NSHAP's Wave II Non-Response Weight Adjustment with Some Responses from Wave I Non-Respondents

Steven Pedlow[1], Colm O'Muircheartaigh[1], and Phil Schumm[2]
[1]NORC at the University of Chicago, 55 E. Monroe Suite 2000, Chicago, IL 60603
[2]Department of Health Studies, The University of Chicago, 5841 S. Maryland Ave. MC 2007, Chicago, IL 60637

Abstract

The National Social Life, Health, and Aging Project (NSHAP) is a longitudinal study of the health of older adults, concentrating on the role of social relationships in the aging process. In 2005 and 2006, NORC and a group of investigators at the University of Chicago interviewed a nationally representative sample of adults aged 57 to 85 for Wave I. In 2010 and 2011, these respondents as well as their spouses or cohabitating romantic partners were interviewed for Wave II.

Unlike most longitudinal surveys, we also attempted to interview in Wave II individuals who were sampled but declined to be interviewed in Wave I. This creates two alternatives in computing a non-response weight adjustment for Wave II. One possibility is that the Wave I non-respondents who responded in Wave II are most like the non-respondents to both waves, and thus should be in the same cells as the non-respondents to both waves, resulting in high weights for these cases and a larger design effect. An alternative is that other characteristics better explain the non-response, which divides the non-respondents to both waves across many cells, resulting in a smaller design effect. This paper evaluates and compares these two alternatives.

Key Words: Weighting Cells, Panel Surveys

1. Introduction

The health of older adults is influenced by many factors. One of the least understood is the role that social relationships—including marital, family, friends and others—play in health and aging. The National Social Life, Health, and Aging Project (NSHAP) is a longitudinal, U.S. population-based study of health and social factors, aiming to understand the well-being of older, community-dwelling Americans by examining the associations among physical health and illness, medication use, cognitive function, emotional health, sensory function, health behaviors, social connectedness, sexuality, and relationship quality. An overview can be found in Lindau et al., 2007. NSHAP provides policy makers, health providers, and individuals with useful information and insights into these factors, particularly those involving social and intimate relationships (see Smith et al., 2009; Cornwell et al., 2008; and Shiovitz-Ezra et al., 2009). Results from the study have implications for clinical practice and developing interventions to improve health as people age. NSHAP is funded by the National Institutes of Health, including the National

Institute on Aging, the Office of Research on Women's Health, and the Office of AIDS Research, with additional financial support provided by NORC.

NSHAP uses a national area probability sample of community residing adults born between 1920 and 1947 (aged 57 to 85 at the time of the Wave I interview), which includes an oversampling of African-Americans and Hispanics. In 2005 and 2006, NORC and investigators at the University of Chicago conducted the first wave of NSHAP, completing 3,005 interviews. In 2010 and 2011, 3,377 interviews were completed for Wave II including Wave I (returning) Respondents, Wave I Non-Interviewed Respondents, and their spouses or cohabiting romantic partners. The sample designs for Waves I and II are described in O'Muircheartaigh et al., 2009 and O'Muircheartaigh et al., 2013. The second wave of NSHAP is essential to understanding how social factors are related to within-respondent changes in health, and to health dynamics within the couple. By collecting a wide range of social and health measures (including a rich panel of biomeasures) from respondents over time, NSHAP provides data that will allow researchers in a number of fields to examine health changes during later life, and how these are related to several specific social processes.

As noted above, NSHAP attempted a Wave II interview with almost all of the eligible households selected for Wave I, including all non-interviews except for hostile refusals. The central question in this paper is whether Wave II interviews with Wave I non-respondents should be separated during the Wave II non-response weighting adjustment, or if Wave I response status should be ignored during the Wave II non-response weighting adjustment. Section 2 describes the Wave II non-response weighting adjustment. Section 3 lists the methodology to determine the answer to the question as well as the 30 variables used in our analyses. Section 4 describes the results of our analyses, and Section 5 summarizes our conclusions and our decision.

2. The Wave II Non-Response Weighting Adjustment

Wave I interviewed 3,005 adults aged 57-85, one per household. There were also 1,012 eligible non-interview households selected for Wave I, including 188 hostile refusals. Due to their age, only 85 percent of Wave I respondents were still alive and eligible (i.e., healthy enough to be interviewed) five years later. Among the 824 non-hostile-refusal cases, the eligibility rate for Wave II was only 75 percent.

Wave II respondents consisted of 3,377 respondents from four different sample types. First, 2,261 of the Wave I respondents were interviewed again. Since only 85 percent remained eligible, this is an 89 percent retention rate. Interviews were also completed with 907 spouses or cohabiting partners of these Wave I respondents, yielding a response rate of 86 percent. The third group of Wave II respondents were 161 Wave I non-respondents, representing a 26 percent response rate. Finally, 48 partners of Wave I non-respondents were interviewed, representing a 65 percent response rate.

For the Wave I non-response weight adjustment, NSHAP used a simple cell sample-based weighting adjustment (Kalton and Kasprzyk, 1986) using two variables and six cells: three age categories (born 1920-1929, born 1930-1938, and born 1939-1947) and two urbanicity categories (urban and rural). This method assumes that once we control for a few key characteristics, non-respondents are like respondents. Response rates were higher for rural respondents and due to health issues, response rates were highest for the youngest age group and lowest for the oldest age group.

For the Wave II non-response weight adjustment, we separated prime respondents and non-respondents attempted in Wave I from the partners added in Wave II since the response mechanisms for the two groups are different. Also, since the partners can be outside the three age categories used in Wave I, we used four age categories (born 1929 and before, born 1930-1938, born 1939-1947, and born 1948 and later).

The major decision that we faced for Wave II was whether to separate the Wave I respondents and Wave I non-respondents into different cells for the non-response weighting adjustment. Thinking about the non-respondents to both waves, if they are more like the Wave I non-respondents who responded in Wave II, then we should separate the Wave I non-respondents from the Wave I respondents. If they are more like the Wave I respondents who did not respond to Wave II, then we don't have a straightforward solution. If they are more like respondents within the same demographic cells, then we should not separate the Wave I non-respondents from the Wave I respondents.

The main advantage to separating the Wave I respondents from the Wave I non-respondents is bias reduction. It seems plausible that the non-respondents to both waves are more like Wave I non-respondents than Wave I respondents. If so, separating the Wave I non-respondents will result in a larger adjustment for them, reducing the non-response bias in the dataset.

At the same time, separating the Wave I respondents and Wave I non-respondents has the disadvantage of increasing the variance of the resulting estimates. This occurs because the Wave I non-respondents, due to their Wave II response rate of only 26 percent, will have a non-response weight adjustment of around 4. This results in an increase in the variability of weights, which increases analysis variances as measured by the design effect.

From a design perspective, since almost all eligibles from Wave I were attempted again in Wave II (only the hostile refusals from Wave I were not attempted), it can be proper to treat Wave II as its own survey. We do not adjust Wave I weights based on Wave II non-responses, though this idea might hold promise if the bias reduction was sufficient. Under this approach, we should not adjust Wave II weights based on Wave I non-responses.

Finally, another theory is that the Wave I non-respondents interviewed in Wave II are similar replacements to the Wave I respondents who drop out. In other words, the theory is that these two groups of one-wave respondents are more similar to each other than to the two-wave respondents and two-wave non-respondents. Our analyses in this paper test these theories to help us decide whether to separate the Wave I respondents and Wave I non-respondents in the Wave II non-response weight adjustment.

3. Variables Used and Methodology

To answer the questions raised above, we used the thirty key variables in Table 1 below that were asked in both Wave I and Wave II. These thirty variables can be divided into four sets. The first set of nine variables contains questions asking if the respondent has difficulty with specific daily activities such as walking, dressing, or driving. The second set of six variables asks about incontinence and other urinary issues. The third set of

twelve variables are questions that ask about the respondent's happiness and feelings over the last week including whether the respondent has had trouble getting going, has been lonely, or has had a poor appetite. Finally, the fourth set of three variables consists of miscellaneous quantitative variables: two ask about the respondent's general (physical or mental) health, while the third asks about the frequency of masturbation. These thirty variables represent a mix of twenty-three categorical (simplified to binary) and seven quantitative variables.

Table 1: The Thirty Variables Used in Our Analyses

Set 1: Daily Activity Difficulty Set (9)

1. Walking Across Room	4. Dressing	7. Using Toilet
2. Walking a Block Outside	5. Eating	8. Driving at Night
3. Getting In and Out of Bed	6. Bathing	9. Driving During Day

Set 2: Incontinence Issues within the last year? (6)

10. Incontinence?	12. Other Urinary Issues	14. Stool Incontinence
11. Number of Episodes	13. Number of Episodes	15. Number of Episodes

Set 3: Happiness and Feelings in the Last Week (12)

16. Mostly Enjoying Life	20. Mostly Happy	24. Restless Sleep
17. People were Unfriendly	21. Lonely	25. Poor Appetite
18. Everything Takes Effort	22. Depressed	26. Felt Disliked
19. Could Not Get Going	23. Sadness	27. General Happiness

Set 4: Miscellaneous Quantitative Variables (3)

28. General Self-Reported Physical Health
29. General Self-Reported Mental Health
30. Frequency of Masturbation

Our first analysis (Section 4.1) is a simple comparison of the variances and coefficients of variation of two different weights: one in which Wave I respondents and Wave I non-respondents are completely separated during the non-response weight adjustment (the "Separation" weight) and one in which Wave I respondents and Wave I non-respondents are not specifically separated (the "No Separation weight") and the cells for the non-response weight adjustment are based only on age, urbanicity, and whether the respondent was attempted in Wave I or is a partner added during Wave II. This analysis also calculates mean-squared errors under both weights. Mean-squared errors are the sum of the bias squared and the variance of estimates. While it is possible to directly calculate estimated variances, actual biases are typically unknown. We know that the variances will be smaller for the "No Separation" weight. It is also reasonable to think the bias might be smaller for the "Separation" weight. For our comparison, we assumed that the "Separation" weight results in unbiased estimates and that the bias of the "No Separation" weight is the difference between the two estimates. This assumption clearly favors the "Separation" weight.

Our second analysis (Section 4.2) compared Wave I respondents and Wave I non-respondents using the thirty Wave II variables. This analysis simply examines whether there are differences between the Wave I respondents and non-respondents in Wave II.

Our third analysis (Section 4.3) tests the theory that the Wave I non-respondents who are interviewed in Wave II are similar to the Wave I dropout respondents who did not complete a Wave II interview. Since the thirty questions in Table 1 were asked in both waves, we can compare Wave I answers for the Wave I dropout respondents with the Wave II answers given by the Wave I non-respondents.

Our fourth and final analysis looks at the differences between two-wave respondents and one-wave respondents for each of the two waves. We already compared Wave I respondents and Wave I non-respondents on Wave II variables in our second analysis (Section 4.2). In this analysis, we compare the Wave II respondents and Wave II non-respondents using the Wave I variables. Our purpose is to examine whether the Wave I differences are similar to the Wave II differences.

4. Analysis Results

4.1 Variability and Mean-Squared Errors of Weights

Table 2 shows some simple comparisons between the two different weights that we considered. The "Separation" weight separates the Wave I respondents and the Wave I non-respondents into different cells for the non-response weight adjustment while the "No Separation" weight allows Wave I respondents and Wave I non-respondents to be in the same cell if they are in the same age, urbanicity, and prime/partner categories.

Table 2: Comparison of Variability for Two Weights

Weight	*Mean*	*Standard Deviation*	*Coefficient of Variation*	*Weighting Design Effect*	*Effective Sample Size*
"Separation"	1.000	0.954	0.954	1.909	1,769
"No Separation"	1.000	0.773	0.773	1.597	2,115

Both weights have been re-scaled to sum to the total sample size of 3,377, so the means of both weights are 1. However, the standard deviation for the "Separation" weight is 23 percent larger. Since the coefficient of variation is the standard deviation divided by the mean, it is equal to the standard deviation. Kish (1965) approximated the design effect due to differential weighting as one plus the square of the coefficient of variation. Table 2 shows that the weighting design effect is therefore estimated to be 19.5 percent higher for the "Separation" weight. Since the effective sample size is the total sample size (3,377) divided by the design effect, the effective sample size due to weighting is 350 less for the "Separation" weight.

While the "No Separation" weights have less variability, they could still result in higher mean-squared errors if they result in a high bias. It is not possible to calculate bias since we do not know the true values for our thirty variables in the population of interest. For the purposes of this analysis, we assume that the "Separation" weight is unbiased and assume that the bias of the "No Separation" weight is the difference in estimates from the two weights. Since this assumption provides a clear advantage to the "Separation" weight, we will be able to determine if the possible bias could outweigh the variance advantage of the "No Separation" weight.

Table 3 below shows the results for this analysis, with the means, standard errors (SE), and mean-squared errors (MSE) under both weights. For each of the thirty variables, the lower mean-squared error is in **bold.**

Table 3: Mean-Squared Error Comparison for Two Weights

Variable	*"Separation" Weight*			*"No Separation" Weight*		
	Mean	*SE*	*MSE*	*Mean*	*SE*	*MSE*
Set 1: Daily Difficulties						
Walking Across Room	0.1311	0.0074	0.000054	0.1313	0.0069	**0.000047**
Walking a Block Outside	0.2486	0.0118	0.000140	0.2484	0.0115	**0.000132**
Getting In and Out of Bed	0.0945	0.0073	**0.000054**	0.0987	0.0068	0.000064
Dressing	0.1354	0.0078	0.000061	0.1357	0.0070	**0.000049**
Eating	0.0320	0.0032	**0.000010**	0.0355	0.0034	0.000024
Bathing	0.0846	0.0075	0.000056	0.0833	0.0067	**0.000046**
Using Toilet	0.0927	0.0078	0.000061	0.0938	0.0068	**0.000048**
Driving at Night	0.3582	0.0127	0.000161	0.3544	0.0118	**0.000154**
Driving During Day	0.0985	0.0091	0.000084	0.1002	0.0081	**0.000069**
Set 2: Incontinence						
Within last year?	0.3968	0.0181	0.000328	0.3995	0.0157	**0.000254**
Number of Episodes	62.1904	4.4715	19.994250	62.0857	4.4088	**19.448642**
Other Urinary Problems	0.2247	0.0142	**0.000201**	0.2300	0.0148	0.000247
Number of Episodes	44.5640	3.6995	**13.686056**	44.8315	4.1195	17.041969
Stool Incontinence	0.1050	0.0096	0.000092	0.1072	0.0084	**0.000075**
Number of Episodes	6.0213	1.7651	3.115638	4.7251	0.9170	**2.520952**
Set 3: Happiness/Feelings						
Mostly Enjoying Life	0.8589	0.0079	**0.000063**	0.8675	0.0073	0.000128
People Were Unfriendly	0.1316	0.0088	**0.000078**	0.1356	0.0082	0.000083
Everything Takes Effort	0.3746	0.0144	0.000208	0.3718	0.0137	**0.000196**
Could Not Get Going	0.4348	0.0120	0.000145	0.4376	0.0115	**0.000141**
Mostly Happy	0.7836	0.0108	**0.000116**	0.7914	0.0099	0.000159
Lonely	0.3163	0.0105	**0.000111**	0.3110	0.0093	0.000115
Depressed	0.3242	0.0124	**0.000153**	0.3152	0.0115	0.000212
Sadness	0.3974	0.0121	**0.000145**	0.3914	0.0107	0.000151
Sleep is Restless	0.4983	0.0131	**0.000172**	0.5067	0.0135	0.000253
Poor Appetite	0.2083	0.0091	**0.000082**	0.2135	0.0086	0.000101
Felt Disliked	0.1069	0.0067	**0.000045**	0.1109	0.0065	0.000059
General Happiness	2.3956	0.0203	0.000414	2.3882	0.0179	**0.000377**
Set 4: Miscellaneous						
General Physical Health	3.2397	0.0361	**0.001302**	3.2558	0.0373	0.001652
General Mental Health	3.6525	0.0277	0.000769	3.6589	0.0269	**0.000767**
Frequency of Masturbation	14.5245	1.9143	3.664483	14.0859	1.8291	**3.537904**

Overall, the "No Separation" weight has a lower mean-squared error for sixteen of the thirty variables even though we assumed the bias for the "Separation" weight was zero. For five of the variables, the estimated standard error is smaller for the "Separation" weight, guaranteeing a lower mean-squared error under our bias assumption. These

results suggest that the variance differential is likely to be larger than the bias differential, indicating that the "No Separation" weight should be preferred.

The preferred weight (by lowest mean-squared error) does differ for the four sets of variables. For the nine Daily Activity Difficulty variables, the "No Separation" weight has the lower mean-squared error for seven variables; the "Separation" weight has the lower mean-squared error only for Getting In and Out of Bed and for Eating. For the six Incontinence variables, the "No Separation" weight has the lower mean-squared error for four variables; the "Separation" weight has the lower mean-squared error only for the two variables on Other Urinary Problems. For the twelve Happiness and Feeling variables, the "Separation" weight has the lower mean-squared error for nine variables; the "No Separation" weight has the lower mean-squared error only for the Everything Takes Effort, Could Not Get Going, and General Happiness variables. Finally, for the three miscellaneous continuous variables, the "No Separation" weight has the lower mean-squared error for two variables; the "Separation" weight has the lower mean-squared error for only the General Physical Health variable.

4.2 Comparison of Wave I respondents and Wave I non-respondents

Our second analysis compares the Wave I respondents (W1R) and Wave I non-respondents (W1NR) on the thirty Wave II variables. We compared prime respondents and partners separately, and the results are shown in Table 4, with p-values less than 0.05 in **bold**.

Since there were fourteen variables in Section 4.1 where the "Separation" weight had a lower mean-squared error, we expected to see larger differences between Wave I respondents and Wave I non-respondents among these fourteen variables. However, we were surprised to find significant differences for only three variables among prime respondents. Wave I respondents had more difficulty Getting In and Out of Bed and Eating, and Wave I respondents were more likely to have Restless Sleep. This means that the Wave I non-respondents were slightly more healthy, but that the two groups were not very different.

For the partners, we found eight significant differences between the Wave I respondent partners and Wave I non-respondent partners. Most notably, the Wave I respondent partners had significantly more difficulty with seven of the nine daily activity variables, with no significant differences only on Getting In and Out of Bed and Driving at Night. Wave I respondent partners were also more likely to Feel Disliked. The partners of Wave I non-respondents were healthier than the partners of Wave I respondents.

There were no significant differences between Wave I respondents and Wave I non-respondents on the Incontinence (even though two of the differences in Number of Episodes seem large) and Miscellaneous Quantitative variables, either among prime respondents or partners.

4.3 Are Wave I non-respondents similar to Wave I respondent dropouts?

In this third analysis, we compare the one-wave completes. Some of the Wave I respondents did not complete a Wave II interview. We classify such cases as Wave II dropouts. Meanwhile, we completed a Wave II interview for some of the Wave I non-respondents. We classify such cases as Wave II refusal conversions. In this analysis, we compare these two groups to determine whether they are similar to each other. If they are

similar, it could be said that the Wave II refusal conversions are similar replacements to the Wave I dropouts. Figure 1 below shows the full response pattern to the two waves, as well as the two groups (marked with diagonal lines) that we are comparing in this analysis. Responses are shown in Green, non-responses are shown in Yellow, and Hostile Refusals are shown in Red.

Table 4: Comparison of Wave I respondents and Wave I non-respondents in Wave II

Variable	*Prime Respondents*			*Partner Respondents*		
	W1R	*W1NR*	*p-value*	*W1R*	*W1NR*	*p-value*
Set 1: Daily Difficulties						
Walking Across Room	0.1469	0.1199	0.3712	0.0919	0.0110	< **0.0001**
Walking a Block Outside	0.2599	0.2370	0.5137	0.2231	0.1016	**0.0114**
Getting In and Out of Bed	0.1110	0.0590	**0.0429**	0.0717	0.0503	0.5771
Dressing	0.1452	0.1267	0.6050	0.1148	0.0110	< **0.0001**
Eating	0.0419	0.0083	<**0.0001**	0.0234	0.0000	< **0.0001**
Bathing	0.0910	0.0924	0.9642	0.0610	0.0138	**0.0006**
Using Toilet	0.1029	0.0839	0.4426	0.0721	0.0150	**0.0059**
Driving at Night	0.3668	0.3760	0.8723	0.3176	0.2552	0.4035
Driving During Day	0.1132	0.0959	0.5296	0.0654	0.0222	**0.0336**
Set 2: Incontinence						
Within last year?	0.3768	0.3274	0.4594	0.4701	0.6174	0.2502
Number of Episodes	61.5652	57.9710	0.8220	61.8942	110.6644	0.1921
Other Urinary Problems	0.2257	0.1754	0.3091	0.2487	0.3196	0.5350
Number of Episodes	46.3051	41.6511	0.7170	40.5553	54.7421	0.5963
Stool Incontinence	0.1044	0.0579	0.1517	0.1232	0.1412	0.8364
Number of Episodes	4.1172	11.6452	0.2998	4.0978	26.3158	0.1991
Set 3: Happiness/Feelings						
Mostly Enjoying Life	0.8766	0.8267	0.1283	0.8472	0.8699	0.6590
People Were Unfriendly	0.1406	0.0939	0.0947	0.1303	0.1106	0.7611
Everything Takes Effort	0.3668	0.3871	0.6646	0.3883	0.2999	0.2684
Could Not Get Going	0.4455	0.4101	0.4522	0.4219	0.3793	0.5792
Mostly Happy	0.7935	0.7450	0.2255	0.7941	0.7997	0.9272
Lonely	0.3343	0.3584	0.6192	0.2315	0.2568	0.7488
Depressed	0.3200	0.3736	0.2367	0.2829	0.4047	0.2103
Sadness	0.3969	0.4247	0.5732	0.3643	0.4417	0.3980
Sleep is Restless	0.5207	0.4166	**0.0495**	0.4831	0.4873	0.9674
Poor Appetite	0.2240	0.1868	0.3180	0.1886	0.1796	0.9020
Felt Disliked	0.1142	0.0720	0.0531	0.1125	0.0441	**0.0064**
General Happiness	2.4162	2.4377	0.8067	2.2914	2.3707	0.5595
Set 4: Miscellaneous						
General Physical Health	3.2532	2.8514	0.4361	3.2774	2.5846	0.4767
General Mental Health	3.6684	2.3953	0.4738	3.6406	2.3371	0.9016
Frequency of Masturbation	13.7733	15.4973	0.7842	13.8879	28.6119	0.5220

For this analysis, we compare the 744 Wave I only respondents using Wave I data to the 161 Wave II only respondents using Wave II data. However, there were five or six years

between the waves, and this results in the Wave II only respondents answering the same questions at older ages.

	Respondents to Both Waves	Wave I Only	Wave II Only	Neither Wave	Hostile Refusals	Total Respondents
Wave I						3,005 in Wave I
Wave II						2,422 in Wave II
Eligibles in Group	2,261	744	161	663	188	

Figure 1: Response Pattern to Waves I and II

Wave I only respondents were aged 57-85 in Wave I while Wave II only respondents were aged 63-91 in Wave II. Therefore, we restricted our analysis to ages 63-85, which reduced the sample sizes to 649 Wave I only respondents and 149 Wave II only respondents.

Table 5 shows that there are eleven statistically significant differences between the two groups (with p-values less than 0.05 in **bold)**. The Wave I only respondents have more difficulty for all nine daily activities and the differences for seven of them are statistically significant. The two driving variables are the only two that are not significant. The Wave I only respondents also had more negative outcomes for all of the Happiness and Feeling variables, but only three (Could Not Get Going, Poor Appetite, and Feel Disliked) were statistically significant. Among the Miscellaneous Quantitative variables, Wave I only respondents had significantly poorer General Physical Health scores. There were no significant differences among the incontinence variables.

This analysis shows that the one-wave respondents are not similar. The Wave I only respondents are less healthy than the Wave II only respondents. This is not surprising since The Wave I only respondents include not just Wave II refusals, but those who are no longer in scope (in other words, those who have died or become too sick to be interviewed). The Wave I only respondents have more negative outcomes on almost all of the health, happiness, and feeling variables, even though many of these are not statistically significant differences.

4.4 Are the Respondent/Non-Respondent Differences Similar?

In our fourth and final analysis, we examine whether the differences between respondents and non-respondents are consistent between the two waves. Figure 2 shows the comparisons in this analysis. We already compared the Wave I non-respondents (C_2 in Figure 2) and the Wave I respondents (A_2 in Figure 2) in Section 4.2. Now, we compare in Table 6 the Wave II non-respondents (B_1 in Figure 2) and the Wave II respondents (A_1 in Figure 2) to see if the differences are similar.

Table 5: Comparison of Wave I only respondents and Wave II only respondents

Variable	*W1R W2NR*		*W1NR W2R*		*Statistics*	
	Mean	*SE*	*Mean*	*SE*	*F-value*	*p-value*
Set 1: Daily Difficulties						
Walking Across Room	0.2089	0.0180	0.1092	0.0305	6.54	**0.0136**
Walking a Block Outside	0.4329	0.0237	0.2163	0.0356	22.73	**<0.0001**
Getting In and Out of Bed	0.1576	0.0183	0.0598	0.0246	11.01	**0.0017**
Dressing	0.2135	0.0237	0.1258	0.0334	4.87	**0.0319**
Eating	0.0685	0.0110	0.0023	0.0023	34.35	**<0.0001**
Bathing	0.1750	0.0175	0.0838	0.0288	6.37	**0.0124**
Using Toilet	0.1523	0.0182	0.0768	0.0260	6.14	**0.0166**
Driving at Night	0.4292	0.0186	0.3650	0.0542	1.23	0.2718
Driving During Day	0.1451	0.0176	0.0866	0.0292	3.15	0.0822
Set 2: Incontinence						
Within last year?	0.3360	0.0381	0.3121	0.0672	0.10	0.7560
Number of Episodes	61.5336	9.8034	49.0723	12.9124	0.59	0.4443
Other Urinary Problems	0.2470	0.0395	0.1766	0.0465	1.19	0.2763
Number of Episodes	49.2224	12.4959	42.2168	11.0963	0.17	0.6825
Stool Incontinence	0.1255	0.0280	0.0608	0.0313	2.25	0.1352
Number of Episodes	10.8307	3.5484	12.2487	7.8140	0.02	0.8776
Set 3: Happiness/Feelings						
Mostly Enjoying Life	0.8139	0.0204	0.8327	0.0304	0.26	0.6109
People Were Unfriendly	0.1579	0.0149	0.0936	0.0302	2.95	0.0918
Everything Takes Effort	0.4521	0.0248	0.3832	0.0481	1.56	0.2169
Could Not Get Going	0.5166	0.0222	0.4007	0.0455	4.27	**0.0441**
Mostly Happy	0.7069	0.0216	0.7462	0.0388	0.67	0.4171
Lonely	0.3702	0.0239	0.3459	0.0477	0.22	0.6432
Depressed	0.3813	0.0180	0.3616	0.0461	0.16	0.6919
Sadness	0.4360	0.0223	0.4186	0.0534	0.09	0.7683
Sleep is Restless	0.5020	0.0227	0.4154	0.0488	2.28	0.1372
Poor Appetite	0.3463	0.0197	0.1856	0.0362	14.88	**0.0003**
Felt Disliked	0.1622	0.0161	0.0761	0.0215	10.67	**0.0020**
General Happiness	2.5129	0.0456	2.4263	0.0846	0.65	0.4251
Set 4: Miscellaneous						
General Physical Health	2.8420	0.0606	3.1726	0.1282	5.17	**0.0273**
General Mental Health	3.5460	0.0534	3.6055	0.0824	0.37	0.5443
Frequency of Masturbation	9.6360	2.9320	16.3200	6.1466	0.91	0.3401

While Table 4 shows that Wave I prime respondents are not very different from Wave I prime non-respondents in Wave II, Table 6 shows that Wave II prime respondents are very different from Wave II prime non-respondents. There are twenty-one statistically significant differences in Table 6 (with p-values less than 0.05 in **bold)**. Wave II prime respondents are significantly less likely to have difficulty with all nine of the daily activities, have significantly more positive outcomes on ten of the twelve happiness/feeling variables, and also report significantly better physical and mental health. Six of the nine variables that are not significantly different are the six incontinence variables that have shown no significant differences in any of our analyses.

The Wave I respondents who dropped out (or died or became too sick to interview) in Wave II are significantly less healthy than the Wave I respondents who were also interviewed in Wave II.

	Respondents to Both Waves	Wave I Only	Wave II Only	Neither Wave	Hostile Refusals	Total Respondents
Wave I	A_1	B_1				3,005 in Wave I
Wave II	A_2		C_2			2,422 in Wave II
Eligibles in Group	2,261	744	161	663	188	

Figure 2: Comparisons of One-Wave Respondents to Two-Wave Respondents

5. Conclusions and Decision

In this paper, we have compared two different weights based on whether the Wave II non-response weighting adjustment separates the Wave I respondents from the Wave I non-respondents (the "Separation" weight) or not (the "No Separation" weight).

Our first analysis showed that the "Separation" weight had a standard deviation (and coefficient of variation) 23 percent larger than the "No Separation" weight, resulting in a weighting design effect that is 19.5 percent larger and a reduced effective sample size of 350 interviews. We also compared mean-squared errors for the two weights. The "No Separation" weight had a lower mean-squared error for sixteen out of thirty variables despite the assumption that the "Separation" weight was unbiased and any difference between the two weights' estimates was assumed to be a bias for the "No Separation" weight.

Our second analysis showed that the Wave I respondents and Wave I non-respondents were significantly different for only three of thirty Wave II variables, but the partners of Wave I non-respondents had significantly more difficulty for seven of the nine daily activity variables.

Our third analysis showed that Wave I non-respondents were healthier in Wave II than Wave II non-respondents were in Wave I. Wave II non-respondents had more negative outcomes for almost all of the variables, even though some were not statistically significant. The Wave II non-respondents had significantly more difficulty for seven of the nine daily activity variables, reported significantly poorer General Physical Health, and had more negative happiness and feeling outcomes for all twelve variables, even though the differences were only statistically significant for three of them.

Finally, while the second analysis showed that Wave I prime respondents and Wave I prime non-respondents were not very different in Wave II, the fourth analysis showed

that Wave II respondents and Wave II non-respondents were very different in Wave I. Wave II non-respondents had poorer health outcomes than Wave II respondents.

Table 6: Comparison of Wave II respondents and Wave II non-respondents

Variable	*Prime Respondents*		
	W2R	*W2NR*	*p-value*
Set 1: Daily Difficulties			
Walking Across Room	0.0776	0.2069	**<0.0001**
Walking a Block Outside	0.1890	0.4195	**<0.0001**
Getting In and Out of Bed	0.0929	0.1477	**0.0025**
Dressing	0.1326	0.2186	**0.0002**
Eating	0.0295	0.0716	**<0.0001**
Bathing	0.0657	0.1753	**<0.0001**
Using Toilet	0.0890	0.1502	**0.0015**
Driving at Night	0.2912	0.3987	**<0.0001**
Driving During Day	0.0396	0.1309	**<0.0001**
Set 2: Incontinence			
Within last year?	0.3845	0.3292	0.1450
Number of Episodes	51.2697	56.2576	0.6094
Other Urinary Problems	0.2262	0.2557	0.5019
Number of Episodes	37.7210	50.2037	0.2973
Stool Incontinence	0.0828	0.1092	0.3707
Number of Episodes	3.1247	8.7956	0.1251
Set 3: Happiness/Feelings			
Mostly Enjoying Life	0.8633	0.8095	**0.0145**
People Were Unfriendly	0.1596	0.1877	0.1545
Everything Takes Effort	0.3917	0.4586	**0.0152**
Could Not Get Going	0.4358	0.5094	**0.0042**
Mostly Happy	0.7757	0.6932	**0.0009**
Lonely	0.2739	0.3879	**<0.0001**
Depressed	0.2886	0.3922	**<0.0001**
Sadness	0.3787	0.4495	**0.0080**
Sleep is Restless	0.5503	0.5296	0.3725
Poor Appetite	0.2071	0.3453	**<0.0001**
Felt Disliked	0.1345	0.1732	**0.0239**
General Happiness	2.3389	2.5219	**0.0018**
Set 4: Miscellaneous			
General Physical Health	3.3880	2.8596	**<0.0001**
General Mental Health	3.8495	3.5437	**<0.0001**
Frequency of Masturbation	16.0072	11.0299	0.1534

Summarizing our evidence, the "No Separation" weight results in a larger effective sample size and slightly outperforms the "Separation" weight on mean-squared error even when we assume the "Separation" weight is unbiased and the difference between estimates for the two weights is the bias for the "No Separation" weight. The Wave I non-responses interviewed in Wave II (refusal conversions) are healthier than the Wave

II non-responses interviewed in Wave I (dropouts), but it is not because the Wave I respondents are different from the Wave I non-respondents in Wave II, it is because the Wave II respondents are much healthier than the Wave II non-respondents in Wave I. Therefore, our conclusion is that the "Separation" weight does not reduce bias while increasing the variability, so the "No Separation" weight is preferred. In fact, our analyses suggest that bias reduction has more promise in separating the Wave II respondents from the Wave II non-respondents in the Wave I non-response weighting adjustment. Of course, these weights have been used by analysts for five years, so it is undesirable to change them now.

NSHAP research data (with no identifiable respondent information) is available to the public. De-identified data from the Wave I and Wave II interviews are available to researchers through the National Archive of Computerized Data on Aging, located within Inter-University Consortium for Political and Social Research (ICPSR). The Wave II data was originally released with "Separation" weights, but these have now been replaced with the "No Separation" weights.

References

Cornwell, B., Laumann, E., and Schumm P. (2008). The Social Connectedness of Older Adults: A National Profile. American Sociological Review, Volume 73, pp. 185-203.

Kalton, G. and D. Kasprzyk (1986). The Treatment of Missing Survey Data. Survey Methodology, Vol. 12, pp. 1-16.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons, Inc.

Lindau, S., Schumm P., Laumann, E., Levinson, W., O'Muircheartaigh, C., and Waite, L. (2007). A Study of Sexuality and Health among Older Adults in the United States. New England Journal of Medicine, Volume 357, pp. 762-774.

O'Muircheartaigh, C., Eckman, S., and Smith, S. (2009). Statistical Design and Estimation for the National Social Life, Health, and Aging Project. Journal of Gerontology, Series B, Volume 64B, pp. i12-i19.

O'Muircheartaigh, C., English, N., Pedlow, S., and Kwok, P. (2013). Sample Design, Sample Augmentation, and Estimation for Wave II of the National Social Life, Health, and Aging Project. Journal of Gerontology, Series B, in press.

Shiovitz-Ezra, S., Leitsch, S., Graber, J., and Karraker, A. (2009). Quality of Life and Psychological Health Indicators in the National Social Life, Health, and Aging Project. Journal of Gerontology, Series B, Volume 64B, pp. i30-i37.

Smith, S., Jaszczak, A., Graber, J., Lundeen, K., Leitsch, S., Wargo, E., and O'Muircheartaigh, C. (2009). Instrument Development, Study Design Implementation, and Survey Conduct for the National Social Life, Health, and Aging Project. In Researching Social Gerontology, Four-Volume Set, edited by Cutchin, M., Kemp, C., and Marshall, V., Volume 4, SAGE Publications: Thousand Oaks, California.