census division								
1 Northeast Region: New England Division								
2 Northeast Region: Middle Atlantic Division	0.17	0.14	0.32	0.15	0.26	0.10	-0.11	0.10
3 Midwest Region: East North Central Division	-0.16	0.15	0.09	0.17	-0.11	0.12	0.04	0.10
4 Midwest Region: West North Central Division	-0.03	0.19	0.10	0.15	-0.02	0.15	-0.10	0.11
5 South Region: South Atlantic Division	-0.19	0.13	0.19	0.14	-0.14	0.09	0.06	0.10
6 South Region: East South Central Division	0.01	0.17	-0.17	0.15	-0.07	0.13	0.02	0.11
7 South Region: West South Central Division	-0.07	0.14	-0.24	0.24	-0.10	0.11	-0.10	0.10
8 West Region: Mountain Division	0.03	0.15	0.00	0.22	-0.14	0.13	-0.02	0.12
9 West Region: Pacific Division	-0.06	0.13	0.17	0.15	-0.06	0.09	0.04	0.10
SP metro / non-metro residence								

	Est	SE	Est	SE	Est	SE	Est	SE
1 Metropolitan								
2 Non-metropolitan	-0.19	0.12	-0.11	0.10	-0.21	0.10	0.00	0.05
CG relationship to SP								
2 Spouse/partner of SP								
3 Children					0.26	0.12	-0.76	0.08
4 Other relatives	0.05	0.10			0.26	0.15	-0.74	0.10
5 Non-relatives	-0.01	0.13			0.01	0.17	-0.70	0.10
SP's income from NHATS (log-transformed)	0.04	0.04			0.25	0.03		
CG in household with SP								
1 YES								
2 NO	0.30	0.10			0.47	0.10	0.02	0.06
CG owns business or farm								
No								
Yes	0.20	0.13	0.11	0.13				
CG has private insurance coverage								
No								
Yes	0.22	0.08	0.04	0.07			0.01	0.04
CG has Medicaid coverage								
No								
Yes	-0.42	0.15	-0.39	0.10			-0.17	0.05
CG owns home								
No								
Yes	0.29	0.08	0.19	0.08			-0.03	0.05
CG has checking acct								
No								
Yes	0.37	0.17	-0.08	0.16			-0.01	0.08
CG has savings account								
No								
Yes	0.10	0.10	0.21	0.11			0.21	0.05
CG has certificate of deposit								
No								
Yes	0.03	0.08	0.12	0.09			0.01	0.06
CG has retirement plan								
No								
Yes	0.48	0.09					0.06	0.05
CG has stocks or mutual funds								
No								
Yes	0.09	0.08	0.06	0.09			0.24	0.05
CG worked for pay last week								
Yes								
No	-0.20	0.09	0.04	0.13			0.03	0.05
Retired	-0.18	0.09	0.07	0.15			-0.05	0.06
CG income from round 11							0.09	0.02
Survey mode								
1 Phone								
2 Web	-0.23	0.07	0.03	0.07	-0.10	0.06	0.14	0.05

Constant	8.09	0.54	10.31	0.41	5.86	0.42	9.93	0.29
Log of sigma	-0.17	0.06	-0.53	0.09	-0.16	0.04	-0.37	0.03
Sigma	0.84	0.05	0.59	0.05	0.85	0.03	0.69	0.02