ABSTRACT

Title of Document:	ENHANCING THE UNDERSTANDING OF THE RELATIONSHIP BETWEEN SOCIAL INTEGRATION AND NONRESPONSE IN HOUSEHOLD SURVEYS
	Ashley E. Amaya, Doctor of Philosophy, 2015
Directed By:	Professor Stanley Presser Joint Program in Survey Methodology

Nonresponse and nonresponse bias remain fundamental concerns for survey researchers as understanding them is critical to producing accurate statistics. This dissertation tests the relationship between social integration, nonresponse, and nonresponse bias.

Using the rich frame information available on the American Time Use Survey (ATUS) and the Survey of Health, Ageing, and Retirement in Europe (SHARE) Wave II, structural equation models were employed to create latent indicators of social integration. The resulting variables were used to predict nonresponse and its components (e.g., noncontact). In both surveys, social integration was significantly predictive of nonresponse (regardless of type of nonresponse) with integrated individuals more likely to respond. However, the relationship was driven by different components of integration across the two surveys.

Full sample estimates were compared to respondent estimates on a series of 40 dichotomous and categorical variables to test the hypothesis that variables measuring social activities and roles would suffer from nonresponse bias. The impact of nonresponse on multivariate models predicting social outcomes was also evaluated. Nearly all of the 40 assessed variables suffered from significant nonresponse bias

resulting in the overestimation of social activity and role participation. In general, civic and political variables suffered from higher levels of bias, but the differences were not significant. Multivariate models were not exempt; beta coefficients were frequently biased. Although, the direction was inconsistent and often small.

Finally, an indicator of social integration was added to the weighting methodology with the goal of eliminating the observed nonresponse bias. While the addition significantly reduced the bias in most instances compared to both the base- and traditionally-weighted estimates, the improvements were small and did little to eliminate the bias.

ENHANCING THE UNDERSTANDING OF THE RELATIONSHIP BETWEEN SOCIAL INTEGRATION AND NONRESPONSE IN HOUSEHOLD SURVEYS

By

Ashley Elaine Amaya

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Distinguished University Professor Stanley Presser, Chair
Professor Katharine Abraham
Associate Professor Jeffrey Harring
Research Associate Professor Jill Montaquila
Research Professor Richard Valliant

Acknowledgements

Thank you JPSM. This program gave me the opportunity to learn and grow in ways that I never expected. Thank you to Rick, Fred, and Frauke who helped make it financially feasible for me to be a full-time student. A special thanks to my committee: Stanley, Jeff, Jill, Katharine, and Rick. I am grateful for your willingness to spend countless hours reading drafts, teaching me analytic techniques, and providing feedback on my ideas.

I am also indebted to the other members of my cohort for all of the times they listened to practice presentations, answered questions, and were willing to grab lunch when I needed a break.

Thank you to my mentors at NORC, especially Felicia LeClere. Felicia encouraged me to go back to school. NORC made it feasible for me to relocate to Maryland and adjusted my workload so that I had the time to study. I am grateful for the nearly eight years of opportunities, experiences, and support that they granted me.

Table of Contents

Acknowledgements ii
Table of Contents iii
List of Tables v
List of Figures viii
Chapter 1: Introduction 1
1.1 The Social Integration Hypothesis 1
1.2 Previous Investigations of Social Integration in Survey Research 5
1.3 Alternative Measures of Social Integration 13
1.4 Research Hypotheses 20
Chapter 2: Data 25
2.1 American Time Use Survey (ATUS) 26
2.2 Survey of Health, Ageing and Retirement in Europe (SHARE) Wave II 34
Chapter 3: Social Integration and Nonresponse 44
3.1 Methods 44
3.1.1 Constructing the Social Integration Measures 44
3.1.2 Testing the Effect of Social Integration on Nonresponse 55
3.2 Results 60
3.2.1 Constructing the Social Integration Measures 60
3.2.2 Testing the Effect of Social Integration on Nonresponse 74
3.3 Discussion 81
Chapter 4: Social Integration and Nonresponse Bias 85
4.1 Methods 85
4.1.1 Univariate Bias 85
4.1.2 Comparative Bias 88
4.1.3 Multivariate Bias 90
4.2 Results 92
4.2.1 Univariate Bias 92
4.2.2 Comparative Bias 98
4.2.3 Multivariate Bias 105
4.3 Discussion 120

Chapter 5: Social Integration and Weighting 121
5.1 Methods 121
5.1.1 Simplifying the Social Integration Indicator 121
5.1.2 Constructing ATUS Weights 128
5.1.3 Constructing SHARE Wave II Weights 130
5.1.4 Evaluating the Effectiveness of the Weights 132
5.2 Results 134
5.2.1 Constructing the ATUS Social Integration Indicator and Weights 134
5.2.2 Constructing the SHARE Wave II Social Integration Indicator and Weights 139
5.2.3 Evaluating the Effectiveness of the Weights 143
5.3 Discussion 156
Chapter 6: Conclusions 160
References 165

List of Tables

Table 1: Demographic Distribution by Sample Type (ATUS) 29

Table 2: Demographic Distribution by Sample Type (SHARE) 36

Table 3: Fit Statistics of Latent Class Analyses by Survey and Number of Classes 63

Table 4: Fit Statistics of Exploratory Factor Analyses by Survey and Number of Classes 68

Table 5: ESEM Factor Loadings (ATUS) 71

Table 6: ESEM Factor Loadings (SHARE) 74

Table 7: Binomial Logit Predicting Response (N=4,821) (ATUS) 76

Table 8: Binomial Logit Predicting Response (N=18,536) (SHARE) 77

Table 9: Multinomial Logistic Model Predicting Nonresponse (N=4,776) (ATUS) 80

Table 10: Differences in Social Activity and Role Estimates by Sample Type (ATUS) . 93

Table 11: Differences in Social Activity and Role Estimates by Sample Type (SHARE)94

Table 12: Differences in Social Activity and Role Distributions by Sample Type (ATUS) 97

Table 13: Differences in Social Activity and Role Distributions by Sample Type (SHARE) 98

Table 14: Logit Predicting Various Social Activities and Social Roles (ATUS) 106

Table 15: Logit Predicting Various Social Activities and Social Roles (SHARE) 111

Table 16: Number of Significant Differences Found in the Multivariate Models in Table 14 by Significance Level and Dependent Variable (ATUS) 116

Table 17: Number of Significant Differences Found in the Multivariate Models in Table 15 by Significance Level and Dependent Variable (SHARE) 116

Table 18: Number of Significant Differences Found in the Multivariate Models in Table 14 by Significance Level and Independent Variable (ATUS) 117

Table 19: Number of Significant Differences Found in the Multivariate Models in Table 15 by Significance Level and Independent Variable (SHARE) 118

Table 20: Summary of Changes in the Level of Significance of the Beta Coefficients Found in the Multivariate Models in Table 14 (ATUS) 119

Table 21: Summary of Changes in the Level of Significance of the Beta Coefficients Found in the Multivariate Models in Table 15 (SHARE) 120

Table 22: Model Fit of Various Reduced LCA Models (ATUS) 135

Table 23: Differences in Social Activity and Role Estimate by Social Integration Category (ATUS).. 138

Table 24: Descriptive Statistics of the Distribution of the Weights by Weighting Scheme (ATUS) 139

Table 25: Model Fit of Various Reduced LCA Models (SHARE)................................ 140

Table 26: Differences in Social Activity and Role Estimate by Social Integration Category (SHARE) 142

Table 27: Descriptive Statistics of the Distribution of the Weights by Weighting Scheme (SHARE)... 143

Table 28: Quality Indicators of Various Social Activities & Roles under Different Weighting Schemes (ATUS) 145

Table 29: Quality Indicators of Various Social Activities & Roles under Different Weighting Schemes (SHARE)... 147

Table 30: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Base Weights (ATUS) 150

Table 31: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Base Weights (SHARE) 150

Table 32: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Traditional Weights (ATUS) 152

Table 33: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Traditional Weights (SHARE)... 152

Table 34: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Full Sample Estimates (ATUS)... 154

Table 35: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Full Sample Estimates (SHARE) 154

Table 36: Bias of Social Activity & Role Estimates Using Traditional Weights Vs. Full Sample Estimates (ATUS)... 155

Table 37: Bias of Social Activity & Role Estimates Using Traditional Weights Vs. Full Sample Estimates (SHARE) 155

List of Figures

Figure 1: CPS Question Wording of Social Activities and Social Roles by Route to Integration (ATUS) 30

Figure 2: SHARE Wave I Question Wording of Social Activities and Social Roles by Route to Integration (SHARE) 38

Figure 3: Structural Equation Model of Social Integration Using Latent Class Analysis (ATUS) 61

Figure 4: Structural Equation Model of Social Integration Using Latent Class Analysis (SHARE) 62

Figure 5: Item Probabilities by Social Integration Class (ATUS) 64

Figure 6: Item Probabilities by Social Integration Class (SHARE) 66

Figure 7: Structural Equation Model of Social Integration Using Exploratory Structural Equation Modeling (ATUS) 70

Figure 8: Structural Equation Model of Social Integration Using Exploratory Structural Equation Modeling (SHARE) 73

Figure 9: Relative Difference in Univariate Estimates ((Respondents-Full Sample)/Full Sample) (ATUS) 100

Figure 10: Relative Difference in Univariate Estimates ((Respondents-Full Sample)/Full Sample) (SHARE) 101

Figure 11: Absolute Difference of Univariate Estimates (Respondents-Full Sample) (ATUS) 103

Figure 12: Absolute Difference of Univariate Estimates (Respondents-Full Sample) (SHARE) 104

Figure 13: Variable Subsets Tested for LCA Used in Weighting (ATUS) 123

Figure 14: Variable Subsets Tested for LCA Used in Weighting (SHARE) 123

Figure 15: Latent Class Model of Social Integration for Weighting (ATUS) 137

Figure 16: Latent Class Model of Social Integration for Weighting (SHARE) 141

Chapter 1: Introduction

Researchers have tried to explain nonresponse since the inception of probability-based surveys. Understanding nonresponse may be useful in order to increase response rates, identify when estimates will suffer from nonresponse bias, and improve weighting methods to correct nonresponse bias. Studies of nonresponse have often been informed by a hypothesis about *social integration* which posits that individuals who are more integrated into society (i.e., individuals who participate in a broad range of social relationships) are more likely to respond to a survey request while individuals who are socially isolated are less likely to participate (Goyder 1987; Groves & Couper 1998).

1.1 The Social Integration Hypothesis

There are three grounds for the hypothesis that social integration is related to survey response. First, individuals wish to fit in with their social groups. Social groups range from formal organizations (e.g., civic organizations) to informal or abstract groups (e.g., neighbors). Regardless of the form, if individuals perceive that survey participation is consistent with the expectations of group members, they will agree to be respondents. Groves and Couper (1998) used this rationale to explain why membership surveys often achieve higher response rates than general population surveys. Additionally, individuals will often behave in the manner they believe other group members would act. If one believes that other members would complete the survey, that will increase the chance that the individual will (Schwartz 1977; Groves, Cialdini, & Couper 1992; Brehm 1993). Second, people may participate to avoid consequences (perceived or real) of not participating (Putnam 1993). Research indicates that individuals who act outside group norms are often treated poorly (Wolfensberger & Tullman 1982). While Wolfensberger and Tullman focus on relatively extreme examples of deviant behavior, individuals who decline to participate in a survey may still face negative consequences. For example, a neighbor may see an interviewer arrive at a sampled individual's door. The sampled individual may choose to participate in order to avoid disapproval from the neighbor. Perceived consequences may be as important as real consequences. Even if the individual would not be reprimanded for failure to participate, the individual may not correctly perceive that. Finally, individuals who are socially integrated internalize a set of social norms learned from their group membership(s) (Schwartz 1977). Failure to comply with these norms may result in a sense of guilt. The individual may weigh such guilt as a negative consequence and factor it into their response decision.

Third, individuals who are more socially integrated may be more likely to feel that their participation in a survey will yield personal or group benefits in the long-term (e.g., safer neighborhoods or more school funding) (Gouldner 1960; Groves et al 1992; Putnam 1993; Goyder, Boyer, & Martinelli 2006). In this way, the social integration hypothesis is linked to social exchange theory, which asserts that individuals weigh social costs and benefits in their decision to acquiesce to or decline a survey request (Dillman, 1978). The social integration hypothesis expands on social exchange theory by explaining who will be more likely to perceive a social benefit. Individuals are more likely to perceive a social benefit or assign a greater weight to the benefit if they feel connected to the group(s) that will profit from the information gained from the survey results. The more socially integrated an individual, the more likely the potential benefits of participation may apply to multiple groups. The individual will give greater weight to the benefits of participation the more groups for which benefits may be bestowed.

Researchers have articulated two primary weaknesses in the social integration hypothesis. First, individuals may participate in relationships that do not assign value to survey participation because expectations regarding survey participation are unknown. As a result, researchers have posited an alternative, narrower explanation: civic engagement theory. Civic engagement theory suggests that individuals who are more active in helping lessen societal ills are more likely to respond to a survey request because they perceive it as a civic duty (Dillman 1978; Goyder 1987; Groves, Singer, & Corning 2000). For example, an individual who volunteers will likely perceive fellow volunteers as people that support community building. While no written policy exists for survey participation, the individual may conclude that group members would agree to a survey request because it would contribute to the greater community.

This civic engagement theory seems too narrow. While it is unlikely that other types of social groups have explicit expectations for survey participation, individuals may still deduce how group members would expect them to behave. Consider the relationship between playing a group sport and survey participation. By playing on a team, the individual is expected to be courteous to individuals on the opposing team even though they are likely strangers and have conflicting goals (both teams want to win). This expectation may carry over to the relationship between the individual and a survey interviewer; the individual feels pressure to be courteous to a stranger who is competing for the individual's time. As a result, the individual will comply with the interviewer's request. An alternative explanation involves the contact hypothesis. Playing on a team may introduce an individual to a diverse range of others, either on the team, on the opposing team, or game officiants. The more exposure one has to other types of people, the more tolerant one will become (Allport 1954).[1] This tolerance may translate into tolerance of the interviewer, recognition of the individuals who the survey will benefit, and desire to help them. Nonetheless, this effect is apt to be weaker than in the volunteering example. As a result, socially integrated individuals will be more likely to respond than isolated individuals, but individuals who are socially integrated via civic or political relationships will be more likely to respond than individuals who are integrated without civic and political relationships.

A second argument against the social integration hypothesis is that individuals may become overextended. The more relationships individuals form, the more stress they will experience, resulting in a withdrawal from society and expectations (Goode 1960; Belle 1982). However, there is little evidence to support this argument. Individuals who hold more roles have been observed to experience more productivity, self-esteem, and happiness (Sieber 1974; Thoits 1983). Moreover, while taking on each additional role increases burden, it also provides additional benefits. Taking on additional roles solidifies one's social status by buffering roles and increases resources that may be used to enhance one's status in another role (Sieber 1974). For example, a teacher may volunteer at the local zoo and use his/her volunteer role to set up a zoo field trip for his/her class. Zoo access was made available via the volunteer role and was leveraged to more easily fulfill the role of teacher.

Similar to the overextension argument, survey methodologists have hypothesized that the busier people are, the less likely they are to participate in a survey. This may be because

[1] The counterargument to the contact hypothesis is that diversity reduces cohesion. This argument is well outlined by Putnam (2007). However, thit counterargument assumes forced integration. That is not the case here. Social activities are chosen, not prescribed.

they are truly busy and, thus, less likely to be home, or it may be because they view themselves as busy (regardless of their actual schedule) and perceive the opportunity cost of participation to be too high. In either event, repeated investigations suggest that busyness is not predictive of nonresponse once demographic variables are taken into account (Couper, Singer, & Kulka 1998; Abraham, Maitland, & Bianchi 2006; Fricker 2007). Instead, Sieber (1974) showed that the more relationships individuals have, the more likely they are to respond to a request that would fulfill expectations of multiple relationships. People can strengthen their status in each role by fulfilling a single request (thereby killing two birds with one stone).

1.2 Previous Investigations of Social Integration in Survey Research

Numerous investigations have tested whether or not socially integrated individuals are more likely to respond to surveys than socially isolated individuals and/or whether or not measures of social roles/activities suffer from nonresponse bias. Most of this research has been conducted using a single proxy indicator or single proxy topic for social integration. For instance, researchers have used measures of volunteering or charitable contributions as a proxy for social integration (Groves et al 2000; Kennickell 2005; Abraham, Helms, & Presser 2009) and found that individuals who made a charitable contribution or volunteered their time were more likely to participate in the survey than individuals who did not. As a result, volunteerism and contribution estimates from the survey were higher than the true population values. The authors inferred that volunteers and those who made charitable contributions were more likely to be integrated, thus integrated individuals were more likely to respond than isolated individuals.

Researchers have also used various politically-related measures with varying results. Couper and his colleagues (1998) predicted 1990 Census mail return rates using an individual's attachment to the polity (measured using two indices: trust in the government and perceived effectiveness of the government) and various proximal measures of individual's perception of the Decennial Census. The indices were not significant once the proximal measures and demographics were included in a multivariate model. Similarly, Billiet, Philippens, Fitzgerald, and Stoop (2005) compared cooperative and reluctant respondents on six measures of trust and participation in politics. The analyses were repeated for five countries in the European Social Survey (ESS). Only five of the 30 comparisons (6 variables*5 countries) were found to be significant at the five percent level. Reluctant Austrians were less likely to participate in politics than cooperative Austrians, and reluctant respondents in four of the five countries were more likely to perceive immigrants as a threat than cooperative respondents. Finally, Knack (1995) used state-level data from the Social Sanctions Survey (SSS) to compare voter turnout rates to 1990 Decennial Census mail return rates. He found a relatively large correlation of 0.56, suggesting that voters were more likely to return their mail Census forms than nonvoters. Other researchers have conducted similar analyses using neighborhood characteristics instead of individual characteristics. Durrant and Steele (2009) argued that urbanicity should be a strong predictor of social integration because urbanites form fewer relationships with neighbors and are more likely to move. They hypothesized that urbanites would be less likely to participate in the United Kingdom Census Link Survey than individuals who lived elsewhere. To test their hypothesis, they regressed whether individuals cooperated, refused, or were not contacted onto two dichotomous indicators:

whether or not they lived in London and whether or not they lived in a rural setting. Londoners were slightly more likely to refuse, but all other indicators were not significant at the 0.05 level. These weak and/or insignificant results are consistent with other findings in the literature using urbanicity as a proxy for social integration (e.g., Groves & Peytcheva 2008). Casas-Cordero Valencia (2010) conducted a similar analysis using a much more comprehensive list of neighborhood characteristics. She created a variety of indices ranging from residential decay to community involvement. Most of the measures were developed by self-reports, but some used interviewer observations or Census tract data. Some area-level differences were significant predictors of cooperation even after controlling for household characteristics, but the direction of the effect was not always consistent. For example, areas where respondents reported that neighbors acted on behalf of the common good had higher levels of survey participation, but respondents who reported sharing norms and values with neighbors had lower levels.

Finally, church attendance has been used as a proxy for social integration. Church attendance estimates collected in the National Election Survey (NES) and the National Survey of Religious Identification were significantly higher than denominations' membership lists. Woodberry (1998) argued that the estimates were upwardly biased due to differential nonresponse. While he did not explicitly attribute his findings to the social integration hypothesis, he did suggest that church attendees were more cooperative than non-attendees because they felt a social obligation to participate.

Taken together, the literature cited thus far suggests mixed results on whether or not integrated individuals are more likely to respond than isolated individuals and whether measures of social activities and roles are likely to suffer from nonresponse bias.

Research on volunteerism, charitable contributions, and church attendance identifies a relationship among integration, nonresponse, and nonresponse bias while analyses that use politically-related and neighborhood variables were more inconclusive. These mixed findings may be the result of using poor proxy measures of integration. Most researchers have used either a single variable (e.g., volunteerism) to measure integration or have used a series of variables that are clustered around a single topic (e.g., political interest). These proxies are not diverse enough to fully capture integration. Using a single variable results in assigning a low integration score to anyone who does not participate in that particular activity. Integration may be accomplished in a variety of ways. Individuals who do not volunteer, for example, may still integrate themselves by playing soccer or attending church. Under the social integration hypothesis, these individuals are likely to respond to the survey but would be labeled as socially isolated if volunteerism was the only measure for integration. A more diverse measure of integration would be better able to capture the concept of integration, allowing a more accurate test of the hypothesis.

The use of single, or single-topic, variables also has the potential to result in spurious conclusions about the correlation between social integration and nonresponse. For example, individuals who volunteer are more likely to participate in a survey (Groves et al 2000; Kennickell 2005; Abraham et al 2009). Because integrated individuals are also more likely to volunteer, researchers have concluded that this finding supports the hypothesis that integrated individuals are more likely to respond. However, volunteering is a measure of civic engagement, a subcategory of social integration. Using a single measure may result in researchers falsely attributing the cause of nonresponse to social integration instead of more narrowly to civic engagement. Alternatively, the previous analyses may be capturing measurement error due to social desirability. Not all of the previously cited literature has true values for respondents and nonrespondents, making it difficult to differentiate between nonresponse bias and measurement error (e.g., Woodberry 1998). Respondents may report higher than average church attendance in order to be perceived favorably by the interviewer; they may not be more likely to actually attend church. This results in higher than expected prevalence rates due to measurement error, not nonresponse bias (Presser & Stinson 1998). Finally, the single and single-topic analyses may measure the effect of topic salience and not social integration. Individuals interested in the survey topic are more likely to participate because their personal benefit (i.e., enjoyment) outweighs their personal cost (i.e., time). Again considering church attendance as the example, individuals who identify with a religion and attend services may find participation in the National Survey of Religious Identification, one of the data sources used in Woodberry's analysis (1998), more interesting than individuals who are not religious. As a result, bias would result from differential topic salience, not social integration. Without a more diverse measure of social integration, determining whether nonresponse and nonresponse bias are functions of integration or an alternative underlying construct is impossible.

Some researchers have used a more diverse set of proxy measures. Most common has been the use of a set of demographic variables that are hypothesized to be linked to social integration. These variables are generally included in a single multivariate regression or in bivariate analyses by response group. Measures often include age, household composition, marital status, presence of children, and renting vs owning. Race and sex may also be considered. Researchers have found some evidence that younger and older individuals, single-person households, renters, minority groups, and men are less likely to participate in surveys than their counterparts, although the evidence is mixed (O'Neil 1979; Goyder 1987; Groves & Couper 1998; Stoop 2005; Abraham et al 2006; Fricker 2007; Durrant & Steele 2009). The authors have concluded that these groups are less likely to cooperate because they are less integrated.

Other survey researchers have incorporated social activities and formal group memberships into their analyses. Stoop (2004) built a logistic regression model using data from the Amenities and Services Utilization Survey (AVO) to compare cooperative respondents to converted refusers and, separately, to compare easy-to-reach respondents to difficult-to-reach respondents. She used five measures of integration: sports activities, use of recreational facilities, cultural participation, media use, and participation in the arts. More culturally active individuals were more difficult to reach and more likely to initially refuse. A similar analysis was conducted on the Dutch Time Use Survey by Van Ingen, Stoop, and Breedveld (2009) in which respondents were compared to individuals who only completed a follow-up interview of critical items. Interviewers asked whether or not the respondent read the newspaper, practiced sports, watched television, was interested in politics, traveled, and/or volunteered. Each variable was found to be a moderately significant predictor of cooperation with participation in each activity resulting in higher probability of cooperation ($p < 0.05$). Investigations have also examined whether or not group membership measures suffer from nonresponse bias. Smith (1984) compared initially cooperative respondents to those who initially refused on how often they visited with friends, visited with neighbors, visited with relatives, went to a bar, visited with parents, and visited siblings. He found only small differences in most cases, with the direction of the bias not consistent across measures.

Although the multivariate social integration proxies are more robust than the univariate proxies, they are not without weaknesses. The cited analyses on nonresponse rates used multivariate models in which response was the dependent variable and each proxy variable (e.g., volunteerism or church attendance) was included as an independent variable. This approach does not adequately account for the covariance between indicators or the interactions that may occur. Each input variable measures the effect of a specific relationship instead of the combined effect of a set of diverse relationships (Berkman & Syme 1979; Bassuk, Glass, & Berkman 1999). Cornwell and Warburton's (2014) research on the integration of shift workers provides an example of why variables should not be considered individually. Shift workers were less likely to participate in some social activities than individuals who worked 9am to 6pm, but they were not found to be less integrated overall. They were only less likely to participate in activities that conflicted with their work schedules. This brought down the proportion of shift worker participation on individual measures but not their combined score. Individual measures of social activities were inadequate to capture the complex relationship of integration and work. In order to minimize the potential for misinterpretation, a better indicator is necessary to assess the relationship between nonresponse, nonresponse bias, and social integration.

Additionally, data on a diverse set of social relationships have frequently not been available for both respondents and nonrespondents. As a result, researchers have not been able to compare respondents to the full sample on these measures to estimate bias.

Instead, they have used late respondents or converted refusers as proxies for nonrespondents (e.g., Van Ingen et al 2009). Similarly, statistical tests have rarely been used to identify whether or not the differences between respondents and the full sample are significant. Instead, researchers have tested for significant differences between respondents and the proxy nonrespondents. This technique informs us as to whether the two groups are different from each other but not whether the estimate among respondents is biased. In order to draw more conclusive inferences, more statistically rigorous analyses should be conducted on data for which roles and activities are available for all sampled individuals.

To our knowledge, only Groves and his colleagues (2000) have addressed these weaknesses. In 1996, the Detroit Area Study (DAS) was administered to 451 individuals in the metro area. Respondents were asked whether or not they had participated in 14 activities or organizations in the last year. These 14 variables were placed into a principal component analysis (PCA) to create a single measure of civic engagement which was used in a regression model to predict response to a follow-up survey. After controlling for incentives and demographics, civic engagement remained significantly predictive of response at the 0.05 level with engaged individuals more likely to respond.

While Groves et al (2000) is an excellent example of using a combined measure to test a hypothesis, it does not adequately test the social integration hypothesis for two reasons. First, Groves and his colleagues sought to measure civic engagement, a subcategory of social integration. Civic engagement only measures activities that seek to address community problems whereas social integration encompasses all social relationships. Second, alternative analytics may be more appropriate. PCA seeks to maximize the variance explained by the variables in the model. The resulting principal components are not theoretically based and are often difficult to interpret in any meaningful way. PCA also assumes a linear relationship between the input variables and the resulting components which may be incorrect (Cumming & Henry 1961). An alternative approach, described in more detail in Chapter 3, is to use latent class analysis (LCA) and confirmatory factor analysis (CFA) to construct measures of social integration and avoid the limitations of PCA.

Overall, survey methodologists have taken an important first step in the investigation of the relationship among social integration, nonresponse, and nonresponse bias. However, the research to date has not moved past the initial questions of whether or not there is a relationship between social integration and nonresponse and what is the resulting bias. In order to better understand the effect of social integration on nonresponse and nonresponse bias, it is important not only to test the general hypotheses, but also to look at its pieces. Namely, do different types of integration have different effects on nonresponse and nonresponse bias? And, do they affect subcategories of nonresponse differently? Only after these questions have been answered, can work be done to develop techniques to apply our knowledge to survey best practices.

1.3 Alternative Measures of Social Integration

More complex, holistic constructs of integration than those used by survey methodologists have been developed in other fields such as sociology, psychology, and epidemiology. In these fields, social integration has been approached in five different ways, measuring social activities, social roles, social networks, self-perception, and social support.

Social activities include participation in group sports, volunteering, attending religious services, and so forth. Constructs of social activities have taken a variety of forms. Glass, de Leon, Marottoli, and Berkman (1999) divided social activities into three sub-indices: social (e.g., church attendance and playing bingo), fitness (exercise-related activities), and productive (e.g., gardening and community work). All three were found to significantly predict mortality rates in elderly individuals. Individuals with higher index scores lived longer than individuals with lower scores. Hanson, Isacsson, Janzon, and Lindell (1989) found similar results using a single index instead of sub-indices. Building upon this methodology, the Social Participation Scale (SPS) combines measures of activities involving intimate interactions, formal organizational activities, active social activities, and passive social activities into a single index. The index predicted whether or not a disabled individual would successfully reintegrate into the community (Yasui & Berven 2009). Paek, Yoon, and Shah (2005) hypothesized that different social activities would have different importance on their outcome variable, newspaper consumption. In order to incorporate differential weights for each of five activities, they performed an exploratory factor analysis (EFA) to create a social activities index. The resulting index was highly significant in predicting newspaper consumption.

As suggested by the above discussion, the methods used to create a social activities index are not standardized, neither in their components nor in the approach used to create the index. Neither of these issues have been perceived as weaknesses in the fields for which these indices have been created. Researchers have argued that differential construction schemes are necessary based on the dependent variable of interest, and indices can be created using a variety of input variables. As long as a diverse set of proxy variables is used and is theoretically or empirically justified, then the set will adequately capture the latent construct of social activity. While no particular set of questions is required, the more (and more diverse) variables that are used, the more stable the measure will be (House & Kahn 1985; Brissette, Cohen, & Seeman 2000).

Diversity in social activities is likely correlated with survey response and nonresponse bias. The more diverse the activities are, the more individuals are exposed to different types of people and situations. This may make them more open and receptive to requests from strangers, including, presumably, an interviewer (Allport 1954). Additionally, more diverse activities increase the chance that individuals perceive that response to an unsolicited request will result in a benefit for members of one of their groups (Gouldner 1960; Dillman 1978; Groves, et al 1992; Putnam 1993; Goyder et al 2006).

Separate from social activities, social roles indices have also been used with great success to measure social integration. Social roles are general personas such as parent or employee that individuals take on as part of their identity. While the exact inputs to the index vary by researcher, most social role indices include all or some of the following: parent, spouse, relative, worker, friend, neighbor, student, church member, volunteer, and/or group member (Brissette et al 2000). Using unweighted combinations of the number of unique roles, researchers have demonstrated that the number of roles is highly correlated with age with older individuals taking on fewer roles than younger adults (Cumming & Henry 1961).

As with social activities indices, social roles indices are not standardized. The roles used in any given analysis vary based on availability and theoretical importance. Standardized weighting techniques for each input have also not been developed and vary across the literature. Despite these limitations, social roles should still be a good indicator of survey response. Individuals with more roles should be more likely to fulfill a survey request. They should perceive more sources of pressure to comply or more potential benefits for individuals that hold the same roles. Compliance may also result in additional rewards and personal benefits from individuals that hold the same roles. The more roles an individual holds, the more opportunity to receive rewards (Sieber 1974; Wolfensberger & Tullman 1982). Perhaps most importantly, individuals use roles to self-identify. Internalizing a role leads to acting in accordance with role expectations in order to avoid disruptions to self-identity (Schwartz 1977).

Researchers have also attempted to use measures of social networks to quantify integration. Surveys that collect information on social networks often include recall name generators in which respondents are asked to name their closest friends/family, their other friends/family, and their acquaintances. The General Social Survey (GSS) and the National Social, Health, Life, and Aging Project (NSHAP) are two examples of surveys that collect measures of social networks in this way. Unfortunately, this method has repeatedly been demonstrated to be a poor measure of social integration for a variety of reasons. First, in an evaluation of the GSS data, Marsden (2003) found large interviewer effects when calculating network size (ρ_{int} = 0.153). Second, name generators result in large measurement error. Respondents are more likely to list family members than friends, coworkers, or other non-kin relationships (Cornwell, Schumm, Laumann, & Graber 2009). This is problematic since family networks alone are not highly predictive of social integration. The way in which individuals act within their family structure and take advice from relatives is often quite different from expectations and actions away from kin (Landecker 1951). Moreover, individuals with larger networks are likely to report more diverse networks than individuals with smaller networks, even if there is no actual difference in diversity (Marin 2004). More generally, social network measures are useful in identifying how tightly individuals are tied to each other, but they are not designed to measure the relationship of an individual to a group or organization. Since our explanation of the effect of social integration on nonresponse and nonresponse bias is based on individual norms as learned from the group, this type of social network variable does not appear appropriate for our analyses.

The fourth approach to measuring integration involves perception. Instead of using objective/factual indicators, one may measure whether the respondent feels integrated in the community. Researchers have used a variety of questions and techniques to create indices of perceived integration. Hanson and colleagues (1989) developed the Malmö Influence, Contact, and Anchorage Measure (MICAM). The MICAM was made up of three components, one of which was an index of perceived integration (labeled "social anchorage" by the authors) comprised of eight questions about how the respondent felt about his[2] community. While the MICAM had little effect on predicting health outcomes, other studies have found the opposite. In a literature review, Uchino (2004) found that 80 percent of articles on the relationship between perceived social integration and mortality found a significant effect. Most of the cited articles used similar, although not exact, measurement techniques as those used to construct the MICAM.

While perceived integration measures are likely predictive of social integration as it relates to mortality, they are unlikely useful in predicting survey response. Measures of

[2] All respondents were male.

perceived integration generally ask about local community integration. This places significant weight on how close the individual feels to neighbors and does not adequately account for other relationships (e.g., coworkers or friends). Additionally, the goal is to measure the probability of an action (i.e., agreeing to cooperate with a survey request), not health. While actions are frequently driven by social norms learned by group membership, it is rare that this link between actions and social norms enters consciousness (Schwartz 1977). If individuals are unaware of the rationale for their actions, then a direct question about perceived integration will be a poor measure of social integration in the context of survey response.

The last approach to measuring integration involves social support. Social support is measured as the frequency of financial or emotional support that individuals receive from the groups to which they belong. The Inventory of Socially Supportive Behaviors (ISSB) is the most common index. It includes 40 questions, each with a five point scale (not at all, once or twice, about once a week, several times a week, or about every day) (Barrera, Sandler, & Ramsay n.d.). This facet of integration measures the extent to which the individual uses the group for personal gain. It does not measure what the individual does in return to support the group. Since survey participation is an action in which the individual is doing a favor for an organization, variables of social support do not measure the relevant part of the individual-group dynamic.

Looking at the various measurement types, variables measuring social activities and social roles are the most relevant for this research. As the previous literature implies, this would suggest the creation of two indices to test the hypotheses focused on here, one for social activities and one for social roles. However, this approach is not conducive to identifying the effect of different routes to integration on the probability of response. An alternative, and more attractive, approach for this research is to use measures of social roles and activities to create subcategories of integration based on the route to integration. Level of integration is a function of the social relationships in which individuals participate. However, there are many paths to integration. One individual may volunteer and vote (political engagement), another may be a church elder and participate in the neighborhood watch (civic engagement), and yet another may frequently socialize with friends and play a group sport (connectedness).

Adler and Kwon (2002) were one of the first to create subcategories grounded in sociological theory. They parsed social integration into the uses of relationships. Some social relationships may promote networking opportunities (e.g., attending work functions) while others may motivate individuals to achieve their goals or best self (e.g., attending religious services) while others may improve individuals' abilities (e.g., taking a class). While the categories laid out by Adler and Kwon may be useful when analyzing the role of social integration on various economic indicators, they are not mutually exclusive nor do they map well onto reasons for nonresponse.

The National Research Council (2014) (NRC) has offered an alternative set of nine routes to integration: political engagement, nonpolitical engagement, connectedness, trust, informed citizenry, confidence in institutions, civil liberties, political polarization, and macro-level cohesion.[3] While each of these routes to integration stresses the importance

[3] The NRC used the term "social integration" instead of "macro-level cohesion." We have avoided this term in order to prevent confusion of the umbrella idea of social integration and the subcategories. Macro-level cohesion includes items such as community-level divorce rates, the level of income inequality, and crime rates.

of general societal norms, each route applies different weights to each norm. For example, conventional political engagement reinforces government legitimacy, working within the legal system, and giving a voice to all individuals while connectedness emphasizes being a good friend and family member and the importance of forming personal bonds with others. Although little research has been done to determine how well this type of categorization creates an overall measure of social integration, it seems best suited to measure social integration in the context of nonresponse. Each concept is measurable using a series of questions that ask about social roles and social activities.

1.4 Research Hypotheses

This dissertation is divided into three primary investigations: social integration and nonresponse, social integration and nonresponse bias, and the incorporation of a social integration measure into weighting methodology.

Consistent with hypotheses laid out in previous research, integrated individuals should be more likely to respond to a survey request than socially isolated individuals (H1a). However, there are many routes one may take to integration. For example, one individual may be politically engaged by boycotting, voting, or running for city council. A different individual may be otherwise connected – playing a team sport and hosting dinner parties. Individuals act in accordance with the expectations of the social groups in which they belong. The expectation of survey participation may be more pronounced among civically- and politically-oriented groups since the request to participate in a survey is most similar to work done in these groups: volunteering of time, helping the greater good, cooperating with a government (in the case of a government-sponsored survey) request. For the purposes of predicting survey nonresponse, not all forms of integration may be equal. Civically and politically active individuals should be more likely to respond than individuals who lack such forms of integration (H1b).

Although nonresponse has two major sources – noncontacts and refusals – integrated individuals should be both more likely to be contacted and less likely to refuse (H1c). Less integrated individuals may be more likely to refuse because they are less likely to perceive the benefits of participation. Although some noncontacts have not made a conscious choice to not participate since they were not reached, many, if not most, are the result of call screening and other avoidance techniques (Dixon & Tucker 2010). As such, most noncontacts are aware of the survey request and have not recognized the benefit of participation. Isolated individuals should be more likely to be noncontacts than integrated individuals, although the magnitude of the difference may be slightly smaller than the difference among refusers given the small proportion of noncontacts that are unaware of the survey request.

Integrated individuals are more likely to participate in any given social relationship (e.g., voting, playing a team sport, attending religious services) compared to isolated individuals. Assuming hypothesis 1a is accurate, integrated individuals are also more likely to respond to a survey request. Thus without effective nonresponse adjustments, individual measures of social activities and social roles should be upwardly biased as a result of differential nonresponse among individuals of various integration levels (H2a). Since not all forms of integration should have the same effect on nonresponse, not all social activity and social role indicators will be equally biased. Instead, variables measuring political and civic activities and roles may suffer from higher levels of nonresponse bias than other activity and role measures (H2b). Individuals who maintain political and civic relationships should be the most likely to respond to a survey request (H1b, previously defined). This should increase the measurable difference of the prevalence of civic and political activities between respondents and the population more so than other types of social activities and roles.

While the bias of prevalence estimates is important to some analysts, many researchers are more interested in predicting participation and role adoption via multivariate models. Bias in univariate estimates does not necessarily imply that multivariate estimates, for instance analyses predicting social roles or activities, will be biased. If the bias of the univariate estimate is consistent within various subgroups, then nonresponse is missing at random (MAR). Under MAR, coefficients associated with the independent variables predicting participation and role adoption will be unbiased (Little & Rubin 2002; Groves 2006). For a simple example, assume a model in which the likelihood of voting is a function of race:

$$\ln\left(\frac{P(Vote)}{P(NoVote)}\right) = \beta_0 + \beta_{White}\, White$$

Assuming isolated individuals are less likely to respond (H1a), then integrated individuals will be overrepresented in the data. Since integrated individuals are more likely to vote, the data would overestimate the proportion of voters, biasing the intercept (β_0). However, if socially isolated Whites were equally as likely as socially isolated individuals in other racial groups to be nonrespondents, then the effect of race on voting (β_1) would be unbiased.

Given the work conducted by Abraham and her colleagues (2009), there was reason to expect the coefficients of independent variables in multivariate models predicting social outcomes would be unbiased (H2c). They analyzed bivariate and multivariate relationships between volunteerism and a host of covariates (e.g., children in the household, marital status, etc.). They created a series of two-way tables and probit models for the full sample and then, separately, for the respondent subsample, concluding that the relationship between volunteerism and its covariates was unbiased since the magnitude of the effects and significance levels were similar for a given table/model. While their analysis goes a long way to suggest that a researcher may use a multivariate model without concern for nonresponse bias, it is not conclusive as they did not test for significance across the samples. And, of course, it was a single study. Additional research is needed to assess the generalizability of their conclusion.

Survey methodologists often use demographic variables in weight construction in an attempt to reduce nonresponse bias. This is only effective if the nonrespondents and respondents are similar within weighting subclasses. Moreover, the variables used in the weighting algorithm must be correlated both with response and with the variables of interest (Little & Vartavarian 2005). If they are not, as is the case among measures of social roles and social activities, the weights will be ineffective (Abraham et al 2006; Abraham et al 2009). Instead of conducting nonresponse adjustment using only demographic variables, one may also include a measure of social integration. This addition should more appropriately correct for nonresponse bias. As a result, the alternative weights should correct nonresponse bias better than both the base-weights (H3a) and the traditional nonresponse-adjusted weights (H3b) and result in unbiased estimates (H3c).

Data from two surveys were used to test these hypotheses: the American Time Use Survey (ATUS) (2012) and the Survey of Health, Ageing and Retirement in Europe

(SHARE) Wave II (2006). Both surveys had measures of social integration for respondents and nonrespondents which allowed for comparisons between respondents and the population. Chapter 2 describes the datasets. Chapters 3 through 5 test the various hypotheses. Finally, Chapter 6 provides a summary of the results and discusses how to further advance our understanding of the link between social integration and nonresponse.

Chapter 2: Data

In order to test the hypotheses, it was necessary to identify a dataset that met the following criteria:

1) had information available for both respondents and nonrespondents for a diverse set of social roles and activities that could be grouped by routes to integration,
2) for which the frame was representative of the general population, and
3) did not have an unusually high response rate.

The first criterion was necessary to allow for comparisons of integration by response outcome. The second criterion was desirable for two reasons. First, using a general population survey maximized the generalizability and applicability of the results to other surveys. Second, rare population surveys are frequently targeted toward groups that are more likely to be homogenous in terms of their level of integration. For example, the National Immunization Survey (NIS), the National Household Education Survey (NHES), and the Los Angeles Family and Neighborhood Survey (LAFANS) all targeted households with children. Being a parent is one role that factors into an individual's level of integration (Brissette et al 2000). Using a homogenous dataset would have limited variability and reduced the likelihood of identifying significant differences between groups. Finally, a dataset was sought that did not have unusually high response rates. High response rates hinder generalizability since most surveys do not achieve response rates in the 80 and 90 percent range. Moreover, the higher the response rate, the less likely the presence of nonresponse bias. Bias is a function of both the nonresponse rate and the difference between respondents and nonrespondents. The higher the response rate, the larger the difference between respondents and nonrespondents must be to bias the estimate.

Ten datasets were evaluated on the three criteria outlined above.[4] Two were selected: the 2012 American Time Use Survey (ATUS) and Wave II of the Survey of Health, Ageing and Retirement in Europe (SHARE).

2.1 American Time Use Survey (ATUS)

ATUS was the most appropriate dataset as it met all three criteria. ATUS drew its sample from former Current Population Survey (CPS) households. The CPS frame was constructed from the 2000 Decennial Census, an area frame, and an update of new housing permits. The CPS sample was a multistage stratified sample covering 824 individual sampling areas. Once a household was sampled, a roster was completed and a household member aged 15 or older was selected to complete the interview.[5] Given the CPS sample design and the exceptionally high response rate (90.6 percent in November 2011), the ATUS sample frame was considered to be representative of the general population (Bureau of the Census 2011).

Data collected from households sampled for the 2012 ATUS were used in all analyses. Two months after "retirement" from CPS (i.e., aging out of the panel after eight data collection waves), ATUS staff randomly drew 2,190 CPS households for ATUS. This occurred monthly throughout 2012 resulting in an ATUS 2012 sample of 26,280 former

---

[4] The 10 datasets were: the American Time Use Survey (ATUS); the British Household Panel Survey (BHPS) Wave II; the European Community Household Panel (ECHP) Wave II; The European Union Statistics on Income and Living Conditions (EU-SILC) Wave II; the General Social Survey (GSS); the Health and Retirement Study (HRS) Wave II; the Los Angeles Family and Neighborhood Survey (LA-FANS) Wave II; the National Social Life, Health, and Aging Project (NSHAP) Wave II; the Panel Survey of Income Dynamics (PSID) Wave II; and the Survey of Health, Ageing and Retirement in Europe (SHARE) Wave II.

[5] For more information on CPS frame construction and sampling, please see Chapter 3 of Technical Paper 66 (Bureau of the Census 2006).

CPS households. A stratified sample of CPS responding households was drawn using a three-stage design meant to correct for CPS's small-state oversample and to introduce an oversample of racial/ethnic minorities and households with children. ATUS then randomly selected an individual within the household 15 years or older from the roster collected in the CPS.

Once an individual was selected, a bilingual prenotification letter was mailed to the individual. Sampled individuals were randomly assigned to reference day, called on the following day, and asked about the reference day's activities. Callbacks were made as necessary over a period of eight weeks.[6] This methodology resulted in an AAPOR Response Rate 2 of 53.2 percent in 2012 (American Association for Public Opinion Research 2009; Bureau of the Census 2014).

In order to test the hypotheses, data on social roles and social activities across diverse routes of integration needed to be available for both nonrespondents and respondents to ATUS. These data were collected in the November 2011 CPS as part of the Civic Engagement (CE) Supplement. At the end of the standard CPS interview, respondents were asked to answer supplemental questions. In November 2011, 81.6 percent of CPS respondents completed the CE supplement. Under ideal conditions, response to both the CPS and the CE Supplement would have been 100 percent. Assuming individuals who are less integrated were also less likely to respond, then the most socially isolated individuals were likely excluded from the analyses since they would have been the individuals that chose not to participate in the CPS and/or the CE Supplement. This may

[6] For the approximately five percent of households that did not have a telephone number available on the frame, prenotification letters were mailed with $40 inactive debit cards and instructions to call the telephone center. Upon call-in, the interviewer provided the PIN to activate the card.

have upwardly skewed the prevalence estimates of the proportion of individuals who participate in various social roles and social activities in the sample frame itself. It may have also moderated the relationship between social integration and response, biasing the analyses downward. However, this means that any findings are conservative; the true, unobserved relationship between social integration, nonresponse, and nonresponse bias is apt to be stronger.

All analyses were limited to 5,150 cases (2,779 ATUS respondents and 2,371 nonrespondents). To be included, an individual had to have completed the CE Supplement (i.e., proxy respondents were excluded) and been the sampled household member for ATUS. In multiple adult households, it was possible that the CPS and CE Supplement respondent was not the same individual sampled for ATUS. In this case, social role and activity data were unknown for the ATUS sampled individual. These cases were excluded from the analyses. The resulting sample used in analysis was disproportionately female, renters, college educated, and non-Hispanic White (Table 1). Included individuals were also less likely to be married, reported a lower household income, were older, and lived in smaller households. The differences between the full 2012 ATUS sample and the sample used in the analyses limits generalization. However, all analyses between respondents and the analytic subsample were unbiased since comparisons were made between the subsample and the subsample ATUS respondents, making the differences between the 2012 full ATUS sample and the analytic subsample irrelevant.

Table 1: Demographic Distribution by Sample Type (ATUS) [7]

	Full 2012 ATUS Sample	Sample Used in Analyses
N	69,655	5,150
Sex		
Male	49.1%	47.7%
Female	50.9%	52.3%
Housing Status		
Own	71.1%	70.0%
Rent	28.9%	30.0%
Education		
Less than HS	20.4%	15.8%
High School	28.3%	28.1%
Some College	26.5%	27.4%
College Degree or More	24.7%	28.7%
Race/Ethnicity		
Hispanic	18.7%	12.1%
Non-Hispanic White	61.5%	69.1%
Non-Hispanic Black	10.9%	12.1%
Non-Hispanic Other	8.9%	6.7%
Marital Status		
Married	53.5%	49.5%
Not Married	46.5%	50.5%
Household Income		
Less than $20,000	15.0%	19.6%
$20,000-$39,999	22.0%	23.9%
$40,000-$59,999	18.2%	17.6%
$60,000-$99,999	23.2%	20.7%
$100,000 or More	21.6%	18.2%
Age (Mean)	36.6	47.8
Household Size (Mean)	3.9	2.8

A total of 18 indicators of social activities and social roles were available for both respondents and nonrespondents, 10 of which measured civic and political activities or roles (Figure 1). Eleven of the variables were dichotomous, while the remaining seven variables were ordinal. In analyses using proportions, categorical variables were collapsed into “never” and “at least once.” When analyses were conducted using the ordinal categories, some categories had to be collapsed due to small cell sizes. The categories used in the analyses are shown in the last column in Figure 1.

[7] Estimates are base-weighted.

Figure 1: CPS Question Wording of Social Activities and Social Roles by Route to Integration (ATUS)

	Label	Question Wording	Original Responses	Analyzed Responses
Other Engagement	Dinner w/ Family	How often did you eat dinner with any of the other members of your household? Basically every day, a few times a week, a few times a month, once a month, less than once a month, or not at all?	Almost daily Few times/week Few times/month Monthly Less than monthly Never	Almost daily Few times/week Few times/month Monthly or less
	Friend/Family	During the last twelve months, how often did you see or hear from friends or family, whether in-person or not? Basically every day, a few times a week, a few times a month, once a month, less than once a month, or not at all?	Almost daily Few times/week Few times/month Monthly Less than monthly Never	Almost daily Few times/week Few times/month Monthly or less
	Parent	Parent (constructed from roster based on presence of children in household)	Parent Not a parent	Parent Not a parent
	Spouse	Spouse (constructed from roster)	Married Not married	Married Not married
	Sports Group	Have you participated in any of these groups during the last 12 months, that is since November 2010. A sports or recreation organization such as a soccer club or tennis club?	Yes No	Yes No
	Neighbor	How often did you talk with any of your neighbors? Basically every day, a few times a week, a few times a month, once a month, less than once a month, or not at all?	Almost daily Few times/week Few times/month Monthly Less than monthly Never	Almost daily Few times/week Few times/month Monthly Less than monthly Never
	Employee	Last week, did you do any work for (either) pay (or profit)?	Yes No Retired Disabled Unable to work	Yes No
	Neighbor Favors	How often did you and your neighbors do favors for each other? By favors we mean such things as watching each other's children, shopping, house sitting, lending garden or household items, and other small acts of kindness? Basically every day, a few times a week, a few times a month, once a month, less than once a month, or not at all?	Almost daily Few times/week Few times/month Monthly Less than monthly Never	Almost daily Few times/week Few times/month Monthly Less than monthly Never

Figure 1: CPS Question Wording of Social Activities and Social Roles by Route to Integration (ATUS) (cont'd)

	Label	Question Wording	Original Responses	Analyzed Responses
Political/Civic Engagement	Talk Politics	How often did you discuss politics with family or friends - basically every day, a few times a week, a few times a month, once a month, less than once a month, or not at all?	Almost daily Few times/week Few times/month Monthly Less than monthly Never	Almost daily Few times/week Few times/month Monthly Less than monthly Never
	Vote	Do you always vote in local elections, do you sometimes vote, rarely vote, or do you never vote?	Always Sometimes Rarely Never	Always Sometimes Rarely Never
	Internet Post	How often, if at all, have you used the Internet to express your opinions about POLITICAL or COMMUNITY issues within the last 12 months? Basically every day, a few times a week, a few times a month, once a month, less than once a month, or not at all?	Almost daily Few times/week Few times/month Monthly Less than monthly Never	Few times/week or more Few times/month Monthly Less than monthly Never
	Contact Official	Please tell me whether or not you have done any of the following in the last 12 months, that is between November 2010 and now. Contacted or visited a public official - at any level of government - to express your opinion?	Yes No	Yes No
	Boycott	Please tell me whether or not you have done any of the following in the last 12 months, that is between November 2010 and now. Bought or boycotted a certain product or service because of the social or political values of the company that provides it?	Yes No	Yes No
	Other Org.	Have you participated in any of these groups during the last 12 months, that is since November 2010. Any other type of organization that I have not mentioned?	Yes No	Yes No
	Religious Org.	Have you participated in any of these groups during the last 12 months, that is since November 2010. A church, synagogue, mosque, or other religious institution or organization, NOT COUNTING your attendance at religious services?	Yes No	Yes No

Figure 1: CPS Wording of Social Activities and Social Roles by Route to Integration (ATUS) (cont'd)

	Label	Question Wording	Original Responses	Analyzed Responses
Political/Civic Engagement (cont'd)	Civic Org.	Have you participated in any of these groups during the last 12 months, that is since November 2010. A service or civic organization such as American Legion or Lions Club?	Yes No	For structural equation models, these variables were combined into a single count variable, "Community," ranging from 0-3. For all other analyses, each variable was measured individually as "Yes"/"No."
	Community Officer	In the last 12 months, that is since November 2010, have you served on a committee or as an officer of any group or organization?	Yes No	
	Community Group	Have you participated in any of these groups during the last 12 months, that is since November 2010. A school group, neighborhood, or community association such as PTA or neighborhood watch group?	Yes No	

The measure of “parent” did not include all parents, only adults who had children under 18 living in the household. As a result, a grandmother whose grand children live with her would have been labeled a parent whereas a mother of an adult child would not. While this deviated from the traditional definition of “parent,” it was appropriate in this context. Being a parent is a societal role and integrates an individual through parent-teacher interactions, parent-parent interactions, and by strengthening bonds with extended family. This effect is generally strongest when the child is a minor and lives at home. Moreover, it was the only measure available.

In addition to the 18 social activity and social role variables, the ATUS frame also included other variables that were used in the analyses, including the sample member’s sex, race/ethnicity, household income, education, home ownership status, and age. Whether or not the household received an incentive in ATUS and the reference day were also used at various points in the analysis.

All analyses were conducted using the ATUS base weights and replicated base weights unless otherwise specified. Base weights were calculated on the full-sample weight after the CPS first-stage adjustment. Because ATUS researchers did not draw a simple random sample from the CPS, adjustments were made to the CPS first-stage adjusted weights to account for the CPS oversample of State Children’s Health Insurance Program (SCHIP) participants and small-state residents along with stratification introduced by ATUS such as household size. In addition to the overall base weight, 160 replicated base weights were available. Replicates were created on the CPS base weights using Fay’s method and carried through to ATUS (Judkins 1990; Rao & Shao 1999).

One limitation of the ATUS base weights was that they did not account for nonresponse to the CE Supplement. Unfortunately, not enough information was available to recreate the base weights factoring in the CE Supplement nonresponse. As a result, the base weights may not have entirely correct for differential probabilities of selection within the analytic subsample. However, it seems unlikely that any remaining differences were large.

Final weights were also available which adjusted for nonresponse to ATUS. Final weights were only used in Chapter 5, where details on their construction may be found.

2.2 Survey of Health, Ageing and Retirement in Europe (SHARE) Wave II

In order to test for generalizability, analyses were replicated using SHARE Wave II data. SHARE is a longitudinal in-person survey which collects information on the health, economic, and social well-being of individuals aged 50 or older and their spouses. In order to be eligible for SHARE Wave I, sampled individuals must have been born before 1955, not have been institutionalized, speak the national language, and have a primary residence in the country from which they were sampled. All eligible individuals within a household were selected along with their spouses, regardless of the spouse's age (Börsch-Supan, Brandt, Hunkler, Kneip, Korbmacher, Malter, Schaan, Stuck, & Zuper 2013). SHARE Wave I was conducted in 2004[8] in 9 countries: Austria, Denmark, France, Germany, Greece, Italy, Spain, Switzerland, and The Netherlands.[9] The Wave I sampling frames varied by country based on availability and coverage. Some countries used national or regional registers of individuals or households while others used telephone

[8] Data collection continued into 2005 for some countries.

[9] SHARE Wave I was also conducted in Belgium, Israel, and Sweden. These countries were excluded from this paper because they operated on different time lags between waves and/or changed sampling designs across waves which would have confounded the analysis.

frames. Due to variation in the sampling frames, sampling strategies also varied by country with some using a simple random sample and others using a multistage clustered design. Finally, some countries did not have age information on the frame and had to conduct a screening interview to determine household eligibility prior to administering the main interview. Despite differences in sampling frames, each had nearly full coverage of the target population (Börsch-Supan & Jürges 2005).

A total of 31,036 sample lines were drawn in Wave I in the nine countries. Interviews were completed with 17,066 individuals from 11,794 households, resulting in an average AAPOR Response Rate 2 of 61.8 percent across countries (household level).[10] The response rate varied by country with most hovering between 50 and 65 percent (Börsch-Supan & Jürges 2005). Attempts were made in 2006-2007 to reinterview the 17,066 Wave I respondents as part of Wave II. As with ATUS, if the most isolated individuals were nonrespondents in Wave I, then measures of social integration constructed from the Wave II frame were positively skewed with less variability than found in the population. Estimates were likely downwardly biased, and the likelihood of identifying significant results was reduced. SHARE was also not representative of the general population since its target population was 50 and older. This restricted the inference that could be made. Analyses were limited to 19,299 of the original 20,449 sampled individuals (12,904 Wave II respondents and 6,395 nonrespondents). In order to be included, an individual had to have completed Wave I (i.e., proxy interviews were excluded) and been born before 1955 (i.e., underage spouses were excluded). Individuals who died prior to Wave

[10] The response rate of 61.8 percent includes Sweden in addition to the nine countries used in analysis. Not enough information was available on the Wave I dataset to recreate the response rate to exclude Sweden.

II were also excluded from analysis. As seen in Table 2, the demographic makeup of the analytic subsample is nearly identical to the full Wave II sample in the nine countries.

Table 2: Demographic Distribution by Sample Type (SHARE)[11]

	Full SHARE Wave II Sample	Sample Used in Analyses
N	20,449	19,299
Sex		
Male	45.2%	44.9%
Female	54.8%	55.1%
Housing Status		
Own	69.0%	69.2%
Rent	31.0%	30.8%
Education		
Primary School or Less	35.3%	33.7%
Some Secondary School	17.0%	17.2%
Secondary School	31.6%	32.5%
First Stage Tertiary or Higher	16.1%	16.5%
Marital Status		
Married/Partnership	65.8%	66.6%
Other	34.2%	33.4%
Country		
Austria	2.2%	2.3%
Germany	29.4%	30.0%
The Netherlands	5.1%	5.1%
Spain	13.7%	13.5%
Italy	21.3%	21.2%
France	20.3%	19.8%
Denmark	1.8%	1.8%
Greece	3.8%	3.8%
Switzerland	2.4%	2.5%
Household Income (€) (Mean)	46,849	47,338
Age (Mean)	66.2	65.7
Household Size (Mean)	2.2	2.2

A total of 12 social activities and roles questions were asked in Wave I and were available for the Wave II sample, including four civic and political engagement measures (Figure 2). All variables were ordinal except for whether or not individuals were married and whether or not they regularly helped another household member. Consistent with

[11] Table 2 is base-weighted.

methods used to analyze the ATUS data, ordinal variables were sometimes collapsed into "never" and "at least once" categories for dichotomous analysis. In other instances, categories were collapsed when cell sizes were small. The categories used for analysis may be found in the last column of Figure 2.

Figure 2: SHARE Wave I Question Wording of Social Activities and Social Roles by Route to Integration (SHARE)

	Label	Question Wording	Original Responses	Analyzed Responses
Political/Civic Engagement	Volunteer	How often in the last four weeks did you do voluntary or charity work?	Almost daily Almost every week Less often Never	At least once Never
	Sick Adult	How often in the last four weeks have you cared for a sick or disabled adult?	Almost daily Almost every week Less often Never	At least once Never
	Community Group	How often in the last four weeks have you taken part in a political or community-related organization?	Almost daily Almost every week Less often Never	At least once Never
	Help Others (derived from 3 questions)	In the last twelve months, have you personally given any kind of help listed on card 28 to a family member from outside the household, a friend, or neighbor?	Yes No	Almost every month or more Less often Never
		Which [other] family member from outside the household, friend or neighbor have you helped [most often] in the last twelve months?	Responses included: Friend (Ex-) colleague Neighbor Other acquaintance	
		In the last twelve months, how often altogether have you given such help to this person? Was it…	Almost daily Almost every week Almost every month Less often Never	
	Training	How often in the last four weeks have you attended an educational or training course? Almost daily, almost every week, less often?	Almost daily Almost every week Less often Never	At least once Never
	Religious Org.	How often in the last four weeks have you taken part in a religious organization? Almost daily, almost every week, less often?	Almost daily Almost every week Less often Never	Almost every week or more Less often Never

Figure 2: SHARE Wave I Question Wording of Social Activities and Social Roles by Route to Integration (SHARE) (cont'd)

	Label	Question Wording	Original Responses	Analyzed Responses
Other Engagement	Spouse/Partner	What is your marital status? Married and living together with spouse; registered partnership; and married, living separated from spouse; never married; divorced; widowed.	Married, living together Registered partnership Married, living separate Never married Divorced Widowed	Married/registered partnership Other
	Contact Parent	During the past 12 months, how often did you have contact with your mother/father, either personally, by phone or mail? Daily, several times a week, about once a week, about every two weeks, about once a month, less than once a month, or never?	Daily Several times a week About once a week About every two weeks About once a month Less than once a month Never	Daily Several times a week About once a week Less often Never
	Contact Child	During the past 12 months, how often did you or your husband/wife/partner have contact with [CHILD], either personally, by phone or mail? Daily, several times a week, about once a week, about every two weeks, about once a month, less than once a month, or never?	Daily Several times a week About once a week About every two weeks About once a month Less than once a month Never	Daily Several times a week Less often Never
	Babysit	On average, how often did you look after the child(ren) of [CHILD] in the last twelve months? Almost daily, almost every week, less often?	Almost daily Almost every week Almost every month Less often Never	Almost daily Almost every week Almost every month Less often Never
	Help HHM	Is there someone living in this household whom you have helped regularly during the last twelve months with personal care, such as washing, getting out of bed, or dressing?	Yes No	Yes No

Figure 2: SHARE Wave I Question Wording of Social Activities and Social Roles by Route to Integration (SHARE) (cont'd)

	Label	Question Wording	Original Responses	Analyzed Responses
Other Engagement (cont'd)	Help Family (derived from 3 questions)	In the last twelve months, have you personally given any kind of help listed on card 28 to a family member from outside the household, a friend, or neighbor?	Yes No	Almost daily Almost every week Almost every month Less often Never
		Which [other] family member from outside the household, friend or neighbor have you helped [most often] in the last twelve months?	The list is too long to include here. Any response was coded as family except those listed under "Help Others"	
		In the last twelve months, how often altogether have you given such help to this person? Was it…	Almost daily Almost every week Almost every month Less often Never	

A few of the social activity and role measures deserve note. First, there are four variables measuring care for others – caring for a sick adult, helping a household member, helping others, and caring for a family member who did not live in the household. To ensure that these variables were measuring different activities, a chi-square test was run on each pair of variables. Since the chi-square statistic was almost guaranteed to be significant given the sample size, Cramer's V was produced to determine the strength of the relationship. No comparison yielded a score higher than 0.16, demonstrating weak relationships and suggesting the variables were measuring different activities (Cohen 1988). As such, caring for a sick adult and helping others were categorized as civic engagement variables since they measured care for non-family members.

Second, individuals were accredited with talking to their children as long as either they or their husband/wife/partner communicated with them. It is also likely that individuals overreported the frequency of communication as they may have summed the number of times that they and their spouse/partner communicated with them. As a result, child communication measure was upwardly biased, although the magnitude of the bias was unknowable given the available data.

Many of the ordinal variables used were created from a combination of questions. In some instances, interviewers asked whether or not an individual participated in an activity. Only if the individual said "yes," were follow-up questions on frequency of participation asked. Individuals who said "no" were coded as "never." In questions regarding frequency of contact with family members, questions were asked about each family member. For example, individuals were asked how often they spoke with their mother separately from how often they spoke with their father. In order to create a single variable for frequency of parent contact, an individual was assigned the most frequent category found across their family members. In the case of parental contact, individuals would have been assigned to "daily" if they spoke to either their mother or father daily. Finally, it was possible that some questions did not apply to some individuals. For example, if individuals lived alone, it was impossible for them to regularly help another household member. In these cases, individuals were coded as "never." This coding mechanism combines individuals who could not have participated in an activity given their living arrangements and those who did not participate for other reasons. However, helping is an activity engaged in by the integrated. Social expectations can only be reinforced if an individual participates.

The goal of using a second dataset was to replicate the analyses conducted using the ATUS data. While this was done in a broad sense, the reader will note that some of the social activities and roles variables asked about in SHARE Wave I are quite different from those available in ATUS. Questions about social activities should reflect a diverse spectrum of activities that may be engaged in by the target population. Given the difference in target populations between ATUS and SHARE, the questions were appropriately different.

Additional socio-demographic variables were also available on the SHARE Wave II frame. These included country, household income, education, sex, age, home ownership status, employment status, household size, and region.

All analyses were conducted using individual-level (as opposed to household-level) base weights. These weights were calculated individually for each country based on the sample design, the probability of selection, and a nonresponse adjustment for Wave I nonresponse. The Wave I nonresponse adjustment was made using calibration, more specifically Deville and Sarndal's (1992) equation 6. In most countries, weights were calibrated to a 4*2 table of age and sex in addition to region. The number of regions varied by country, and region was not used at all for France, Denmark, or Switzerland. Details on individual country probability of selection calculations may be found in Börsch-Supan & Jürges (2005). For more information on calibration, see Section 5.1.3.

In order to conduct the analyses, it was necessary to create replicate weights. A series of pseudo-strata and clusters were created before replicate weights could be constructed. In most cases, the original strata were preserved. In rare cases, sample sizes were small within a given stratum, and strata were collapsed. Two clusters were created within each stratum. Where possible, existing clusters were maintained. In some cases, clusters were collapsed. In cases where a simple random sample was drawn within a stratum, households were randomly assigned to one of two clusters. In total, 72 pseudo-strata and 144 clusters were created.[12] WesVar 5.1 was used to create 72 replicates using jackknife repeated replication (JRR).

Final weights were also available which adjusted for nonresponse in Wave II. Final weights were only used in the analysis found in Chapter 5, where details on their construction may be found.

[12] The number of pseudo-strata by country was: Austria (8), Germany (10), The Netherlands (1), Spain (6), Italy (15), France (24), Denmark (1), Greece (6), and Switzerland (1).

Chapter 3: Social Integration and Nonresponse

This chapter tests the first three hypotheses outlined in Chapter 1:

H1a: Integrated individuals should be more likely to respond to a survey request than socially isolated individuals.

H1b: Civically and politically active individuals should be more likely to respond than individuals integrated in other ways.

H1c: Integrated individuals should be more likely to respond to a survey request than socially isolated individuals, regardless of the type of nonresponse (e.g., noncontact or refusal).

3.1 Methods

3.1.1 Constructing the Social Integration Measures

Before the hypotheses could be tested, a strong measure of social integration had to be created. Two types of structural equation models (SEM) were used to create such a measure: latent class analysis (LCA) and confirmatory factor analysis (CFA). SEM is used when the concept of interest cannot be measured directly. Instead, data are collected on a variety of measures (i.e., endogenous variables) that are correlated with the unmeasured, latent construct. In the case of ATUS, 18 endogenous variables were used while 12 variables were used for analysis of the SHARE data.[13] For both ATUS and SHARE, if the variable was originally ordinal, it was included using the categories identified in the last columns of Figures 1 and 2, respectively. If it was originally dichotomous, it was included as dichotomous.[14]

---

[13] Three of the 18 ATUS variables were collapsed into one nominal variable, "community," before being included in the SEMs.

[14] An alternative would to have used the dichotomous version of each endogenous variable. Since LCA does not distinguish between ordinal and nominal variables, using the dichotomous version would have

Generally speaking, LCA is used to identify groups or classes of individuals who have similar sets of values among all of the endogenous variables. Once classes are identified, individuals receive a posterior probability of belonging to each class based on their responses to the endogenous variables. Posterior probabilities were insufficient for this analysis. Individuals needed to be assigned to a single class. A categorical variable was created in which the categories were each of the identified classes. Individuals were assigned to a class based on their posterior probabilities. While multiple methods exist on how to use posterior probabilities to assign class membership, modal assignment was used here. That is, an individual was assigned to the class for which their posterior probability was highest. More details on this approach, including a discussion on the pros and cons, may be found in Section 3.1.2.

MPlus 7.11 was used to construct the LCA using the expected-maximization (EM) algorithm with ML estimation with robust standard errors. This optimization approach was used because it results in smaller standard errors of the parameters when accounting for complex sample designs (Muthén 2004). The number of random start values and the number of final stage optimizations were increased from the default of 20 to 100 and from four to 10, respectively. This change allowed for a more thorough investigation of the model fit. All other MPlus default settings were used. In addition to the model parameters already described, this included the use of logit model parameterization and a maximum of 10 iterations. Other parameterizations available allow for variable interactions and other flexibility that was not needed in this analysis. Similarly, 10

guaranteed no illogical groupings occurred. However, ordinal variables also provide the opportunity to identify more fine-tuned differences in the data. Regardless, the data suggested little difference between an LCA using both ordinal and dichotomous indicators and one in which only dichotomous variables were used.

iterations of the EM algorithm was sufficient to allow convergence and did not require exorbitant computing power.

The LCA was constructed based on equations for marginal, joint, and posterior probabilities. Let *c* represent the latent categorical variable of social integration with *K* classes. *K* is set by the researcher. Details on choosing a value for *K* are below. Let *J* represent the number of endogenous variables, each with M_j response categories. *J* was 16 and 12 for ATUS and SHARE, respectively. The marginal probability for any endogenous variable, x_j, being equal to response m_j is:

$$P(x_j = m_j) = \sum_{k=1}^{K} (P(c = k) * P(x_j = m_j \mid c = k))$$

The marginal probabilities ($P(x_j = m_j)$) were known at the outset, but the components were not since *c* was unknown. For example, the probability of always voting (*P(Vote = always)*) was 0.36, the proportion of individuals in the ATUS sample who reported always voting.

The joint probability for all endogenous variables with a given combination of response values is:

$$P(x_1, x_2, ..., x_J) = \sum_{k=1}^{K} (P(c = k) * P(x_1 \mid c = k) * P(x_2 \mid c = k) * ... * P(x_J \mid c = k)) ,$$

which assumes conditional independence. Using the marginal and joint probabilities, Bayes' theorem suggests that the posterior probabilities may be written as:

$$P(c = k \mid x_1, x_2, ..., x_J) = \frac{P(c = k) * P(x_1 \mid c = k) * P(x_2 \mid c = k) * ... * P(x_J \mid c = k)}{P(x_1, x_2, ..., x_J)}$$

A posterior probability is the probability of belonging to class *k* given an individual's combination of responses to the *J* endogenous variables. The posterior probabilities were unknown since *c* was unknown.

In order to determine the posterior probabilities, the EM algorithm was used to compute expectations for *c* in order to maximize the expectation of the log-likelihood of belonging to class *k*. Generally speaking, maximum likelihoods are identified by taking the derivatives of the likelihood function and solving for the unknowns, e.g., beta coefficients in a regression model. Unfortunately, this approach is not possible here since class membership is unobserved. As a results, the equations above do not provide enough information in conjunction to directly solve. The EM algorithm is an alternative, iterative approach to identifying the solution. The likelihood function is written:

$$\log L = \sum_{k=1}^{K} \sum_{j=1}^{J} n_{x_1, x_2, \ldots, x_J, c} \log(P(c = k) * P(x_1 \mid c = k) * P(x_2 \mid c = k) * \ldots * P(x_J \mid c = k)) + \alpha$$

where

$n_{x_1, x_2, \ldots, x_J, c}$ = unknown parameter with a given combination of endogenous variables in a given class

In the E-step, the values of $n_{x_1, x_2, \ldots, x_J, c}$ are found using the observed data and parameter estimates:

$$\bar{n}_{x_1, x_2, \ldots, x_J, c} = n_{x_1, x_2, \ldots, x_J} * \frac{P(c = k) * P(x_1 \mid c = k) * P(x_2 \mid c = k) * \ldots * P(x_J \mid c = k)}{P(x_1, x_2, \ldots, x_J)}$$

The M-step maximizes the log-likelihood over the parameters using the $\arg\max(\bar{n}_{x_1, x_2, \ldots, x_J, c})$. The process is then repeated until it converges or the maximum number of iterations is reached.

While class parameters and class assignment are based on the algorithm, the number of classes, *K*, is set by the researcher. In order to determine the optimal number of social integration categories, *K* was initially set to 2. A step-wise approach was used to assess whether or not the specification of additional classes improved the model. Models with various classes were compared with four criteria in mind. In order of importance, the final model needed to be theoretically justified, have sufficient fit, have fair verisimilitude, and be replicable. All else being equal, the smaller class model was preferred. Sociological theory was used to assess whether the number of classes and the grouping of cases were logical.

Entropy was used to determine goodness of fit and verisimilitude. Entropy may be calculated as:

$$E = 1 - \frac{-\sum_{i=1}^{N}\sum_{k=1}^{K}(P_i(c = k) * \log(P_i(c = k)))}{N * \log K}$$

where *N* is the total sample size and *i* represents each individual in the sample. Entropy may range from zero to one with one meaning perfect class separation. Entropy approaches one as the posterior probabilities for individuals get closer to one and zero, suggesting less measurement error in the model, i.e., less chance of assigning an individual to one class when he/she actually belongs to a different class. Entropy may be used both to assess the quality of a given model as well as compare across models where a higher entropy value may indicate a stronger model. In order to further compare models with differing numbers of classes, the Bayesian Information Criterion (BIC) and Vuong-Lo-Mendell-Rubin (VLMR) Likelihood Ratio Test (LRT) were used. The BIC was calculated for all models. A lower value indicated a better-fitting model. As with other

LRTs, VLMR tests whether a model with *K* classes has a significantly improved model fit compared to the model with *K*-1 classes. VLMR is more accurate than other LRT methods for testing goodness of fit among SEMs because it does not assume the log likelihood follows a chi-squared distribution (Asparouhov & Muthén 2012). A *p*-value of 0.05 or less indicated that more categories significantly improved the model fit.

Finally, in order to ensure the model was replicable, the sample was randomly split into two groups. For ATUS, the model was fit to a random 25 percent of the sample. The resulting parameter estimates were then fixed and used to predict class membership of the remaining 75 percent of the sample.[15] If the model fit similarly well, it was considered to be replicable.

Once a final value for *K* was selected, the model was rerun on the entire sample, producing posterior probabilities of class membership for each individual. The same approach was used for analyzing SHARE, but the sample was randomly split into two halves.

Separate from the overall measure of integration, it was necessary to develop variables representing different routes to integration in order to test the specific effect of civic and political engagement on nonresponse (H1b). In this scenario, it was more appropriate to collapse endogenous variables which were manifestations of similar routes to integration into single factors. An exploratory factor analysis (EFA) was constructed and used to inform a final CFA model. CFA identifies the correlations among endogenous variables and uses them to create continuous latent variables (i.e., factors). CFA requires a pre-

[15] Preliminary and exploratory analyses were limited to the validation portion of the sample. A larger sample size was necessary to allow for adequate cell counts, so the sample was split 25/75 instead of 50/50. The differences in group sizes did not affect model selection.

existing theoretical framework that dictates which endogenous variables load onto which factors and how many factors should be created. Since previous work on measuring social integration was not standardized, no preexisting model existed for constructing routes to integration. EFA does not assume a preexisting model structure. Instead, it allows all endogenous variables to map onto every factor and provides information on which variables are significantly related to each factor. EFA also allows the number of factors to vary and provides comparative fit statistics for models that differ on the number of factors created.

An EFA was run first to inform the creation of a more constrained model. The EFA was reviewed to identify how many factors should be used and which endogenous variables should load onto which factors. The CFA was then run constraining the model to the number of factors and loadings determined by the EFA. Each individual was assigned a score for each factor. Factors in this case represented different routes to integration (e.g., civic or family). The endogenous variables used were the activities and roles relevant to each route.

Specific to this analysis, all endogenous variables were first included in an exploratory factor analysis (EFA) run in MPlus 7.11. The model was limited to the same random subsample used to test the LCA. MPlus default settings were used, including the maximum likelihood estimator with robust standard errors and a geomin rotation which allowed the factors to covary. The math notation for the model specification is similar to that used in the CFA and is described later in this section.

Five models were run for both ATUS and SHARE, specifying one through five factors. The resulting EFAs were compared on two criteria: sociological theory and model fit. As with the LCA, preexisting sociological theory and empirical research into social integration was used to determine if the endogenous variables that were found to have significant loadings on a given factor should be measuring the same route to integration. The confirmatory fit index (CFI) and the root mean square error of approximation (RMSEA) were used to determine model fit. The CFI compares the fitted model to the null model. CFI may range from zero to one with a high score demonstrating strong model fit. The RMSEA compares the implied covariance of the factor loadings to the sample covariance. A RMSEA less than 0.05 is desired to conclude the model has a strong model fit (Browne & Cudeck 1993). Other fit statistics were available but not used. For example, many researchers use chi-squared as a goodness of fit statistic. However, chi-squared is a poor indicator of model fit when sample sizes are large, as was the case here. Chi-square sums the differences between the observed and expected value of the latent variable for each individual. As the sample sizes increase, the sum also increases, but the chi-squared distribution to assess significance remains unchanged. As a result, chi-squared tests are almost assured to be significant with large sample sizes (Markland 2007; Miles & Shevlin 2007).

In addition to providing information to settle on the appropriate number of factors, the EFA was also used to determine which endogenous variables should load onto which factors. Any endogenous variable that had a significant and positive factor loading (regardless of size) as included for a given factor as long as it made theoretical sense. For example, if having dinner with family had a significant and positive factor loading on the civic and political engagement factor, it would not have been mapped to the factor because there is no theoretical reason having family dinner is a result of being civically or politically engaged. However, if voting had a positive and significant loading on the same factor, it would have been included.

The CFA was run, specifying the desired number of factors and which endogenous variables were to be loaded onto which factors. All endogenous variables were loaded onto at least one factor. As with the previous models, MPlus 7.11 was used. Because all of the endogenous variables were binary or ordinal, the model was specified to use a polychoric correlation matrix instead of a standard correlation matrix. The standard matrix underestimates correlations of categorical variables. The polychoric matrix uses an alternative estimator to correct for the underestimation (Lee, Poon, & Bentler 1995; UCLA Institute for Digital Research and Education n.d.). Polychoric correlations are useful when the measure of interest has a continuous underlying structure as was the case with most of the endogenous variables. For example, individuals on the ATUS frame were asked how often they had posted a political comment to the internet in the past 12 months. If they posted daily for a year, they would have a continuous value of 365 and would have mapped their response to "everyday." The model also included some binary variables that did not have a continuous underlying structure (e.g., single vs. married). While polychoric correlation matrices were not developed for this type of variable, they have been demonstrated to still work well in factor analysis (Rhumtella, Brosseau-Liard, & Savalei 2012).

Using the polychoric structure, the dichotomous and ordinal variables were transformed into continuous variables. Let x_j represent each of the J categorical variables. Each variable had M_j response categories. The underlying continuous variable, x_j^*, was calculated as:

$$x_j = m_j \text{ if } \tau_{j,m_j} < x_j^* < \tau_{j,m_{j+1}}$$

where $\tau_{j,1} = -\infty$ and $\tau_{j,M_{j-1}} = \infty$.

The CFA model may be written as a series of linear regressions predicting the underlying continuous endogenous variables:

$$\mathbf{x}^* = \mathbf{\Lambda}_{\mathbf{x}}\boldsymbol{\eta} + \boldsymbol{\delta}$$

where

$\mathbf{x}^*$ = J by 1 vector of the underlying continuous values of the endogenous variables

$\mathbf{\Lambda_x}$ = J by H matrix of the factor loadings, $\lambda_{x\eta}$. The factor loadings were set to 0 where a relationship was not defined. H represents the number of factors.

$\boldsymbol{\eta}$ = H by 1 vector of the factors

The implied covariance matrix of $\mathbf{x}^*$ is defined as:

$$\mathbf{\Sigma}(\mathbf{\Lambda}_{\mathbf{x}}, \mathbf{\Phi}) = \mathbf{\Lambda}_{\mathbf{x}}\mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{x}}^{\mathbf{T}} + \mathbf{\Theta}$$

where $\mathbf{\Phi}$ is the H*H covariance matrix of the factors, and $\mathbf{\Theta}$ is the fixed J by J polychoric correlation matrix of the error terms defined as:

$$\mathbf{\Theta} = \mathbf{I} - diag\ (\mathbf{\Lambda\Phi\Lambda}^{\mathbf{T}})$$

The EM algorithm was specified with ML estimation and robust standard errors was specified as was a delta parameterization. Under this parameterization, scale factors for the latent variable of the endogenous variables are parameters in the model, but the residual variances for the continuous latent variable are not. The delta parameterization has been demonstrated to be superior than the alternative, theta parameterization in some circumstances and is recommended as the preferred option (Asparouhov & Muthén 2002).

Because the original endogenous variables were dichotomous and ordinal and the latent variables were continuous, it was necessary to specify numerical integration. While other approaches were available, Monte Carlo was used with 5,000 points of integration. This approach reduces the processing time while making little difference to the outcome (Muthén & Muthén n.d.).

Three criteria were used to evaluate the CFA. The final model was required to have sufficient fit, have fair verisimilitude, and be replicable. Similar to the EFA, model fit was assessed by the CFI and RMSEA. RMSEA was also used to evaluate verisimilitude (Preacher, Zhang, Kim & Mels 2013). CFI and RMSEA cannot be computed for models which use numerical integration. In order to calculate these fit statistics, a second CFA was run using weighted least squares (WLS) instead of ML estimation, eliminating the need for numerical integration. The WLS estimator should not be used to produce factor loadings for models with few factors, but it is sufficient to produce pseudo fit indices for the preferred model in which the ML estimator was used (Muthén & Muthén n.d.). The model using WLS was also used to compute modification indices using a Lagrange multiplier. This test calculates the minimum amount that the chi-square value may decrease if a given change is made to the model specification. In the event that the recommended alteration would make a large change to model fit and assuming it made theoretical sense, the model was rerun and reevaluated. Finally, replicability was tested by running the model on the other portion of the sample, fixing the factor loadings to the values identified from the initial portion of the sample. Consistent fit indices suggested replicability.

Once the model had been assessed and determined final, the CFA using the ML estimator was rerun on the full sample, resulting in each individual being assigned a factor score for each factor.

3.1.2 Testing the Effect of Social Integration on Nonresponse

A series of logit models were used to test hypotheses 1a through 1c. Hypothesis 1a stated that integrated individuals should be more likely to respond to a survey request than socially isolated individuals. A dichotomous indicator for response to the target survey (i.e., ATUS or SHARE Wave II) was regressed onto the overall social integration variable produced by the LCA along with socio-demographic covariates. For ATUS, the model took the form:

$$\ln\left(\frac{P(\mathrm{Re}\ sponse)}{P(Nonrespons\ e)}\right) = \beta_0 + \boldsymbol{\beta}_{\mathbf{c}}\mathbf{C} + \beta_{Sex}Sex + \boldsymbol{\beta}_{\mathbf{Race}}\mathbf{Race} + \boldsymbol{\beta}_{\mathbf{Inc}}\mathbf{Inc} + \boldsymbol{\beta}_{\mathbf{Educ}}\mathbf{Educ}$$
$$+ \beta_{Age}Age$$

where

C = (*K*-1)*1 vector of 0/1 dummy variables representing each of the social integration classes. The social isolation class was the reference group.

Sex = 0/1 dummy variable representing the individual's sex. Male was the reference group.

Race = 3*1 vector of 0/1 dummy variables representing the individual's race/ethnicity (Hispanic, non-Hispanic Black, or non-Hispanic other). Non-Hispanic White was the reference category.

Inc = 4*1 vector of 0/1 dummy variables representing the household's gross annual income ($20,000-$39,999; $40,000-$59,999; $60,000-$99,999; or $100,000 or more). Less than $20,000 was the reference category.

Educ = 3*1 vector of 0/1 dummy variables representing the individual's highest education (high school diploma, some college, a college degree or more). Less than high school diploma was the reference category.

Age = continuous indicator of the individual's age

For SHARE, the model took the form:

$$\ln\left(\frac{P(Response)}{P(Nonresponse)}\right) = \beta_0 + \boldsymbol{\beta}_{\mathbf{C}}\mathbf{C} + \beta_{Sex}Sex + \boldsymbol{\beta}_{\mathbf{Country}}\mathbf{Country} + \beta_{Inc}Inc + \boldsymbol{\beta}_{\mathbf{Educ}}\mathbf{Educ} + \beta_{Age}Age$$

where

Country = 8*1 vector of 0/1 dummy variables representing the individual's country of residence. Austria was the reference category.

Inc = continuous variable representing the household's gross annual income in Euros.

Educ = 3*1 vector of 0/1 dummy variables representing the individual's highest education (some secondary school, secondary school diploma, or first stage tertiary or higher). Primary school or less was the reference category.

All other variables were previously defined. Interaction terms between social integration and the socio-demographic variables were also tested for thoroughness.

As indicated in Section 3.1.1, the LCA did not produce a single indicator of class membership, but instead produced a set of posterior probabilities for each individual. The posterior probabilities represented the chance of belonging to each of the different classes given the individual's set of responses on the endogenous variables. Before the regression model could be run, a single variable of social integration had to be created. There are many techniques available to assign an individual to a class. Most obviously, one could use modal assignment where an individual is assigned to the class in which his/her probability is highest. Alternatively, one could use pseudo-class assignment. A random number is generated for each individual. The posterior probabilities are used to create bounds. For example, an individual may have probabilities of 0.1, 0.7, and 0.2 for the first, second, and third classes, respectively. If the random number is between 0-0.1, the individual would be assigned to the first class, between 0.1-0.8, the second class, and 0.8-1.0, the third class. Unfortunately, modal assignment does not account for measurement error in the class assignment while the pseudo-class approach biases the beta coefficients of the regression model toward zero (Asparouhov & Muthén 2012).

A superior approach, and the one used in this analysis, is the three-step proportional ML approach. It accounts for measurement error in class assignment and is least likely to overestimate the variance of the beta coefficients (Vermunt 2010; Asparouhov & Muthén 2012; Bakk, Tekle, & Vermunt 2013). In the first step, individuals were assigned to the class in which their posterior probability was the highest. In the second step, the measurement error for class assignment was estimated. Let **G** be a K*K matrix of the average conditional probabilities of the true class membership, T, of an individual given their assigned class, A:

$$g_{k_T, k_a} = P(T_{k_T} \mid A_{k_A})$$

Measurement error was calculated using the conditional probabilities from the matrix:

$$ME_{k_T,k_A} = \ln\left(\frac{P(T_{k_T} \mid A_{k_T})}{P(T_K \mid A_{k_T})}\right)$$

In the final step, each individual was assigned to a class using modal assignment. The resulting social integration variable was converted into *K*-1 dummy variables. The dummy variables were then included in the regression model and weighted using the log ratios calculated in the second step to account for measurement error in class assignment (Asparouhov & Muthén 2012). The posterior probabilities obtained from the LCA in MPlus were input into Latent Gold 5.1 to run the second and third steps.

Logistic regression was also used to test whether individuals who were integrated by participating in civic and political activities were more likely to respond than individuals who were integrated in other ways (H1b). The continuous factors created by the CFA were included as independent variables, along with other covariates, to predict response. For ATUS, the model took the form:

$$\ln\left(\frac{P(\mathit{Response})}{P(\mathit{Nonresponse})}\right) = \beta_0 + \boldsymbol{\beta}_{\boldsymbol{\eta}}\boldsymbol{\eta} + \beta_{Sex}Sex + \boldsymbol{\beta}_{\mathbf{Race}}\mathbf{Race} + \boldsymbol{\beta}_{\mathbf{Inc}}\mathbf{Inc} + \boldsymbol{\beta}_{\mathbf{Educ}}\mathbf{Educ} + \beta_{Age}Age$$

where $\boldsymbol{\eta}$ is the same H*1 vector of the factors previous specified in Section 3.1.1. The model for SHARE may be written as:

$$\ln\left(\frac{P(\mathit{Response})}{P(\mathit{Nonresponse})}\right) = \beta_0 + \boldsymbol{\beta}_{\boldsymbol{\eta}}\boldsymbol{\eta} + \beta_{Sex}Sex + \boldsymbol{\beta}_{\mathbf{Country}}\mathbf{Country} + \beta_{Inc}Inc + \boldsymbol{\beta}_{\mathbf{Educ}}\mathbf{Educ} + \beta_{Age}Age$$

The models were run in SAS 9.4 using proc surveylogistic. The three-step proportional ML approach was not necessary since the latent variables were continuous and measurement error was naturally accounted for in the model.

To further understand the relationship between response and social integration, nonrespondents were divided into subgroups based on type of nonresponse (H1c). A multinomial logit was run to compare each subcategory of nonresponse to the respondents, controlling for the categorical social integration variable and other covariates. This analysis was limited to ATUS data because subcategories of nonresponse were not publically available for SHARE Wave II. The following five response/nonresponse categories were used: completed interview, noncontact – inadequate contact information, noncontact – no contact attempt made, noncontact – generic, and refusal. Individuals with bad contact information (e.g., non-working telephone number) were classified as having inadequate contact information. Individuals who moved out of the household or were absent from the household for some other reason were not attempted by ATUS. The final, generic, noncontact group included all cases where the interviewer was not able to reach the individual at any time during the data collection period. These categories were consistent with those used by Abraham and her colleagues (2006) with one exception. Abraham included two additional categories: ineligible and other nonresponse. Combined, only 1.0 percent of sampled individuals fell into these additional categories (n = 50). Given the sample size, meaningful conclusions about these two groups were not possible. Thus ineligibles and "other" nonrespondents were dropped from this analysis. The model took the form:

$$\ln\left(\frac{\mathbf{P(Response)}}{\mathbf{P(Nonresponse)}}\right) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_C \mathbf{C} + \boldsymbol{\beta}_{Sex} \mathit{Sex} + \boldsymbol{\beta}_{Race} \mathbf{Race} + \boldsymbol{\beta}_{Inc} \mathbf{Inc} + \boldsymbol{\beta}_{Educ} \mathbf{Educ} + \boldsymbol{\beta}_{Age} \mathit{Age}$$

where

$\ln\left(\frac{P(\text{Response})}{P(\text{Nonresponse})}\right)$ = 4*1 vector of the log odds of each nonresponse subcategory compared to response

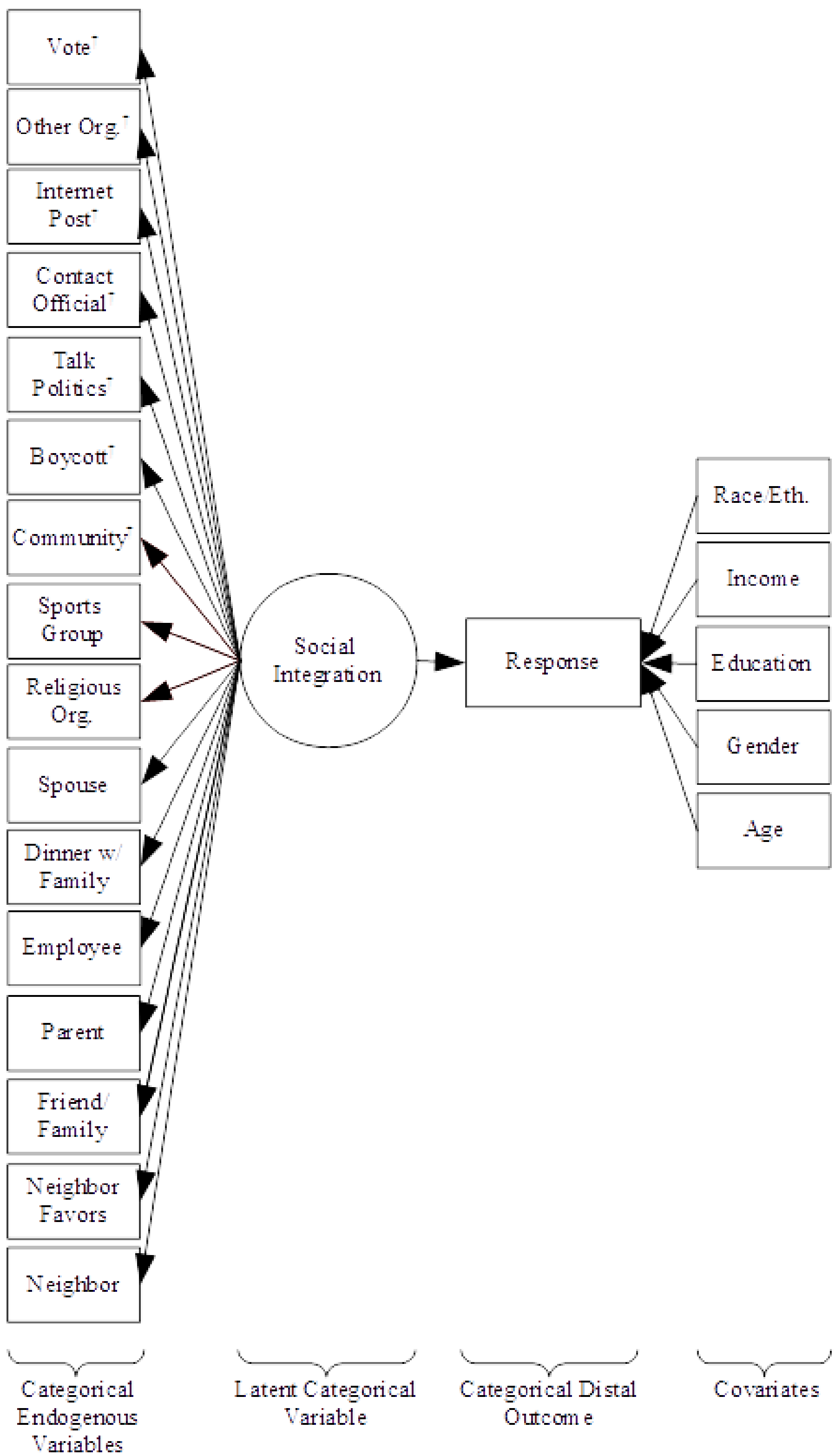
The beta matrices were all 4*m where m is the number of dummy variables used for a given covariate.

3.2 Results

3.2.1 Constructing the Social Integration Measures

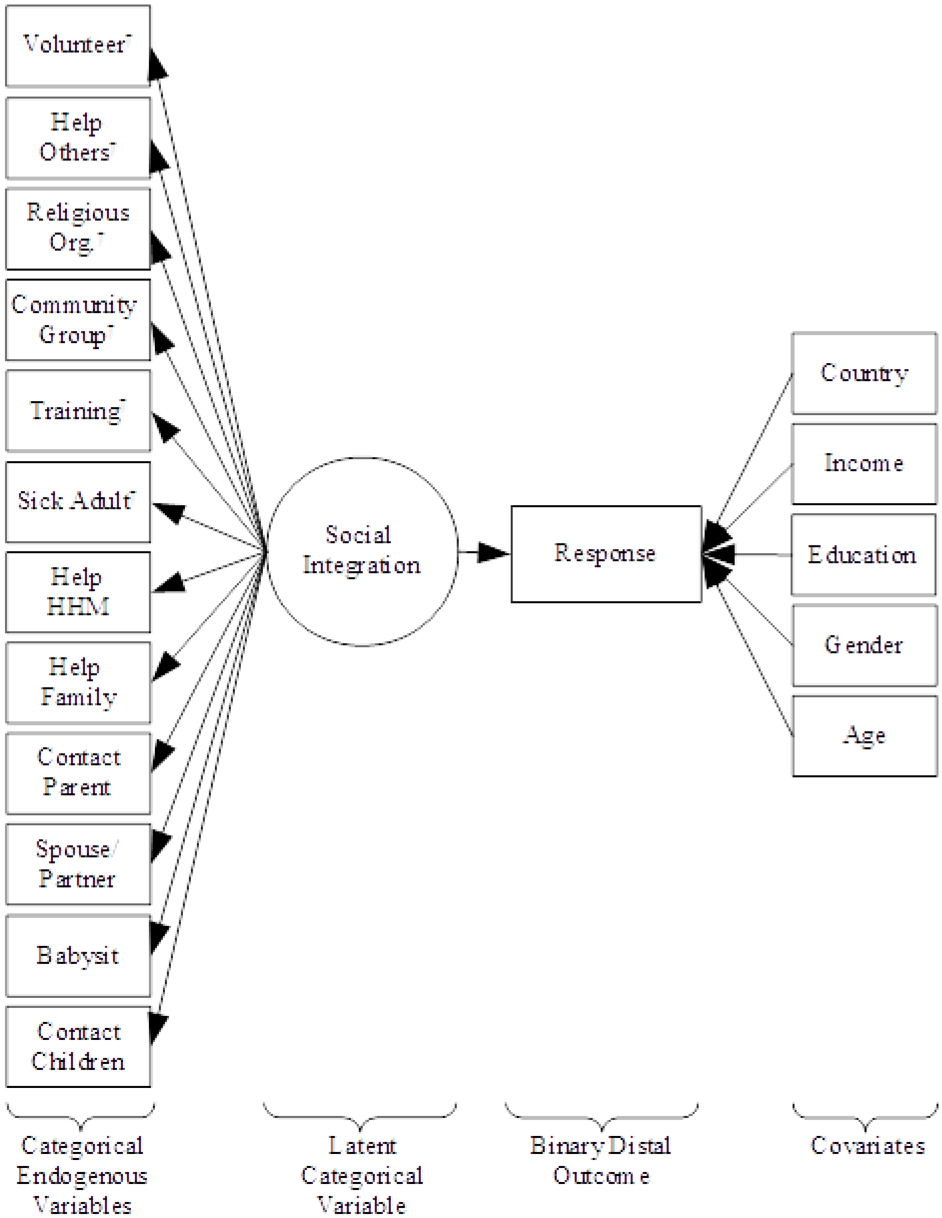
Figures 3 and 4 display a pictorial representation of the LCA for ATUS and SHARE, respectively. All of the endogenous variables were mapped onto the latent categorical construct of social integration. Two and three-class models were created for both surveys on a random subset of the sample (n = 1,226 and 9,286 for ATUS and SHARE, respectively). The models were compared by survey on the four criteria outlined in Section 3.1.1: theory, model fit, verisimilitude, and replicability.

Figure 3: Structural Equation Model of Social Integration Using Latent Class Analysis (ATUS)



† Identifies civic and political activities and roles.

Figure 4: Structural Equation Model of Social Integration Using Latent Class Analysis (SHARE)



† Identifies civic and political activities and roles.

Based on the random one-quarter ATUS sample, both the two and three-class models were theoretically justified. The two-class model suggested an integrated and isolated group in which item probabilities of frequent social interactions were higher for all categories in the integrated group (model not shown). Item probabilities are the probabilities of having a particular response option to an endogenous variable given assignment to a particular class. For example, the item probability of always voting was 57.8 percent in the integrated class but only 21.8 percent in the isolated class. Individuals in the isolated class from the two-class model were split into two groups in the three-class model, resulting in an integrated group, a moderately isolated group, and a completely isolated group (model not shown). In addition to being theoretically grounded, both models fit moderately well and had fair verisimilitude with entropy equal to 0.758 and 0.787 for the two and three-class models, respectively (Table 3). While both entropy and BIC improved slightly with the addition of a third class, the changes were small, and the VLMR LRT suggested that the three-class model did not provide a significantly better fit ($p = 0.663$). The two-class model parameter estimates were used to model the remaining three quarters of the ATUS sample.[16] The model fit similarly well, suggesting that the two-class model was replicable. As the two-class model met all four criteria, it was chosen as the final LCA for ATUS.

Table 3: Fit Statistics of Latent Class Analyses by Survey and Number of Classes

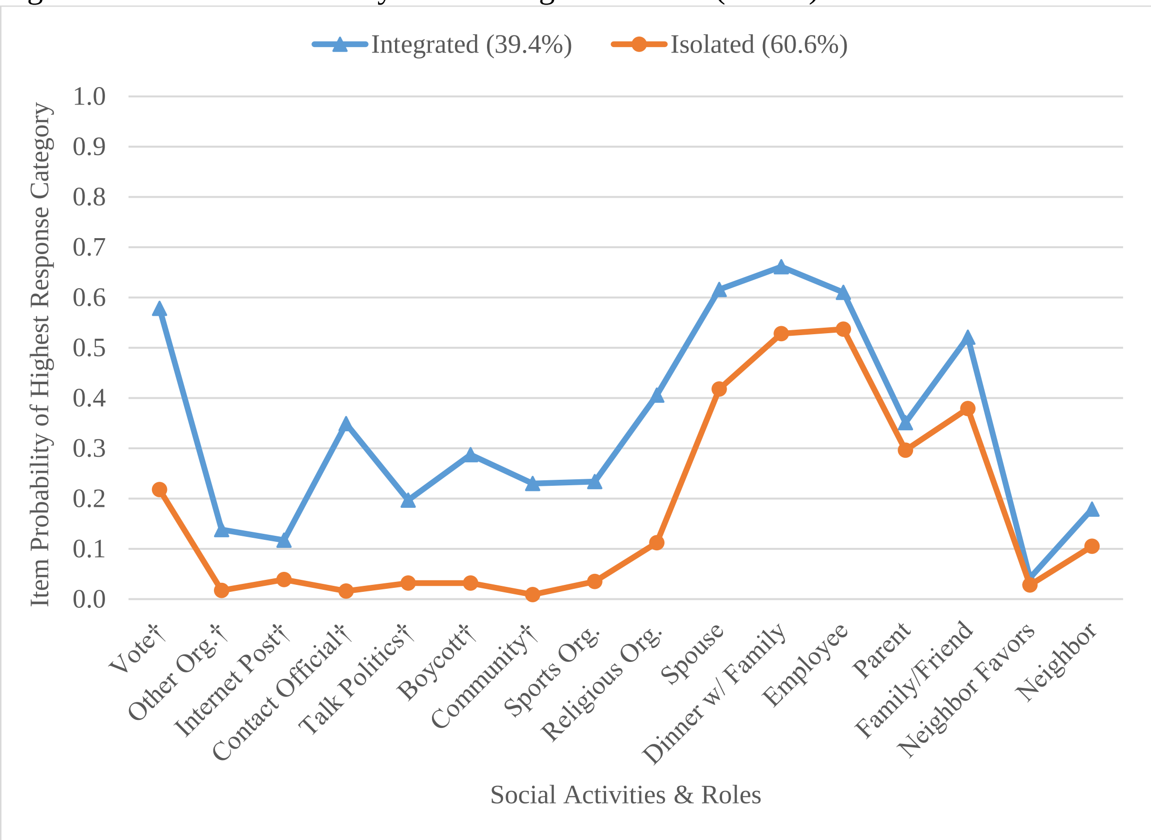
	ATUS		SHARE	
	2 Classes	3 Classes	2 Classes	3 Classes
N	1,226		9,286	
Entropy	0.758	0.787	0.634	0.617
BIC	34,091	33,981	108,404	107,228
VLMR LRT *p*-value	N/A	0.663	N/A	0.595

The two-class model was rerun using the full sample. Of the 4,820 individuals included in the analysis, 39.4 percent were assigned higher probabilities of falling into the integrated class and 60.6 percent were assigned higher probabilities of falling into the

[16] The split samples summed to 4,820, 330 less than the 5,150 ATUS sample. Cases were dropped from this analysis if they were missing data for one or more of the endogenous variables.

isolated class. The final model had an entropy of 0.729. Figure 5 displays the item probabilities for the most active response option by variable and class (in the case of voting, for example, the most active option was "always"). The integrated class had higher item probabilities for the most active response category for every variable. Overall, the two classes were generally well separated with much higher item probabilities in the integrated class than the isolated class. The differences were largest for the variables related to civic and political activities and roles, but differences were also large for belonging to sports and religious organizations. The differences between classes in probabilities for children in the household (i.e., parent) and doing favors for neighbors were smallest.

Figure 5: Item Probabilities by Social Integration Class (ATUS)



† Identifies civic and political activities and roles.

The SHARE data also suggested a two-class model with an integrated and isolated class. Using the random half sample, the item probabilities for the most frequent activity category were higher for the integrated class in nearly all instances (model not shown). Frequent participation in a religious organization was not affected by class membership (0.058 for both classes). Helping friends was more common in the isolated group (0.056 and 0.072 in the integrated and isolated groups, respectively), but the difference was not statistically significant. Unlike the ATUS three-class model, the three class SHARE model created three categories that represented different routes to integration (model not shown). However, the entropy of the three-class model was lower than the two-class model, and the VLMR LRT suggested that the three-class model did not significantly improve model fit (p = 0.595) (Table 3). The parameter estimates from the two-class model were used to estimate the remaining half (n = 9,347) of the sample. The resulting model fit well, and the two-class model was determined to be replicable.

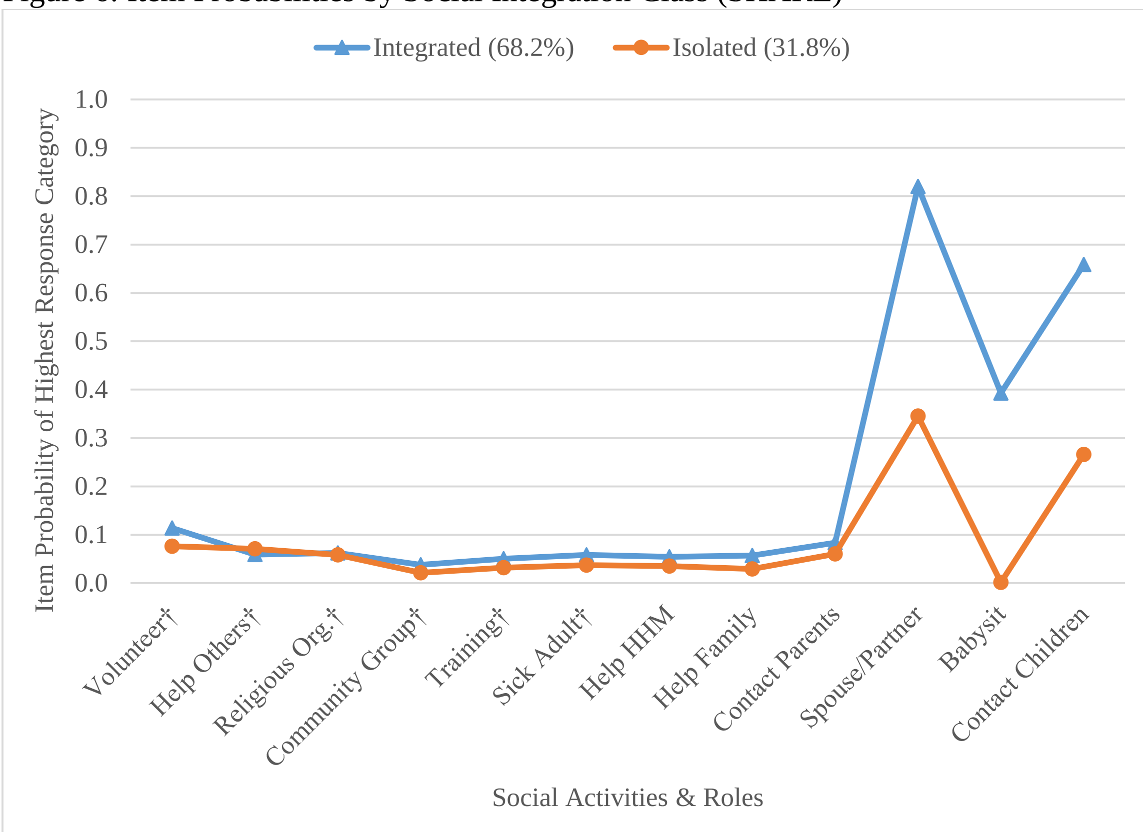
Given that the two-class model met all four criteria, it was rerun on the full dataset. Of the 18,633 cases, 68.2 percent had higher posterior probabilities of falling into the integrated class while the remaining 31.8 percent were more likely to belong to the isolated class.[17] This was nearly opposite of the distribution found in ATUS where only 39.4 percent of individuals were more likely to fall into the integrated class. The final model had an entropy value of 0.605. This was lower than the ATUS model and lower than preferred, but it still suggested a fair model fit.

The integrated class was more likely to participate at the most frequent rate in all activities and roles except for helping others (Figure 6). Individuals in the isolated class

[17] 666 cases were dropped from analysis due to item missingness on one or more of the endogenous variables.

were more likely to help a friend, coworker, or neighbor (0.059 vs. 0.071), although the difference was not significant. While the integrated class was more likely to participate in activities and roles and do so more frequently, the differences in item probabilities were small in nearly all cases. This was not surprising given the entropy score. Only the probabilities of helping with grandchildren (i.e., babysit), staying in close communication with children, and being married or having a registered partnership were noticeably different across classes.

Figure 6: Item Probabilities by Social Integration Class (SHARE)



† Identifies civic and political activities and roles.

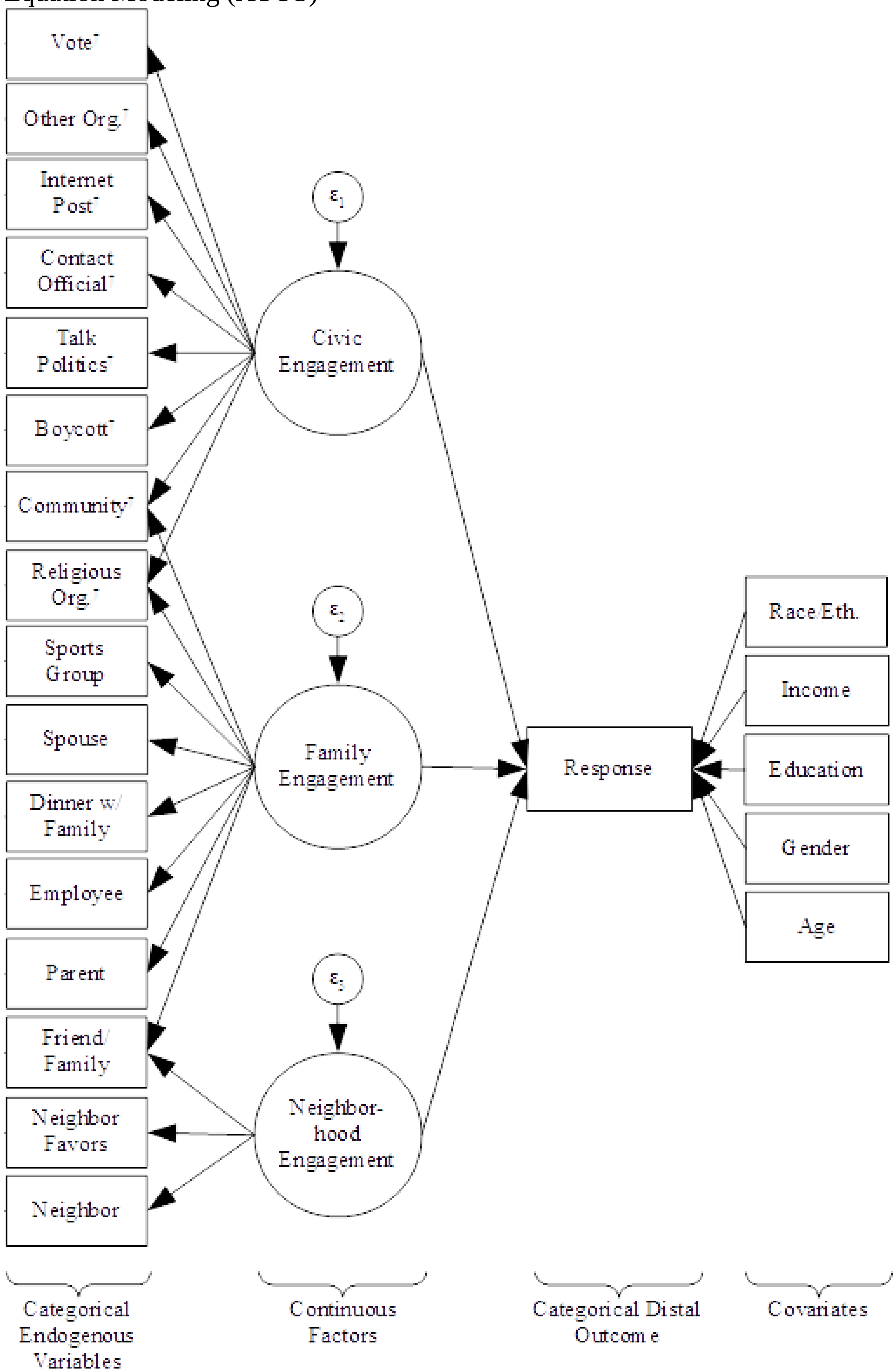
Independently of the LCA, a CFA was constructed for each survey. The first step in defining the model was to run a series of EFAs, varying the number of factors. In order to test for replicability later, all EFAs were performed on a subset of the sample. Each EFA was evaluated on the basis of theory and model fit. The results of the EFAs in both surveys suggested a three-factor model. For both ATUS and SHARE, the one- and two-factor models did not have sufficient fit. The CFI was low, and, in the case of ATUS, the RMSEA was not below the 0.05 threshold (Table 4). The higher-factor models in both surveys all met fit statistic criteria, and the addition of each additional factor significantly improved model fit. However, only the three-factor models were consistent with existing theory on routes to integration. For example, the SHARE four factor model indicated that one factor should have been created to measure just the frequency of contact with individuals' parents – a result difficult to explain theoretically.

Table 4: Fit Statistics of Exploratory Factor Analyses by Survey and Number of Classes

	ATUS					SHARE				
	1 Factor	2 Factors	3 Factors	4 Factors	5 Factors	1 Factor	2 Factors	3 Factors	4 Factors	5 Factors
N			1,226					9,286		
CFI	0.684	0.810	0.935	0.960	0.986	0.552	0.679	0.839	0.952	0.977
RMSEA	0.069	0.058	0.037	0.031	0.021	0.034	0.032	0.025	0.016	0.012
X^2 *p*-value (model comparison)	N/A	<0.0001	<0.0001	<0.0001	<0.0001	N/A	<0.0001	<0.0001	<0.0001	<0.0001

Based on the results of the EFA, a three-factor CFA was specified for ATUS in which eight variables were mapped onto a civic engagement factor, eight variables were mapped onto a family engagement factor, and three variables were mapped onto a neighborhood engagement factor (Figure 7). “Community” and “religious org.” mapped onto both “civic engagement” and “family engagement.” “Community” was constructed by collapsing three variables, including whether or not an individual belonged to a civic organization. Individuals who belonged to the Elks Club and those who participated in the PTA would both have positive values. Attending religious-affiliated activities includes couples counseling with a rabbi which would have been associated with family engagement and activities such as organizing a volunteer event through the church which would be more of a civic activity. Talking to friends or family also mapped onto two factors: family and neighborhood. When responding to the question, individuals could have been referencing communication with family or nearby friends, making it relevant to both factors.

Figure 7: Structural Equation Model of Social Integration Using Exploratory Structural Equation Modeling (ATUS)



† Identifies civic and political activities and roles.

With a CFI of 0.898 and RMSEA of 0.043 the model was considered to have sufficient fit and fair verisimilitude. Modification indices were reviewed to determine if any alterations could be made to improve model fit, but none of the recommended changes were theoretically justified. The factor loadings were fixed, and the three factor CFA was run on the remaining sample. The model fit was similar, and the model was determined to be replicable. Given that the three factor CFA met all criteria, it was rerun on the full sample (Table 5), resulting in an RMSEA of 0.040 and CFI of 0.885.

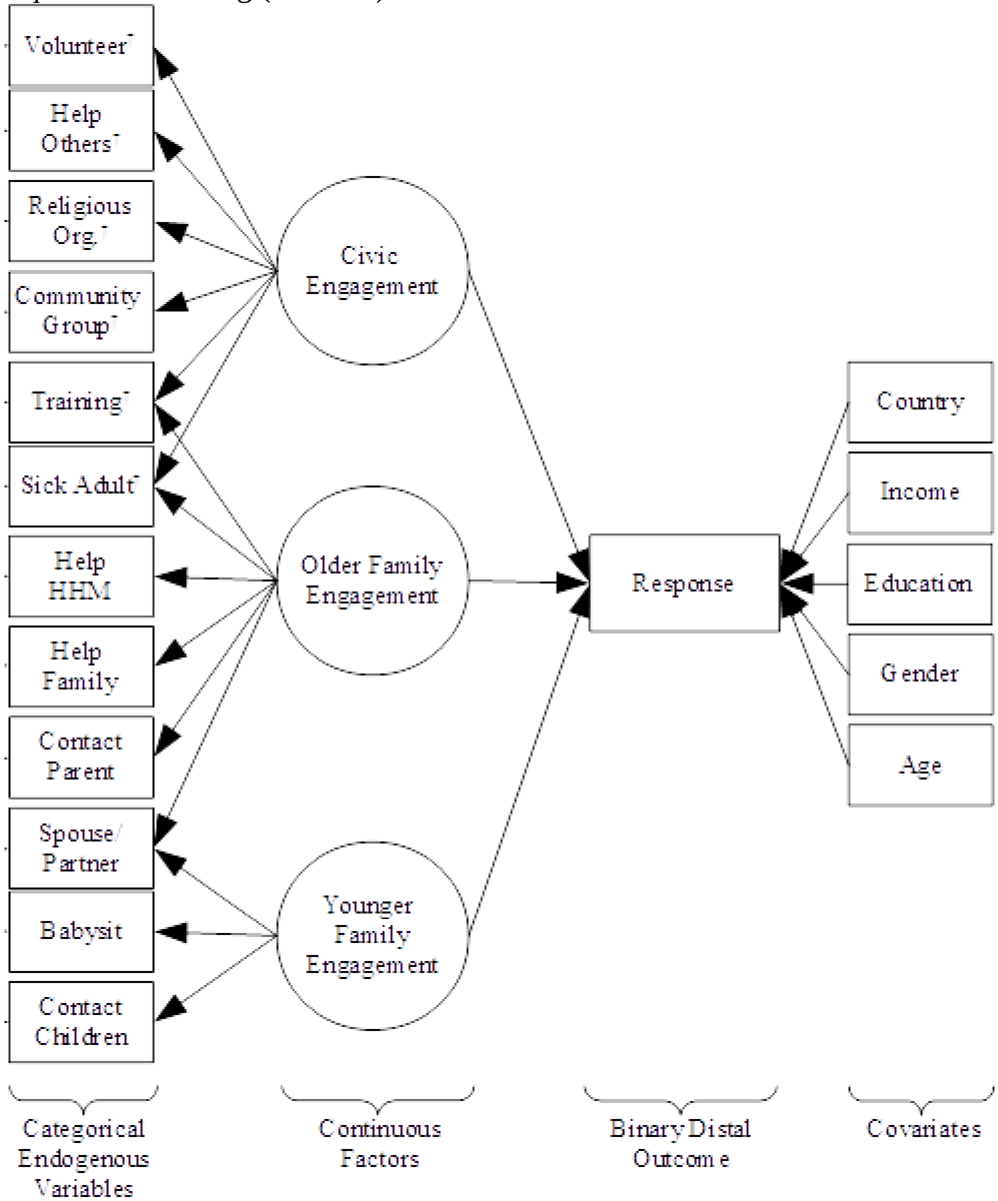
Table 5: CFA Factor Loadings (ATUS)

	Loading	s.e.
Civic Engagement		
Vote†	1.00	0.00
Other Org. †	1.08	0.10
Internet Post†	0.93	0.07
Contact Official†	2.32	0.23
Talk Politics†	1.17	0.09
Boycott†	1.38	0.14
Community†	1.16	0.08
Religious Org.†	0.70	0.06
Family Engagement		
Community†	0.80	0.09
Religious Org.†	0.50	0.10
Sports Group	1.00	0.00
Spouse	1.99	0.36
Dinner w/ Family	3.70	1.11
Employee	0.32	0.06
Parent	1.62	0.25
Friend/Family	0.24	0.05
Neighborhood Engagement		
Friend/Family	1.00	0.00
Neighbor Favors	6.55	1.07
Neighbor	5.14	0.50

† Identifies civic and political activities and roles.

A three factor CFA was also run on the SHARE subsample, but the factors were different from those identified for ATUS. In the case of SHARE, six variables were mapped to civic engagement, six were mapped to a factor measuring engagement with older family members, and three were mapped to a factor of engagement with younger family (Figure 8). The older family engagement factor referred primarily to how much individuals engaged with their parents. The younger family engagement factor primarily measured interactions with children and grandchildren. Being married or in a registered partnership loaded on both family factors. Participating in a training and caring for a sick adult loaded on two factors as well, civic engagement and older family engagement. Whether or not an individual participated in a training encompassed a large number of activities, some of which could be relevant to multiple routes of integration. For example, taking a CPR class may have been required to volunteer, placing it as a civic activity, or it could have been taken so that the individual could better care for a sick parent. The question pertaining to caring for a sick adult did not require the individual to disclose whether the adult was a family member or someone else. Without such distinction, the variable could have been relevant to either factor.

Figure 8: Structural Equation Model of Social Integration Using Exploratory Structural Equation Modeling (SHARE)



† Identifies civic and political activities and roles.

The resulting model was of sufficiently good fit and verisimilitude with a CFI of 0.840 and RMSEA of 0.024. As with ATUS, modification indices were reviewed, but no changes to the model structure were made. The factor loadings were applied to the other half of the sample. The resulting model had a similar fit, so the CFA was determined replicable. The three factor CFA was rerun on the full sample, resulting in a CFI of 0.854 and RMSEA of 0.021 (Table 6).

Table 6: CFA Factor Loadings (SHARE)

	Loading	s.e.
Civic Engagement		
Volunteer†	1.00	0.00
Help Others†	0.56	0.06
Religious Org.†	0.51	0.05
Community Group†	0.84	0.07
Training†	0.66	0.07
Sick Adult†	0.62	0.08
Older Family Engagement		
Training†	0.43	0.08
Sick Adult†	1.00	0.00
Help HHM	0.32	0.05
Help Family	1.93	0.20
Contact Parent	1.16	0.14
Spouse/Partner	0.33	0.05
Younger Family Engagement		
Spouse/Partner	1.22	0.10
Babysit	1.00	0.00
Contact Children	2.08	0.28

† Identifies civic and political activities and roles.

3.2.2 Testing the Effect of Social Integration on Nonresponse

Tables 7 and 8 display the results of the models used to test hypotheses 1a and 1b for ATUS and SHARE, respectively. The first model in each table includes the categorical measure of social integration created in the LCA. The beta coefficients for the integration variable are significant for both the ATUS and SHARE models (p<0.0001 for both surveys), providing support for hypothesis 1a that socially integrated individuals are more likely to respond than socially isolated individuals. In the case of ATUS, the probability of responding increased by 0.379 for integrated individuals compared to isolated individuals, all else being equal. For SHARE, the magnitude of the difference was even larger with a 0.513 increase. While the beta coefficients suggest that social integration had a large and significant effect on the probability to respond, the overall model fit for both surveys was poor. The models for both surveys only explain one percent of the variance in response.

A base model limited to the social integration variable and a model that included interaction terms between the social integration and the covariates were also run for each survey (not shown). The coefficients for the variable of interest were similar, and none of the interaction effects was significant. In addition to the variables of interest, the remaining variables trended in the expected direction in ATUS. SHARE covariates deviated from findings in the previous literature, but none of the deviations were significant.

Table 7: Binomial Logit Predicting Response (N=4,821) (ATUS)

	LCA			CFA Base Model			CFA w/ Interactions		
	β	s.e.	*p*-value	β	s.e.	*p*-value	β	s.e.	*p*-value
Intercept	**-0.913**	0.141	<0.0001	**-0.682**	0.148	<.0001	**-0.638**	0.146	<.0001
Socially Integrated (ref=Isolated)	**0.379**	0.084	<0.0001						
Social Integration Subcategories									
Civic				**0.207**	0.042	<.0001	**0.208**	0.044	<.0001
Family				0.050	0.057	0.383	0.054	0.060	0.370
Neighbor				-0.040	0.102	0.695	-0.082	0.107	0.446
Civic*Family							-0.003	0.054	0.956
Civic*Neighbor							-0.105	0.078	0.178
Family*Neighbor							**-0.334**	0.140	0.017
Civic*Family*Neighbor							-0.003	0.090	0.977
Female (ref=Male)	0.130	0.084	0.120	0.118	0.082	0.151	0.117	0.082	0.152
Education (ref=LT HS)									
High School	0.033	0.114	0.770	0.026	0.116	0.823	0.026	0.114	0.820
Some College or AA	0.213	0.131	0.100	0.165	0.121	0.172	0.172	0.121	0.157
College Graduate	**0.323**	0.134	0.016	**0.284**	0.136	0.037	**0.295**	0.136	0.030
Income (ref=LT $20k)									
$20,000-$39,999	0.015	0.090	0.870	0.010	0.092	0.915	0.003	0.093	0.973
$40,000-$59,999	0.085	0.116	0.460	0.052	0.119	0.664	0.041	0.121	0.735
$60,000-$99,999	**0.267**	0.102	0.009	0.219	0.124	0.078	0.212	0.128	0.097
$100,000 or more	0.181	0.117	0.120	0.152	0.133	0.253	0.152	0.134	0.257
Race (ref=non-Hisp. White)									
non-Hispanic Black	**-0.275**	0.076	0.0003	**-0.271**	0.078	0.001	**-0.267**	0.078	0.001
Hispanic	-0.195	0.099	0.050	-0.103	0.104	0.320	-0.105	0.107	0.328
non-Hispanic Other	-0.100	0.152	0.510	-0.039	0.161	0.810	-0.037	0.164	0.824
Age	**0.014**	0.002	<0.0001	**0.014**	0.003	<.0001	**0.014**	0.003	<.0001
Pseudo R^2		0.01			0.05			0.05	
BIC		26,845			6,531			6,552	

Table 8: Binomial Logit Predicting Response (N=18,536) (SHARE)

	LCA			CFA Base Model			CFA w/ Interactions		
	β	s.e.	*p*-value	β	s.e.	*p*-value	β	s.e.	*p*-value
Intercept	**0.869**	0.196	<0.0001	**1.027**	0.214	<.0001	**1.032**	0.213	<.0001
Socially Integrated (ref=Isolated)	**0.513**	0.062	<0.0001						
Social Integration Subcategories									
Civic				**0.238**	0.030	<.0001	**0.249**	0.033	<.0001
Older Family				-0.055	0.044	0.213	-0.040	0.048	0.412
Younger Family				**0.439**	0.047	<.0001	**0.435**	0.045	<.0001
Civic*Older Family							-0.041	0.034	0.230
Civic*Younger Family							0.029	0.045	0.511
Older Family*Younger Family							-0.043	0.063	0.492
Civic*Older Family*Younger Family							0.014	0.052	0.796
Female (ref=Male)	0.028	0.031	0.380	0.022	0.031	0.481	0.025	0.031	0.421
Education (ref=Primary School or Less)									
Some Secondary School	-0.011	0.075	0.890	-0.033	0.074	0.656	-0.037	0.074	0.622
Secondary School	-0.031	0.062	0.620	-0.060	0.063	0.339	-0.065	0.063	0.301
First Stage Tertiary or Higher	0.102	0.084	0.230	0.021	0.085	0.807	0.017	0.085	0.839
Income (per 1,000€)	0.0003	0.0002	0.220	**-0.0005**	0.0002	0.023	**-0.001**	0.000	0.021
Country (ref=Austria)									
Germany	**-0.951**	0.109	<0.0001	**-0.913**	0.104	<.0001	**-0.912**	0.104	<.0001
The Netherlands	**-0.453**	0.130	0.0005	**-0.484**	0.151	0.001	**-0.484**	0.151	0.001
Spain	**-0.648**	0.147	<0.0001	**-0.639**	0.150	<.0001	**-0.637**	0.151	<.0001
Italy	-0.231	0.130	0.076	-0.241	0.132	0.067	-0.239	0.132	0.069
France	**-0.362**	0.116	0.002	**-0.360**	0.113	0.002	**-0.360**	0.114	0.002
Denmark	0.251	0.205	0.220	0.332	0.231	0.150	0.334	0.231	0.147
Greece	**0.356**	0.070	<0.0001	**0.279**	0.068	<.0001	**0.274**	0.068	<.0001
Switzerland	-0.097	0.066	0.140	**-0.123**	0.057	0.029	**-0.121**	0.057	0.034
Age	-0.003	0.002	0.220	0.001	0.003	0.840	0.001	0.003	0.752
Pseudo R^2		0.01			0.04			0.04	
BIC		94,911			23,874			23,910	

Separate from the overall effect of integration, response was also regressed on the route to integration in order to test hypothesis 1b, civically integrated individual should be more likely than other integrated individuals to respond to a survey request. Three routes were identified in the CFA for ATUS: civic engagement, family engagement, and neighborhood engagement. These three routes were included in two regression models. Only main effects were accounted for in the first model (second model in Table 7) while the second model (third model in Table 7) included interaction terms among the three routes. In both models, being integrated through civic activities significantly increased the likelihood that an individual would response to ATUS ($p<0.0001$ for both models). The main effects of the other routes did not significantly influence response. The interaction terms were also generally not significant with the exception that scoring high on both family and neighborhood engagement reduced the probability that an individual would respond ($p = 0.017$). Given the number of significance tests run, this may be a chance finding.

Similar models were constructed using the SHARE data. As in ATUS, individuals who were civically engaged were more likely to participate in SHARE Wave II than individuals who were not civically active ($p<0.0001$). While the models suggested that civic engagement was important, it was not the route that was most important in predicting response. Individuals who were integrated by engaging with younger family members were more likely to respond than individuals who did not interact with younger family ($p<0.0001$). The effect of young family engagement was nearly double that of civic engagement. As a result, no support was found for hypothesis 1b among the SHARE data.

The coefficient for older family engagement (i.e., individuals who communicated with their parents or took care of sick family members) was negative. Individuals who engaged older family were less likely to respond. However, the coefficient was not significant. The interaction terms varied in their direction, but none was significant. Finally, analysis was conducted on the ATUS dataset to test hypothesis 1c, the relationship between integration and nonresponse was independent of the type of nonresponse. Table 9 displays the results of the multinomial logit. Isolated individuals were significantly more likely to be nonrespondents, regardless of the type of nonresponse. The largest difference between isolated and integrated individuals was in the probability of contact not being attempted. Integrated individuals had a 0.745 lower probability of "no contact attempt made," all else being equal. ATUS staff did not attempt to make contact if the individual had moved from the original CPS address. This finding is consistent with the hypothesis that individuals who are not tied to a community are more mobile. The magnitude of the effect of integration on the other categories of nonresponse was smaller but remained significant, finding support for hypothesis 1c.

Table 9: Multinomial Logistic Model Predicting Nonresponse (N=4,776) (ATUS)

	Refusal			General Noncontact			Inadequate Contact Information			No Contact Attempt Made		
	β	s.e.	*p*-value	β	s.e.	*p*-value	β	s.e.	*p*-value	β	s.e.	*p*-value
N	910			547			245			414		
Intercept	-0.432	0.226	0.056	-0.023	0.178	0.900	**-1.337**	0.325	<0.0001	-0.415	0.235	0.078
Socially Integrated (ref=Isolated)	**-0.240**	0.100	0.016	**-0.383**	0.120	0.001	**-0.370**	0.185	0.045	**-0.745**	0.151	<0.0001
Female (ref=Male)	-0.029	0.097	0.760	-0.226	0.119	0.057	-0.106	0.188	0.570	**-0.353**	0.142	0.013
Education (ref=LT HS)												
High School	-0.161	0.169	0.340	-0.170	0.155	0.270	0.352	0.277	0.200	0.181	0.123	0.140
Some College or AA	-0.250	0.154	0.100	-0.301	0.180	0.095	0.002	0.281	0.990	-0.154	0.196	0.430
College Graduate	**-0.354**	0.176	0.044	**-0.573**	0.161	0.0004	0.172	0.309	0.580	**-0.372**	0.181	0.039
Income (ref=LT $20k)												
$20,000-$39,999	0.063	0.136	0.650	-0.021	0.118	0.860	-0.300	0.164	0.067	0.118	0.201	0.560
$40,000-$59,999	0.056	0.147	0.700	-0.238	0.183	0.190	0.225	0.195	0.250	-0.216	0.206	0.290
$60,000-$99,999	0.070	0.146	0.630	**-0.411**	0.146	0.005	-0.488	0.290	0.092	**-0.649**	0.248	0.009
$100,000 or more	0.235	0.151	0.120	**-0.505**	0.186	0.007	-0.187	0.242	0.440	**-0.754**	0.208	0.0003
Race (ref=non-Hisp. White)												
non-Hispanic Black	0.144	0.104	0.170	**0.265**	0.115	0.021	**0.627**	0.193	0.001	**0.438**	0.150	0.003
Hispanic	-0.047	0.145	0.740	0.199	0.163	0.220	0.775	0.179	<0.0001	0.299	0.173	0.085
non-Hispanic Other	-0.162	0.194	0.400	0.091	0.161	0.570	0.425	0.327	0.190	**0.557**	0.213	0.009
Age	**-0.007**	0.002	0.004	**-0.018**	0.004	<0.0001	**-0.021**	0.005	<0.0001	**-0.020**	0.004	<0.0001
Pseudo R^2	0.01											
BIC	38,418											

3.3 Discussion

The results from both surveys demonstrated support for hypothesis 1a – integrated individuals were more likely to become respondents than were isolated individuals. Yet the regression models for both surveys had an extremely poor fit. In other words, differences in integration only accounted for a very small proportion of the variation in the response decision. Moreover, the ATUS logit model that included the continuous factors provided evidence that integration was a necessary, but not sufficient condition, to response. The relationship between social integration and response was driven entirely by civic engagement. Individuals who were integrated via other routes were not more likely to respond. This finding undermined the underlying premise of the second hypothesis (H1b) which assumed that all routes to integration would be significant but civically engaged individuals would be most likely to respond. An anomaly also existed in the ATUS models; the interaction between family and neighborhood engagement was negative and significant suggesting that individuals who were close to their family and neighbors were less likely to participate. This result may have been a result of chance, but it is consistent with Casas-Cordero Valencia's (2010) findings that individuals who reported sharing norms and values with neighbors had lower levels of survey participation. More research is necessary to uncover whether this effect is real and, if so, to explain the finding.

The results from SHARE pointed to a different conclusion for hypothesis 1b. While civic engagement significantly increased the probability of response, it did not have as large of an effect as younger family engagement.

The discrepancies between ATUS and SHARE were not limited to the conclusions drawn about hypothesis 1b. The SHARE LCA had a lower entropy score, a different distribution of sample by class, and small differences in the item probabilities between classes compared to the large item probability differences identified in the ATUS data. Similarly, the type of routes toward integration identified in the CFAs varied by dataset. While both identified a civic or political engagement factor, the ATUS data allowed for measurement of family and neighborhood engagement while the SHARE data implied two types of family engagement. There are several potential explanations for these differences. First, it is plausible that fewer isolated individuals were in the SHARE Wave II sample than the ATUS sample. SHARE Wave I, the source of the SHARE Wave II sample, achieved a much lower response rate (61.8 percent) than the CPS and CE Supplement (73.9 percent), the source of the ATUS sample. As a result, many more of the isolated individuals were likely nonrespondents to SHARE Wave I and unobserved, shifting the sample distribution. Second, the target populations were different. ATUS was a general population survey while SHARE was a survey of individuals 50 years old or older. The SHARE data was more homogenous on variables likely correlated with integration, namely age. This results in less variability and more extreme differences are necessary to observe statistical significance. Third, the questions may have mattered. Previous research suggested that the diversity of questions, not the questions themselves, were critical to formulating a measure of social integration (House & Kahn 1985; Brissette et al. 2000). The questions used to construct the LCA for each survey were similarly diverse but covered very different topics. It is possible that the questions, not just the diversity of them, influenced the quality and results of the models. Unfortunately, these proposed explanations must be answered by future research as none were testable given the data available.

In addition to unpacking the social integration variable into factors, nonresponse was also broken down into its components to determine whether or not social integration differentially affected different types of nonresponse. Analysis conducted on ATUS found support for hypothesis 1c that no such differences existed. Isolated individuals were more likely to refuse and were less likely to be contacted, regardless of the type of noncontact. While the effect of integration was significant across all types of nonresponse, the magnitude of the effect was approximately double among individuals for whom no contact attempt was made compared to the other nonresponse groups. A contact attempt was not made when the individual had moved. It was hypothesized that the effect of integration would be largest among refusals since some noncontacts were unaware of the survey request and were not making an informed choice. However, these data suggest a stronger link between integration and mobility and nonresponse. Additional research may be warranted to further understand such a relationship.

Lastly, it is worth noting that, in addition to testing the hypotheses, the CFAs allowed us to uncover some differences between the likely intention of the researchers who constructed the questionnaire and the interpretation by the sample members. The National Research Council's report (2014, Table 2-1) implied that each question asked in the CE Supplement should map onto one route to integration. However, the ATUS model suggested otherwise with community involvement, religious organization participation, and communication with friends or family mapping onto multiple routes. In some situations, an activity or role was correlated with multiple routes to integration. The same insight into the intention of the SHARE researchers was not available, but similar overlaps between endogenous variables and routes to integration were observed. Before digging deeper to uncover the source of the discrepancies across surveys, it may be worthwhile to take a step back and reexamine the wording of the questions on social activities and roles.

Chapter 4: Social Integration and Nonresponse Bias

This chapter tests the middle three hypotheses outlined in Chapter 1:

H2a: Univariate estimates of social activities and social roles should be upwardly biased.

H2b: Variables measuring political and civic activities and roles should suffer from higher levels of nonresponse bias than other social activity and role variables.

H2c: Coefficients of independent variables in multivariate models used to predict social activities and roles should be unbiased.

4.1 Methods

4.1.1 Univariate Bias

Both ATUS and SHARE Wave II were used to test the hypothesis that univariate estimates of social activities and roles should be upwardly biased (H2a). All measures of social activities and roles identified in Figures 1 and 2 were used. This included 18 variables from ATUS (11 dichotomous and seven categorical) and 12 from SHARE (six dichotomous and six categorical). The categorical variables were evaluated using both the coding scheme outlined in the last columns of Figures 1 and 2 and as the dichotomies "never" vs. "at least once."

A simple *t*-test or chi-squared test was not appropriate to test the differences between the full sample and the respondents because the two samples were not independent. In order to account for the covariance between the full sample and the respondents, replication was used to adjust the standard deviations. For dichotomous variables, both the absolute and relative differences of the proportion between the respondents and the full sample were calculated for each of the *I* base-weighted replicates (*I* = 160 and 72 replicates for ATUS and SHARE, respectively). The standard deviation of the absolute difference across all replicates may be written as follows:

$$d_i = p_{ri} - p_{fi}$$

$$s_d = \sqrt{\frac{1}{I}\left(\sum_{i=1}^{I}(d_i - (p_r - p_f))^2\right)}$$

where

d_i = the absolute difference between the estimate among respondents and the estimate among the full sample for replicate *i*

p_{ri} = the proportion of respondents that participated in the social activity/holds the social role for replicate *i*

p_{fi} = the proportion of the full sample that participates in the social activity/holds the social role for replicate *i*

s_d = the standard deviation of the absolute differences of the estimate across all replicates

The *t*-statistic was calculated using the standard deviation of the difference across all replicates:

$$t_{I-1} = \frac{p_r - p_f}{s_d}$$

Similarly, the equations for adjusting the standard deviations of the relative difference are:

$$d_i^* = \frac{p_{ri} - p_{fi}}{p_{fi}}$$

$$s_d^* = \sqrt{\frac{1}{I}\left(\sum_{i=1}^{I}\left(d_i^* - \frac{p_r - p_f}{p_f}\right)^2\right)}$$

$$t_{I-1}^* = \frac{p_r - p_f}{s_d^*}$$

where the "*" indicates the statistics calculated from the relative difference.

Replication was also used to test the distributions of the categorical variables. Similar to a traditional chi square test, the difference of the proportion of cases that fell into each response category was calculated between the respondents (i.e., observed) and the full sample (i.e., expected) and divided by the adjusted variance. This set up the following equation for each variable, *j*:[18]

$$\chi^2_{m-1} = (\mathbf{p}_\mathbf{r} - \mathbf{p}_\mathbf{f})^T (\mathbf{V}(\mathbf{p}_\mathbf{r} - \mathbf{p}_\mathbf{f}))^{-1} (\mathbf{p}_\mathbf{r} - \mathbf{p}_\mathbf{f})$$

Where $(\mathbf{p}_\mathbf{r} - \mathbf{p}_\mathbf{f})$ is a M_j*1 vector where M_j is the number of response categories for variable *j*. The variance is a M_j*M_j matrix defined as:

$$\mathbf{V}(\mathbf{p}_\mathbf{r} - \mathbf{p}_\mathbf{f}) = \frac{1}{I}\sum_{i=1}^{I}((\mathbf{p}_\mathbf{ri} - \mathbf{p}_\mathbf{fi})(\mathbf{p}_\mathbf{ri} - \mathbf{p}_\mathbf{fi})^\mathrm{T})$$

Given the number of tests performed (43 across both samples, all variables, and all tests), two were expected to be significant by chance. To minimize the potential of making a Type I error, the false discovery rate (FDR) was used to adjust the conclusions drawn from the *p*-values (Benjamini & Hochberg 2000). This approach is similar to a Bonferroni adjustment in that it accounts for the increased chance of committing a Type I error as the number of tests increases. However, a Bonferroni adjustment increases Type II errors and may cause researchers to draw different conclusions solely based on the

[18] The complete equation notation would include a subscript *j* on each variable. It has been dropped to simplify the notation.

number of tests run. The FDR does not have either of these weaknesses. While FDR originally assumed independent samples across tests (i.e., tests run on mutually exclusive sets of cases), simulation studies suggest that the FDR is just as effective when this assumption is violated, as was the case here (Benjamini & Yekutieli 2001).

4.1.2 Comparative Bias

In order to examine Hypothesis 2b – estimates of civic and political activities and roles should be more biased than estimates of other social activities and roles – the dichotomous form of the variables (18 in ATUS and 12 in SHARE) was sorted by the magnitude of both absolute and relative bias and the results reviewed

In order to statistically test the hypothesis, the difference of the bias among each combination of two variables was compared. Replication was used to control for the covariance between the full sample and respondents. The equations used to perform the *t*-tests were similar to those used to assess univariate bias, with one exception. In the univariate analysis, the *t*-tests were conducted on the difference between the full sample and the respondents, d_i. In this analysis, the *t*-tests were used to test the difference of the absolute and relative biases between two variables. The difference of the absolute bias may be calculated as:

$$h_i = d_{ji} - d_{j'i}$$

$$s_h = \sqrt{\frac{1}{I}\left(\sum_{i=1}^{I}\left(h_i - (d_j - d_{j'})\right)^2\right)}$$

$$t_{I-1} = \frac{d_j - d_{j'}}{s_h}$$

and of the relative bias as:

$$h_i^* = d_{ji}^* - d_{j'i}^*$$

$$s_h^* = \sqrt{\frac{1}{I}\left(\sum_{i=1}^{I}\left(h_i^* - (d_j^* - d_{j'}^*)\right)^2\right)}$$

$$t_{I-1}^* = \frac{d_j^* - d_{j'}^*}{s_h^*}$$

where d^*_{ji} ($d^*_{j'i}$) is the difference in an estimate between the respondent sample and the full sample for the variable *j* (*j'*) in replicate *i*. If the test was significant and if d^*_j-$d^*_{j'}$ was greater than zero then it was concluded that variable *j* was significantly more biased than variable *j'*. A total of 153 ($_{18}C_2$) comparisons were made in ATUS and 66 ($_{12}C_2$) in SHARE. The FDR was calculated and used to control for Type I error.

While repeated pairwise comparisons were used to test hypothesis 2b, it would also have been possible to use ranking procedures to order the variables and test for "ties" in the level of bias. This approach was deemed inferior for this analysis for three reasons. First, the use of replicates made it easier to control for covariance between the full sample and respondents. Second, most ranking procedures rely on one of two assumptions, neither of which applied in this situation. In some types of ranking the groups being ranked are expected to be mutually exclusive (e.g., the ordering of state math scores). In other cases, reviewers are expected to rank a series of items and the ranking procedure identifies commonalities in orders across reviewers (e.g., several people order reasons for switching jobs from most important to least important). Third, ranking procedures would not have provided a complete picture. In assigning ranks, variables are first ordered from most biased to least biased. A ranking statistic is calculated to determine if the second variable on the list is significantly less biased from the first variable. Assuming it is not, it would be considered a tie. The third variable is then compared to the first variable. If it is different, the first two variables would be assigned first rank, and the third variable would be assigned third rank. A test would not be conducted to determine if the second variable is more biased than the third. Pairwise comparisons were more thorough because all pairs of variables were compared.

4.1.3 Multivariate Bias

In order to test hypothesis 2c, a series of multivariate regression models were constructed in which the independent variables remained the same and the dependent variable changed. There are numerous multivariate regression models that predict a social activity or role. The literature was reviewed to identify a model commonly used by researchers and feasible given the data available from the CPS or SHARE Wave I (i.e., ATUS and SHARE Wave II frame data, respectively). Sociologists, economists, and others frequently use logit models to predict social activities and roles as a function of a variety of demographic variables (e.g., Levin-Waldman 2013; McCabe 2013; Wemlinger & Kropf 2013). A model similar to those identified in the literature was used for this analysis. For ATUS, the model took the following form:

$$\begin{aligned} SocialVariable = {} & \beta_0 + \beta_{Own} Own + \boldsymbol{\beta}_{\mathbf{Race}} \mathbf{Race} + \boldsymbol{\beta}_{\mathbf{Educ}} \mathbf{Educ} + \beta_{Spouse} Spouse \\ & + \beta_{Sex} Sex + \beta_{Age} Age + \beta_{Emp} Employee + \beta_{Parent} Parent + \boldsymbol{\beta}_{\mathbf{Inc}} \mathbf{Inc} \end{aligned}$$

where

Own = 0/1 dummy variable indicating home ownership. Renters were the reference group.

The same variables were not available for SHARE, so the model was slightly adjusted:

$$\begin{aligned} SocialVariable = {} & \beta_0 + \beta_{Own} Own + \boldsymbol{\beta}_{\mathbf{Country}} \mathbf{Country} + \boldsymbol{\beta}_{\mathbf{Educ}} \mathbf{Educ} + \beta_{Spouse} Spouse \\ & + \beta_{Sex} Sex + \beta_{Age} Age + \beta_{Emp} Employee + \beta_{HHSize} HHSize + \beta_{Inc} Inc \end{aligned}$$

where

HHSize = number of persons living in the household

The dependent variable was the dichotomous version of each of the social activities and roles with the exception of activities and roles that were included as independent variables in the model (e.g., spouse). Models were built using 15 different dependent variables from ATUS and 11 from SHARE. The model was run twice for each dependent variable, once using data from the full sample and once limiting the analysis to respondents.

The model constructed from the full sample was then compared to the model constructed using the respondents on two metrics. First, *t*-tests were used to determine whether complementary beta coefficients were significantly different from one another. The difference of the complementary beta coefficients was calculated for each replicate:

$$b_{xi} = \beta_{rxi} - \beta_{fxi}$$

where *x* represents each of the independent variables in the model. The standard deviation of the difference and the *t*-statistic were calculated similarly to the other analyses using replication:

$$s_{bx} = \sqrt{\frac{1}{I}\left(\sum_{i=1}^{I}(b_{xi} - (\beta_{rx} - \beta_{fx}))^2\right)}$$

$$t_{I-1} = \frac{\beta_{rx} - \beta_{fx}}{s_{bx}}$$

A total of 255 comparisons (17 coefficients * 15 models) were made in ATUS and 209 (19 coefficients * 11 models) in SHARE. FDR was used to account for multiple tests. Second, a qualitative analysis was undertaken to determine whether complimentary models were likely to result in similar conclusions. Complementary beta coefficients were compared on their significance levels. Coefficients that had the same or similar significance levels would likely be interpreted similarly. For example, if age was significant at the 0.001 level in both the full sample model and the respondent model, then one would likely draw a similar conclusion about the effect of age on the outcome variable, regardless of the difference in the magnitude.

4.2 Results

4.2.1 Univariate Bias

Tables 10 and 11 display the level of nonresponse bias identified in the social activity and role variables in ATUS and SHARE, respectively. Bias was observed in the expected direction for 27 of the 30 dichotomous variables, lending support for hypothesis 2a. These findings held even after applying the FDR. In many cases, the differences were large with up to a 23.6 percent relative and 4.2 percentage point absolute change in the estimate, but some of the differences were small and ignorable. For example, 98.0 percent of the ATUS sample communicated with friends or family. Limiting the analysis to respondents increased the estimate to 98.4 percent. Although both the relative and absolute differences in the estimate were significant ($p = 0.0002$), the magnitude of the bias was so small as to be of little concern to practitioners. Similarly, the SHARE estimate of individuals who help household members was 0.04 percentage points higher among respondents. Given the small standard deviation of the estimate, the difference was significant ($p = 0.013$) but not meaningfully so.

Table 10: Differences in Social Activity and Role Estimates by Sample Type (ATUS)

	N		Sample Type		Relative Difference		Absolute Difference	
	Full Sample	Respondents	Full Sample	Respondents	Value	*p*-value	Value	*p*-value
Parent	5,150	2,779	31.6%	29.0%	**-8.04%**	<0.0001	**-2.54**	<0.0001
Dinner w/ Family	5,009	2,732	76.3%	74.8%	**-2.01%**	<0.0001	**-1.54**	<0.0001
Employee	5,148	2,778	56.5%	56.3%	-0.46%	0.249	-0.26	0.248
Family/Friend	4,925	2,707	98.0%	98.4%	**0.40%**	0.0002	**0.39**	0.0002
Neighbor	4,925	2,705	89.2%	91.1%	**2.12%**	<0.0001	**1.90**	<0.0001
Neighbor Favors	4,895	2,689	67.2%	70.6%	**5.07%**	<0.0001	**3.40**	<0.0001
Community Group†	4,989	2,723	17.4%	18.2%	**5.07%**	0.001	**0.88**	0.001
Vote†	5,035	2,745	72.8%	76.7%	**5.40%**	<0.0001	**3.93**	<0.0001
Talk Politics†	4,921	2,697	76.0%	80.3%	**5.56%**	<0.0001	**4.23**	<0.0001
Spouse	5,150	2,779	49.5%	52.9%	**6.83%**	<0.0001	**3.38**	<0.0001
Internet Post†	4,975	2,714	28.0%	30.2%	**7.95%**	<0.0001	**2.22**	<0.0001
Religious Org.	4,978	2,719	22.6%	25.9%	**15.03%**	<0.0001	**3.39**	<0.0001
Boycott†	4,996	2,728	13.0%	15.2%	**17.60%**	<0.0001	**2.28**	<0.0001
Sports Group	4,989	2,723	11.2%	13.3%	**18.99%**	<0.0001	**2.13**	<0.0001
Contact Official†	5,007	2,734	14.6%	17.4%	**19.34%**	<0.0001	**2.82**	<0.0001
Community Officer†	4,983	2,723	13.0%	15.6%	**19.65%**	<0.0001	**2.56**	<0.0001
Other Org.†	4,983	2,721	6.4%	7.7%	**21.28%**	<0.0001	**1.35**	<0.0001
Civic Org.†	4,987	2,722	8.8%	10.8%	**23.55%**	<0.0001	**2.07**	<0.0001

† Identifies civic and political activities and roles.

Table 11: Differences in Social Activity and Role Estimates by Sample Type (SHARE)

	N		Sample Type		Relative Difference		Absolute Difference	
	Full Sample	Respondents	Full Sample	Respondents	Value	*p*-value	Value	*p*-value
Help HHM	19,139	12,873	5.2%	5.2%	**0.76%**	0.013	**0.04**	0.013
Contact Parent	18,900	12,681	25.2%	25.6%	**1.69%**	<0.0001	**0.43**	<0.0001
Contact Children	19,146	12,868	85.3%	87.5%	**2.61%**	<0.0001	**2.23**	<0.0001
Spouse	19,229	12,890	66.6%	68.5%	**2.79%**	<0.0001	**1.86**	<0.0001
Religious Org.†	19,091	12,864	9.5%	10.1%	**6.37%**	<0.0001	**0.61**	<0.0001
Help Family	19,097	12,864	20.4%	21.7%	**6.60%**	<0.0001	**1.35**	<0.0001
Sick Adult†	19,091	12,864	5.2%	5.5%	**7.06%**	<0.0001	**0.36**	<0.0001
Help Others†	19,097	12,864	10.9%	11.7%	**7.98%**	<0.0001	**0.87**	<0.0001
Babysit	19,109	12,852	26.7%	28.9%	**8.10%**	<0.0001	**2.16**	<0.0001
Training	19,090	12,863	4.3%	4.7%	**9.55%**	<0.0001	**0.41**	<0.0001
Volunteer†	19,090	12,863	10.1%	11.6%	**14.93%**	<0.0001	**1.51**	<0.0001
Community Group†	19,089	12,863	3.2%	3.8%	**16.80%**	<0.0001	**0.54**	<0.0001

† Identifies civic and political activities and roles.

Three variables, all in ATUS, did not behave as expected. The proportion of parents, employees, and individuals who have dinner with family were all lower among respondents than for the full sample, although being employed was not significant. The bias on the proportion of parents was especially large (31.6 percent vs. 29.0 percent among the full sample and respondents, respectively). This observation was inconsistent with previous research that found that parents are more likely to respond (for a summary of the literature, see Groves & Couper 1998). Thus it seemed possible that the ATUS finding was an artifact of the selection criteria (individuals had to be both the CE Supplement respondent and sampled for ATUS) resulting in the inclusion of a disproportionate number of single adult households and, consequently, single parent households. Single parent households may be busier or otherwise different from dual parent households, reducing their propensity to respond. However, the proportion of parents among respondents remained significantly lower than the full sample, even after controlling for the number of adults in the household (results not shown).

The categorical analysis for both surveys mimicked the findings of the dichotomous analysis. All 13 of the categorical distributions were significantly different between the full sample and respondents (Tables 12 and 13 for ATUS and SHARE, respectively). Twelve of the distributions were in the expected direction with respondents reporting more frequent participation than the full sample. Having dinner with family (ATUS) was the only exception. Respondents had more extreme behavior – reporting both higher frequency of having family dinner almost daily and a higher frequency of never having dinner with family. While the chi-squared tests were significant in all comparisons across both surveys, the differences observed in SHARE were, in some cases, smaller than those identified in ATUS. As with the dichotomous analysis, one should factor in the magnitude of the differences along with the significance levels. For example, SHARE respondents reported more frequent contact with their parents than the full sample. However, the differences were no larger than 0.43 percentage points in any given response category, suggesting inconsequential bias. Overall, there was support for hypothesis 2a among categorical measures.

Table 12: Differences in Social Activity and Role Distributions by Sample Type (ATUS)

		N		Sample Type		Absolute Difference	X^2 (*p*-value)
		Full Sample	Resp.	Full Sample	Resp.		
Dinner w/ Family	Monthly or Less	5,009	2,732	25.3%	26.4%	1.11	**42.1** (<0.0001)
	Few Times/Month			3.6%	3.0%	-0.64	
	Few Times/Week			13.6%	12.6%	-0.92	
	Almost Daily			57.6%	58.0%	0.45	
Friend/ Family	Monthly or Less	4,925	2,707	7.9%	7.3%	-0.61	**22.0** (<0.0001)
	Few Times/Month			12.3%	11.8%	-0.54	
	Few Times/Week			36.0%	35.9%	-0.05	
	Almost Daily			43.7%	44.9%	1.19	
Neighbor	Never	4,925	2,705	10.8%	8.9%	-1.90	**95.7** (<0.0001)
	Less than Monthly			9.3%	8.6%	-0.71	
	Monthly			9.0%	9.1%	0.12	
	Few Times/Month			24.2%	25.0%	0.77	
	Few Times/Week			33.3%	35.0%	1.71	
	Almost Daily			13.4%	13.4%	0.01	
Neighbor Favors	Never	4,895	2,689	32.8%	29.4%	-3.40	**134.4** (<0.0001)
	Less than Monthly			20.8%	22.2%	1.42	
	Monthly			13.3%	14.3%	1.00	
	Few Times/Month			19.3%	20.6%	1.27	
	Few Times/Week			10.5%	10.4%	-0.06	
	Almost Daily			3.3%	3.1%	-0.23	
Vote†	Never	5,035	2,745	27.2%	23.3%	-3.93	**204.0** (<0.0001)
	Rarely			9.6%	9.3%	-0.30	
	Sometimes			27.2%	27.5%	0.26	
	Always			36.0%	39.9%	3.97	
Talk Politics†	Never	4,921	2,697	24.0%	19.7%	-4.23	**235.0** (<0.0001)
	Less than Monthly			15.4%	15.5%	0.11	
	Monthly			10.8%	11.2%	0.38	
	Few Times/Month			20.4%	21.5%	1.08	
	Few Times/Week			19.8%	21.6%	1.84	
	Almost Daily			9.6%	10.4%	0.81	
Internet Post†	Never	4,975	2,714	72.0%	69.8%	-2.22	**55.9** (<0.0001)
	Less than Monthly			11.3%	12.6%	1.36	
	Monthly			4.5%	5.0%	0.52	
	Few Times/Month			5.4%	6.1%	0.68	
	Few Times/Week or More			6.9%	6.5%	-0.33	

† Identifies civic and political activities and roles.

Table 13: Differences in Social Activity and Role Distributions by Sample Type (SHARE)

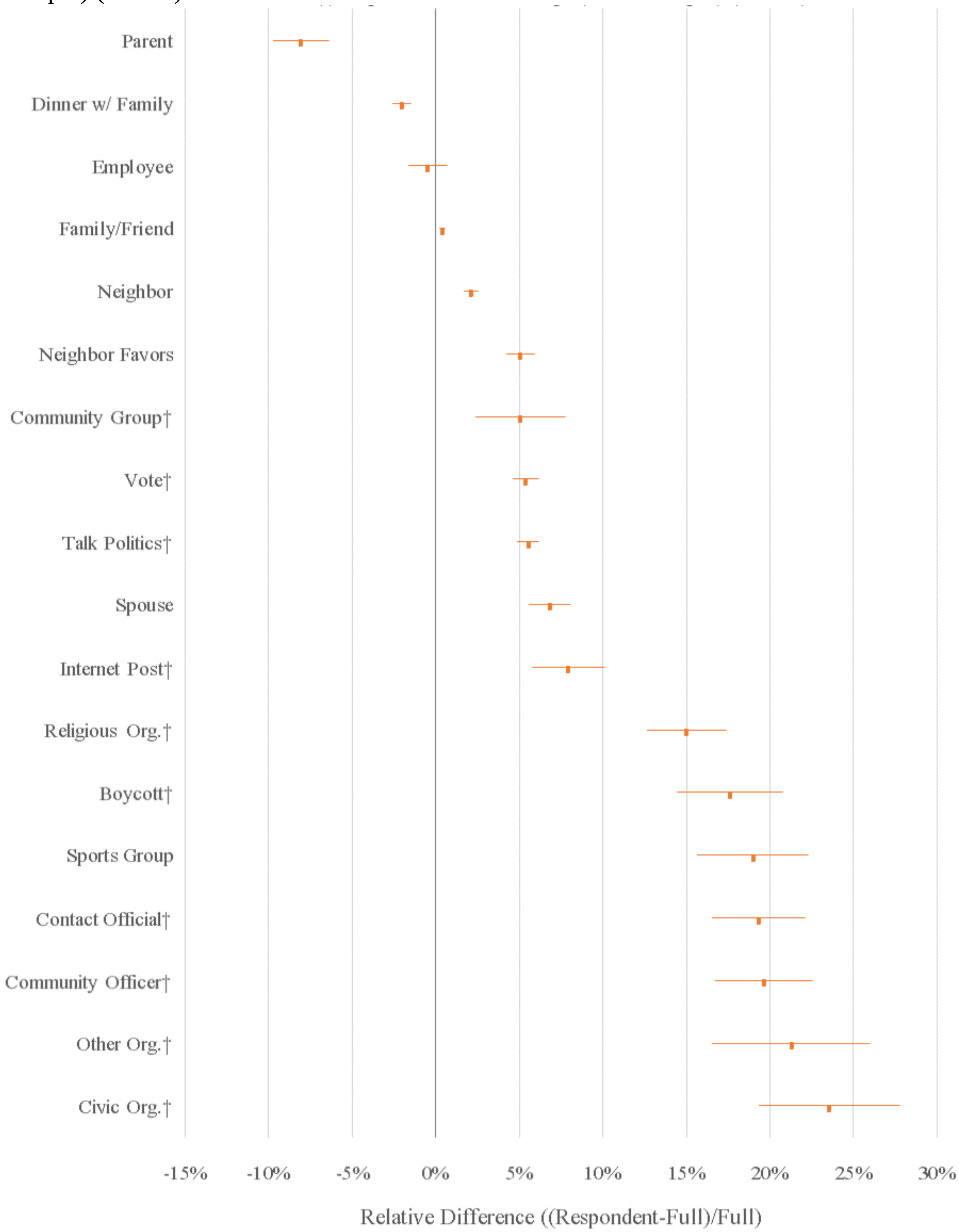
		N		Sample Type		Absolute Difference	X^2 (*p*-value)
		Full Sample	Resp.	Full Sample	Resp.		
Contact Parent	Never	18,900	12,681	74.8%	74.4%	-0.43	**237.7** (<0.0001)
	Every 2 Weeks or Less			4.7%	4.8%	0.05	
	About Once per Week			5.7%	5.7%	0.08	
	Several Times per Week			7.3%	7.2%	-0.03	
	Daily			7.5%	7.9%	0.32	
Contact Children	Never	19,146	12,868	14.7%	12.5%	-2.23	**3,618.2** (<0.0001)
	Weekly or Less			13.2%	12.2%	-0.98	
	Several Times per Week			18.9%	19.0%	0.08	
	Daily			53.2%	56.3%	3.13	
Religious Org.†	Never	19,091	12,864	90.5%	89.9%	-0.61	**575.4** (<0.0001)
	Less than Weekly			3.5%	3.8%	0.30	
	Almost Every Week or More			6.0%	6.3%	0.31	
Help Family	Never	19,097	12,864	79.6%	78.3%	-1.34	**1,605.3** (<0.0001)
	Less than Monthly			5.0%	5.4%	0.33	
	Almost Every Month			4.1%	4.3%	0.15	
	Almost Every Week			6.4%	6.7%	0.26	
	Almost Daily			4.8%	5.4%	0.61	
Help Others†	Never	19,097	12,864	89.1%	88.3%	-0.87	**827.1** (<0.0001)
	Less than Monthly			4.6%	5.0%	0.40	
	Almost Every Month or More			6.3%	6.8%	0.47	
Babysit	Never	19,109	12,852	30.0%	29.0%	-1.05	**1,383.8** (<0.0001)
	Less than Monthly			6.9%	7.5%	0.59	
	Almost Every Month			4.8%	4.9%	0.06	
	Almost Every Week			8.3%	9.0%	0.76	
	Almost Daily			50.0%	49.6%	-0.37	

† Identifies civic and political activities and roles.

4.2.2 Comparative Bias

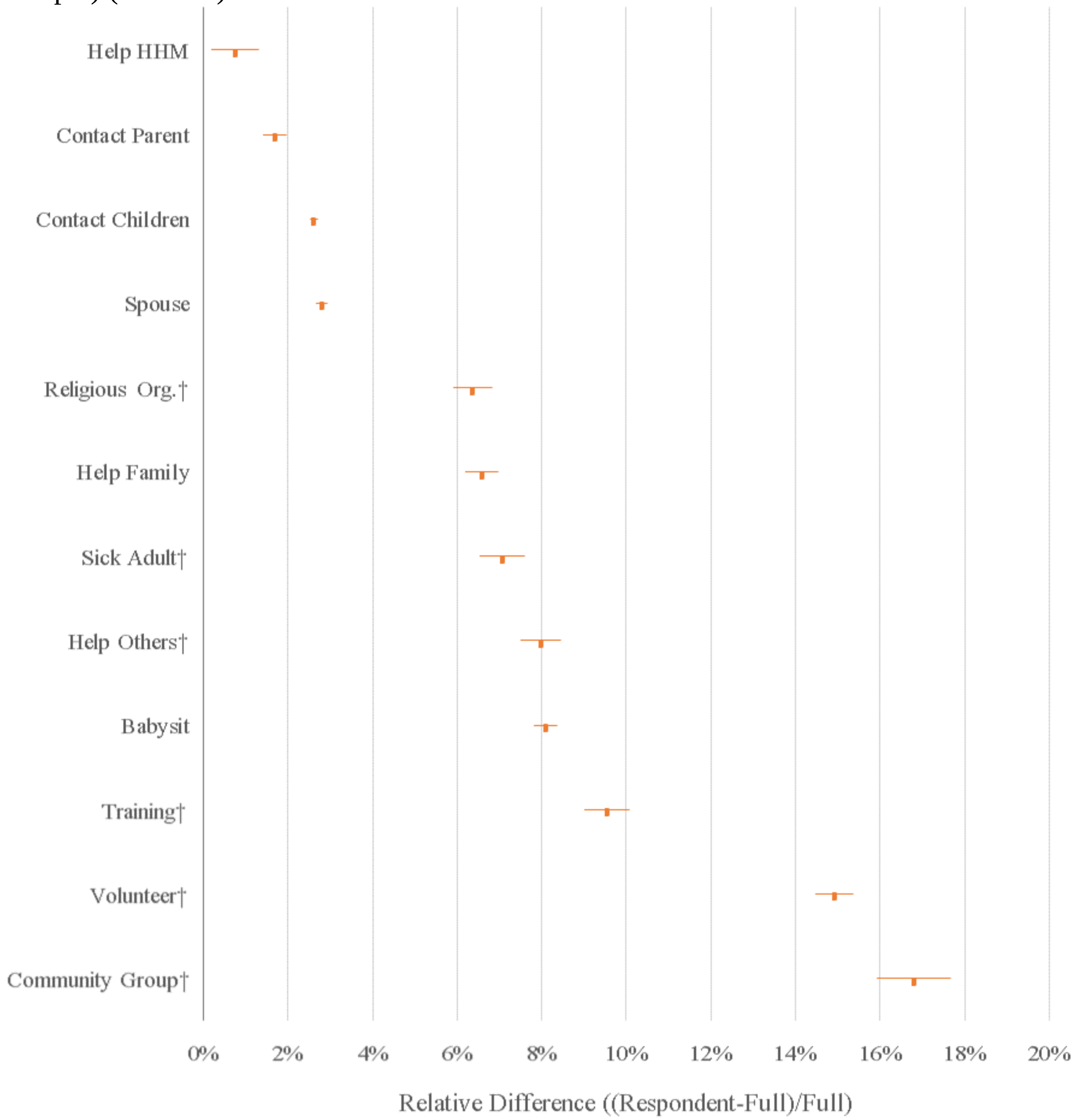
In order to test hypothesis 2b, variables were ordered by both their relative and absolute differences. For ATUS, variables generally clustered into three levels of relative bias: less than five percent, five to 10 percent, and more than 10 percent with five, six, and seven variables falling into each of the respective categories (Figure 9). Variables from SHARE also fell into three general levels, although the categories were different. Four variables showed less than a four percent bias; six variables suffered from six to 10 percent relative bias, and 2 variables 14 to 18 percent (Figure 10). In both surveys, all of the civic and politically oriented variables fell into the two higher categories. However, the confidence intervals of the estimates of the civic and political activities and roles frequently overlapped with other variables. For example, the confidence interval of ATUS's community organization variable was not significantly different from three of the eight variables that measured something other than a civic or political activity or role. Significance tests performed on the differences between variables (results not shown) were consistent with the pictorial representation. These results showed that while civic and politically oriented variables trended toward higher levels of relative bias, the differences were not statistically significant.

Figure 9: Relative Difference in Univariate Estimates ((Respondents-Full Sample)/Full Sample) (ATUS)



† Identifies civic and political activities and roles.

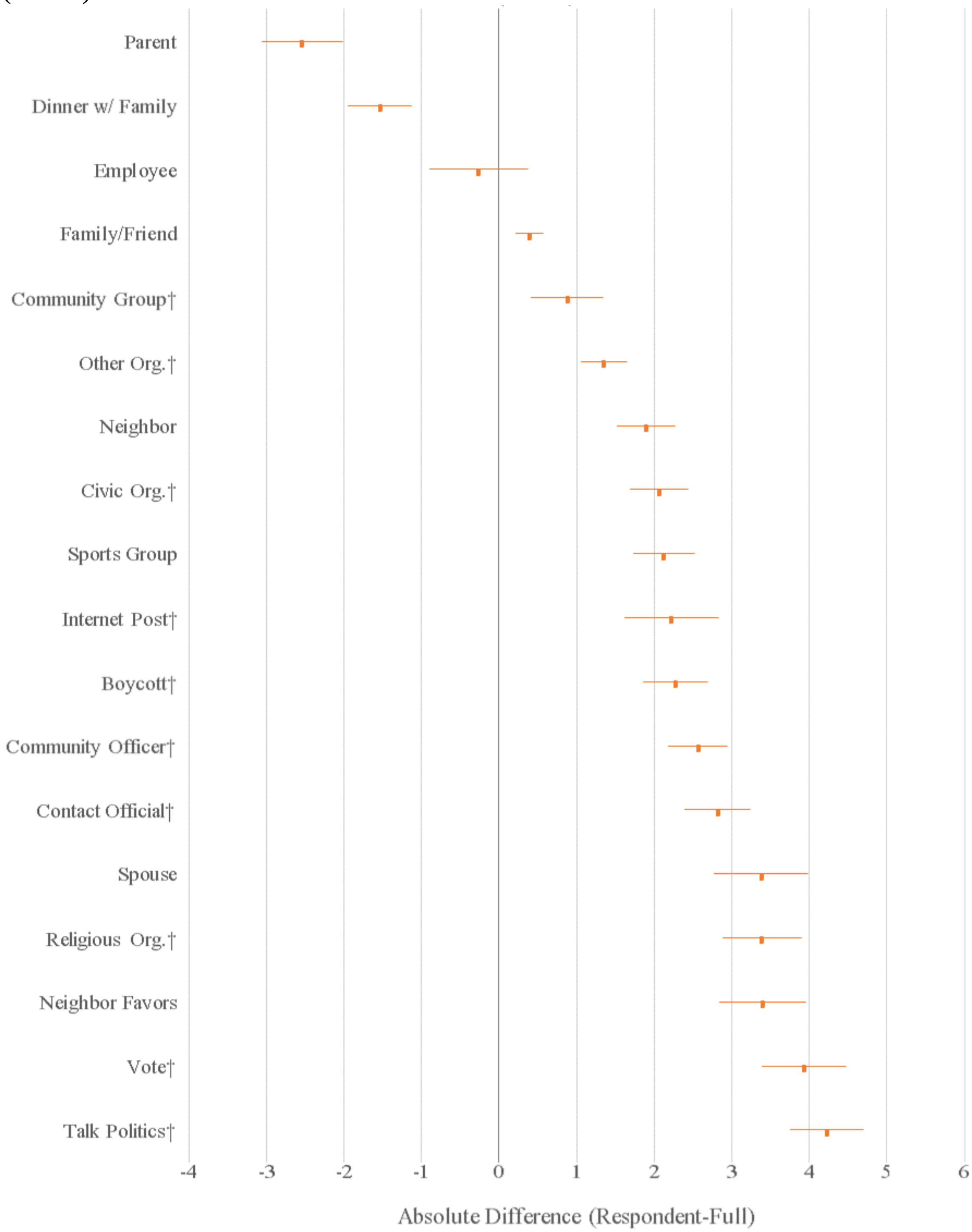
Figure 10: Relative Difference in Univariate Estimates ((Respondents-Full Sample)/Full Sample) (SHARE)



† Identifies civic and political activities and roles.

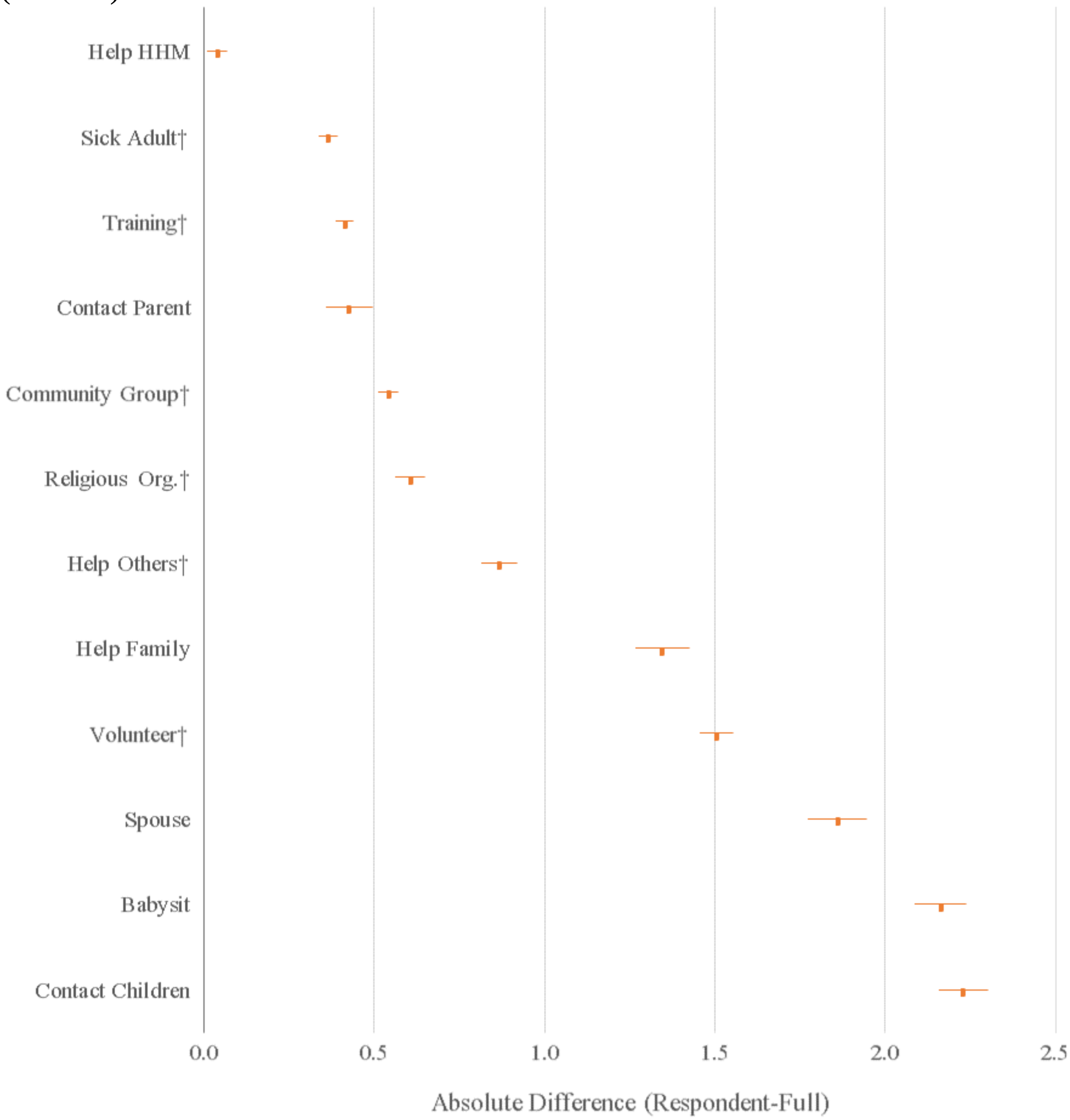
Reordering the data by absolute difference told a different story. The civic and political variables were interspersed throughout the range of values, suggesting no relationship between civic engagement and level of bias (Figures 11 and 12 for ATUS and SHARE, respectively). The discrepancies between the findings of the relative and absolute analysis were a result of the differences in the point estimates across item types. Civic and political activities and roles were generally less common than other types of activities and roles. As a result, similar absolute differences resulted in larger relative differences among the civic and political measures. In evaluating the hypothesis, more weight was given to the relative analysis. A one percentage point change in the estimate for an activity with 98 percent rate is less important to researchers than a one percentage point difference for an activity with a 10 percent participation rate. Overall, there was some evidence for general support of hypothesis 2b, but findings were not significant nor consistent between analysis method.

Figure 11: Absolute Difference of Univariate Estimates (Respondents-Full Sample) (ATUS)



† Identifies civic and political activities and roles.

Figure 12: Absolute Difference of Univariate Estimates (Respondents-Full Sample) (SHARE)



† Identifies civic and political activities and roles.

4.2.3 Multivariate Bias

To test Hypothesis 2c – coefficients of independent variables in multivariate models used to predict social activities and roles should be unbiased – logit models were estimated in both ATUS and SHARE (Tables 14 and 15, respectively). A total of 464 comparisons were made between coefficients of models that used the full sample and those that used just the respondents. Given the large number of comparisons, a total of 23 comparisons would have been significant at the five percent level by chance. However, 246 (53.0 percent, after applying the FDR) of the beta coefficients in the respondent models were significantly different from their full sample counterparts. The proportion of significant differences was much smaller in the ATUS models (38.0 percent) than in SHARE's (82.3 percent). As with the previous analyses, much of the difference in the number of significant findings was the result of tighter confidence intervals for the SHARE coefficients. Among significant differences, the magnitude of the difference was frequently smaller in SHARE than in ATUS. Even after considering the small magnitudes of some of the differences, there was still ample evidence to suggest nonresponse bias was not limited to univariate analyses.

Table 14: Logit Predicting Various Social Activities and Social Roles (ATUS)

	Vote†			Community Group†			Civic Org.†		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
N	5,034	2,744		4,988	2,722		4,986	2,721	
Intercept	-0.314 *	-0.325	-0.012	-1.734 ***	-1.930 ***	-0.196 ‡	-3.715 ***	-3.416 ***	**0.299** **
Home Owner	0.316 ***	0.254 ***	**-0.063** **	0.134 **	0.141 *	0.007	0.221 *	0.240 *	0.019
Race/Ethnicity (ref=NH White)									
NH Black	0.838 ***	0.874 ***	0.036	0.319 ***	0.233	-0.086	-0.141	-0.291	-0.150 ‡
Hispanic	-0.589 ***	-0.709 ***	**-0.120** **	-0.108	-0.173	-0.065	-0.319	-0.336	-0.017
NH Other	-0.695 ***	-0.627 **	0.068	-0.401 **	-0.353 *	0.048	-0.086	0.107	**0.193** **
Education (ref=LT HS)									
High School	-0.385 ***	-0.357 ***	0.029	-0.324 ***	-0.165	**0.159** ***	-0.175	-0.147	0.028
Some College	0.144 *	0.207 *	0.063 ‡	-0.028	0.060	**0.087** **	0.381 **	0.345 *	-0.036
College Degree or More	0.540 ***	0.522 ***	-0.018	0.530 ***	0.658 ***	**0.128** ***	0.120	0.113	-0.007
Married	-0.150 ***	-0.145 *	0.005	-0.022	-0.089	**-0.067** **	0.016	-0.027	-0.043
Female	0.043	-0.034	**-0.076** ***	0.020	0.033	0.012	-0.019	-0.056	-0.037
Age	0.022 ***	0.024 ***	0.002	0.000	0.003	0.003	0.015 **	0.012	-0.003
Employed	0.031	0.019	-0.012	-0.058	-0.129	**-0.071** **	0.107	0.078	-0.030
Children in Household	0.018	0.002	-0.016	0.612 ***	0.674 ***	**0.061** *	-0.032	-0.072	-0.039
Income (ref=LT $20k)									
$20,000-$39,999	-0.192 **	-0.244 *	-0.052	-0.281 ***	-0.394 ***	**-0.114** **	-0.004	-0.050	-0.047
$40,000-$59,999	0.231 **	0.162	-0.069	0.135	0.118	-0.016	0.126	0.191	0.065
$60,000-$99,999	0.190 *	0.130	-0.060	0.277 **	0.269 **	-0.009	0.087	0.065	-0.022
$100,000 or More	0.402 ***	0.503 ***	0.101	0.277 **	0.251 *	-0.025	0.087	0.083	-0.004

Table 14: Logit Predicting Various Social Activities and Social Roles (ATUS) (cont'd)

	Sports Group			Religious Org.†			Other Org.†		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
N	4,988	2,722		4,977	2,718		4,982	2,720	
Intercept	-1.961 ***	-1.634 ***	**0.327** ***	-1.737 ***	-1.567 ***	0.170 ‡	-3.799 ***	-3.891 ***	-0.092
Home Owner	0.098	0.112	0.014	0.221 ***	0.153 *	**-0.068** ***	0.130	0.066	-0.064 ‡
Race/Ethnicity (ref=NH White)									
NH Black	-0.007	0.039	0.047	0.212 *	0.226 *	0.014	-0.396 *	-0.612 *	**-0.217** **
Hispanic	-0.345 **	-0.440 *	-0.095	-0.199	-0.109	0.090 ‡	-0.077	0.218	**0.294** ***
NH Other	0.038	0.106	0.068	-0.112	-0.127	-0.016	0.139	0.132	-0.008
Education (ref=LT HS)									
High School	-0.346 ***	-0.447 ***	-0.101 ‡	-0.386 ***	-0.314 ***	0.073 ‡	-0.681 ***	-0.793 ***	-0.111
Some College	0.000	0.012	0.011	0.039	0.048	0.009	0.263 *	0.334 *	0.071
College Degree or More	0.454 ***	0.455 ***	0.001	0.303 ***	0.309 ***	0.006	0.555 ***	0.883 ***	**0.328** ***
Married	-0.013	-0.067	**-0.054** *	0.196 ***	0.198 ***	0.002	0.035	-0.074	**-0.108** ***
Female	0.051	-0.022	**-0.073** ***	0.110 *	0.110 *	0.000	0.090	0.076	-0.014
Age	-0.010 **	-0.012 *	-0.002	0.007 *	0.008	0.001	0.010 *	0.015 **	0.005 ‡
Employed	0.077	-0.015	**-0.092** ***	-0.012	-0.072	**-0.060** **	0.055	0.019	-0.037
Children in Household	0.415 ***	0.410 ***	-0.004	0.132 *	0.170 **	0.038	-0.247 **	-0.177	0.071 ‡
Income (ref=LT $20k)									
$20,000-$39,999	-0.315 *	-0.144	**0.171** ***	0.014	0.032	0.018	-0.233	-0.223	0.009
$40,000-$59,999	0.192	0.188	-0.003	0.020	-0.058	-0.078 ‡	-0.084	-0.174	-0.090
$60,000-$99,999	0.326 ***	0.277	-0.049	0.059	0.004	-0.055	0.376 **	0.213	**-0.163** **
$100,000 or More	0.378 ***	0.408 **	0.030	0.054	0.215	**0.161** ***	0.254	0.172	-0.082

Table 14: Logit Predicting Various Social Activities and Social Roles (ATUS) (cont'd)

	Committee Officer†			Friend/Family			Neighbor		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
N	4,982	2,722		4,924	2,706		4,924	2,704	
Intercept	-3.008***	-2.732***	**0.276****	4.076***	4.337***	0.261	1.710***	1.518***	-0.192‡
Home Owner	0.337***	0.278**	-0.059‡	0.206*	0.223	0.018	0.305***	0.210**	**-0.095****
Race/Ethnicity (ref=NH White)									
NH Black	0.185	0.1669	-0.018	0.231	0.127	-0.104	-0.105	0.128	**0.233*****
Hispanic	-0.527	-0.644	-0.117	-0.526***	-1.081**	**-0.555****	-0.224	-0.220	0.004
NH Other	-0.009	0.1886	**0.197*****	-0.298	0.936	**1.235****	-0.033	-0.236	**-0.203****
Education (ref=LT HS)									
High School	-0.429***	-0.452***	-0.023	-0.368	-0.371	-0.004	0.062	-0.185	**-0.247*****
Some College	0.113	0.1049	-0.008	0.023	0.035	0.011	0.034	0.310	**0.276*****
College Degree or More	0.405***	0.4215***	0.017	0.337	0.226	-0.111	-0.008	-0.007	0.001
Married	0.041	0.019	-0.022	0.314*	0.515	**0.201***	0.129*	0.060	-0.069‡
Female	0.030	0.0155	-0.014	0.177	-0.010	**-0.187****	0.013	0.085	**0.072****
Age	0.011**	0.0103**	-0.001	-0.002	0.001	0.004	0.006	0.013**	**0.007*****
Employed	0.154**	0.1246*	-0.029	-0.041	0.050	0.091	-0.069	-0.134	**-0.064***
Children in Household	0.083	0.0715	-0.012	0.337*	0.190	-0.146	0.104	0.020	-0.084‡
Income (ref=LT \$20k)									
\$20,000-\$39,999	-0.154	-0.13	0.024	-0.053	-0.035	0.018	-0.283**	-0.240	0.044
\$40,000-\$59,999	0.109	0.009	-0.100‡	0.140	0.344	0.204	0.131	0.131	0.001
\$60,000-\$99,999	0.163	0.123	-0.040	0.315	-0.087	**-0.401*****	0.099	0.046	-0.053
\$100,000 or More	0.441***	0.4351**	-0.005	0.081	-0.196	-0.277	0.236	0.530*	**0.294*****

Table 14: Logit Predicting Various Social Activities and Social Roles (ATUS) (cont'd)

	Contact Official†			Boycott†			Internet Post†		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
N	5,006	2,733		4,995	2,727		4,974	2,713	
Intercept	-2.580 ***	-2.409 ***	0.170 ‡	-2.524 ***	-2.294 ***	**0.230** *	-0.849 ***	-0.722 **	0.128 ‡
Home Owner	0.278 ***	0.307 ***	0.029	0.040	0.010	-0.029	0.029	-0.069	**-0.098** ***
Race/Ethnicity (ref=NH White)									
NH Black	-0.063	-0.175	-0.113 ‡	-0.198	-0.325 *	-0.127 ‡	0.023	-0.173	**-0.196** ***
Hispanic	-0.387	-0.512 **	-0.125	-0.452 *	-0.285	**0.167** **	-0.102	-0.035	0.067
NH Other	0.053	0.307	**0.254** ***	0.077	0.052	-0.025	-0.107	0.047	**0.154** **
Education (ref=LT HS)									
High School	-0.516 ***	-0.628 ***	**-0.112** *	-0.366 ***	-0.421 ***	-0.055	-0.316 ***	-0.212	**0.104** **
Some College	0.166	0.223	0.057	0.183 *	0.276 *	**0.093** **	0.250 ***	0.273 **	0.024
College Degree or More	0.525 ***	0.620 ***	**0.095** **	0.554 ***	0.655 ***	**0.101** **	0.413 ***	0.492 ***	**0.078** **
Married	0.003	-0.076	**-0.079** ***	-0.041	-0.072	-0.031	0.005	-0.044	**-0.049** **
Female	0.034	-0.018	**-0.052** **	0.025	0.071	0.046 ‡	-0.023	-0.072	**-0.050** **
Age	0.005	0.004	-0.002	0.001	-0.001	-0.002	-0.007 **	-0.007	0.000
Employed	-0.014	-0.029	-0.015	0.015	0.034	0.019	-0.022	-0.063	**-0.042** **
Children in Household	-0.062	-0.086	-0.024	-0.043	-0.037	0.006	-0.031	-0.044	-0.013
Income (ref=LT $20k)									
$20,000-$39,999	-0.213 *	-0.238	-0.025	-0.211	-0.210	0.001	0.007	-0.045	-0.052
$40,000-$59,999	0.079	0.094	0.015	0.074	0.185	**0.111** **	0.117	0.211 *	**0.094** **
$60,000-$99,999	0.128	0.117	-0.010	0.302 **	0.215	**-0.087** **	0.120	0.138	0.018
$100,000 or More	0.263 *	0.270 *	0.007	0.122	-0.013	**-0.135** **	0.198 *	0.245 *	0.047

Table 14: Logit Predicting Various Social Activities and Social Roles (ATUS) (cont'd)

	Talk Politics†			Dinner w/ Family			Neighbor Favors		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
N	4,920	2,696		5,008	2,731		4,894	2,688	
Intercept	0.726 ***	0.900 ***	0.174 ‡	5.941 ***	6.087 ***	0.146	0.235	0.146	-0.089
Home Owner	0.181 ***	0.136 *	-0.046	0.328 ***	0.330 *	0.002	0.372 ***	0.426 ***	**0.055** **
Race/Ethnicity (ref=NH White)									
NH Black	0.289 **	0.300 *	0.012	-0.087	-0.115	-0.027	-0.044	0.019	0.063
Hispanic	-0.222 **	-0.337 *	**-0.116** **	0.001	0.202	**0.201** **	-0.101	-0.130	-0.029
NH Other	-0.472 ***	-0.336	**0.136** *	0.461 *	0.412	-0.049	-0.195	-0.271	-0.077
Education (ref=LT HS)									
High School	-0.386 ***	-0.338 ***	0.048	0.256 **	0.306 *	0.050	-0.011	-0.139	**-0.128** ***
Some College	0.204 ***	0.229 **	0.025	-0.147	-0.118	0.029	0.018	0.246 *	**0.228** ***
College Degree or More	0.607 ***	0.559 ***	-0.048	-0.772 ***	-0.745 ***	0.027	0.111	0.164	0.053
Married	-0.003	0.011	0.014	1.938 ***	2.223 ***	**0.285** ***	0.127 ***	0.069	**-0.058** **
Female	-0.015	-0.075	**-0.060** **	0.153 *	0.199 *	0.046 ‡	-0.001	-0.022	-0.020
Age	0.004	0.004	0.001	-0.051 ***	-0.052 ***	-0.001	0.005	0.007	0.002
Employed	0.111 *	0.056	**-0.055** *	-0.295 ***	-0.285 **	0.010	-0.061	-0.107 *	-0.046 ‡
Children in Household	-0.063	-0.051	0.012	1.334 ***	1.277 ***	-0.057	0.186 ***	0.173 *	-0.013
Income (ref=LT $20k)									
$20,000-$39,999	-0.212 *	-0.278 *	-0.066	-0.386 **	-0.462 **	-0.076	-0.038	0.048	**0.086** **
$40,000-$59,999	-0.051	0.053	**0.104** *	0.094	0.042	-0.052	-0.119	-0.163	-0.044
$60,000-$99,999	0.313 **	0.281	-0.031	0.345 *	0.439 *	0.093	-0.001	-0.058	-0.057
$100,000 or More	0.482 ***	0.499 ***	0.017	0.799 ***	0.788 ***	-0.011	0.223 **	0.308 **	0.085 ‡

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ‡ no longer significant after FDR adjustment (only applies to "Difference" column)

† Identifies civic and political activities and roles.

Table 15: Logit Predicting Various Social Activities and Social Roles (SHARE)

	Volunteer†			Sick Adult†			Community Group†		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
Intercept	-0.725	-0.400	**0.325** ***	-1.361 ***	-1.341 **	0.020	-1.893 ***	-1.543 **	**0.350** ***
Home Owner	0.187 ***	0.142 **	**-0.044** ***	0.040	0.011	**-0.029** ***	0.099	0.002	**-0.097** ***
Country (ref=Austria)									
Germany	-0.101	0.015	**0.115** ***	0.015	0.111	**0.096** ***	-0.268 *	-0.233	**0.036** **
The Netherlands	0.959 ***	0.937 ***	**-0.022** ***	0.322 *	0.325 *	0.003	-0.274 **	-0.344 ***	**-0.070** ***
Spain	-1.091 ***	-1.032 ***	**0.059** ***	-0.547 **	-0.551 *	-0.004	-0.624 ***	-0.517 *	**0.107** ***
Italy	-0.115	-0.019	**0.096** ***	-0.510 ***	-0.699 ***	**-0.188** ***	-0.250	-0.046	**0.204** ***
France	0.553 ***	0.572 ***	**0.019** ***	0.273 *	0.342 **	**0.069** ***	-0.066	0.003	**0.069** ***
Denmark	0.624 ***	0.472 **	**-0.153** ***	-0.377	-0.400	-0.024	-0.084	-0.375 *	**-0.291** ***
Greece	-1.145 ***	-1.231 ***	**-0.085** ***	0.047	0.040	-0.007	0.429 **	0.371 *	**-0.057** ***
Switzerland	0.581 **	0.536 **	**-0.045** ***	0.398 ***	0.422 ***	**0.024** ***	0.712 ***	0.719 ***	0.007
Education (ref=Primary School or Less)									
Some Secondary School	-0.224 **	-0.309 ***	**-0.086** ***	-0.159	-0.231 *	**-0.072** ***	-0.022	-0.013	0.009
Secondary School	0.090	0.087	-0.003	0.125	0.088	**-0.036** ***	-0.017	0.113	**0.129** ***
First Stage Tertiary or Higher	0.773 ***	0.786 ***	**0.012** **	0.355 ***	0.425 ***	**0.070** ***	0.944 ***	0.875 ***	**-0.069** ***
Married	0.038	0.023	**-0.014** ***	0.060	0.034	**-0.025** ***	0.097	0.131	**0.034** ***
Female	-0.023	-0.057	**-0.034** ***	0.316 ***	0.289 ***	**-0.027** ***	-0.351 ***	-0.390 ***	**-0.039** ***
Age	-0.025 ***	-0.027 ***	**-0.001** ***	-0.028 ***	-0.028 ***	0.000	-0.023 ***	-0.025 **	**-0.003** ***
Employed	-0.278 ***	-0.298 ***	**-0.020** ***	-0.173 **	-0.227 *	**-0.054** ***	0.146 **	0.168 *	**0.023** ***
HH Size	-0.047	-0.112	**-0.064** ***	0.024	0.036	**0.012** **	-0.043	-0.081	**-0.038** ***
Income (per 1,000€)	0.000	0.000	**0.000** ***	0.001	0.001 *	**0.000** *	0.000	0.000	**0.000** ***

Table 15: Logit Predicting Various Social Activities and Social Roles (SHARE) (cont'd)

	Babysit			Help HHM			Help Family		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
Intercept	0.258	0.398	**0.140** ***	-5.777 ***	-5.766 ***	0.010	2.342 ***	2.312 ***	-0.030
Home Owner	0.104 **	0.057	**-0.047** ***	-0.088	-0.085	0.003	0.126 ***	0.069	**-0.057** ***
Country (ref=Austria)									
Germany	-0.074	0.041	**0.115** ***	0.075	0.250	**0.175** ***	0.065	0.114	**0.049** ***
The Netherlands	0.328 ***	0.316 *	**-0.012** *	-0.105	-0.250	**-0.145** ***	0.500 ***	0.501 ***	0.001
Spain	-0.233 ***	-0.267 **	**-0.034** ***	0.421 ***	0.230	**-0.190** ***	-0.524 ***	-0.422 ***	**0.102** ***
Italy	-0.405 ***	-0.377 ***	**0.027** ***	0.265 **	0.393 ***	**0.129** ***	-0.110	-0.145	**-0.034** ***
France	0.243 **	0.255 ***	**0.012** **	0.032	0.058	**0.025** **	0.054	0.065	**0.011** *
Denmark	0.791 ***	0.649 ***	**-0.142** ***	-0.925 ***	-0.960 *	-0.035	0.344	0.360	0.016
Greece	-0.283 ***	-0.283 ***	0.000	-0.143	-0.062	**0.082** ***	-0.271 **	-0.351 **	**-0.079** ***
Switzerland	-0.395	-0.375	**0.020** ***	0.150	0.127	**-0.023** **	0.154	0.183	**0.029** ***
Education (ref=Primary School or Less)									
Some Secondary School	0.009	0.014	0.005	-0.053	0.005	**0.058** ***	-0.060	-0.032	**0.028** ***
Secondary School	0.006	-0.005	**-0.011** **	0.049	0.082	**0.033** ***	0.115 **	0.091	**-0.024** ***
First Stage Tertiary or Higher	-0.199 **	-0.196 **	0.004	-0.112	-0.280 *	**-0.167** ***	0.269 ***	0.260 ***	-0.009 ‡
Married	0.533 ***	0.542 ***	**0.009** ***	0.267 ***	0.331 ***	**0.064** ***	0.147 ***	0.147 ***	0.000
Female	0.115 ***	0.117 ***	0.002	0.141 **	0.155 **	**0.014** ***	0.075 **	0.070 **	**-0.005** *
Age	-0.019 ***	-0.019 ***	0.000	0.021 ***	0.020 **	0.000	-0.058 ***	-0.055 ***	**0.003** ***
Employed	-0.336 ***	-0.355 ***	**-0.019** ***	-0.269 ***	-0.260 **	0.008	-0.036	-0.041	**-0.005** *
HH Size	-0.205 ***	-0.237 ***	**-0.032** ***	0.427 ***	0.377 ***	**-0.050** ***	-0.076	-0.087 *	**-0.012** ***
Income (per 1,000€)	0.000	0.000	**0.000** ***	0.001 ***	0.001 ***	**0.000** ***	0.001 *	0.001 **	**0.000** ***

Table 15: Logit Predicting Various Social Activities and Social Roles (SHARE) (cont'd)

	Contact Parent			Training†			Help Others†		
	Full	Respond.	Diff.	Full	Respond.	Diff.	Full	Respond.	Diff.
Intercept	10.522***	10.399***	**-0.123*****	0.362	0.175	**-0.187*****	0.245	0.070	**-0.175*****
Home Owner	0.027	0.016	**-0.010*****	0.194**	0.170**	**-0.024*****	0.051	0.006	**-0.045*****
Country (ref=Austria)									
Germany	-0.219***	-0.176*	**0.043*****	0.001	0.075	**0.074*****	0.177	0.251	**0.074*****
The Netherlands	-0.191***	-0.212**	**-0.021*****	0.363***	0.253***	**-0.110*****	0.540***	0.506***	**-0.034*****
Spain	0.208**	0.182*	**-0.025*****	-0.500	-0.487	0.013	-0.997***	-0.859***	**0.138*****
Italy	0.150*	0.093	**-0.056*****	-1.229***	-1.195***	**0.034****	0.044	0.105	**0.061*****
France	0.333***	0.242***	**-0.091*****	-0.212	-0.316*	**-0.105*****	0.127	0.067	**-0.060*****
Denmark	-0.447**	-0.400**	**0.047*****	0.269**	0.237	**-0.032*****	0.666***	0.614***	**-0.053*****
Greece	0.275***	0.242**	**-0.033*****	-0.110	-0.130	**-0.019*****	-0.665***	-0.738***	**-0.073*****
Switzerland	0.099	0.221	**0.122*****	1.651***	1.746***	**0.095*****	0.489***	0.516***	**0.026*****
Education (ref=Primary School or Less)									
Some Secondary School	-0.031	-0.070	**-0.039*****	-0.457***	-0.554***	**-0.097*****	0.000	0.011	**0.011***
Secondary School	0.109*	0.084	**-0.025*****	0.194*	0.173	**-0.020*****	0.099	0.144*	**0.045*****
First Stage Tertiary or Higher	0.310***	0.316***	0.006	1.124***	1.154***	**0.031*****	0.285***	0.243*	**-0.042*****
Married	0.140***	0.203***	**0.063*****	-0.027	-0.098	**-0.072*****	-0.146***	-0.167*	**-0.021*****
Female	0.032	0.047	**0.015*****	0.200***	0.164**	**-0.036*****	-0.114**	-0.168***	**-0.054*****
Age	-0.190***	-0.187***	**0.003*****	-0.059***	-0.055***	**0.003*****	-0.036***	-0.031***	**0.005*****
Employed	0.046	0.048	0.003	0.234***	0.222***	**-0.012****	-0.202***	-0.174***	**0.028*****
HH Size	-0.024	-0.029	**-0.005***	0.018	0.076	**0.059*****	-0.067	-0.080	**-0.013****
Income (per 1,000€)	0.001**	0.001*	**0.000*****	0.000	0.000	**0.000*****	0.000	0.000	0.000

Table 15: Logit Predicting Various Social Activities and Social Roles (SHARE) (cont'd)

	Religious Org.†			Contact Children		
	Full	Respond.	Diff.	Full	Respond.	Diff.
Intercept	-3.432 ***	-3.052 ***	**0.380** ***	-1.532 ***	-1.873 ***	**-0.341** ***
Home Owner	0.135 *	0.133 *	-0.001	0.054	-0.003	**-0.057** ***
Country (ref=Austria)						
Germany	-0.250 *	-0.285 *	**-0.035** ***	-0.184	0.002	**0.186** ***
The Netherlands	-0.028	-0.014	**0.014** *	0.092	0.109 *	**0.018** ***
Spain	0.077	0.074	-0.003	-0.413 ***	-0.431 ***	-0.018 ‡
Italy	-0.817 ***	-0.803 ***	0.014	-0.456 ***	-0.281 *	**0.175** ***
France	-0.820 ***	-0.826 ***	-0.006	0.150	0.055	**-0.095** ***
Denmark	-0.810 ***	-0.683 ***	**0.127** ***	0.707 ***	0.529 ***	**-0.178** ***
Greece	1.635 ***	1.612 ***	**-0.023** ***	0.021	-0.024	**-0.045** ***
Switzerland	0.178 *	0.110	**-0.067** ***	0.047	0.022	**-0.025** ***
Education (ref=Primary School or Less)						
Some Secondary School	-0.014	-0.019	-0.006	0.009	-0.002	-0.011
Secondary School	-0.056	-0.100	**-0.045** ***	-0.037	0.027	**0.064** ***
First Stage Tertiary or Higher	0.414 ***	0.483 ***	**0.069** ***	-0.082	-0.121	**-0.039** **
Married	-0.001	-0.009	**-0.008** *	0.566 ***	0.540 ***	**-0.026** ***
Female	0.279 ***	0.262 ***	**-0.018** ***	0.334 ***	0.328 ***	-0.006
Age	0.017 ***	0.013 *	**-0.004** ***	0.026 ***	0.031 ***	**0.005** ***
Employed	-0.061	-0.076	**-0.015** ***	-0.028	-0.034	-0.006
HH Size	0.076 *	0.068	**-0.008** *	0.906 ***	0.971 ***	**0.065** ***
Income (per 1,000€)	0.000	-0.001	**0.000** ***	-0.001 *	0.000	**0.000** ***

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ‡ no longer significant after FDR adjustment (only applies to "Difference" column)

† Identifies civic and political activities and roles.

The beta coefficients were examined in a variety of ways in an attempt to identify a pattern of bias. First, the direction of the bias was examined. However, the direction of the bias was inconsistent across variables and surveys. This was not unexpected since the relationship between the dependent and independent variable could have been positive or negative, and the effect of nonresponse could have accentuated or deemphasized the relationship.

Second, the number of significantly different betas was examined by dependent variable (Tables 16 and 17 for ATUS and SHARE, respectively). Models predicting some social activities and roles were more likely to result in significantly different coefficients between samples. For example, the ATUS model predicting civic organization membership had many more unbiased coefficients than the model predicting communication with a neighbor (18 vs. eight unbiased coefficients). Unfortunately, no clear pattern emerged on the type of dependent variable that suffered from biased coefficients.

Table 16: Number of Significant Differences Found in the Multivariate Models in Table 14 by Significance Level and Dependent Variable (ATUS)

	p-value			
	<0.001	<0.01	<0.05	n.s.
Civic Org.†	0	2	0	15
Dinner w/ Family	1	1	0	15
Committee Officer†	1	1	0	15
Vote†	1	2	0	14
Religious Org.†	2	1	0	14
Friend/Family	1	3	1	12
Talk Politics†	0	2	3	12
Contact Official†	2	2	1	12
Neighbor Favors	2	3	0	12
Other Org.†	3	2	0	12
Community Group†	2	4	1	10
Boycott†	0	6	1	10
Internet Post†	2	7	0	8
Neighbor	5	3	1	8

† Identifies civic and political activities and roles.

Table 17: Number of Significant Differences Found in the Multivariate Models in Table 15 by Significance Level and Dependent Variable (SHARE)

	p-value			
	<0.001	<0.01	<0.05	n.s.
Sick Adult†	11	1	1	6
Religious Org.†	11	0	3	5
Babysit	11	2	1	5
Help HHM	12	2	0	5
Help Family	11	0	3	5
Contact Children	14	1	0	4
Contact Parent	16	0	1	2
Community Group†	16	1	0	2
Sports Group	17	0	0	2
Help Others†	16	1	1	1
Training†	16	2	0	1
Volunteer†	17	1	0	1

† Identifies civic and political activities and roles.

Lastly, bias was examined by independent variable (Tables 18 and 19 for ATUS and SHARE, respectively). For ATUS, marital status was found to be biased more often than other variables, with eight significantly different coefficients. Age and having children in the household was least likely to have a biased coefficient with only one significantly different coefficients across the 15 models. The number of significant differences was more consistent across the SHARE independent variables. All variables were observed to have between seven and 11 significant differences. As with the previous analysis, no clear patterns were identified.

Table 18: Number of Significant Differences Found in the Multivariate Models in Table 14 by Significance Level and Independent Variable (ATUS)

	p-value			
	<0.001	<0.01	<0.05	n.s.
Intercept	1	2	1	11
Home Owner	2	3	0	10
Race/Ethnicity (ref=NH White)				
NH Black	2	1	0	12
Hispanic	1	5	0	9
NH Other	2	4	1	8
Education (ref=LT HS)				
High School	3	1	1	10
Some College	2	2	0	11
College Degree or More	2	3	0	10
Married	3	3	2	7
Female	2	5	0	8
Age	1	0	0	14
Employed	1	3	2	9
Children in Household	0	0	1	14
Income (ref=LT $20k)				
$20,000-$39,999	1	2	0	12
$40,000-$59,999	0	2	1	12
$60,000-$99,999	1	2	0	12
$100,000 or More	2	1	0	12

Table 19: Number of Significant Differences Found in the Multivariate Models in Table 15 by Significance Level and Independent Variable (SHARE)

	p-value			
	<0.001	<0.01	<0.05	n.s.
Intercept	8	0	0	3
Home Owner	9	0	0	2
Country (ref=Austria)				
Germany	10	1	0	0
The Netherlands	7	0	2	2
Spain	8	0	0	4
Italy	9	1	0	1
France	7	2	1	1
Denmark	8	0	0	3
Greece	9	0	0	2
Switzerland	9	1	0	1
Education (ref=Primary School or Less)				
Some Secondary School	6	0	1	4
Secondary School	9	1	0	1
First Stage Teriary or Higher	7	2	0	3
Married	9	0	1	1
Female	8	0	1	2
Age	8	0	0	3
Employed	6	1	1	3
HH Size	7	2	2	0
Income (per 1,000€)	9	0	1	1

It was possible for a coefficient to have a significantly different magnitude between models, but have similar significance levels. For example, the effect of being Hispanic on voting was significantly larger in the ATUS respondent model than the full sample model (-0.71 vs. -0.59, respectively), but both coefficients were significant at the 0.001 level. By contrast, although the magnitude of the effect of non-Hispanic black on predicting membership to a community organization was unchanged across the full sample and respondent samples, the coefficient was significant at the 0.001 level in the full model but was not significant in the respondent model. In the first example, a researcher would likely draw the same conclusion about the effect of being Hispanic on voting even though the magnitude of the effect was biased. In the second example, a researcher would likely have drawn a different conclusion of the effect of being non-Hispanic black even though the magnitude of the effect was unbiased. Across both surveys, the significance level of the coefficients was observed to be consistent across the full sample and respondent models 75.2 percent (n=349) of the time (Tables 20 and 21 for ATUS and SHARE, respectively). This pattern was similar across the two surveys. Among the 115 coefficients that changed significance level between models, the vast majority of differences (70.4 percent) were marginal. That is, they changed from one level of significance (e.g., 0.01) to a neighboring level (e.g., 0.05 or 0.001). Overall, the coefficients of the multivariate models were biased, resulting in no support for hypothesis 2c. While there was no statistical support for the hypothesis, there was evidence that the bias was unlikely to change the interpretation of the models since the significance level of the coefficients was often unchanged between models.

Table 20: Summary of Changes in the Level of Significance of the Beta Coefficients Found in the Multivariate Models in Table 14 (ATUS)

		Full Sample			
		n.s.	$p<0.05$	$p<0.01$	$p<0.001$
Respondents	n.s.	131	12	8	6
	$p<0.05$	5	12	11	5
	$p<0.01$	2	2	4	10
	$p<0.001$	0	0	0	47

Table 21: Summary of Changes in the Level of Significance of the Beta Coefficients Found in the Multivariate Models in Table 15 (SHARE)

		Full Sample			
		n.s.	*p*<0.05	*p*<0.01	*p*<0.001
Respondents	n.s.	81	9	2	5
	p<0.05	9	3	6	8
	p<0.01	0	2	5	12
	p<0.001	0	0	5	81

4.3 Discussion

Of the 507 tests performed in this chapter to identify nonresponse bias, 308 (60.7 percent) yielded significant results, far too many to be attributed to chance. Nonresponse bias existed in both surveys, in univariate estimates (H2a), and in multivariate models (H2c). What differentiates the 308 biased statistics from the remaining 199? Univariate estimates of civic and political indicators trended toward higher levels of bias, but the differences were not significant and the trend did not hold for the multivariate analyses (H2b). Further evaluation of the independent variables in the logit models was also unsuccessful in uncovering a pattern.

The multivariate results were somewhat at odds with those of Abraham and her colleagues (2009). Within every subgroup analyzed, they found respondents were more likely to report volunteerism, suggesting multivariate models would be unbiased. However, the magnitude of the difference between respondents and nonrespondents varied by subgroup. When statistical tests were performed on the multivariate models in ATUS and SHARE, the variation by subgroup was large enough to bias the coefficients of the models. But, consistent with Abraham et al (2009), the bias was often small and unlikely to alter interpretations.

Chapter 5: Social Integration and Weighting

This chapter tests the last three hypotheses outlined in Chapter 1:

H3a: The addition of a social integration indicator into the weighting methodology should significantly reduce nonresponse bias associated with prevalence estimates of social activities and roles when compared to base-weighted estimates.

H3b: The addition of a social integration indicator into the weighting methodology should significantly reduce nonresponse bias associated with prevalence estimates of social activities and roles when compared to traditionally-weighted estimates.

H3c: The addition of a social integration indicator into the weighting methodology should eliminate nonresponse bias associated with prevalence estimates of social activities and roles.

5.1 Methods

5.1.1 Simplifying the Social Integration Indicator

In order to test our hypotheses, a measure of social integration was incorporated into the nonresponse adjustment of the weighting procedures for both ATUS and SHARE Wave II. Ideally, the weighting process would include the latent categorical variable constructed in Section 3.2.1 (Figures 3 and 4). However, that latent variable was created using 18 endogenous variables for ATUS and 12 for SHARE Wave II. Even if the incorporation of social integration in weight construction eliminated bias, it would be unrealistic to expect it to be adopted. Researchers would have to add all endogenous variables to their survey, significantly increasing the length and cost.

To be more realistic, a reduced LCA was constructed, resulting in a simplified social integration indicator for use in weight construction. Similar to the difference between a reduced and full regression model, a reduced LCA includes fewer endogenous variables. Unfortunately, as endogenous variables are dropped from the LCA, posterior probabilities and class assignment change. An individual that had a higher probability of being in the integrated class under the full model may be assigned a higher probability of being isolated under the reduced model, resulting in a different modal class assignment. It was necessary to identify the subset of endogenous variables that would minimize the change to the model results. No systematic method exists, so four different approaches were tested. Under the first method, variables with high standardized factor loadings in the CFA in Section 3.2.1 (Figures 7 and 8) were included in the LCA. The cutoff value differed by survey. The average factor loadings were much higher in ATUS, justifying higher cutoffs. Two factor loading cutoffs were used for each survey. The cutoffs and resulting variables selected are shown in Figures 13 and 14.

Figure 13: Variable Subsets Tested for LCA Used in Weighting (ATUS)

Standardized Loadings		Absolute Difference	Relative Difference	Civic Engagement
≥0.75	≥0.65	≥0.25	≥0.85	Single-Loading Variables
Contact Official†	Contact Official†	Contact Official†	Contact Official†	Vote†
Dinner w/ Family	Dinner w/ Family	Boycott†	Boycott†	Other Org.†
Neighbor Favors	Neighbor Favors	Vote†	Other Org.†	Internet Post†
Neighbor	Neighbor	Religious Org.†	Community†	Contact Official†
	Talk Politics†			Talk Politics†
	Boycott†			Boycott†

Figure 14: Variable Subsets Tested for LCA Used in Weighting (SHARE)

Standardized Loadings		Absolute Difference	Relative Difference	Civic Engagement
≥0.65	≥0.55	≥0.25	≥0.50	Single-Loading Variables
Sick Adult†	Sick Adult†	Babysit	Babysit	Volunteer †
Contact Children	Contact Children	Contact Children	Contact Children	Religious Org.†
Help Family	Help Family	Spouse	Spouse	Community Group†
	Community Group†			Help Others†
	Contact Parents			

A second method of reducing the model was to select variables in which the absolute difference in item probabilities between classes in the full LCA was high. For both surveys, the cutoff was set to 0.25. Four ATUS variables had absolute differences larger than 0.25 while SHARE had three. The third approach was similar to the second, except it used relative difference to select variables. As with the factor loadings method, the cutoff varied by survey given larger relative differences in ATUS. Six ATUS variables were selected for inclusion while three were selected for SHARE. The selected SHARE variables were the same under the second and third method.[19]

In the last method, civic or political engagement variables that only loaded onto the civic and political engagement factor in the CFA in Chapter 3 were selected. This resulted in a total of six variables selected for ATUS and 4 for SHARE. A more thorough investigation would have included all civic and political engagement variables, but this would have resulted in nine variables selected in ATUS. While significantly reduced from the original 18 endogenous variables, nine variables is still a large number to add to a survey for weighting purposes.

An LCA was produced under each method using all sampled individuals previously included in analyses (n = 5,150 and 19,299 for ATUS and SHARE, respectively). The resulting models were evaluated in four steps. First, the appropriate number of classes for each model was determined using the same approach used in Section 3.1.1. Different class models were compared on entropy, BIC, and VLMR LRT. The reduced model should have two classes in order to be comparable to the full model. Therefore, only models in which two classes were identified were placed in the second step, assessing

[19] Since the model was unchanged from the second method, the results are only displayed once in Section 5.2.1 under the heading "Absolute Difference."

model fit. Any models with exceptionally low entropy were considered to have insufficient fit and were excluded from additional steps. Third, the modal class assignments determined by each of the remaining reduced models were compared to those of the full model from Section 3.1.2. For each model, the false positive (FPR), false negative (FNR), and accuracy rates were calculated:

$$FPR = \frac{FP}{FP + TN}$$

$$FNR = \frac{FN}{FN + TP}$$

$$Accuracy = \frac{TN + TP}{N}$$

where

FP = the number of false positives, individuals placed in the integrated class in the reduced model that were not placed in the integrated class in the full model

FN = the number of false negatives, individuals placed in the isolated class in the reduced model that were placed in the integrated class in the full model

TP = the number of true positives, individuals who were placed in the integrated class in the reduced model and the full model

TN = the number of true negatives, individuals who were placed in isolated class in the reduced model and the full model

N = sample size

The lower the FPR and FNR and higher the accuracy rate, the more closely the reduced model mimicked the full model assignments.

A single, two-class model was selected after the third step. It was required to have sufficient entropy and replicability to the full model. In order for any weighting variable to have a significant and positive effect on bias, it must be strongly correlated with both response and the variable of interest (Little & Vartavarian 2005). Therefore, the selected model was run on the respondents, and modal class assignments were produced. Chi-square tests were performed to identify any relationship between the resulting social integration variable and the social activity and roles estimates. A chi-square test was used as opposed to a correlation coefficient because the data violated several correlation assumptions (e.g., normality). If a relationship between the social integration measure and the social activity and roles estimates existed, the chi-squared tests should have been significant. If they were not, a different two-class model was chosen and the process was repeated.

Once a final subset of variables was determined, the LCA was rerun for each survey to create population totals. For ATUS, this involved using all CE Supplement participants (*n* = 42,073), regardless of whether or not they were sampled for ATUS. The model accounted for the CE Supplement's complex sample design and used the CE Supplement's final, nonresponse-adjusted weights. For SHARE, no external control totals were available. The reduced LCA was run on the SHARE Wave I sample by country (*n* = 20,449). The resulting final class counts from the estimated models were used as the population totals. In both ATUS and SHARE, the population totals used were not independent from the respondents. Moreover, the population totals were created from sample surveys. Sampling variance based off of incomplete survey data is often underestimated, and the weighted survey estimates likely suffered from some level of bias. In the case of ATUS, the dataset used to produce the control totals was significantly larger than ATUS (8.2 times larger). As a result, the negative effects of using survey data for population totals should have been greatly reduced (Dever 2008). Regardless, the effectiveness of the alternative weights is likely overstated, especially in the case of SHARE. As a result, the findings should be interpreted with some caution.

The final reduced LCA was rerun for the ATUS and SHARE Wave II respondents (n = 2,779 and 12,904, respectively). Individuals were assigned to an integration category using modal class assignment. Modal assignment places each individual into the class in which their posterior probability is highest. This approach to class assignment differs from Chapter 3 where a three-step ML approach was applied. While the ML approach is superior in its ability to account for measurement error in class assignment, it was not feasible for this analysis. ML can only be applied in regression. Nonresponse adjustments were made using raking (for ATUS) and calibration (for SHARE Wave II), not regression. Because modal assignment does not effectively account for measurement error, variances are typically underestimated. This was of limited concern for this analysis since variance estimates were calculated across replicates, introducing variability at a later point in the analysis.

In previous analyses, cases that were missing values for one or more endogenous variables were dropped from analysis. In order to construct complete weights, these cases could not be dropped and, instead, had to be imputed. A total of 94 (3.4 percent) ATUS respondents and 248 (1.3 percent) SHARE Wave II respondents were assigned to a social integration class using nearest neighbor imputation. Neighbors were identified using all endogenous variables for which the respondent had values. Only the class assignment, not the posterior probabilities, was imputed.

5.1.2 Constructing ATUS Weights

In order to test the hypotheses, two sets of final, nonresponse-adjusted weights had to be created. First, traditional weights were reconstructed, then the alternative weights were created. To isolate the effect of social integration on the weighting technique, it was critical that the only difference between the traditional and alternative weights was the incorporation of the social integration variable. In reviewing the documentation, it was discovered that it would be impossible to use the exact same steps previously used to construct the traditional weights found on the public use file. The preexisting final weights were created by varying the population control by replicate. Unfortunately, CPS replicates (i.e., the source of the population totals) were not publically available. The preexisting weights were also applied to all 2012 ATUS respondents and did not account for CE Supplement nonresponse. Finally, they were created by month, not year. The number of cases used in this analysis was too small to allow for raking by month. Given these limitations to replication, the traditional weights were recreated at the annual level by fixing the population totals and limiting the cases to the 2,779 respondents used in this dissertation. Weighted point estimates were compared using the original weights found on the public use file and the reconstructed traditional weights. Differences were small (less than 0.1 percent), so the deviations from the original weighting scheme were deemed acceptable.

Both the traditional and the alternative weights were constructed using a four step process. In the first step, base weights were created that took into account final weights for CPS and the sampling probabilities for ATUS. Since the CPS had an overall final weight and 160 replicate final weights, ATUS also had an overall base weight and 160 replicate base weights. The original CPS replicates were created using Fay's method (Judkins 1990; Rao & Shao 1999). The same replicates were carried through from CPS, and adjustments were made to each replicate (Tupek 2004). The ATUS base weights were publically available and were used, unchanged, in this analysis.

In the second step, a non-interview adjustment was made on the base weights. In this step, a seven by two table was created to account for all combinations of reference day (Monday-Sunday) and incentive (offered vs. not offered). The sum of the base weights was calculated for all eligible respondents. In the publically available weights, eligible respondents included all 2012 ATUS eligible respondents. For this dissertation, eligible respondents were subset to those who had who had completed the CE Supplement. An adjustment factor was then computed by cell by dividing the base-weighted count of eligibles by the base-weighted count of interviews. Respondents' base weights were then inflated by the adjustment factor.

Third, a raking adjustment was used to further adjust for nonresponse. Under the standard technique, respondents were placed into a table of age by sex by race by Hispanicity. Population control totals were obtained from the CPS. One iteration of raking was conducted. The resulting values were placed into a second table of sex by education by presence of children in the household. One iteration of raking was conducted on this table. The resulting values were placed back in the first raking table and the process was repeated for a total of six iterations (Tupek 2004). This step was conducted overall and for each of the 160 replicates.

A third table was added to construct the alternative weights. This table only included the control totals for social integration. The interaction effects between social integration and various demographic variables on response were reviewed, but none were significant. As a result, no additional variables were added to the third table. The results from the first iteration of the second raking table were placed into the third table and raked to the population control totals for social integration. The results were placed back in the first table and the process was repeated for a total of six iterations. The addition of this third raking table was the only difference between the traditional and alternative weights. Finally, ATUS weights were adjusted for day in which the respondent was asked to record his/her activities. Individuals were divided into three categories based on their assigned reference day: Monday-Friday, Saturday, and Sunday. A factor for each category was calculated in order to account for differential nonresponse by day of the week. All days in 2012 were divided across each of the three categories. A factor was calculated for each category by dividing the product of days in the category and total interviews by the number of interviews in the category. Respondents' nonresponse-adjusted weight was multiplied by the appropriate factor. The result was the final weight.

5.1.3 Constructing SHARE Wave II Weights

As with ATUS, two sets of weights had to be constructed for SHARE Wave II – a traditional set and an alternative set. The publically available final, nonresponse-adjusted weights were not replicated. Moreover, they included proxy respondents that were excluded from these analyses. Traditional weights were reconstructed to account for replication and were limited to the cases used in this dissertation.

In Wave I, base weights were calculated by country based on the inverse probabilities of selection (Börsch-Supan & Jürges 2005). The base weights were publically available and were not altered for this analysis. However, replicates were not available on the public use file. The Wave I base weights were replicated in a similar fashion as described in Section 2.2.

Once the base weights were replicated, Deville and Sarndal's (1992) calibration techinique, case #6, was applied to the overall base weight and each of the 72 replicates. Calibration seeks to make the smallest adjustments to the base weights while still achieving the population distribution of the variables used in the nonresponse adjustment. For both sets of weights, each country was calibrated to age (four categories) by sex control totals, creating eight control totals per country. Some countries were also calibrated to regional population counts. The preexisting weights found on the public use file used region for Austria, Denmark, Greece, Italy, Spain, and The Netherlands. When weights were reconstructed for this analysis, region was only used to calibrate Austria, Greece, Italy, Spain, and The Netherlands. Denmark's weights did not converge for some replicates given the number of regions (n = 15) and the small sample sizes per replicate. In order to create the alternative weights, social integration was added as an additional calibration variable for all countries. No other differences existed between the methods used to reconstruct the traditional weights and those used to construct the alternative weights.

Under the basic calibration technique, weights may be negative and vary widely. Deville and Sarndal's (1992) case #6 sets upper and lower boundaries for the weights to control variance and ensure non-negative weights. In order to determine appropriate bounds, 23 upper bounds and 21 lower bounds were tested by country and by replicate. The bounds chosen were the ones that minimized the standard deviation of the given weight. Once the bounds were determined, the calibration equation was run by country and replicate, and final, nonresponse-adjusted weights resulted.

5.1.4 Evaluating the Effectiveness of the Weights

Before testing the hypotheses directly, descriptive statistics of various social activities and roles were calculated using the base weights, the traditionally-constructed weights, and the alternative final weights. The variance, absolute bias, root mean squared error (RMSE), and design effect were calculated for each dichotomous social activity and role variable that was not used in the construction of the reduced social integration indicator or otherwise used for weighting. Bias was calculated across replicates in the same manner as defined in Section 4.1.1. The remaining quality indicators were calculated using the standard equations:

$$s^2 = \frac{p_r(1-p_r)}{n-1}$$

$$RMSE = \sqrt{s^2 + (p_r - p_f)^2}$$

$$d_{eff} = \frac{s^2}{s_{SRS}^2}$$

where s_{SRS}^2 is the variance assuming a simple random sample.

While the above statistics provided a bigger picture of the effect of the weights on point estimates, the remainder of the analysis focused on a statistical evaluation of the bias component. In order to test hypotheses 3a, the relative bias of the base-weighted estimates was compared to the relative bias of the alternatively-weighted estimates using similar equations to those found in Section 4.1.2. As with previous analyses, the standard deviations were calculated on the replicated weights. In order to find support for the hypothesis that the alternatively-weighted estimates should suffer from less bias than the base-weighted estimates, the *t*-statistics had to be significant and the bias associated with the alternative weights had to be less than the bias associated with the base weights.

This analysis was repeated to compare the traditionally-weighted estimates to the alternatively-weighted estimates to test whether the addition of a social integration indicator in the weighting approach significantly reduced bias compared to traditional methods (H3b). To find support for this hypothesis, the bias identified under the alternative weights had to be significantly less than that identified under the traditional weights.

Support for hypotheses 3a and 3b did not ensure that bias was eliminated, only that it was lessened. In order to determine whether the addition of a social integration variable to the weighting approach eliminated bias, the alternatively-weighted estimates were compared to the sample estimates.[20] Both the relative and absolute difference was calculated using the same equations from Section 4.2.1. *T*-tests were run on the replicated differences. Significant differences suggested that the alternative weights did not correct for all of the bias. However, tests that were not significant were inconclusive. The traditional weights could be sufficient to eliminate the bias. Therefore, the traditionally-weighted estimates were also compared to the sample estimates. If the alternative weights corrected for the bias, but the traditional weights did not, then it could be concluded that the correction was the result of the inclusion of the social integration variable.

As with previous analyses, findings were adjusted using FDR.

[20] The sample estimates were base-weighted.

5.2 Results

5.2.1 Constructing the ATUS Social Integration Indicator and Weights

Five different subsets of ATUS variables were tested to identify a reduced LCA that would closely replicate the modal assignment of the full model created in Chapter 3 and would retain sufficient model fit. A two-class model was necessary in order to maintain consistency with the full model. In four of the five subsets, the three-class model did not significantly improve model fit (Table 22). A three-class model was superior only in the LCA that used variables with standardized factors larger than 0.75 (VLMR $p = 0.029$). However, the BIC was nearly unchanged and the entropy decreased in the three-class model, suggesting a two-class model may be sufficient.

Table 22: Model Fit of Various Reduced LCA Models (ATUS)

	Standardized Loadings				Absolute Difference		Relative Difference		Civic Engagement	
	≥0.75		≥0.65		≥0.25		≥0.85		Single-Loading Variables	
	2-Class	3-Class	2-Class	3-Class	2-Class	3-Class	2-Class	3-Class	2-Class	3-Class
Entropy	0.771	0.722	0.745	0.698	0.646	0.644	0.702	0.741	0.650	0.657
BIC	44,448	43,930	64,609	64,048	25,154	25,167	16,794	16,835	47,055	46,766
VLMR	0.029		0.104		0.197		0.739		0.098	
False Positive	0.372		0.379		0.021		0.026		0.078	
False Negative	0.186		0.129		0.550		0.536		0.281	
Accuracy	0.693		0.709		0.792		0.794		0.850	

Limiting the focus to the various two-class models, one can see that the LCAs that used variables with high standardized loadings had the best model fit (entropy = 0.745 or higher). While not as high, all other two-class models had sufficient model fit with entropy scores of at least 0.646.

Next, cases' modal assignments under the various two class reduced models were compared to their modal assignments under the full model. The models that included variables with high standardized loadings had the highest false positive rates and the lowest accuracy rates while the absolute and relative difference models suffered from false positive rates over 50 percent. The civic engagement model most closely resembled the full model with 85.0 percent of individuals coded into the same modal class as the full model. While the false positive rate was slightly higher than the absolute and relative magnitude models, the difference was small.

Ultimately, the model that used six civic engagement variables was selected to form the reduced model used in weighting (Figure 15). It identified two classes, and had sufficient model fit. In order to maximize the potential that it would be effective at reducing bias, a series of chi-square tests were performed (Table 23). The distribution of each of the 11 social activity and role variables was significantly different among the integrated class compared to the isolated class. Given this finding, the model was deemed sufficient and finalized.

Figure 15: Latent Class Model of Social Integration for Weighting (ATUS)

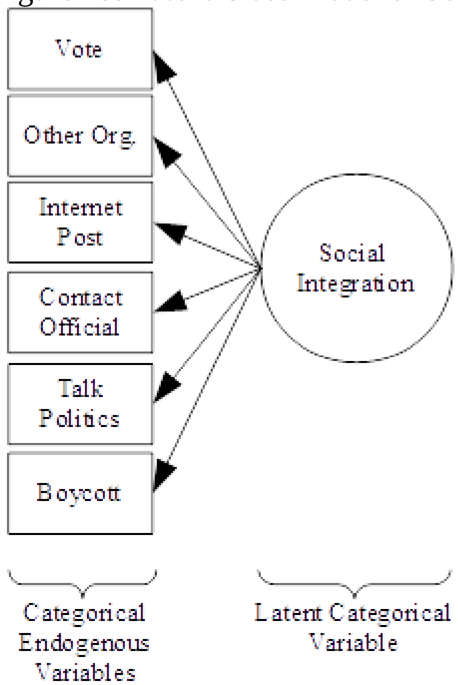


Table 23: Differences in Social Activity and Role Estimate by Social Integration Category (ATUS)

		N	Social Integration		X^2
			Integrated	Isolated	(*p*-value)
Dinner w/ Family	Monthly or Less	2,732	62.1%	55.7%	**10.0** (0.019)
	Few Times/Month		12.1%	13.0%	
	Few Times/Week		1.8%	3.6%	
	Almost Daily		24.0%	27.7%	
Employee	No	2778	40.6%	45.5%	**7.0** (0.008)
	Yes		59.4%	54.5%	
Friend/ Family	Monthly or Less	2,707	3.8%	9.4%	**34.5** (<0.0001)
	Few Times/Month		11.1%	12.2%	
	Few Times/Week		37.3%	35.2%	
	Almost Daily		47.8%	43.2%	
Neighbor	Never	2,705	3.7%	11.9%	**63.0** (<0.0001)
	Less than Monthly		7.2%	9.4%	
	Monthly		8.7%	9.3%	
	Few Times/Month		25.8%	24.5%	
	Few Times/Week		40.1%	32.1%	
	Almost Daily		14.5%	12.8%	
Neighbor Favors	Never	2,689	3.5%	2.8%	**133.7** (<0.0001)
	Less than Monthly		13.1%	8.9%	
	Monthly		24.4%	18.4%	
	Few Times/Month		19.0%	11.6%	
	Few Times/Week		23.4%	21.5%	
	Almost Daily		16.6%	36.8%	
Community Group†	No	2,723	70.3%	88.3%	**138.8** (<0.0001)
	Yes		29.7%	11.7%	
Spouse	No	2,779	41.7%	50.2%	**17.3** (<0.0001)
	Yes		58.3%	49.8%	
Religious Org.†	No	2,719	64.1%	79.7%	**100.3** (<0.0001)
	Yes		35.9%	20.9%	
Sports Group	No	2,723	78.1%	91.6%	**95.4** (<0.0001)
	Yes		21.9%	8.4%	
Community Officer†	No	2,723	71.4%	91.8%	**209.1** (<0.0001)
	Yes		28.6%	8.2%	
Civic Org.†	No	2,722	81.5%	93.5%	**99.6** (<0.0001)
	Yes		18.5%	6.5%	

† Identifies civic and political activities and roles.

The final model was run on the full CE Supplement. Among the 42,073 cases used to form the population totals, 30.5 percent were assigned to the integrated category. The final reduced model was also run on the ATUS respondents, resulting in 33.9 percent being assigned to the integrated class.

Traditional and alternative weights were constructed for the 2,779 respondents used in this analysis. While the weight assigned to a given respondent varied widely, the distribution of the weights was similar for both schema (Table 24). Both sets of weights were right-skewed with a long upper tail. The alternative weights also had a range approximately five percent larger than the traditional weights.

Table 24: Descriptive Statistics of the Distribution of the Weights by Weighting Scheme (ATUS)

	Traditional Weights	Alternative Weights	Absolute Difference
Minimum	1,120,565	1,056,688	205
Maximum	129,769,151	136,960,815	10,420,649
Median	9,991,135	9,840,171	410,876
Mean	13,498,221	13,498,221	640,643
Standard Deviation	12,898,493	12,953,452	750,665

5.2.2 Constructing the SHARE Wave II Social Integration Indicator and Weights

As with ATUS, before the weights could be constructed, a reduced LCA had to be created. A total of four models with various subsets of variables were tested to determine which subset resulted in a model most similar to the full LCA in Chapter 3. For each subset, the number of classes was determined. The two-class model was sufficient for the large absolute differences subset and the civic engagement subset (Table 25). In both cases, the VLMR *p*-value was not significant and the BIC was nearly unchanged between the two and three-class models. A two-class model also resulted from the LCA that used variables with standardized factor loadings of 0.65 or higher. However, this was because a higher class model did not converge. The LCA using variables with standardized loadings of 0.55 or higher did not converge under any number of classes. While an investigation could have been conducted to determine the source of nonconvergence, it was unnecessary. The models that did converge were of sufficient fit (entropy ≥ 0.663).

Table 25: Model Fit of Various Reduced LCA Models (SHARE)

	Standardized Loadings				Absolute Difference		Civic Engagement	
	≥0.65		≥0.55		≥0.25		Single-Loading Variables	
	2-Class	3-Class	2-Class	3-Class	2-Class	3-Class	2-Class	3-Class
Entropy	0.868	Did not converge	Did not converge	Did not converge	0.663	0.546	0.715	0.750
BIC	87,859				88,285	88,201	44,899	44,889
VLMR	N/A				N/A	0.394	N/A	0.577
False Positive	0.798				0.359		0.914	
False Negative	0.213				0.001		0.133	
Accuracy	0.647				0.913		0.680	

Each of the three two-class reduced LCAs were compared to the full LCA from Chapter 3. The absolute difference LCA was far superior to the other two models. The modal class assignment was the same as the full model 91.3 percent of the time, and the false negative rate was nearly zero (0.001). While the false positive rate was higher than desired (0.359), it was still lower than either of the other two models.

The absolute difference LCA was chosen to create an indicator of social integration to be used in weighting (Figure 16). Before finalizing the model choice, a series of chi-square tests were run to compare the distribution of social activity and role participation between integration classes (Table 26). Only four of the nine comparisons were significant at the 0.05 level, with a fifth (sick adult) marginally significant (p = 0.086). Given the lack of relationship between the social integration measure and the remaining four variables of interest, it was suspected that the chosen model would not produce a social integration variable capable of eliminating bias. Chi-square tests were conducted using the social integration measures created by the other two two-class models. Unfortunately, the relationships were not improved. Ultimately, the absolute difference LCA was chosen, recognizing that it would likely be insufficient to eliminate bias on all nine variables.

Figure 16: Latent Class Model of Social Integration for Weighting (SHARE)

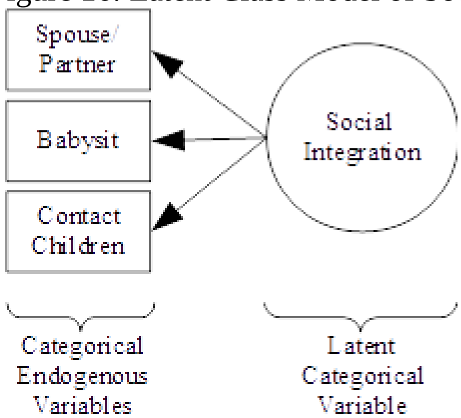


Table 26: Differences in Social Activity and Role Estimate by Social Integration Category (SHARE)

		N	Social Integration		X^2
			Integrated	Isolated	(*p*-value)
Volunteer†	Never	12,863	88.3%	88.8%	1.0
	At Least Once		11.7%	11.2%	(0.319)
Sick Adult†	Never	12,864	94.6%	93.9%	3.1
	At Least Once		5.4%	6.1%	(0.086)
Community Group†	Never	12,863	96.1%	96.9%	0.5
	At Least Once		3.9%	3.1%	(0.488)
Help HHM	Never	12,873	94.7%	95.1%	0.0
	At Least Once		5.3%	4.9%	(0.995)
Help Family	Never	12,864	5.6%	4.3%	**59.1**
	Less than Monthly		7.1%	4.4%	(<0.0001)
	Almost Every Month		4.6%	2.7%	
	Almost Every Week		5.6%	4.0%	
	Almost Daily		77.1%	84.6%	
Contact Parent	Never	12,681	8.0%	7.0%	**29.2**
	Every 2 Weeks or Less		7.8%	4.2%	(<0.0001)
	About Once per Week		6.1%	4.0%	
	Several Times per Week		4.9%	3.8%	
	Daily		73.2%	81.0%	
Training†	Never	12,863	4.7%	5.0%	**9.3**
	At Least Once		95.3%	95.0%	(0.002)
Help Others†	Never	12,864	6.3%	9.3%	**42.9**
	Less than Monthly		4.8%	6.1%	(<0.0001)
	Almost Every Month or More		88.9%	84.6%	
Religious Org.†	Never	12,864	6.3%	6.6%	1.1
	Less than Weekly		3.7%	4.3%	(0.574)
	Almost Every Week or More		90.0%	89.1%	

† Identifies civic and political activities and roles.

The final reduced model was run on the Wave II sample (i.e., the Wave I respondents) to create population control totals. On average, 81.4 percent of sample was assigned to the integrated class although there was some variation by country. The final reduced model was also run on the SHARE Wave II respondents, resulting in 85.5 percent on average being assigned to the integrated class. As with the population totals, this varied somewhat by country.

Traditional and alternative weights were constructed for the 12,904 SHARE Wave II respondents. The traditional and alternative weights were similar in several regards (Table 27). Both skewed right with a long upper tail. They also had similar descriptive statistics such as their minimum and maximum values. However, an individual's value could vary widely between the two weights with a median change of 117 and an average change of 400. In other words, an average individual's weight changed by 5.5 percent (400/7,314) between the two weighting schema.

Table 27: Descriptive Statistics of the Distribution of the Weights by Weighting Scheme (SHARE)

	Traditional Weights	Alternative Weights	Absolute Difference
Minimum	552	559	0
Maximum	55,618	55,036	11,770
Median	4,144	4,101	117
Mean	7,314	7,314	400
Standard Deviation	7,371	7,553	970

5.2.3 Evaluating the Effectiveness of the Weights

While the primary purpose of this chapter is to test the effectiveness of alternative weighting techniques to reduce nonresponse bias, the standard errors, RMSE, and design effects were reviewed under different weighting approaches to assess their effects on the variance of social activities and roles estimates. Tables 28 and 29 display the results for ATUS and SHARE Wave II, respectively. Four estimates are provided for each variable. They include the estimate of the full sample (i.e., the "true" value), the base-weighted estimate among respondents, the traditionally-weighted estimate among respondents, and the alternatively-weighted estimate among respondents. The remaining statistics were calculated using the latter three weighting approaches. For both ATUS and SHARE, the standard errors and design effects were larger under the nonresponse-adjusted weighting schemes than the base-weighted estimates. This was expected since nonresponse adjustments typically increase the variance of the weights and, as a result, the variance of the estimate. More importantly, the addition of a social integration indicator into the weighting methodology did not change either the standard errors or design effects for most variables.

Table 28: Quality Indicators of Various Social Activities & Roles under Different Weighting Schemes (ATUS)

	N		Estimate				Standard Error		
	Full Sample	Resp.	Full Sample	Base Weights	Traditional Weights	Alternative Weights	Base Weights	Traditional Weights	Alternative Weights
Dinner w/ Family	5,009	2,732	76.3%	74.8%	75.7%	75.7%	0.008	0.009	0.008
Employee	5,148	2,778	56.5%	56.3%	60.5%	60.3%	0.010	0.012	0.012
Family/Friend	4,925	2,707	98.0%	98.4%	98.3%	98.2%	0.002	0.003	0.003
Neighbor	4,925	2,705	89.2%	91.1%	89.6%	89.5%	0.007	0.009	0.010
Neighbor Favors	4,895	2,689	67.2%	70.6%	67.1%	66.6%	0.009	0.012	0.012
Community Group†	4,989	2,723	17.4%	18.2%	17.5%	17.1%	0.008	0.008	0.008
Spouse	5,150	2,779	49.5%	52.9%	47.6%	47.5%	0.013	0.014	0.014
Religious Org.†	4,978	2,719	22.6%	25.9%	24.2%	23.8%	0.008	0.008	0.008
Sports Groups	4,989	2,723	11.2%	13.3%	13.5%	13.1%	0.008	0.008	0.008
Community Officer†	4,983	2,723	13.0%	15.6%	15.1%	14.6%	0.006	0.007	0.007
Civic Org.†	4,987	2,722	8.8%	10.8%	10.1%	9.9%	0.008	0.010	0.009

Table 28: Quality Indicators of Various Social Activities & Roles under Different Weighting Schemes (ATUS) (cont'd)

	Absolute Difference			RMSE			Design Effect		
	Base Weights	Traditional Weights	Alternative Weights	Base Weights	Traditional Weights	Alternative Weights	Base Weights	Traditional Weights	Alternative Weights
Dinner w/ Family	-1.54	-0.61	-0.65	0.017	0.011	0.011	0.92	1.08	1.05
Employee	-0.26	3.93	3.80	0.010	0.041	0.040	1.16	1.68	1.64
Family/Friend	0.39	0.26	0.23	0.005	0.004	0.004	1.04	1.27	1.34
Neighbor	1.90	0.38	0.24	0.020	0.010	0.010	1.81	2.57	2.63
Neighbor Favors	3.40	-0.07	-0.50	0.035	0.012	0.013	1.02	1.69	1.73
Community Group†	0.88	0.12	-0.22	0.012	0.008	0.008	1.17	1.20	1.21
Spouse	3.38	-1.88	-1.94	0.036	0.024	0.024	1.75	2.29	2.30
Religious Org.†	3.39	1.61	1.23	0.035	0.018	0.014	0.86	0.89	0.88
Sports Groups	2.13	2.26	1.91	0.023	0.024	0.021	1.42	1.50	1.49
Community Officer†	2.56	2.00	1.56	0.026	0.021	0.017	0.77	1.13	1.08
Civic Org.†	2.07	1.34	1.08	0.022	0.017	0.014	1.86	2.79	2.76

† Identifies civic and political activities and roles.

Table 29: Quality Indicators of Various Social Activities & Roles under Different Weighting Schemes (SHARE)

	N		Estimate				Standard Error		
	Full Sample	Resp.	Full Sample	Base Weights	Traditional Weights	Alternative Weights	Base Weights	Traditional Weights	Alternative Weights
Volunteer†	19,089	12,862	10.1%	11.6%	11.9%	11.9%	0.005	0.006	0.006
Sick Adult†	19,090	12,863	5.2%	5.5%	5.8%	5.8%	0.003	0.003	0.003
Community Group†	19,088	12,862	3.2%	3.8%	4.0%	4.0%	0.003	0.003	0.003
Help HHM	19,138	12,872	5.2%	5.2%	5.1%	5.1%	0.002	0.003	0.003
Help Family	19,096	12,863	20.4%	21.7%	23.0%	22.8%	0.008	0.008	0.008
Contact Parent	18,899	12,680	25.2%	25.6%	28.9%	28.8%	0.005	0.006	0.005
Training†	19,089	12,862	4.3%	4.7%	5.3%	5.3%	0.004	0.005	0.005
Help Others	19,096	12,863	10.9%	11.7%	12.3%	12.4%	0.006	0.007	0.007
Religious Org. †	19,090	12,863	9.5%	10.1%	9.9%	10.0%	0.005	0.005	0.005

Table 29: Quality Indicators of Various Social Activities & Roles under Different Weighting Schemes (SHARE) (cont'd)

	Absolute Difference			RMSE			Design Effect		
	Base Weights	Traditional Weights	Alternative Weights	Base Weights	Traditional Weights	Alternative Weights	Base Weights	Traditional Weights	Alternative Weights
Volunteer†	1.51	1.81	1.78	0.526	0.569	0.563	3.47	3.97	3.90
Sick Adult†	0.36	0.61	0.61	0.271	0.288	0.290	1.81	1.96	1.99
Community Group†	0.54	0.79	0.77	0.268	0.291	0.289	2.54	2.82	2.79
Help HHM	0.04	-0.03	-0.03	0.241	0.263	0.253	1.52	1.83	1.70
Help Family	1.35	2.59	2.39	0.750	0.829	0.836	4.25	4.99	5.11
Contact Parent	0.43	3.79	3.65	0.479	0.551	0.544	1.53	1.86	1.82
Training†	0.41	0.99	0.95	0.433	0.472	0.476	5.34	5.67	5.81
Help Others	0.87	1.45	1.56	0.594	0.684	0.669	4.38	5.56	5.29
Religious Org. †	0.61	0.44	0.46	0.468	0.515	0.493	3.10	3.81	3.49

† Identifies civic and political activities and roles.

While the alternative weights had similar effects on the design effect and standard errors between surveys, the effect on the RMSE differed. In ATUS, the RMSE was generally lower for both types of nonresponse-adjusted weights compared to the base-weighted estimates. A higher RMSE was only observed on two variables, one of which had a higher value only under the traditional weight. Since the RMSE is a function of both variance and bias, this finding provided initial evidence that both nonresponse adjustments reduced bias more than they increased the variance. Unfortunately, the RMSE increased for all nonresponse-adjusted estimates compared to the base-weighted estimates in SHARE Wave II. This could have been because the nonresponse adjustments failed to reduce the bias more than they increased the variance. Unfortunately, looking at Table 27, this was not the case. The absolute difference, i.e., bias, increased when either nonresponse adjustment was made.

To investigate the effect of the alternative weights on bias in more detail, the bias of the base-weighted estimates was statistically compared to the bias under the alternative weights. Tables 30 and 31 display the results. The case counts and estimates are the same as those displayed in Tables 28 and 29 for ATUS and SHARE Wave II, respectively. The relative difference is the difference between the weighted estimate and the full sample divided by the full sample estimate. The test statistic was calculated on the absolute difference of the relative difference.

The alternative weights significantly reduced bias for nine of the 11 ATUS variables. While the reduction was small in some cases, most variables were considerably improved. For example, the relative bias of being a community officer decreased from 19.65 to 11.95 percent, shifting the estimate by 1.0 percentage points. Between the two estimates that were not significantly improved, the change in the proportion of individuals who belong to a sports group trended in the right direction but was not significant. Whether or not the respondent was employed was the only statistic in which bias increased under the alternative weights.

Oppositely, the SHARE Wave II alternative weights increased the bias in seven of the nine social activity and roles variables. Only estimates on helping household members and attending religious organization events were improved. Where improvements were observed, the magnitude was small, only 0.1 percentage point for each variable. However, where the alternative weights increased the bias, the magnitude was often large. For example, the estimate of the proportion of respondents who communicate with their parents increased from 25.6 percent under the base weights to 28.8 under the alternative weights, increasing the relative bias from 1.7 to 14.5 percent.

Under ATUS, there was support for hypothesis 3a that alternatively-weighted estimates are less biased, but the SHARE results suggested the opposite conclusion. However, this analysis does not isolate whether the changes (for better or worse) in the estimates were the result of the inclusion of the social integration indicator or the result of the traditional nonresponse adjustment.

Table 30: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Base Weights (ATUS)

Variable	N		Estimate			Relative Difference		Difference of the Relative Difference	
	Full Sample	Respondents	Full Sample	Base Weights	Alternative Weights	Base Weights	Alternative Weights	Value	p-value
Dinner w/ Family	5,009	2,732	76.3%	74.8%	75.7%	-2.01%	-0.85%	**1.16**	0.007
Employee	5,148	2,778	56.5%	56.3%	60.3%	-0.46%	6.73%	**7.19**	<0.0001
Family/Friend	4,925	2,707	98.0%	98.4%	98.2%	0.40%	0.23%	**-0.17**	0.034
Neighbor	4,925	2,705	89.2%	91.1%	89.5%	2.12%	0.26%	**-1.86**	<0.0001
Neighbor Favors	4,895	2,689	67.2%	70.6%	66.6%	5.07%	-0.75%	**-5.81**	<0.0001
Community Group†	4,989	2,723	17.4%	18.2%	17.1%	5.07%	-1.30%	**-6.37**	<0.0001
Spouse	5,150	2,779	49.5%	52.9%	47.5%	6.83%	-3.93%	**-10.76**	<0.0001
Religious Org.†	4,978	2,719	22.6%	25.9%	23.8%	15.03%	5.46%	**-9.57**	<0.0001
Sports Groups	4,989	2,723	11.2%	13.3%	13.1%	18.99%	17.10%	-1.89	0.199
Community Officer†	4,983	2,723	13.0%	15.6%	14.6%	19.65%	11.95%	**-7.70**	0.0003
Civic Org.†	4,987	2,722	8.8%	10.8%	9.9%	23.55%	12.34%	**-11.21**	0.0001

† Identifies civic and political activities and roles.

Table 31: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Base Weights (SHARE)

Variable	N		Estimate			Relative Difference		Difference of the Relative Difference	
	Full Sample	Respondents	Full Sample	Base Weights	Alternative Weights	Base Weights	Alternative Weights	Value	p-value
Volunteer†	19,089	12,862	10.1%	11.6%	11.9%	14.93%	17.66%	**2.73**	<0.0001
Sick Adult†	19,090	12,863	5.2%	5.5%	5.8%	7.06%	11.78%	**4.72**	<0.0001
Community Group†	19,088	12,862	3.2%	3.8%	4.0%	16.80%	23.87%	**7.07**	<0.0001
Help HHM	19,138	12,872	5.2%	5.2%	5.1%	0.76%	-0.59%	**-1.35**	<0.0001
Help Family	19,096	12,863	20.4%	21.7%	22.8%	6.60%	11.71%	**5.11**	<0.0001
Contact Parent	18,899	12,680	25.2%	25.6%	28.8%	1.69%	14.50%	**12.81**	<0.0001
Training†	19,089	12,862	4.3%	4.7%	5.3%	9.55%	22.02%	**12.47**	<0.0001
Help Others	19,096	12,863	10.9%	11.7%	12.4%	7.98%	14.36%	**6.38**	<0.0001
Religious Org.†	19,090	12,863	9.5%	10.1%	10.0%	6.37%	4.88%	**-1.49**	<0.0001

† Identifies civic and political activities and roles.

To determine whether the social integration variable had a significant effect on the estimates and to test hypothesis 3b, the alternatively-weighted estimates were compared to the traditionally-weighted estimates. Tables 32 and 32 are laid out similarly to Tables 30 and 31, respectively. In both surveys, the alternative weights generally reduced the amount of bias compared to the traditional weights. In ATUS, nine of the 11 social activity and role estimates were significantly different between the two weighting schemes. However, only seven of the nine variables were improved. The bias was exacerbated by the alternative weights for estimating favors for the neighbors and belonging to a community group. For the seven variables for which bias was significantly reduced, the reductions were small.

The SHARE data told a similar story. The alternative weights significantly reduced the bias compared to the traditional weights in five of the nine estimates. A sixth variable was unchanged, and the remaining three estimates were significantly more biased under the alternative weighting scheme. While the differences were significant, they were small. Across the two surveys, there was some evidence to support hypothesis 3b, but the alternative weights neither consistently nor considerably improve the estimates.

Table 32: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Traditional Weights (ATUS)

Variable	N		Estimate			Relative Difference		Difference of the Relative Difference	
	Full Sample	Respondents	Full Sample	Traditional Weights	Alternative Weights	Traditional Weights	Alternative Weights	Value	p-value
Dinner w/ Family	5,009	2,732	76.3%	75.7%	75.7%	-0.80%	-0.85%	-0.05	0.125
Employee	5,148	2,778	56.5%	60.5%	60.3%	6.95%	6.73%	**-0.22**	0.001
Family/Friend	4,925	2,707	98.0%	98.3%	98.2%	0.26%	0.23%	**-0.03**	0.001
Neighbor	4,925	2,705	89.2%	89.6%	89.5%	0.42%	0.26%	**-0.16**	0.0003
Neighbor Favors	4,895	2,689	67.2%	67.1%	66.6%	-0.10%	-0.75%	**-0.65**	0.0001
Community Group†	4,989	2,723	17.4%	17.5%	17.1%	0.72%	-1.30%	**-2.01**	<0.0001
Spouse	5,150	2,779	49.5%	47.6%	47.5%	-3.79%	-3.93%	**-0.13**	0.037
Religious Org.†	4,978	2,719	22.6%	24.2%	23.8%	7.13%	5.46%	**-1.68**	0.0001
Sports Groups	4,989	2,723	11.2%	13.5%	13.1%	20.14%	17.10%	**-3.05**	0.0001
Community Officer†	4,983	2,723	13.0%	15.1%	14.6%	15.35%	11.95%	**-3.40**	0.0001
Civic Org.†	4,987	2,722	8.8%	10.1%	9.9%	15.26%	12.34%	**-2.92**	0.0001

† Identifies civic and political activities and roles.

Table 33: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Traditional Weights (SHARE)

Variable	N		Estimate			Relative Difference		Difference of the Relative Difference	
	Full Sample	Respondents	Full Sample	Traditional Weights	Alternative Weights	Traditional Weights	Alternative Weights	Value	p-value
Volunteer†	19,089	12,862	10.1%	11.9%	11.9%	17.98%	17.66%	**-0.32**	<0.0001
Sick Adult†	19,090	12,863	5.2%	5.8%	5.8%	11.79%	11.78%	-0.01	0.427
Community Group†	19,088	12,862	3.2%	4.0%	4.0%	24.58%	23.87%	**-0.71**	<0.0001
Help HHM	19,138	12,872	5.2%	5.1%	5.1%	-0.50%	-0.59%	**-0.09**	0.020
Help Family	19,096	12,863	20.4%	23.0%	22.8%	12.68%	11.71%	**-0.97**	<0.0001
Contact Parent	18,899	12,680	25.2%	28.9%	28.8%	15.06%	14.50%	**-0.56**	<0.0001
Training†	19,089	12,862	4.3%	5.3%	5.3%	22.75%	22.02%	**-0.72**	<0.0001
Help Others	19,096	12,863	10.9%	12.3%	12.4%	13.36%	14.36%	**1.00**	<0.0001
Religious Org.†	19,090	12,863	9.5%	9.9%	10.0%	4.67%	4.88%	**0.21**	<0.0001

† Identifies civic and political activities and roles.

The analyses thus far were conducted to determine whether the alternative weights were an improvement over other weighting approaches. However, a final analysis was necessary to test hypothesis 3c and determine whether the alternative weights successfully eliminated bias. Tables 34 and 35 show the results of the comparisons for ATUS and SHARE Wave II, respectively. For both surveys, the alternative weights eliminated bias for only 25 percent of the variables (5 out of 20). All five of the corrected variables were in ATUS. The continuing presence of bias among the SHARE variables was not a surprise given the previous finding that the weighted estimates were more biased than the base weighted estimates.

This comparison did not isolate whether the correction was the result of the addition of the social integration measure or whether the traditional weights also eliminated bias among these five variables. To be certain, the traditional weights were compared to the full sample (Tables 36 and 37 for ATUS and SHARE, respectively). Beginning with ATUS, the traditional and alternative weights eliminated bias for a similar number of variables. The traditional weights corrected the nonresponse bias in four of the 11 variables. While the traditional weights did not correct for the bias associated with communicating with friends and family, the difference was only significant at the 0.05 level and substantively ignorable (p = 0.032, relative difference = 0.3 percent). Therefore, while there was limited support for hypothesis 3c as it was stated, the elimination of bias was not the result of the addition of the social integration indicator.

Table 34: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Full Sample Estimates (ATUS)

Variable	N		Estimate		Relative Difference		Absolute Difference	
	Full Sample	Respondents	Full Sample	Alternative Weights	Value	p-value	Value	p-value
Dinner w/ Family	5,009	2,732	76.3%	75.7%	-0.85%	0.063	-0.65	0.063
Employee	5,148	2,778	56.5%	60.3%	6.73%	<0.0001	**3.80**	<0.0001
Family/Friend	4,925	2,707	98.0%	98.2%	0.23%	0.056	0.23	0.056
Neighbor	4,925	2,705	89.2%	89.5%	0.26%	0.259	0.24	0.259
Neighbor Favors	4,895	2,689	67.2%	66.6%	-0.75%	0.160	-0.50	0.159
Community Group†	4,989	2,723	17.4%	17.1%	-1.30%	0.262	-0.22	0.262
Spouse	5,150	2,779	49.5%	47.5%	-3.93%	<0.0001	**-1.94**	<0.0001
Religious Org.†	4,978	2,719	22.6%	23.8%	5.46%	0.002	**1.23**	0.002
Sports Groups	4,989	2,723	11.2%	13.1%	17.10%	<0.0001	**1.91**	<0.0001
Community Officer†	4,983	2,723	13.0%	14.6%	11.95%	<0.0001	**1.56**	<0.0001
Civic Org.†	4,987	2,722	8.8%	9.9%	12.34%	0.001	**1.08**	0.0004

† Identifies civic and political activities and roles.

Table 35: Bias of Social Activity & Role Estimates Using Alternative Weights Vs. Full Sample Estimates (SHARE)

Variable	N		Estimate		Relative Difference		Absolute Difference	
	Full Sample	Respondents	Full Sample	Alternative Weights	Value	p-value	Value	p-value
Volunteer†	19,089	12,862	10.1%	11.9%	17.66%	<0.0001	**1.78**	<0.0001
Sick Adult†	19,090	12,863	5.2%	5.8%	11.78%	<0.0001	**0.61**	<0.0001
Community Group†	19,088	12,862	3.2%	4.0%	23.87%	<0.0001	**0.77**	<0.0001
Help HHM	19,138	12,872	5.2%	5.1%	-0.59%	0.009	**-0.03**	0.009
Help Family	19,096	12,863	20.4%	22.8%	11.71%	<0.0001	**2.39**	<0.0001
Contact Parent	18,899	12,680	25.2%	28.8%	14.50%	<0.0001	**3.65**	<0.0001
Training†	19,089	12,862	4.3%	5.3%	22.02%	<0.0001	**0.95**	<0.0001
Help Others	19,096	12,863	10.9%	12.4%	14.36%	<0.0001	**1.56**	<0.0001
Religious Org.†	19,090	12,863	9.5%	10.0%	4.88%	<0.0001	**0.46**	<0.0001

† Identifies civic and political activities and roles.

Table 36: Bias of Social Activity & Role Estimates Using Traditional Weights Vs. Full Sample Estimates (ATUS)

Variable	N		Estimate		Relative Difference		Absolute Difference	
	Full Sample	Respondents	Full Sample	Traditional Weights	Value	p-value	Value	p-value
Dinner w/ Family	5,009	2,732	76.3%	75.7%	-0.80%	0.073	-0.61	0.073
Employee	5,148	2,778	56.5%	60.5%	6.95%	<0.0001	**3.93**	<0.0001
Family/Friend	4,925	2,707	98.0%	98.3%	0.26%	0.032	**0.26**	0.032
Neighbor	4,925	2,705	89.2%	89.6%	0.42%	0.146	0.38	0.146
Neighbor Favors	4,895	2,689	67.2%	67.1%	-0.10%	0.445	-0.07	0.445
Community Group†	4,989	2,723	17.4%	17.5%	0.72%	0.368	0.12	0.368
Spouse	5,150	2,779	49.5%	47.6%	-3.79%	<0.0001	**-1.88**	<0.0001
Religious Org.†	4,978	2,719	22.6%	24.2%	7.13%	0.0002	**1.61**	0.0001
Sports Groups	4,989	2,723	11.2%	13.5%	20.14%	<0.0001	**2.26**	<0.0001
Community Officer†	4,983	2,723	13.0%	15.1%	15.35%	<0.0001	**2.00**	<0.0001
Civic Org.†	4,987	2,722	8.8%	10.1%	15.26%	<0.0001	**1.34**	<0.0001

† Identifies civic and political activities and roles.

Table 37: Bias of Social Activity & Role Estimates Using Traditional Weights Vs. Full Sample Estimates (SHARE)

Variable	N		Estimate		Relative Difference		Absolute Difference	
	Full Sample	Respondents	Full Sample	Traditional Weights	Value	p-value	Value	p-value
Volunteer†	19,089	12,862	10.1%	11.9%	17.98%	<0.0001	**1.81**	<0.0001
Sick Adult†	19,090	12,863	5.2%	5.8%	11.79%	<0.0001	**0.61**	<0.0001
Community Group†	19,088	12,862	3.2%	4.0%	24.58%	<0.0001	**0.79**	<0.0001
Help HHM	19,138	12,872	5.2%	5.1%	-0.50%	0.026	**-0.03**	0.026
Help Family	19,096	12,863	20.4%	23.0%	12.68%	<0.0001	**2.59**	<0.0001
Contact Parent	18,899	12,680	25.2%	28.9%	15.06%	<0.0001	**3.79**	<0.0001
Training†	19,089	12,862	4.3%	5.3%	22.75%	<0.0001	**0.99**	<0.0001
Help Others	19,096	12,863	10.9%	12.3%	13.36%	<0.0001	**1.45**	<0.0001
Religious Org. †	19,090	12,863	9.5%	9.9%	4.67%	<0.0001	**0.44**	<0.0001

† Identifies civic and political activities and roles.

The traditionally-constructed SHARE Wave II weights did an even poorer job of correcting the bias. All of the nine variables remained bias, although the effect was marginal and small for helping household members (p-value = 0.026, relative difference = 0.5 percent). Among the remaining eight variables, the absolute difference was small, but the relative bias was quite large with nearly all variables suffering from a 10-20 percent change in the estimate. These findings suggested that the traditional weights were not sufficient to correct for nonresponse bias found in social activities and roles variables, but the alternative weights were also insufficient.

None of the significance levels were changed after applying the FDR.

5.3 Discussion

The alternative weights had very different effects when applied to the two surveys. In ATUS, the alternative weights significantly reduced bias compared to the base weights, lending support for hypothesis 3a. In most cases, they also reduced bias compared to the traditional weights (H3b). However, the improvement over the traditionally-weighted estimates was small and substantively ignorable. To be sure, the alternative weights corrected the bias on 5 of the 11 ATUS variables (H3c), but the elimination of bias could not be attributed to the inclusion of the social integration indicator. The traditional weights also corrected the bias on the same variables.

The alternative weights increased the amount of bias observed compared to the base-weighted SHARE estimates resulting in the rejection of hypothesis 3a. Calibrating to age, sex, and (in some cases) region, as was done in both nonresponse adjustments, was identified to be the source of the increase in the bias. While the inclusion of a social integration measure markedly reduced the damage done by calibrating to these demographic totals (H3b), the correction was not large and not large enough to eliminate the bias (H3c).

Altogether, a nonresponse adjustment is necessary to correct for the bias identified in Chapter 4. Unfortunately, neither the traditional nor alternative weights constructed in this dissertation were sufficient to eliminate bias. A minimal reduction in bias was achieved by the alternative weights in comparison to the traditional weights, but the small improvement, combined with the lack of consistent findings, the increase in cost from adding questions to the survey, and the additional computational effort required to construct the alternative weights suggests that this method is not worthwhile.

The alternative weighting schemes used here may have failed to more completely correct for nonresponse bias for several reasons. First, the reduced LCA may not have produced a strong indicator of integration. Assuming the full LCA from Chapter 3 was the correct model (an overstatement), the reduced ATUS model deviated in modal class assignment 15.0 percent of the time while the SHARE model deviated 8.7 percent of the time. This measurement error introduced noise in the model and may have reduced the effectiveness of the social integration indicator.

Second, while social integration was found to be highly predictive of nonresponse in Chapter 3, it only explained approximately one percent of the variation in each survey. The relationship between integration and nonresponse may not be strong enough to be useful in weighting.

Third, in the case of ATUS, the demographic variables used in the traditional weights were correlated with the social integration measure. While social integration may be a better variable to use in weighting (i.e., more strongly related to both nonresponse and the variables of interest), the demographic variables may be sufficient. This chapter did not tackle whether or not the initial weighting variables were appropriate to use for weight construction. In order to isolate the effect of the adding an integration measure, the original weighting schemes and the variables used in them were assumed to be appropriate. A more thorough investigation may take a step back and reconsider the entire weighting approach, recognizing that using non-traditional variables (i.e., non-demographic variables such as a measure of social integration) may be a superior approach.

The variables used in weight construction may similarly need to be reevaluated for SHARE. The traditional weights were calibrated to age and sex and (in some cases) region control totals. The use of these variables increased the observed level of bias in nearly every variable, suggesting that they were poor variables to include in the nonresponse adjustment. A reevaluation of the variables used in weight construction may identify that a social integration measure would be useful in reducing bias either on its own or in combination with other variables.

Finally, in the case of SHARE, the social integration indicator was not correlated with five of the nine variables of interest. In order for weights to effectively reduce bias, the variables used in the weighting algorithm must be correlated both with response and with the variables of interest (Little & Vartavarian 2005). A strong integration indicator that was created using a small number of variables, had a fair model fit, replicated the full model modal assignments, and was strongly correlated to both response and the variables of interest was impossible given the social activity and roles measures available. As mentioned in Chapter 3, it may be worthwhile to reevaluate the notion that any diverse set of variables may be used to construct an integration indicator. If a better set of variables could be identified for the LCA, then the incorporation of the resulting variable may be more effective at reducing bias.

If future research can refine the social integration measure and weighting technique, it would be useful to investigate whether a revised approach cannot only reduce bias among social activity and roles variables but also among other types of variables. For example, voting has been demonstrated to be predictive of a variety of health outcomes (Subramanian, Huijts, & Perkins 2009; Shin & McCarthy 2013), and incorporation of voting into the weighting process has been demonstrated to reduce bias among health indicators (Peytchev 2015). More research is necessary to identify which variables to use and for what variables bias may be eliminated.

Chapter 6: Conclusions

Researchers seek to collect information on social activities and roles in a variety of surveys. These variables are used to predict health outcomes, election outcomes, and track cultural changes over time. However, analyses conducted on ATUS and SHARE have demonstrated consistent and frequently large levels of nonresponse bias on such measures. Individuals who were more socially integrated, i.e., participated in more social activities and took on more social roles, were more likely to respond. As a result, prevalence estimates were overestimated. This bias was most evident in univariate analyses, but it often persisted in multivariate models predicting participation in social activities and roles. In the multivariate models, the beta coefficients were frequently biased, although the direction of the bias was inconsistent and often small suggesting that the interpretation of the models may be unaffected in most instances.

Traditional weighting techniques that used demographic variables to correct nonresponse bias were ineffective at eliminating the bias associated with most univariate estimates. While an alternative weighting approach that incorporated a social integration indicator into the nonresponse adjustment slightly reduced bias compared to the traditional weights, the reduction was not large nor sufficient to eliminate bias.

Both the components of integration and the components of nonresponse were also examined. While integration was predictive of nonresponse in both surveys, the details were inconsistent. Only civically engaged individuals were significantly more likely to respond to ATUS, suggesting that individuals integrated through other routes are not more likely to respond than isolated individuals. While civic engagement was also a predictor of SHARE response, it was neither the only nor the largest predictor.

Individuals who socialized with younger family members were also more likely to respond than those that did not.

While this research furthered the understanding of nonresponse and its relationship with social integration, there is much more to be learned. Some additional research is feasible using the existing datasets. First, in Chapter 4, bias was evaluated for univariate and multivariate models. Using similar techniques used in that chapter, one may also investigate bias of bivariate relationships. Bivariate models were examined by Abraham and her colleagues (2009), but no statistical tests were conducted and analysis was limited to one survey.

Second, the evaluation of bias in this dissertation focused on social activities and roles. Much research has suggested a link between integration and other outcomes such as health (e.g., Hanson et al 1989; Uchino 2004; Peytchev 2015). Given the wealth of health data available on SHARE, additional analysis could be conducted to determine whether health indicators suffer from nonresponse bias, whether the bias is caused by higher levels of nonresponse among the socially isolated, and whether the alternative weights created in Chapter 5 can reduce or eliminate such bias.

Third, the alternative weights were only evaluated on their ability to correct for bias in univariate analyses. No analysis was conducted to determine whether the alternative weights could be applied to correct for the nonresponse bias of the beta coefficients identified in the Chapter 4 multivariate models. Based on the existing analysis in Chapter 5, the addition of a social integration variable into the weighting algorithm had little value add. However, if the weights were found to significantly improve the accuracy of the multivariate models, this conclusion may need to be reassessed.

Finally, one could attempt to build a different weighting algorithm. The alternative weights were constructed by copying the existing weighting approach and adding to it. This was appropriate to test the hypotheses, but it may not be the most effective way of integrating an integration indicator into the nonresponse adjustment. When choosing variables for weight construction, the researcher typically identifies all variables for which he/she has population totals and hypothesizes may be related to both response and the outcomes of interest. These variables may be included in a regression model or a classification tree. The significant variables would be selected for inclusion in weight construction. In parallel, the researcher must also choose a nonresponse method such as raking or propensity modeling. It may be possible to improve the effectiveness of the weights, if the variables used and the nonresponse adjustment approach were reconsidered in a new context, one in which the inclusion of a social integration measure is possible.

Separate from feasible analyses using ATUS and SHARE, more research is necessary to construct an improved measure of integration. In Chapter 3, it was hypothesized that the discrepancies in the findings between the two surveys were the results of differences in the social activities and roles used to construct a measure of integration. Not only were the activities and roles themselves different in many instances, but the question wording also differed when measuring the same activity or role (e.g., spouse). A further evaluation of the questions may lend some insight into the different findings between surveys, further isolate the route(s) to integration that are most predictive of nonresponse, and identify a standardized set of questions that may be used to measure integration either as an end in itself or as an improved indicator for use in nonresponse adjustment.

In order to accomplish these goals, two avenues of research may be undertaken. First, the measures collected must be tested. This research used the social activity and roles variables that were available. However, the National Research Council (2014) hypothesized nine subcategories of integration, suggesting there are many more variables that may be useful or necessary. These subcategories have yet to be tested together in a single survey. If they were, exploratory models may be used to identify which variables are significant and have large effects on the construction of a social integration measure. This may be done with two related, but separate goals in mind: the construction of an integration measure in its own rite and the construction of a component or combination of components of integration that are predictive of nonresponse. The creation of an overarching integration measure may be useful across industries while the latter goal would focus on the explanation of nonresponse and isolate the pieces of social interactions that affect the decision to participate in a survey. For both goals, it may be possible that several combinations of variables are sufficient. If this is the case, then it would be useful to identify a theme and set of best practices. For example, each model should include a civic engagement measure, a family measure, and a friend measure. Also related to the measures collected is the importance of response options. This dissertation used a combination of dichotomous and ordinal variables. Some research has suggested that the frequency of participation has not been important in producing strong integration indicators (Cumming & Henry 1961). The importance of frequency of participation has not been tested in the context of survey participation. As the number of interactions increases, the frequency in which social expectations are being reinforced also increases. However, the perceived importance of the interaction may be more important than the frequency and more weight may need to be placed on perceived integration questions. An investigation of dichotomous and ordinal versions of the same question may be useful to identify the importance of frequency. Similarly, models which include or exclude perception measures should be evaluated to identify whether they enhance, subsume, or otherwise alter the effects of other activity and role measures.

The second avenue of research should focus on application to survey practice. Several arguments were made in Chapter 5 to explain the failure of the alternative weights to eliminate nonresponse bias. Some of these surrounded the quality of the integration measure. A similar weighting technique as used in this dissertation may be improved once improved measures are identified.

Instead of using integration in the weighting scheme to correct for nonresponse bias, an alternative approach would be to use it to prevent bias. Targeted advance letters, incentives, and interviewing staffing techniques have all been used in an attempt to increase the response rates among underrepresented groups (e.g., de Leeuw, Callegaro, Hox, Korendijk, & Lensvelt-Mulders 2007; Chmura & Yancey 2011). It may be possible to use similar techniques to increase the probability of response among isolated individuals. Of course, to accomplish this, information on the level of integration of the individuals would need to be available on the frame. This is unlikely for most cross-sectional surveys, but may be more feasible on follow-up waves of longitudinal surveys or on surveys which use list frames.

Regardless of which next steps are taken, one thing is clear. Nonresponse and nonresponse bias are a problem and more research is necessary to understand, prevent, and correct it.

References

Abraham, Katharine G, Aaron Maitland, and Suzanne M Bianchi. 2006. "Nonresponse in the American Time Use Survey: Who is Missing from the Data and How Much Does It Matter?" *Public Opinion Quarterly* 676-703.

Abraham, Katharine, Sara Helms, and Stanley Presser. 2009. "How Social Processes Distort Measurement: The Impact of Survey Nonresponse on Estimates of Volunteer Work in the United States." *American Journal of Sociology* 1129-1165.

Adler, Paul S, and Seok-Woo Kwon. 2002. "Social Capital: Prospects for a New Concept." *The Academy of Management Review* 17-40.

Allport, Gordon W. 1954. *The Nature of Prejudice.* Reading: Addison-Wesley.

American Association of Public Opinion Research. 2009. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* http://www.aapor.org/Content/NavigationMenu/ResourcesforResearchers/StandardDefinitions/StandardDefinitions2009new.pdf.

Asparouhov, T, and B Muthén. 2012. *Using Mplus TECH11 and TECH14 to Test the Number of Latent Classes.* https://www.statmodel.com/examples/webnotes/webnote14.pdf.

Asparouhov, Tihomir, and Bengt Muthén. 2002. *Latent Variable Analysis with Categorical Outcomes: Multiple-group and Growth Modeling in MPlus.* Los Angeles: Muthén & Muthén.

Bakk, Zsuzsa, Fetene B Tekle, and Jeroen K Vermunt. 2013. "Estimating the Association between Latent Class Membership and External Variables Using Bias-adjusted Three-step Approaches." *Sociological Methodology* 272-311.

Barrera, M, I M Sandler, and T B Ramsay. n.d. *Inventory of Socially Supportive Behaviors (ISSB) (Long and Short Form).* Accessed November 29, 2014. http://www.midss.org/content/inventory-socially-supportive-behaviors-issb-long-and-short-form.

Bassuk, Shari S, Thomas A Glass, and Lisa F Berkman. 1999. "Social Disengagement and Incident Cognitive Decline in Community-dwelling Elderly Persons." *Annals of Internal Medicine* 165-73.

Belle, Deborah E. 1982. "The Impact of Poverty on Social Networks and Supports." *Marriage and Family Review* 89-103.

Benjamini, Yoav, and Daniel Yuktieli. 2001. "The Control of the False Discovery Rate in Multiple Testing under Dependency." *Annals of Statistics* 1165-88.

Benjamini, Yoav, and Yosef Hochberg. 2000. "On the Adaptive Control of the False Discovery Rate in Multiple Testing with Independent Statistics." *Journal of Educational Behavioral Statistics* 60-83.

Berkman, Lisa F, and S Leonard Syme. 1979. "Social Networks, Host Resistance, and Mortality: a Nine-year Follow-up Study of Alameda County Residents." *American Journal of Epidemiology* 186-204.

Billiet, Jaak, Michel Philippens, Rory Fitzgerald, and Ineke Stoop. 2005. "Estimation of Response Bias in the European Social Survey: Using Information from Reluctant Respondents in Round One." *Paper presented at the 60th AAPOR Conference.* Miami Beach. 1-28.

Börsch-Supan, A, and H Jürges. 2005. *The Survey of Health, Ageing, and Retirement in Europe - Methodology.* Mannheim: MEA.

Börsch-Supan, A, M Brandt, C Hunkler, T Kneip, J Korbmacher, F Malter, B Schaan, S Stuck, and S Zuber. 2013. "Data Resource Profile: the Survey of Health, Ageing, and Retirement in Europe (SHARE)." *International Journal of Epidemiology* 992-1001.

Brehm, John. 1993. *The Phantom Respondents: Opinion Surveys and Political Representation.* Ann Arbor: The University of Michigan Press.

Brissette, Ian, Sheldon Cohen, and Teresa E Seeman. 2000. "Measuring Social Integration and Social Networks." In *Social Support Measurement and Intervention*, by Sheldon Cohen, Lynn G Underwood and Benjamin H Gottlieb, 53-85. New York: Oxford University Press.

Browne, M W, and R Cudeck. 1993. "Alternative Ways of Assessing Model Fit." In *Testing Structural Equation Models*, by K Bollen and J Long, 136-62. Newbury Park: Sage.

Bureau of the Census. 2014. *American Time Use Survey User's Guide: Understanding ATUS 2003-2013.* http://www.bls.gov/tus/atususersguide.pdf.

Bureau of the Census. 2011. *Current Population Survey, November 2011 Civic Engagement Supplement File Technical Documentation CPS--11.* http://www.census.gov/prod/techdoc/cps/cpsnov11.pdf.

Bureau of the Census. 2006. *Design and Methodology: Current Population Survey Technical Paper 66.* http://www.census.gov/prod/2006pubs/tp-66.pdf.

Casas-Cordero Valencia, Carolina. 2010. *Neighborhood Characteristics and Participation in Household Surveys.* Doctoral Dissertation.

Chmura, Lukasz, and Tracie Yancey. 2011. "Enhancing Address-based Sampling with Contingent Incentives and Sample Indicators." *Presentation at the 66th Annual American Association for Public Opinion Research Conerence.* Phoenix.

Cohen, Jacob. 1988. "Set Correlation and Contingency Tables." *Applied Psychological Measurement* 425-34.

Cornwell, Benjamin, and Elizabeth Warburton. 2014. "Work Schedules and Community Ties." *Work and Occupations* 139-74.

Cornwell, Benjamin, L Philip Schumm, Edward O Laumann, and Jessica Graber. 2009. "Social Networks in the NSHAP Study: Rationale, Measurement, and Preliminary Findings." *Journal of Gerontology* i47-i55.

Couper, Mick P, Eleanor Singer, and Richard A Kulka. 1998. "Participation in the 1990 Decennial Census: Politics, Privacy, Pressures." *American Politics Quarterly* 59-80.

Cumming, Elaine, and William E Henry. 1961. *Growing Old: The Process of Disengagement.* New York: Basic Books Inc.

de Leeuw, Edith, Mario Callegaro, Joop Hox, Elly Korendijk, and Gerty Lensvelt-Mulders. 2007. "The Influence of Advance Letters on Response in Telephone Surveys: A Meta-analysis." *Public Opinion Quarterly* 413-43.

Dever, Jill A. 2008. "Sampling Weight Calibration with Estimated Control Totals." Doctoral Dissertation.

Deville, Jean-Claude, and Carl-Erik Sarndal. 1992. "Calibration Estimators in Survey Sampling." *Journal of the American Statistical Association* 376-82.

Dillman, Don A. 1978. *Mail and Telephone Surveys: The Total Design Method.* New York: Wiley & Sons.

Dixon, John, and Clyde Tucker. 2010. "Survey nonresponse." In *Handbook of Survey Research*, by Peter V Marsden and James D Wright, 593-630. Bingley: Emerald Group Publishing Limited.

Durrant, Gabriele B, and Fiona Steele. 2009. "Multilevel Modelling of Refusal and Non-contact in Household Surveys: Evidence from Six UK Government Surveys." *Journal of the Royal Statistical Society: Series A* 361-81.

Fricker, Scott S. 2007. "The Relationship between Response Propensity and Data Quality in the Current Population Survey and the American Time Use Survey." Doctoral Dissertation.

Gans, Herbert J. 1962. *The Urban Villagers: Group and Class in the Life of Italian-Americans.* New York: The Free Press.

Glass, Thomas A, Carlos Mendes de Leon, Richard A Marottoli, and Lisa F Berkman. 1999. "Population Based Study of Social and Productive Activities as Predictors of Survival Among Elderly Americans." *BMJ* 478-83.

Goode, William J. 1960. "A Theory of Role Strain." *American Sociological Review* 483-96.

Gouldner, Alvin W. 1960. "The Norm of Reciprocity: A Preliminary Statement." *American Sociological Review* 161-78.

Goyder, John. 1987. *The Silent Minority.* Boulder: Westview Press.

Goyder, John, Luc Boyer, and Guil Martinelli. 2006. "Integrating Exchange and Heuristic Theories of Survey Nonresponse." *Bulletin de Methologie Sociologique* 28-44.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 646-75.

Groves, Robert M, and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias." *Public Opinion Quarterly* 167-89.

Groves, Robert M, and Mick Couper. 1998. *Nonresponse in Household Interview Surveys.* New York: Wiley-Interscience.

Groves, Robert M, Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly* 299-308.

Groves, Robert M, Robert B Cialdini, and Mick P Couper. 1992. "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly* 475-95.

Hanson, Bertil S, Sven-Olof Isacsson, Lars Janzon, and Sven-Eric Lindell. 1989. "Social Network and Social Support Influence Mortality in Elderly Men." *American Journal of Epidemiology* 100-11.

House, James S, and Robert L Kahn. 1985. "Measures and Concepts of Social Support." In *Social Support and Health*, by Sheldon Cohen and S Leonard Syme, 83-108. New York: Academic Press, Inc.

Judkins, David R. 1990. "Fay's Method for Variance Estimation." *Journal of Official Statistics* 223-39.

Kennickell, Arthur. 2005. "Darkness Made Visible: Field Management and Nonresponse Bias in the 2004 SCF." *Proceedings of the Joint Statistical Meetings.* 3187-94.

Knack, Stephen. 1995. "Civic Norms, Social Sactions, and Voter Turnout." *Rationality and Society* 133-56.

Landecker, Werner S. 1951. "Types of Integration and Their Measurement." *American Journal of Sociology* 332-40.

Lee, Sik-Yum, Wai-Yin Poon, and P M Bentler. 1995. "A Two-stage Estimation of Structural Equation Models with Continuous and Polytomous Variables." *British Journal of Mathematical and Statistical Psychology* 339-58.

Levin-Waldman, Oren M. 2013. "Income, Civic Participation and Achieving Greater Democracy." *Journal of Socio-Economics* 83-92.

Little, R JA, and D B Rubin. 2002. *Statistical Analysis with Missing Data.* New York: John Wiley.

Little, Roderick J, and Sonya Vartivarian. 2005. "Does Weighting for Nonresponse Increase the Variance of Survey Means." *Survey Methodology* 161-8.

Marin, Alexandra. 2004. "Are Respondents More Likely to List Alters with Certain Characteristics? Implications for Name Generator Data." *Social Networks* 289-307.

Markland, David. 2007. "The Golden Rule is That There are No Golden Rules: A Commentary on Paul Parrett's Recommendations for Reporting Model Fit in Structural Equation Modelling." *Structural Equation Modeling* 851-8.

Marsden, Peter V. 2003. "Interviewer Effects in Measuring Network Size Using a Single Name Generator." *Social Networks* 1-16.

Marsden, Peter V, and Noah E Friedkin. 1994. "Network Studies of Social Influence." In *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*, by Stanley Wasserman and Joseph Galaskiewicz, 3-25. Thousand Oaks: Sage Publications.

McCabe, Brian J. 2013. "Are Homeowners Better Citizens? Homeownership and Community Participation in the United States." *Social Forces* 929-54.

Miles, Jeremy, and Mark Shevlin. 2007. "A Tie and a Place for Incremental Fit Indices." *Structural Equation Modeling* 869-74.

Muthén, B O. 2004. *Mplus Technical Appendices, Version 3.* Los Angeles: Muthén & Muthén.

Muthén, Linda K, and Bengt O Muthén. n.d. *Mplus Statistical Analysis with Latent Variables User's Guide, Version 7.* Los Angeles: Muthén & Muthén.

National Research Council. 2014. *Civic Engagement and Social Cohesion: Measuring Dimensions of Social Capital to Inform Policy.* Washington, DC: The National Academies Press.

O'Neil, Michael J. 1979. "Estimating the Nonresponse Bias Due to Refusals in Telephone Surveys." *Public Opinion Quarterly* 218-32.

Paek, Hye-Jin, So-Hyang Yoon, and Dhavan V Shah. 2005. "Local News, Social Integration, and Community Participation: Hierarchical Linear Modeling of Contextual and Cross-level Effects." *Journalism and Mass Communication Quarterly* 587-606.

Peytchev, Andy. 2015. "Weighting Adjustments Using Substantive Survey Variables." *Presented at the 70th Annual American Association for Public Opinion Research Conference.* Hollywood.

Preacher, Kristopher J, Guangjian Zhang, Cheongtag Kim, and Gerhard Mels. 2013. "Choosing the Optimal Number of Factors in Exploratory Factor Analysis: A Model Selection Perspective." *Multivariate Behavioral Research* 28-56.

Presser, Stanley, and Linda Stinson. 1998. "Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance." *American Sociological Review* 137-45.

Putnam, Robert D. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-first Century." *Scandinavian Political Studies* 137-74.

—. 1993. *Making Democracy Work.* Princeton: Princeton University Press.

Rao, J N K, and J Shao. 1999. "Modified Balance Repeated Replication for Complex Survey Data." *Biometrika* 403-15.

Rhemtulla, Mijke, Patricia Brosseau-Liard, and Victoria Savalei. 2012. "When Can Categorical Variables be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods Under Suboptimal Conditions." *Psychological Methods* 354-73.

Roberts, Keith A. 1978. "Toward a Generic Concept of Counter-Culture." *Sociological Focus* 111-26.

Schwartz, Shalom H. 1977. "Normative Influences on Altruism." In *Advances in Experimental Social Psychology*, by Leonard Berkowitz, 222-75. New York: Academic Press.

Shaw, Clifford, and Henry McKay. 1972. *Juvenile Delinquency and Urban Areas.* Chicago: The University of Chicago Press.

Shin, Michael E, and Willian J McCarthy. 2013. "The Association Between County Political Inclination and Obesity: Results from the 2012 Presidential Election in the United States." *Prev Med* 721-4.

Sieber, Sam D. 1974. "Toward a Theory of Role Accumulation." *American Sociological Review* 567-78.

Smith, Tom W. 1984. "Estimating Nonresponse Bias with Temporary Refusals." *Sociological Perspectives* 473-489.

Stoop, Ineke A L. 2004. "Surveying nonrespondents." *Field Methods* 23-54.

—. 2005. *The Hunt for the Last Respondent.* The Hague: Social and Cultural Planning Office.

Subramanian, S V, Tim Huijts, and Jessica M Perkins. 2009. "Association Between Political Ideology and Health in Europe." *European Journal of Public Health* 455-7.

Thoits, Peggy A. 1983. "Multiple Identities and Psychological Well-being: A Reformulation and Test of the Social Isolation Hypothesis." *American Sociological Review* 174-87.

Tupek, Alan R. 2004. "Revised Weighting Specifications for The American Time Use Survey (Doc.#ATUS-06-R)." October 5.

Uchino, Bert N. 2004. *Social Support and Physical Health: Understanding the Health Consequences of Relationships.* New Haven: Yale University Press.

UCLA: Institute for Digital Research and Education. n.d. *How Can I do CFA with Binary Variables?* Accessed November 4, 2014. http://www.ats.ucla.edu/stat/stata/faq/binary_cfa.htm.

Van Ingen, Erik, Ineke Stoop, and Koen Breedveld. 2009. "Nonresponse in the Dutch Time Use Survey: Strategies for Response Enhancement and Bias Reduction." *Field Methods* 69-90.

Vermunt, Jeroen K. 2010. "Latent Class Modeling with Covariates: Two Improved Three-step approaches." *Political Analysis* 450-69.

Wemlinger, Elizabeth, and Martha Kropf. 2013. "Not a Suburban Soccer Mom: Political Engagement among Lower SES Women." *Poverty and Public Policy* 48-66.

Wolfensberger, Wolf, and Stephen Tullman. 1982. "A Brief Outline of the Principle of Normalization." *Rehabilitation Psychology* 131-45.

Woodberry, Robert D. 1998. "When Surveys Lie and People Tell the Truth: How Surveys Oversample Church Attenders." *American Sociological Review* 119-22.

Yasui, Naoko Yura, and Norman L Berven. 2009. "Community Integration: Conceptualisation and Measurement." *Disability and Rehabilitation* 761-71.