ABSTRACT

Title of Document:	THE RELATIONSHIP BETWEEN RESPONSE PROPENSITY AND DATA QUALITY IN THE CURRENT POPULATION SURVEY AND THE AMERICAN TIME USE SURVEY
	Scott S. Fricker, Ph.D. 2007
Directed By:	Professor Roger Tourangeau, Director of the Joint Program in Survey Methodology

An important theoretical question in survey research over the past fifty years has been: How does bringing in late or reluctant respondents affect total survey error? Does the effort and expense of obtaining interviews from difficult to contact or reluctant respondents significantly decrease the nonresponse error of survey estimates? Or do these late respondents introduce enough measurement error to offset any reductions in nonresponse bias? This dissertation attempted to address these questions by examining nonresponse and data quality in two national household surveys—the Current Population Survey (CPS) and the American Time Use Survey (ATUS). Response propensity models were first developed for each survey, and busyness and social capital explanations of nonresponse were evaluated in light of the results. Using respondents' predicted probability of response, simulations were carried out to examine whether nonresponse bias was linked to response rates. Next, data quality in each survey was assessed by a variety of indirect indicators of

response error—e.g., item missing data rates, round value reports, interview-reinterview response inconsistencies, etc.—and the causal roles of various household, respondent, and survey design attributes on the level of reporting error were explored. The principal analyses investigated the relationship between response propensity and the data quality indicators in each survey, and examined the effects of potential common causal factors when there was evidence of covariation. The implications of the findings from this study for survey practitioners and for nonresponse and measurement error studies are discussed.

THE RELATIONSHIP BETWEEN RESPONSE PROPENSITY AND DATA QUALITY IN THE CURRENT POPULATION SURVEY AND THE AMERICAN TIME USE SURVEY

By

Scott S. Fricker

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
2007

Advisory Committee:
Professor Roger Tourangeau, Chair
Professor Stanley Presser
Professor Suzanne Bianchi
Associate Professor Fred Conrad
Dr. N. Clyde Tucker

Dedication

For my family, who offered me unconditional love and support throughout the course of this dissertation. To my wife, Buckley, who shared with me all the sacrifices and emotional burdens involved, and saw me through to the end with her patience, strength, and friendship. To my kids, Kelly and Amelia, who are growing into such wonderful human beings—thank you for daily reminding me what is really important in life. To my parents, who gave me the tools to succeed, and my brother for always being on the other end of the phone when I needed him. To Ira and Janet, for their tremendous generosity and encouragement. To each of you, I extend my deepest appreciation and love.

Acknowledgements

I owe thanks to many people whose assistance was indispensable in completing this project. First among these is Roger Tourangeau, advisor and committee chair, for persevering with me throughout the time it took me to complete this research and write the dissertation. His analytic and editorial advice was invaluable, and his continual willingness to share both his vast knowledge and great humor was an honor and an inspiration to me. In addition to Roger, I also want to thank the other members of my committee—Stan Presser, Fred Conrad, Clyde Tucker, and Suzanne Bianchi. They not only spent considerable time evaluating the design, data, and analyses of the research in this dissertation, but also asked excellent questions that pushed my own thinking and improved my work. I thank the faculty and my fellow graduated students at the Joint Program in Survey Methodology at the University of Maryland and the University of Michigan for providing a rigorous and invigorating intellectual environment. In particular, the time I spent in the Ph.D. seminar was the most enjoyable and rewarding experience of my academic life. I would also like express my deep appreciation for the generous guidance I received from Frauke Kreuter, and the programming assistance I received from Ting Yan. Finally, I want to thank my colleagues at the Bureau of Labor Statistics for first introducing me to the field of survey methodology almost 10 years ago, and for their support and encouragement ever since. I would like to specifically thank Clyde Tucker (again), Bill Mockovak, John Dixon, Steve Miller, Anne Polivka, Brian Meekins, and Dori Allard for repeatedly taking time out of their busy schedules to answer my questions and share their wisdom.

Table of Contents

I. Nonresponse and Data Quality in Household Surveys

1.1 Introduction

The common wisdom about sample surveys is that their inferential value is jeopardized by nonresponse, and response rates are often used as an indicator of survey quality. Concerns about falling response rates in sample surveys over the past few decades have stimulated the development of theories about survey participation decisions in the hope of identifying and countering the social and cognitive causes of nonresponse. Survey practitioners have integrated these concepts into their survey design and procedure decisions, which often involve extraordinary efforts to minimize nonresponse rates (e.g., use of advance letters, incentives, rigorous interviewer training, numerous callback attempts, refusal conversions, etc.). The usefulness of the nonresponse rate as a predictor of nonresponse bias has been called into question, however, by several recent studies that showed little change in survey estimates as a function of response rates. (e.g., Curtin, Presser, and Singer, 2000; Keeter, Miller, Kohut, Groves, and Presser, 2000; 2006; Merkle and Edelman, 2002). The implication of these studies is not that nonresponse bias does not exist, but rather that for response rates to be an effective indicator of nonresponse bias the underlying causes of survey participation must be correlated with the variables of interest in the survey.

Another source of error in survey estimates that has received considerable attention in recent years is measurement or response error—the difference between the value of a characteristic given by a respondent and the unknown but true value of that characteristic. Response errors are thought to be caused by characteristics of respondents—their knowledge, ability, and motivation to answer survey questions fully and accurately—and those of the interviewer, questionnaire, and survey design. For the most part, survey researchers have neglected studying how factors influencing nonresponse also affect response error, largely because the two types of errors often have been assumed to be independent of one another. However, the results of several studies challenge this assumption and suggest that the quality of respondents' answers may be related to their likelihood of participation (e.g., Bollinger and David, 2001; Olson, 2006)

1.2 Overview of this Dissertation

This dissertation examines these two phenomena—survey nonresponse and poor data quality—in two national household surveys—the Current Population Survey (CPS) and the American Time Use Survey (ATUS). I begin in Chapter 2 by reviewing theories of survey nonresponse and studies of nonresponse bias in greater detail, and then briefly discuss the empirical evidence for the relationship between nonresponse propensity and survey data quality.

In Chapter 3, I examine the correlates of nonresponse in the CPS and ATUS building on previous research, and in particular focus on two competing explanations of nonresponse—busyness and social integration or social capital. The busyness hypothesis suggests that the probability of responding to survey requests is getting lower because people are busier, more time-stressed, and are less interested in taking the time to cooperate than they were in the past. According to the busyness hypothesis, characteristics that reduce discretionary time and/or increase subjective time-pressure will reduce survey participation. Social integration or social capital notions of survey participation, by contrast, suggest that individuals who are involved in rich social networks develop norms of cooperation and trust that will decrease their likelihood of nonresponse. On the basis of these hypotheses and findings from the broader nonresponse literature, I develop logistic regression models that predict sample members' response propensities in each survey from indicators of respondent busyness and social capital, as well as other demographic and survey process measures. I then review the busyness and social capital hypotheses in light of the results of these analyses. The final analytic step in Chapter 3 examines the effects of removing high nonresponse propensity cases on survey estimates in each survey, with specific attention given to estimates related to the underlying constructs of busyness and social capital (i.e., those that should be the most susceptible to nonresponse bias).

Chapter 4 examines response quality in the CPS and ATUS. In particular, I look at a number of indicators of data quality that have been relatively understudied in the literature—item missing data rates, respondents' use of round values reports on questions involving continuous variables, response inconsistencies between interviews in the CPS, and the amount of reporting and the absence of certain types of reporting in the ATUS diary. Guided by the literature on the causes of measurement error, I then explore the causal roles that various household, respondent, and survey design attributes may have on the level of reporting error. Finally, I assess the relationship between the individual data quality indicators within each survey and between surveys, and then examine whether these relationships change in the presence of various causal factors.

Chapter 5 utilizes the nonresponse propensity scores and data quality indicators developed in Chapters 3 and 4 to examine the association between nonresponse propensity and data quality in the CPS and ATUS. I begin by looking at the data quality indicators from each survey across propensity strata to assess the relative size and direction of the association. I then explore the extent to which the covariance (where it exists) results from variables that are causally related to both response propensity and data quality. I then repeat these analyses using alternative indicators of nonresponse propensity (e.g., refusal conversion status) and observed nonresponse in the CPS, and end the chapter with a discussion of the implications of these findings for nonresponse reduction efforts in these two surveys, and for studies of nonresponse bias in general. Chapter 6 extends this discussion and provides a summary of the major findings and contributions of this dissertation.

II. Sources and Implications of Survey Nonresponse

2.1 Introduction

Since their inception, scientific surveys have relied almost exclusively on the voluntary participation of sample members, and survey researchers and practitioners long have had to contend with the issue of survey nonparticipation. In recent years, however, a number of studies have shown that there has been a steady decline in survey participation over the last couple of decades (e.g., Groves and Couper, 1998; Harris-Kojetin and Tucker, 1999; de Leeuw and de Heer, 2002; Tortora, 2004). As Tourangeau (2004) has argued, all three sources of nonresponse—failure to contact sample members, to persuade them to take part, and to accommodate their limited abilities to complete the survey—have gotten worse over the last decade.

Concerns about falling response rates have raised questions about the value of sample surveys. Declining response rates threaten the validity of the data collected by reducing sample sizes and increasing the likelihood of bias in the sample. In response, many survey organizations have established minimum response rate criteria in hopes of obtaining more accurate estimates and increasing the public's confidence in the value of their surveys. In the federal statistical system, the Office of Management and Budget and the U.S. Office of Federal Statistical Policy and Standards instructed federal agencies sponsoring data collections that the quality of data obtained may not be sufficient if the response rate falls below 80 percent (OMB, 2006).

Private survey firms have lagged behind in the publication of response rates in the surveys they conduct, though a growing number are responding to response rate calculation and disclosure guidelines set forth by the American Association for Public Opinion Research (e.g., Pew Research Center, Henry J. Kaiser Family Foundation, and The Washington Post) (AAPOR, 2000). And, perhaps surprisingly given the perception that low response rates may be a barrier to peer-reviewed publication, many scholarly journals that routinely publish survey research do not have policies regarding the publication of response rate information (Johnson and Owens, 2003).

2.2 Effects of Nonresponse

Response rates often are seen as the best single indicator of the quality of a survey, but they are only indirect indicators of the risk of nonresponse error. Nonresponse affects the variance of the estimates (by reducing the sample size) and, to the extent that there are differences between the respondents and nonrespondents on the variable of interest, it affects bias as well. The usual expression for the nonresponse bias of the (unadjusted) respondent mean is

$$Bias(\overline{y_r}) = (1-\overline{p})E(\overline{y_r}-\overline{y_n}) \qquad (1)$$

where $\overline{p}$ is the expected probability of responding, and $\overline{y_r}$ and $\overline{y_n}$ are the means of respondents and nonrespondents on the survey variable, respectively . As Equation 1 indicates, bias results if survey nonresponse is not random—i.e., if those who respond differ from those who do not on the survey variables—the results of the survey may be biased. To the extent that respondents differ from nonrespondents, surveys with low response rates will be unrepresentative of the population under study. Although we can measure the impact of nonresponse on demographic characteristics (e.g., by comparing the sample's distributions on demographic variables to the CPS or some other benchmark), its impact on survey estimates of interest is typically unknown.

Because response rates are used as an indirect indicator of survey quality, survey organizations spend a good deal of effort and money in an attempt to improve response rates. Advance letters, use of incentives, refusal conversion training, and rigorous callback schedules are some of the typical methodological enhancements employed to counter falling response rates. Recent papers by Keeter, Miller, Kohut, Groves, and Presser (2000) and Curtin, Presser, and Singer (2000) suggest that the higher response rates achieved with these methods may not lead to reductions in nonresponse bias, and that bringing in excluded groups (e.g., late responders and initial refusers) may not affect the estimates appreciably.

Keeter and his colleagues, for example, compared the results of a standard survey using an at-home sample and a 5-day period fielding period with results from a rigorous survey conducted over an 8-week period with random selection of a household respondent. The authors found that increasing the number of callbacks and extending the survey's field period, as well as implementing refusal conversion techniques and offering cash incentives, produced a significantly higher overall response rate for the rigorous survey than the standard survey (60.6% vs. 36.0%), and led to some significant differences in the demographic variables across the two groups. However, there was little evidence to suggest that nonresponse error was related in any systematic way to this difference in response rates. The low response rates in the standard survey apparently did not lead to major statistical biases; relatively few statistically significant differences in the distribution of respondents' attitudes or knowledge emerged between the standard and rigorous surveys, and even fewer of the differences were practically significant (see also the recent replication by Keeter and his colleagues, 2006).

The results of Curtin et al. (2000) similarly call into the question the idea that lower response rates necessarily produce more biased survey estimates. Utilizing call records from 17 years of the Survey of Consumer Attitudes, the authors retroactively excluded particularly effortful cases (i.e., respondents who required refusal conversion or multiple callbacks). This allowed comparison of estimates from the full dataset to those based on lower response rates, and provided a means of assessing the impact of nonresponse reduction efforts on nonresponse bias. Consistent with Keeter et al. (2000), the authors found very little change in the estimates as a function of response rates.

These and similar findings (e.g., Groves et al., 2004; Merkle and Edelman, 2002) contradict the generally held view that nonresponse reduction procedures help to reduce nonresponse bias (at least in a relative sense), and seem to suggest that such efforts are inefficient (and perhaps unnecessary). Information on the magnitude of nonresponse bias is rarely available, however, and thus nonresponse rates remain widely used as indicators of the potential for bias and of the quality of the data collection procedures. Lower response rates with household survey requests in the U.S. and abroad present significant data quality and cost concerns, and considerable theoretical and empirical work has been devoted toward a better understanding of the causes of survey participation.

2.3 Theories of Survey Response

As conceptualized by Groves and Couper (1998), completing interviews requires two things: (1) locating and gaining access to sample households (contact), and (2) gaining respondents' consent to conduct interviews (cooperation). (In some surveys, a third requirement also is important—that respondents have the physical and mental capacity needed to do the survey.) Contact and cooperation both are affected by the survey field operations and characteristics of the sample members.

Contactability

Contactability is directly influenced by such survey design features as the number and timing of contact attempts, as well as household features such as at-home patterns, residential mobility, and barriers to accessibility (e.g., gated communities, answering machines). Households in which eligible respondents are frequently at home during daytime and early evening hours are more likely to be contacted than those that are at home less often. At-home patterns are largely determined by employment and child care responsibilities, and thus rural households, those with young children, elderly or unemployed adults, and those without privacy impediments (e.g., caller ID)—those who are more likely to be at home at the time the interviewer calls or more likely to answer their telephones—have higher contact rates than households without these characteristics (Groves, Wissoker, Greene, McNeeley, and Montemarano, 2001; Link and Oldendick 1999; Presser and Singer, 2007; Tuckel and O'Neill, 2002).

Cooperation

The characteristics of households and respondents that affect cooperation are more complex and elusive. The propensity to cooperate is affected by features of the survey design (e.g., advance warning of the survey request, survey mode), the characteristics of the sample member (demographic variables often serve as proxies for hypothesized underlying psychological predispositions), and relatively stable features of the environment (e.g., neighborhood characteristics). In addition, cooperation is greatly affected by factors related to the interaction between sample member and interviewer. Both the potential respondent and the interviewer bring various background characteristics (and often conflicting goals) that affect how they behave in the interaction. Thus, the sample person's perception of the legitimacy of the survey request, interest in the topic, privacy/confidentiality concerns, judgment about the response burden, and other factors affecting cooperation are negotiated in part through the interaction with the interviewer (e.g., Campanelli, Sturgis, and Purdon, 1997; Groves, Presser and Dipko, 2004; Singer, Van Hoewyk, and Neugebauer, 2003).

The first 50 years of literature on survey response provided comprehensive reviews of the social-psychological principles relevant to the decision to cooperate with a survey request (e.g., Deming, 1947; Dillman, Gallegos, and Frey 1976; Morton-Williams, 1993; Groves, Cialdini and Couper, 1992). More recently, researchers have developed a number of theoretical explanations of survey response that examine the potential respondent's decision within the conceptual framework of *utility theory* (e.g., see Read, 2004, for an extended discussion of utility theory).

2.3.1 Utility and Social Exchange Theory

Utility theory emphasizes the cost-benefit calculations that individuals use to determine survey participation. Utility theory implies that when individuals are faced with the decision about whether to participate in a survey, they consider both the expected costs and benefits (or utility) of participation, and choose to complete the survey only if the benefits outweigh the costs. The seminal work on survey nonresponse by Groves and Couper (1998) identified many of the factors that affect the perceived utility of participation. The costs of survey participation arise from the missed chance to do some other valued activity, the temporal and cognitive burdens of comprehension and response to the questions, and the potential for embarrassment or unwanted disclosure of responses. Benefits may come in the form of monetary compensation, the attainment of social acceptance, or in the feelings of doing one's civic duty (Groves and Couper, 1998). According to this framework, the decision to cooperate with a survey request is based on an analysis of the exchange of costs and benefits and is motivated by an individual's expectation that cooperation will result in positive subjective rewards.

The leverage-saliency theory developed by Groves, Singer, and Corning (2000) provides a detailed illustration of the principles guiding an expected utility model of decision making regarding survey participation. Leverage saliency theory suggests that different features of the survey request have different "leverage" in the decision process of different individuals. The relevance of these leverage values varies depending on how salient these features are in the survey request. For example, Groves et al. (2000) found that when civic-minded sample members were offered incentives for participation there was no impact on cooperation, presumably because for these individuals the intrinsic value of survey participation was already enough reason to cooperate. When sample members with low levels of civic engagement were offered the same incentive, however, there was a significant, positive impact on cooperation. In a more recent test of the hypotheses derived from leverage-saliency theory, Groves, Presser, and Dipko (2004) found that people predisposed to be interested in a particular survey topic were more likely to cooperate with a survey request. The results of this study were somewhat mixed, however, and they underline the fact that the decision to cooperate is made on the basis of influences that are diverse and often hard to operationalize.

The expected reward (utility) of survey cooperation may be derived from sources that are less immediate and direct than the proximate features of the survey. S*ocial exchange theory* (SET) extends standard utility theory beyond its standard, short-term specifications by stating that decisions also are based on emergent properties of interactions involving the exchange of often intangible and long-term social commodities, such as the expectation of reciprocated trust or norms of obligation (Blau, 1964; Cook, 2000; Thibaut and Kelly, 1959). When social exchanges are on-going (e.g., as they are between an individual and the Federal government), and/or when there is uncertainty about the relative contributions of each party, SET says that people will avoid incurring obligations and have a tendency to help others.

The ideas of SET have been used to explain cooperation with surveys, most notably by Dillman (1978; 2000). Dillman (1978) argues that the propensity to cooperate is greatest when a person trusts that over time the expected rewards of responding will outweigh the expected costs. Building respondent trust and confidence in a rewarding exchange, according to Dillman, can be accomplished by tailoring the survey to the potential respondent, increasing survey design features that emphasize legitimacy and minimize response burden, and by offering 'benefits' that activate norms of reciprocity or obligation, (e.g., small monetary incentives; noting the intrinsic rewards of representing one's group or contributing to research in general).

2.3.2 Social Capital

The extent to which norms of reciprocity, trust, and cooperation are evoked will also depend on structural features of the social environment. Individuals develop shared values, norms, and expectations in response to their everyday experience with the world around them, and these help guide interpersonal behavior. Social capital is the term that is commonly used to refer to the reserves of social trust that people accrue through productive interactions and that facilitate cooperation and collective action (Putman, 2000). Like other forms of capital, social capital can be understood as an asset, one that increases the amount or likelihood of future cooperative behavior.

Social capital has received considerable attention in recent years, but not surprisingly given its multi-dimensional nature, no single definition of the construct has emerged (see, e.g., Adler and Kwan, 2002, for a discussion of the numerous definitions found in the literature). In general, however, social capital can be viewed as consisting of two distinguishable components. The structural component refers to the various types of social organizations that can contribute to cooperation, such as formal and informal associations (e.g., work, school, church, service providers, sets of friends, etc.). These social networks facilitate cooperative behavior by creating and disseminating norms and expectations, thereby establishing patterns of interactions that are predictable, productive, and positive. The cognitive component refers to the norms, attitudes, and values that are shared and reinforced within these social organizations or the larger civic culture. These in turn lead to the development of attitudes of interpersonal trust, reciprocity, and generosity (Uphoff, 2000).

Social capital theories offer insight about the social and psychological processes of interpersonal relationships that foster trust and cooperation. Moreover, they provide a framework for assessing the direction and degree of potential change in social behavior brought about by levels of social capital. Much depends, however, on the way in which social capital is operationalized. For example, Putnam (1995) views the main source of social capital as residing in civil society, specifically in the number and density of community groups (e.g., civic organizations), and argues that more is better—that is, more community groups lead to more social capital, civic engagement, interpersonal trust, and norms of reciprocity. In support of this idea, he found that falling levels of association membership in the US and Italy have had powerful negative effects on voter turnout, newspaper readership, and confidence in public institutions. Although Putnam's pioneering theoretical and analytic work on social capital underscores the important consequences of the relations between people and organizational entities in society, his conceptualization ignores other potential features of the construct (e.g., informal social networks; the existence and impact of interaction with differentially powerful associations), as well as the possibility for negative effects of social capital (e.g., when increased in-group bonding interferes with bridging or uniting larger communities). This may explain why other research has failed to find a relationship between membership in associations and an individual's level of trust in others (e.g., Li, Pickles, and Savage, 2005; Stolle, 2001). More nuanced approaches to social capital take into account a wider diversity of social networks—including both formal and informal associations, and outgroup contacts that also may serve as agents of socialization—as well as the social forces that affect the interrelationships between these components (e.g., Portes and Landolt, 2000).[1] These more comprehensive conceptualizations of social capital seem to reflect the multi-dimensional, interrelated nature of society, but they also raise difficult measurement issues. Because many of the constructs of social capital are inherently abstract (e.g., trust, group identity), their operationalization inevitably involves the use of indirect indicators that are open to conceptual debate.

For example, a central tenet of the social capital approach is that cooperative behavior is a product of network richness and cohesion, but there remains a debate about how social capital should be assessed and whether effects occur at the individual or the community level.[2] Social capital can be approached from both levels of analysis, but one must explain how each level relates to the mechanisms of trust, cooperation and reciprocity. Structural attributes of the community that have

[1] By 'outgroup members' I mean those belonging to a different racial or socio-demographic group than one's own.

[2] Similar ideas are expressed elsewhere in the literature. For example, s*ocial isolation theory* posits that the socially isolated individual with poor social networks will have underdeveloped norms of trust and reciprocity that result in low rates of cooperation (see, e.g., Couper, Singer, and Kulka, 1997). Research on the social psychology of helping behavior also underscores the importance of social identity and cohesion. For example, individuals are more likely to help members of their own group and less likely to engage in prosocial behavior with outgroup members (e.g., Gaertner, 1973).

been examined for their effect on cohesion and social capital include community instability, racial heterogeneity, and socioeconomic inequality (e.g., Bellair, 1997). Recent studies in Britain have found that racially diverse communities have less developed social networks and lower levels of interpersonal trust than homogeneous communities (Alesina and La Ferrara, 2000; Costa and Kahn, 2003). Letki (2005) and others (e.g., Li et al., 2003) argue convincingly, however, that the relevant dimension for social capital development and cooperative behavior is not racial diversity but rather economic inequality between groups. Inequality creates social divisions that affect the quality of interactions and can leave those who are disadvantaged feeling alienated and unwilling to cooperate in arrangements that they believe maintain the status quo (Blau, 1964). Confirming this hypothesis, Letki (2005) found that low socio-economic status (relative to others in the neighborhood) was associated with reduced levels of organizational involvement, network attachments, trust, and cooperative behavior, whereas racial diversity had no significant effect on these outcomes once neighborhood status was controlled for.

Individual-level attributes also are likely to affect the accumulation of social capital. For example, both education and employment provide individuals with expanded network opportunities, greater exposure to norms of participation and reciprocity, and increased efficacy (e.g., Heyneman, 1998). Similarly, one's embeddedness in a community is related to both the opportunity to form social relationships and to develop pro-social norms. Indeed, Hyman and Wright (1971) and Perkins, Brown, and Taylor (1996) found that length of residence and home ownership are associated with gains in social capital in the form of increased community participation and interpersonal trust. Attributes of the family also are expected to affect individuals' networks and norms. Spouses and children enlarge and diversify one's pool of social contacts and may engender values of caring and cooperation. Finally, age and disability may limit one's capacity to form rich social networks and undermine efficacy, but may also contribute to a sense of obligation to organizations that provide needed services.

2.3.3 Busyness

An alternative to social capital conceptualizations about survey participation is that cooperation simply is a product of how busy a person is. In America at least, individuals feel that they are busier than they have ever been (Robinson and Godbey, 1997). And, although time-diary data suggests that this perception is inaccurate, several societal shifts in the last decades may contribute to perceptions of busyness. First, there is a greater participation by women in the labor force and single working mothers than ever before, both of which may contribute to women feeling especially harried. Second, the last decade has seen an exponential growth in the number of unwanted contacts by telemarketers, pollsters, email spammers and others. There is a general consensus among legitimate survey methodologists that people feel bombarded by such contacts and have taken steps to reduce unwanted contact. For example, there has been a rise in the number of technological countermeasures designed to filter and insulate people from bothersome contacts (e.g., using caller IDs and answering machines to screen calls; spam blockers to filter mass emails). Furthermore, individuals may develop rules for quickly dispatching calls from people or organizations that they do not immediately recognize. All of these factors point to the possibility that propensities to respond to survey requests are lower because people are busier (or perceive themselves to be), and if contacted, are less interested in taking the time to cooperate (Abraham, Maitland, and Bianchi, 2006).

Regardless of the theory (or theories) one adheres to, work in the psychological literature alerts us to the fact that choices often are not made on a systematic, thoughtful basis. Rather, individuals often use cognitive shortcuts or heuristics in making decisions. For example, the decision to cooperate with a survey may hinge on initial impressions of the interviewer or the sponsor's affiliation (Groves and Couper, 1998). In fact, because potential respondents in most survey situations lack the interest, knowledge, or time necessary to systematically calculate the personal costs and benefits of participation, decisions often will be based on heuristic cues made salient at the time of request. Indeed, there is ample empirical evidence that decision-making in surveys frequently is driven not by careful deliberation but by 'surface' features of a survey request that evoke norms of reciprocity, consistency, or deference to authority (Groves, Cialdini, and Couper, 1992).

2.4 Stochastic Conception of Nonresponse

This has implications for our conceptualization on nonresponse bias in a survey estimate. Equation 1 (above) implies that nonresponse is deterministic—that is, a person either is a respondent or a nonrespondent. As the proceeding discussion makes clear, however, the decision to participate in a survey is susceptible to a variety of contextual influences and is unlikely to remain the same over time (or theoretically, over different realizations of the same survey design). Rather, people will choose to participate in some instances and not in others, with their decision based on whichever survey features and judgment criteria are salient at the time. From this perspective, an individual can have varying probabilities or *propensities* of being a respondent or nonrespondent, and these propensities may or may not be related to the survey variable. Thus, an alternative formula of nonresponse bias is the following:

$$(2) \qquad E\left(\overline{Y_r} - \overline{Y_n}\right) = \left[\frac{\sigma_{yp}}{\hat{p}}\right]$$

where $\overline{Y_r}$ is the mean of the variable of interest for respondents, $\overline{Y_n}$ is the mean for the full sample, and σ_{yp} is the covariance between the survey variable, *y*, and the response propensity, *p* (Bethlehem, 2002). This formula treats nonresponse as a stochastic outcome and suggests that survey error will occur to the extent that the factors influencing response propensity are related to the survey variable of interest.

2.5 Methods of Assessing Nonresponse Error

Because direct information about nonrespondents is seldom available, indirect analytical approaches typically must be employed to assess nonresponse error. One approach has been to compare estimates from the full survey dataset to estimates derived from a dataset in which hard-to-reach respondents have been removed (Curtin et al., 2000). An increasingly common variant of this method is to compare those who respond late in the survey field period to those who respond early. This method is based on a continuum-of-resistance model (Fitzgerald and Fuller, 1982; Lin and Schaeffer, 1995), in which an individual's propensity to respond is inferred from the level of effort required to obtain their participation, and an assumption is made that the lower a respondent's propensity (i.e., the more difficult they are to get) the more similar they are to nonrespondents.

Propensity scores are calculated using logistic regression models and are the estimated probability that a person will respond, given a vector of observed covariates (e.g., number of contacts, refusal conversion attempts, interviewer workload, and when available, demographic characteristics of the individual, etc.). Propensity score methods utilize indicators of respondent resistance and seek to identify groups of respondents that are similar to actual nonrespondents on the full range of available variables.

Recent studies that looked at the effects of bringing in hard-to-reach and/or reluctant respondents have found some evidence of differences in the composition of the sample as a function of effort. For example, Sangster (2003) found that cases requiring multiple contacts were significantly different from those requiring fewer contacts. Bates and Creighton (2000) found that late responders were more likely to be from higher-income households than early responders. Other predictors of nonresponse in their study (e.g., age, urbanicity), however, failed to support the continuum-of-resistance model. Finally, Lin and Schaeffer (1995), using external validation data, found only mixed support for the assumptions that (a) effortful respondents are like hard core nonrespondents, or (b) bringing in these difficult respondents had any appreciable effect on the magnitude of nonresponse biases in estimates of child support. In general, then, the empirical literature on this approach to assessing and controlling for nonresponse bias is mixed with respect to whether or not efforts to contact and/or persuade the most difficult respondents are effective in terms of reducing survey error.

How might we explain these equivocal results? As discussed earlier, nonresponse error depends on the relationship between the survey variables of interest and the cause of nonresponse. When attributes of the survey (e.g., survey sponsor or mode) or the respondent (e.g., gender, household composition) are causally linked both to the propensity to respond and the responses on the survey variables of interest, error will result (*cf.*, Groves, 2006) The following examples illustrate this relationship. A common finding in the literature is that older adults are less likely than younger adults to be noncontacts. Age also is likely to be correlated with many survey variables of interest (e.g., expenditure patterns, participation in leisure or volunteer activities, etc.), and so without statistical correction for differential nonresponse by age, the estimates of these variables would be biased. The value on the survey variable of interest may play a direct causal role on contact, as well, such as when the variable of interest (e.g., paid work or travel) is likely to be causally related to response propensity (via the increased likelihood of noncontact). Indeed, there is evidence to suggest that time-use surveys are particular vulnerable to bias in this way. Mulligan, Schneider and Wolfe (2001), for example, found that estimates of work hours were significantly underestimated due to higher nonresponse by individuals working longer hours.

2.6 Response Process and Response Errors

According to Tournageau's (1984) theory of survey response, there are four cognitive components to answering a survey question: comprehending the question; retrieving relevant information from memory; integrating information to arrive at a judgment; and formulating and editing a response. Respondent and survey design attributes can affect each stage of the response process, influencing the level of task difficulty, the level of effort that respondents give to answering the questions, and their motivation to provide accurate answers. In addition, certain respondent characteristics (e.g., age, education level) may affect their ability to answer fully and accurately. According to the theory of *satisficing* (Krosnick, 1999), respondents often shortcut the cognitive processing needed to generate an optimal answer and instead settle for a merely satisfactory response in order to minimize the psychological costs of accurate reporting (Tourangeau, 1984). According to this theory, weak *satisficing* occurs when respondents go through each component of the response formation process, but devote reduced effort to all or some of the stages. Strong *satisficing* occurs when respondents do not engage at all in the recall or integration processes but still attempt to provide answers that are acceptable or seem reasonable.

A number of empirical studies support the notion of satisficing in various forms. For example, Holbrook, Green, and Krosnick (2003) found that "don't know" responses and other item nonresponse were more likely to occur in telephone than in-person surveys, and that the most pronounced mode differences were for the least educated. Similarly, telephone respondents provided less differentiated responses than in-person respondents to groups of items that had identical response options, suggesting that telephone respondents were unable or unwilling to process these items as carefully as the respondents to the face-to-face survey. Other satisficing behaviors include the rounded reports of income or wages, "age heaping" or "time heaping" around regular intervals, and unusually abbreviated reports (in terms of number of items reported or the interview duration).

Apart from the limited cognitive abilities of older or less educated respondents, what other factors might induce satisficing? One candidate is "busyness," since busy respondents who are contacted may not be willing to take time to provide careful answers. Social capital may be another possibility. Individuals high in social capital are civically and socially engaged and have attitudes of trust and reciprocity gained through informal interactions with organizations and individuals. According to Putman (2000), a high degree of social capital encourages collaboration with friends and strangers. Thus, in some instances we might expect that individuals with high social capital will attempt to help the interviewer by engaging in more effortful processing. Another possibility is that the desire of high social capital respondents to help interviewers may induce demand characteristics (i.e., giving the answer that they think the interviewer wants) that lead to poor data quality.

2.7 The Relation Between Response Propensity and Data Quality

The effort that individuals give during the survey response process and the quality of their data also may be related to their propensity to respond to the survey request. Changes in propensity may lead to changes in the underlying motivation to respond fully and accurately. For example, increasing a reluctant respondent's trust in the legitimacy of the survey organization or the confidentiality of data may increase his or her likelihood of cooperating *and* of providing careful and truthful answers. By contrast, there may be situations in which the underlying motivation to respond and to respond accurately are related, but survey features designed to increase response propensities (e.g., making additional contacts) do not yield decreases in measurement error. For example, social capital may promote norms of cooperation that affect both amenability to the survey request *and* accuracy of reports. Individuals with high social capital may not only be more willing to take part, but more willing to expend the effort to do a good job.

Although the relationship between nonresponse and measurement error has been relatively understudied, there are a few examples in the literature. For example, Cannell and Fowler (1963) found that respondents who responded at the end of the survey field period provided less accurate reports of their hospital stays than those who responded earlier.[3] More recently, Bollinger and David (1999) utilized matched records from a panel survey to check the accuracy of survey respondents' reported participation in the Food Stamp program. They found that cooperative respondents (i.e., those who had positive values on a cooperativeness latent variable) were less likely to miss an interview and more likely to accurately report program participation than uncooperative respondents.

Besides studies that have examined the relationship between direct measures of reporting accuracy and response propensity, there have been efforts to look at an expanded set of data quality indicators. Many of the indirect indicators examined in

[3] The untested presumption was that early responders were more amenable to the survey request. It is worth noting that another causal mechanism also may have come into play—respondents reporting later in the survey period may have been hampered by recall deficits due to a longer recall period than earlier respondents.

these studies are related to the satisficing behaviors outlined above. In one study of the relationship of propensity to data quality, Triplett, Blair, Hamilton, and Kang (1996) found that initial refusers were more likely than initial cooperators to have higher levels of item nonresponse and to give shorter, less informative responses to open ended questions. Similarly, Friedman and Clusen (2003) found that late responders were more likely than early responders to provide DK responses and to skip questions altogether. In these studies measurement error is an act of resistance. By contrast, in one of the few studies that have looked at nonresponse reduction methods and measurement error, Willimack, Schuman, and Lepkowski (1995) found that respondents who received incentives were more complete and used more words in their answers to substantive open-ended questions.

III. Response Propensities in the Current Population Survey and the American Time Use Survey

3.1 Introduction

In the last fifteen years, survey organizations have become increasingly concerned about falling survey response rates because of the potential threat nonresponse poses to the validity of survey estimates. This concern has stimulated more careful study of the causes of nonresponse in hope of reducing nonresponse and improving post-survey adjustment procedures. Recent empirical studies demonstrate that survey participation may be influenced by the survey protocol, structural features of the social environment in which the survey takes place, and characteristics of sample members, in particular their decision-making rules (e.g., Dillman, 2000; Groves et al., 1992; Groves and Couper, 1998; Groves et al., 2000). This literature also points to a variety of factors that may contribute to nonresponse trends, including concerns about privacy and confidentiality (e.g., Singer et al., 1993), survey burden (e.g., Apodaca et al., 1998), time-related stresses and demands (Campanelli et al., 1997; Couper, 1995), and declining levels of civic engagement (e.g., Groves et al., 2000). Chapter 2 identified several theoretical frameworks for nonresponse that offer causal mechanisms to account for these diverse findings. This chapter will explore two of these causal mechanisms, busyness and social capital.

3.1.1 Busyness

The busyness hypothesis is that response propensities are getting lower because people are busier and have less time for surveys than they did in the past. Moreover, the extent to which people feel time-related stress ultimately may be more important to the decision to participate than the actual amount of free time they have. According to the busyness hypothesis, factors that reduce discretionary time or increase subjective time-pressure will have negative effects on response (Abraham et al., 2006; Groves and Couper, 1998).

The evidence that busy people participate less than others in surveys is mixed. Support for the busyness hypothesis comes in part from analyses of what nonrespondents say. Studies that have implemented nonresponse follow-up procedures or recorded respondent reactions in the first few moments of the survey interview commonly find that busyness is the most frequent reason given for nonresponse (e.g., Abraham et al., 2002; Burton et al., 2004; Couper, 1997; Mertler, 2003). For example, Couper (1997) found that more than 20 percent of those contacted for a personal-visit survey said that they were "too busy" to participate, and those individuals were more likely to refuse the survey request than those who did not mention time constraints. Interestingly, in this study respondents' claims of being "too busy" were unrelated to the number of hours they worked or the presence of children in the household, two factors expected to impose time constraints. Though Couper (1997) did not directly examine the relationship between these respondent characteristics and nonresponse, other studies have looked at similar indicators of discretionary time and failed to find evidence of a clear relationship to response propensity, or found instead that the busiest people actually may be the *most* likely to cooperate (*e.g.,* Abraham et al*.,* 2006; Groves and Couper, 1998, p. 122; Stoop, 2005). By contrast, work by Smith (1983; 1984) suggests that nonresponse may be affected more by respondents' subjective sense of time pressure than objective indicators of busyness. Response rates in his study fell as respondents' feelings of being rushed increased, but did not change as a function of objective measures of discretionary time.

Additional support for the busyness hypothesis comes from studies that have examined nonresponse in populations where the amount of individuals' discretionary time is known in advance. For example, Drago et al. (1999) compared survey response rates across teachers working in schools that varied in the amount of time-related stressors. They found that busy teachers were the least likely to cooperate with the survey request (see also Mertler, 2003). Similarly, surveys of health care professionals have consistently found that physicians' demanding work schedules reduce survey participation (e.g., Kellerman and Herold, 2001; Price, 2000). The actual or perceived burden of the survey itself also interacts with the perceived absence of discretionary time to affect nonresponse. The burden of lengthier surveys and the perceived costs of longitudinal surveys both have been shown to reduce response rates (*e.g.,* Apodaca et al., 1998; *cf.*, Bogen, 1996; McFarlane, 2006), especially for those under heightened time pressure (e.g., Asch et al., 1997). The proposed explanation is that longer survey instruments and/or recurring interviews in panel surveys represent potential burdens which discourage participation, and that these costs interact with time constraints to further reduce response rates.

These studies have several implications for the role of busyness in survey participation decisions. First, sample person's explanations for nonresponse—in particular, protests of being "too busy"—may simply be a polite way of declining the survey request rather than a reflection of sample person's true situation (Couper, 1997). Second, busyness has been assessed in several ways—by measuring individual-level covariates of discretionary time, by aggregate-level or role-based variations in time constraints (e.g., urban vs. rural, executive vs. service occupation, male vs. female), and by indicators of feelings of time-related stress. Each of these factors may have an impact on survey participation, but socio-demographic indicators of discretionary time may not always map well to respondents' subjective feelings of time pressure.[4] The experience of time depends on the number and variety of time-uses, socio-cultural and environmental circumstances, and an individual's ability to allocate and coordinate these factors. In the absence of direct measures of subjective time pressure, assessment of busyness must be based on careful selection of individual- and aggregate-level covariates of discretionary time. Finally, these studies suggest that the salience of the survey topic and perceptions of burden can magnify or dampen the effects of busyness.

3.1.2 Social Capital

An alternative hypothesis suggests that cooperation with survey requests is mediated by pro-social norms that develop through experience in social networks. According to the social capital hypothesis, individuals who are socially integrated are likely to develop community attachments and habits of trust and cooperation that will stimulate them to participate in surveys. Respondent characteristics and features of the environment that increase the likelihood of forming rich and cohesive social networks therefore would be expected to have positive effects on survey cooperation. Conversely, factors that interfere with the development of social and community

[4] In the U.S., for example, there is an apparent disparity between individuals' heightened perceptions of time pressure and their actually having more free time on average (Robinson and Godbey, 1999).

integration would be expected to be negatively related to survey response (e.g., Groves and Couper, 1998).

The social capital hypothesis receives support from studies in the fields of survey research, economics, sociology, and psychology. As discussed in Chapter 2, definitions of social capital vary within this literature, but there is reasonable agreement on the factors that contribute to the formation of social capital. Both individual factors—such as education, employment, citizenship, home ownership, marriage and children—and community variables—such as relative economic well-being, racial homogeneity, and socio-economic stability—have been identified as determinants of social integration and shown to be associated with gains in pro-social norms and behaviors (e.g., Bellair, 1997; Glaeser et al., 2002; Letki, 2005; Perkins et al., 1996; Vigdor, 2004). In one of the few studies that had direct measures of both social network participation and norms of trust and cooperation, Brehm and Rahn (1997) found that individuals who were more socially integrated possessed levels of social trust that extended beyond their own social networks. Many more studies have been conducted which examine the impact of network associations and social integration on various forms of broader civic participation. These studies provide empirical evidence that as individuals' social network resources or community activities increase, so too does their likelihood of making charitable donations (Brooks, 2005), being politically engaged (e.g., La Due Lake and Huckfeldt, 1998), and cooperating with survey requests (e.g., Loosveldt and Carton, 2002; Voogt and Saris, 2003).

3.1.3 Chapter Overview

The analyses presented in this chapter examine the correlates of nonresponse in two national household surveys, the Current Population Survey (CPS) and the American Time Use Survey (ATUS). Data from each survey are used to develop response propensity models based on the theoretical notions of busyness and social capital to predict the likelihood of nonresponse. In the next section of this chapter, I provide an overview of the CPS and ATUS and briefly review studies that have looked at nonresponse in these surveys. I will discuss the methodology by which the survey datasets used in these analyses were created and describe how variables relevant to the busyness and social capital hypotheses were operationalized. I then will examine overall nonresponse trends separately for the CPS and ATUS, and present bivariate and multivariate analyses of the relationships between the predictor variables and nonresponse in each survey. Finally, I will examine the effects of excluding low propensity respondents on estimates of key survey statistics in each survey.

3.2 CPS and ATUS: Description and Review of Recent Empirical Research

3.2.1 Current Population Survey (CPS)

The CPS is a monthly household survey conducted by the Bureau of the Census for the Bureau of Labor Statistics that serves as the primary labor force survey in the United States. The inferential population for the CPS is the approximately 105 million households in the United States and the civilian, non-institutional population residing in those households. Each month, the CPS surveys approximately 60,000 households in 792 sample areas across the country on issues such as employment, earnings, and hours worked. CPS interviews are conducted during the calendar week that contains the 19th of the month. The reference period for questions related to labor force activities covers the week prior to the interview—the week that contains the 12th of the month.

Each CPS household is sampled on a rotational basis so that any given month includes eight different rotation groups. Households within a given rotation group are sampled for four consecutive months, are out of the sample for eight months, and then return to the sample for another four consecutive months. Typically, the first and fifth wave interviews are conducted in face-to-face interviews and the other waves are conducted by telephone. This rotation pattern makes it possible to match information on households monthly across their entire CPS life cycle by using the household's month-in-sample (MIS) number, and the household and individual identifiers provided by the CPS. A more thorough explanation regarding the CPS sample design is provided in BLS Technical Paper 63RV (BLS, 2002).

A number of studies have looked at the influences of nonresponse in CPS. Seminal work conducted by Groves and Couper (1998) matched nonrespondent cases in data pooled from six household surveys in the U.S. (including the CPS) to the decennial census data. Their results lend more support to the social capital hypothesis than to the busyness hypothesis. For example, they found that indicators of social integration—multi-person households, households with small children, low population density, high percent of persons under 20 years old in the neighborhood—had positive effects on survey cooperation. The finding that households headed by older persons had lower refusal rates was inconsistent with much of the previous literature and with theories of social integration. For the study's two main measures of time limitations—number of working adults in the household and hours away from home—no significant effects on cooperation (given contact) were found.[5]

Studies by Dixon and Tucker (2000) and Harris-Kojetin and Tucker (1998) extend the analysis of CPS nonresponse to include many of the same household and geographic variables examined by Groves and Couper but add individual-level variables. Consistent with the social capital hypothesis, these studies report lower survey participation in urbanized and dense areas and in single-person households, households without children, renters, and unmarried individuals. Unlike Groves and Couper (1998), Dixon and Tucker (2000) describe large effects for gender and race, with significantly more nonresponse for males than females and for blacks than non-blacks. In addition, Dixon and Tucker (2000) did find some support for the busyness hypothesis—respondents who worked longer hours tended to cooperate less—but the effect was very small. The busyness and social capital variables in this study had very similar effects on refusal and noncontact.

3.2.2 American Time Use Survey (ATUS)

The ATUS is a cross-sectional, computer-assisted telephone survey that is carried out by the Bureau of the Census for the Bureau of Labor Statistics. Its primary purpose is to provide national estimates of how Americans spend their time.

[5] Groves and Couper (1998) report some exceptions to these findings, however. For example, their measures of discretionary time were positively related to contact propensity, as predicted by the busyness hypothesis. Similarly, two variables thought to decrease social integration (neighborhood crime rates and urbanicity) were found to have no effect on survey cooperation rates once household-level control variables were entered into their model (pg. 124; 183).

The ATUS sample is drawn from CPS households that have completed their eighth CPS interview. A single household member from each responding CPS household is randomly selected to participate in the ATUS interview two months after the eighth CPS interview. The designated person is assigned a specific reporting day of the week (e.g., Monday); substitutions are not allowed either for the designated ATUS respondent or for the assigned reporting day. If the interview cannot be completed on the designated day during the first week of the interviewing period, subsequent interview attempts are made on the designated day each week for up to eight weeks.

Because the ATUS is a relatively new survey, only a handful of studies to date have looked at the correlates of ATUS nonresponse. In January of 2004, the Census Bureau administered a small-scale response analysis survey of ATUS respondents and nonrespondents, with the goal of better understanding sample members' reasons for their response decisions. O'Neill and Sincavage (2004) report that the two most frequent explanations given for ATUS nonresponse were unfavorable past experience with CPS and lack of time to complete the ATUS survey. The first explanation underscores the potential negative effects of perceived longitudinal burden stemming from the unique ATUS sampling procedure; the second appears to support the busyness hypothesis. As noted earlier, however, caution must be used when making inferences from respondent verbatims.

The results of two recent studies that used multivariate approaches to ATUS nonresponse suggest that many of the same factors that influence CPS survey participation also influence ATUS response decisions. Applying propensity score modeling to ATUS nonresponse using a variety of household correlates, O'Neill and

Dixon (2005) found that nonresponse was lowest for white, married, and older sample members, and for households where there was at least one relative present. In a more direct test of the busyness and social integration hypotheses, Abraham et al. (2006) found that their indicators of busyness (e.g., hours worked, spouses hours worked, and the presence of children) generally were either unrelated to response propensities or produced results opposite from those predicted by the busyness hypothesis. For example, overall response rates and refusals were highest for individuals who worked the most hours, for married people whose spouses worked very long hours, and for households were young children were present. As Abraham et al. (2006) note, the finding that the presence of children raised cooperation rates is more consistent with a social integration hypothesis (p. 21). Groves and Couper (1998) treated this variable as a proxy measure of social integration in their models, and found results similar to Abraham and her colleagues. By contrast, Abraham and colleagues found ample evidence corroborating the social integration hypothesis—individuals who were employed, married, well-educated, home owners, and living in non-urban areas had higher response rates. They also reported higher response rates for older, Hispanic, white, and high-income individuals.

3.2.3 Summary of CPS and ATUS Nonresponse Literature

The evidence from studies that have examined determinants of nonresponse in the CPS and ATUS is fairly consistent. The measures of busyness that have been used in these studies either were unrelated to nonresponse or in fact were associated with higher response rates. The one conflicting finding was that longer hours worked was associated with higher nonresponse in the CPS but lower nonresponse in ATUS. On the other hand, measures of social integration did produce effects on nonresponse in the direction predicted by the social capital hypothesis. Table 1 summarizes the findings from the studies reviewed in this section, and will serve as a starting point for the development of my own models of CPS and ATUS nonresponse. In the next section, I will discuss the CPS and ATUS data files used for the analyses in the remainder of this chapter, and describe how the dependent measures and busyness and social capital variables were operationalized.

Table 1. Summary of Findings on the Effects of Busyness, Social Capital, and Demographic Control Variables on Response Rates in CPS and ATUS[†] [*]

		CPS	ATUS
Busyness	Hours worked	↓	↑
	# of working adults	-	n/a
	Hours away from home	-	n/a
	Spouses hours worked	n/a	↑
Social Capital	# of HH members	↑	n/a
	Presence of relatives/non-fam	↑	↑
	Marriage	↑	↑
	Presence of small children	↑	↑
	Employment	↑	↑
	Education	mixed	↑
	Home ownership	↑	↑
	Urbanicity	↓	↓
	Density	↓	n/a
	% youth in area	↑	n/a
Demographic	Age	↑	↑
	Gender (F)	↑	-
	Hispanic	mixed	mixed
	Race (white)	↑	↑
	HH Income	↑	↑

[†]Sources: Groves and Couper (1998); Harris-Kojetin and Tucker (1998); Dixon and Tucker (2000); O'Neill and Dixon (2005); Abraham *et al.* (2006)

[*]The effects are represented as follows: ↑ (increase in nonresponse), ↓ (decrease in nonresponse), - (no effect found), 'n/a' (variable not studied), or 'mixed' (contradictory findings in literature).

3.3 Dataset Creation

3.3.1 CPS

The data files for the CPS household nonresponse analyses cover a two and a half year period from May, 2001 to October, 2003. They included CPS households that had been in sample the full eight months and that were eligible to be sampled in the ATUS in 2003. To create a longitudinal data file of CPS sample units, records from each month were matched based on household id, person id, month-in-sample (MIS), and year.

For the purposes of these analyses, households that were ineligible to participate in the CPS in any round by virtue of being vacant, demolished, nonresidential, etc. were excluded. In addition, I excluded households in which all the previous month's residents had moved and been replaced by an entirely different group of residents. There were two reasons for excluding these replacement households. First, the impact of movers on CPS nonresponse and labor force estimates has been examined elsewhere (e.g., Dixon, 2000). Second, it made little sense to model household panel nonresponse if the households from which the predictors were derived were not the same households that provided the dependent measure.

The resulting CPS dataset contained information on 251,000 individuals. Nonresponse in the CPS is a household-level phenomenon, however, so I created a household-level dataset that included basic household information—e.g., the number of household members by age group, race, employment status and hours worked per week, education level, occupational prestige, etc.—and information about the main household respondent. The *main respondent* was the household member who was the most frequent CPS respondent over the household's eight waves. Over 95% of households had a person who responded to the CPS four or more times, and many of them responded to all eight interviews.[6] After collapsing to the household level, the resulting CPS household data file had 99,135 records. Approximately 2% of these cases were coded as nonrespondents for all eight rounds. Almost no data were available for these records, and they were excluded from most of the analyses. Thus, unless otherwise specified, the CPS dataset that was used for the present analyses contained household-level and main respondent information for 97,053 cases.

3.3.2 ATUS

To create the ATUS dataset used in these analyses, records from the ATUS public-use files were merged with the ATUS Call History File. These files contained information about the respondent (e.g., updated demographic and labor force data), the household (e.g., composition, demographics, weight), the time-use activities of the designated ATUS respondent, the interview process (e.g., interview outcome codes), and ATUS call histories (e.g., outcome codes for individual call attempts). Records were matched based on household id and person id. The data files used for the ATUS household nonresponse analyses were selected to cover January – December 2003. The ATUS dataset then was matched to the CPS data file. The resulting ATUS dataset contained 30,760 observations.[7] Cases that were coded as

[6] If two or more people in the household responded an equal number of times, person-level data was retained for whichever person was the most closely related to the CPS reference person.
[7] This number of cases differs from the total number of ATUS cases sampled in 2003—38,941. The reduction in sample size here is due to cases that were ineligible in any CPS month-in-sample or that were CPS replacement households, which were excluded from both data files.

ineligible to participate in ATUS (because they moved out between the time of their last CPS interview and their scheduled ATUS interview) or of unknown eligibility (because no valid telephone number exists) were additionally excluded from the data file, resulting in a final ATUS data file with 25,778 records.

3.4 Variable Selection

Both busyness and social capital explain the effects of individual, household, and community level factors on nonresponse, though they point to quite different causal mechanisms. In the present analyses, the individual and household-level data were extracted directly from the CPS and ATUS data files; information about the community, however, was collected at the county-level through external sources. The definition of community as county is somewhat unsatisfying since the purpose of these measures was to serve as proxies for features of a small locality. In most cases, counties may be too large to capture this information accurately; ideally, neighborhood or block-level variables would be available to portray the community within which a household is located. However, public-use data files typically do not contain block-level information, so county-level data was the best measure available for community characteristics.

For both the CPS and ATUS datasets, two groups of variables were selected as indicators of busyness and social capital. Two additional groups of variables also were included in these analyses to account for interview process variables and respondent demographic attributes.

The dependent variable for CPS and ATUS analyses was overall unit nonresponse status, not the specific type of nonresponse (noncontact, noncooperation, and “other noninterviews,” which mainly include cases that could not be interviewed due to health or language barriers). The analysis thus does not take into account important differences that may exist in the underlying causes of the different components of nonresponse (e.g., Curtin *et al.*, 2005; Groves and Couper, 1998), and is a potential limitation of these analyses. Two factors mitigate this risk, however, and justify the use of a single overall measure of nonresponse in this study. First, the most significant problem in both CPS and ATUS is noncooperation. Refusals account for nearly seventy percent of all noninterviews in CPS (averaging across months-in-sample) and about sixty percent of noninterviews in ATUS. By comparison, only about twenty percent of noninterviews in either survey are noncontacts. The relative disparity between the number of refusals and noncontacts suggests that survey resistance is the biggest contributor to nonresponse in these surveys. Second, noncontact is in many instances a masked form of resistance. For example, two-thirds of ATUS noncontacts are missed callbacks—sample members who are contacted initially but who ask to be called back and then can never be reached. Given the extraordinary efforts to contact sample members (e.g., in about ten percent of ATUS cases, more than twenty-five call attempts are made to reach the designated respondent), it seems likely that some noncontact actually indicates the filtering out of unwanted intrusion. More generally, the use of technology by sample members for screening calls (e.g., caller ID, answering machines) suggests that noncontacts are overestimated in both surveys. Given these considerations, I decided to use the overall nonresponse status in these analyses rather than examining the different components of nonresponse separately.

3.4.1 CPS Variables

Since the vast majority of CPS households respond to all eight interviews, the dependent variable in the CPS analyses was nonresponse *at any wave* in months three through eight. Where possible, I created household and respondent-level predictors for the CPS analyses by collapsing data from the first two CPS waves. If data were missing for a household from the first month-in-sample interview due to unit nonresponse, information from the second month-in-sample interview was used to populate the variable. Data on the number of personal contact attempts is recorded every month, but is most meaningful during wave 1 when the majority of CPS interviews are conducted by face-to-face field visits. Finally, some predictor variables used in these analyses indicate the occurrence of an event in either of the first two months-in-sample—e.g., was the household a nonrespondent in either month one or month two? A list of the variables included in the CPS analyses is presented in Table 2.

Table 2. Current Population Survey (CPS) Variables

Type of Variable	Description	Response Categories
Process	Region of Country	1 = NE 2 = Midwest 3 = South 4 = West
	Season in Which MIS 1 and 2 Interviews Were Conducted	1= Winter (Nov – Feb) 2 = Summer (Jun – Aug) 3 = Spring or Fall
	Nonresponse in either MIS 1 or 2?	0 = No 1 = Yes
	Item Nonresponse on Family Income Question MIS 1?	0 = No 1 = Yes
	# of Person Contacts Made in MIS 1	0 – 9
Demographics	Age of Main Respondent	0 - 90
	Sex of Main Respondent	1 = Male 2 = Female
	Race of Main Respondent	1= White 2 = Black 3 = Asian 4 = Other
	Hispanic Origin of Main Respondent	0 = Not of Hispanic Origin 1 = Hispanic Origin
	Household Family Income	1 = Less than $30,000 2 = $30,000 - $75,000 3 = More than $75,000

Table 2 (continued).

Type of Variable	Description	Response Categories/Ranges
Busyness	Proportion of Adult Household Members Who Work	1 = No adults work 2 = At least 1 works and 1 does not 3 = All adults working
	Hours Worked Last Week for Main Respondent	0 - 120
	Do All Working Household Members Usually Work More Than 40 Hours Per Week?	1 = No 2 = Yes
	Type of Job of Main Respondent	1 = Not in Labor Force or Unemp. 2 = Executive/Professional 3 = Service 4 = Construction/Production
	Median Commute Time in County	10.2 – 38.9
	Metropolitan Status	1= Metropolitan Area 2 = Not metro, adjacent 3 = Not metro, not adjacent 4 = Rural
	Population Density of County	0 – 66,934
Social Capital	Marital Status of Main Respondent	1 = Married 2 = Not Married
	Presence of Children Under the Age of 6	0 = No children under 6 1 = At least one child under 6
	Presence of a relative or other non-family in the household	0 +
	Household Size	0 - 16
	Citizenship Status of Main Respondent	1 = Native US Citizen 2 = Naturalized US Citizen 3 = Not a US Citizen
	Employment Status of Main Respondent	1 = Employed 2 = Unemployed 3 = Not in Labor Force
	Home Ownership	1 = Rent/Other 2 = Own

Table 2 (continued).

Type of Variable	Description	Response Categories/Ranges
Social Capital	Percent Change in the Number of Business Establishments in the County, 1998 – 99	-10.3% – 22.2%
	Percent of Adults in County with HS education	49.4% – 99.5%
	Diversity Index in County, 1999	1.0 – 77.0%
	Median Family Income in Census Tract	0 - $150,000
	Income Inequality Index for County, 1999	30.9% – 58.4%
	Percent Population Change in Young Adults (Age 18 – 34) in County, 1998 - 1999.	-9.9% – 6.2%
	Unemployment Rate in County, 1999	0 – 33.3%
	Violent Crime Arrests in County, 1999	0 – 1,547

The first column of Table 2 identifies four groups of variables: process; demographic; busyness; or social capital. Each group in the table contains only those variables identified though preliminary modeling to have a significant bivariate relation to CPS nonresponse. When theoretically appropriate, I recoded the variables to ensure adequate sample in each cell, minimize skewness, and limit multicollinearity.

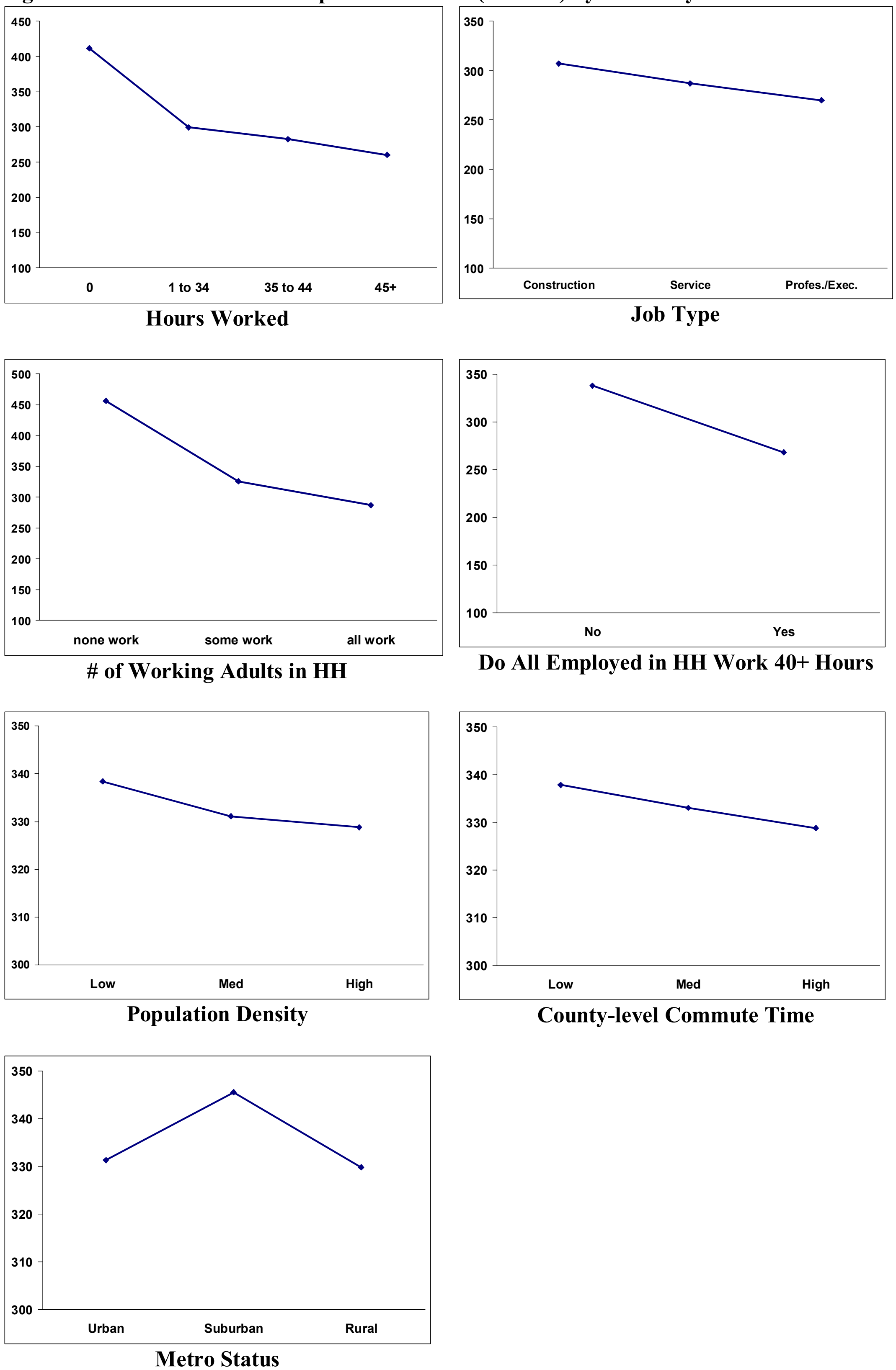
The process variables in Table 2 are perhaps best defined by what they are not—they are not unambiguously and distinctly related to only one of the two theoretical constructs—busyness or social capital. Instead, they are more generally associated with the interview process and indicators of reluctance. For example, region is a feature of the sample design that potentially relates to the quality of the interviewing corps as well as the attributes of the sample households, whereas season may affect ease of contact. Based on previous literature (*e.g.*, Groves and Couper, 1998), response rates were expected be lowest in the most populous regions of the country (i.e., the Northeast and West). The time of year of the first CPS interview was also predicted to noncontact, with winter and summer months exhibiting the highest nonresponse rates. The three remaining process variables—nonresponse in months-in-sample one or two, item nonresponse to the first month-in-sample family income question, and the number of personal contact attempts made in month-in-sample one—are indicators of reluctance and were expected to be positively related to CPS nonresponse in subsequent waves.

In addition to the process variables, the analysis included age, sex, race, Hispanic origin, and household income. These variables are ubiquitous covariates in most nonresponse analyses and could not be clearly identified as indicators unique to either busyness or social capital, so they are presented in the table separately. I expected that nonresponse would be higher for males than females, for younger individuals than older individuals, for racial and ethnic minorities than whites, and for families with greater household income than for poorer households.

The third group of variables was intended to be indicators of busyness. The busyness hypothesis predicts that refusal propensities should be greater on average for the employed than the unemployed and should increase with the number of hours worked per week. Four indicators of the impact of work are shown in the Table 2—the proportion of working adults in the household, whether all working adults work more than 40 hours per week, the number of hours worked by main respondent, and the occupational prestige of the main respondent. According to the busyness hypothesis, nonresponse should increase as a function of all four of these variables. I also included three community-level indicators of busyness: a county-level measure of median travel time to work, metropolitan status, and population density. Although the effects of urbanicity (metropolitan status) and population density on nonresponse sometimes have been attributed to their impact on social connectedness, congestion in dense, urban areas also may have negative effects on discretionary time. Since there is no clear empirical evidence about the mechanism responsible for urbanicity and density effects, all three of these variables are included here as measures of overall time pressure, and were expected to increase nonresponse under the busyness hypothesis. (Urbanicity and population density can also be considered social capital variables, but are listed only once in Table 2).

To assess the construct validity of the busyness indicators, I examined their relationship to a measure of leisure time based on ATUS reports from individuals who responded to both surveys. Figure 1 shows that leisure time consistently declines as busyness increases, providing at least some indirect evidence that these variables are measuring the concept they were intended to measure.

Figure 1. Mean Leisure Time Reported in ATUS (minutes) by CPS Busyness Indicators



The fourth set of variables shown in Table 2 reflects factors thought to affect amount of social capital. Various respondent- and household-level factors should be positively correlated with social capital—being married, being a citizen, having a job, living in a large household, living with young children, having non-family members in household, and owning a home.

I also examined a number of community-level variables that could potentially affect social integration. According to the social capital approach, structural volatility in the community weakens interpersonal ties and lowers trust and cooperation (e.g., Guest et al., 2006; Kubrin and Weitzer, 2003), so I included the annual percent changes in business establishments and in the young adult population as area-level measures of community stability. Measures of racial heterogeneity, urbanicity, population density, and crime rates also were included as indicators of community fragmentation; they were expected to reduce survey participation. Finally, I included measures of employment rates, median family income, income equality, and educational achievement as proxies for community socio-economic well-being. The social capital hypothesis suggests that declines in these factors will reduce social integration and survey response propensity.

Handling Item Nonresponse in CPS Variables

Some of the variables in the evaluated CPS dataset had missing data due to item nonresponse. The missing items included both numerical data such as median family income in the Census tract, and categorical data such as main respondent race and Hispanic origin. In general, the amount of CPS item nonresponse was small, but it varied by item. For example, item missing rates for race and educational attainment were less than one percent, about nine percent for hours worked, and almost thirteen percent for median family income. To adjust for any item nonresponse biases, missing values were imputed using a multiple imputation method described by Raghunathan, Lepkowski, and Van Hoewyk (2001) using IVEware software (Raghunathan, Solenberger, Van Hoewyk, 1998). IVEware fits sequential regression models to the values of observed and imputed data, using the variable to be imputed as the outcome variable and the others variables as predictors. The variables used in the imputation procedure are listed in Table 3. The analyses of the CPS and ATUS datasets used the imputed variables.

Table 3. Variables Included in the Imputation of CPS Data

Name of Imputed Variable	Label and Values	Number Imputed (% Imputed)
Race	White	573 (95.0)
	Black	1 (0.2)
	Asian	6 (1.0)
	Other	23 (3.8)
Hispanic	Hispanic	68 (6.8)
	Non-Hisp	924 (93.2)
Education	Less than HS	91 (14.1)
	HS only	196 (30.3)
	Some college	190 (29.4)
	BA/BS	111 (17.2)
	Post-graduate degree	58 (9.0)
Usual Hours Worked	0 – 132	8517 (8.8)
Hrs Worked Last Week	0 – 198	2035 (2.1)
Home Ownership	Rent	418 (23.0)
	Own	1403 (77.0)
Family Income	Less than $30,000	313 (18.1)
	$30,000 – $75,000	724 (42.0)
	$75,000 +	689 (39.9)
Median Family Income (Tract)	$0 - $150,000	12,362 (12.7)
% HS Diploma, 25 yrs +	49.4 – 96.8	1423 (1.5)
# of Establishments, % change	-10.3 – 22.2	1179 (1.2)
Number of Violent Crime Arrests	0.2 – 1547.4	3504 (3.6)
Covariates Used for Imputation		
Age	Household Size	Region
Sex	Family Type	Urbanicity
Citizenship	Personal Contacts Round 1	Diversity Index
Marital Status	NR in Round 1 or 2	% Young Adult Change
Employment Status	Per Capita Income in County	GINI Index

3.4.2 ATUS Variables

The dependent variable for the ATUS analyses was the interview outcome (i.e., response vs. nonresponse). The respondent, household, and community predictors of nonresponse in the ATUS interview were the same as those discussed above for the CPS analyses, with four exceptions. Three ATUS interview process variables were added (the number of call attempts made to ATUS sample members over the eight week ATUS fielding period; an indication of whether the designated ATUS respondent was the same person identified as the wave eight CPS respondent; the time of day during which the majority of ATUS call attempts were made) as well as an indicator of CPS nonresponse during rounds three through eight.[8] It was predicted that ATUS nonresponse would increase with the number of call attempts necessary to complete the ATUS interview, when the ATUS and CPS respondents were different, when the majority of call attempts were made in the afternoon or evening, and when nonresponse occurred in any of the last six CPS waves.

3.5 Results of CPS Nonresponse Analyses

3.5.1 Examination of CPS Nonresponse Across CPS Waves

As can be seen in Table 4, the completion rates for each month generally are high, ranging from about 94.5% in months one and five, to 96.4% in other waves (American Association for Public Opinion Research's [AAPOR] response rate 6). These rates are several percentage points higher than the rates reported by the BLS. The main reason for this is that the rates reported here were calculated after removing replacement households and households that were ineligible in any CPS month-in-sample. Table 4 shows the familiar finding that CPS response is worst in round one (when noncontact is highest), and again in round five when the household is

[8] The use of the overall CPS response propensity score resulting from multivariate logistic regression models run on the CPS predictors (discussed below) was evaluated as an alternative to the raw measure of response status in CPS MIS 3 – 8, but was found to be less predictive. Therefore, the raw measure was used instead. This measure of CPS nonresponse in rounds three through eight replaced the measure of CPS nonresponse in rounds one or two.

Table 4. CPS Completion Rates by Month-in-Sample[†]

Month-in-Sample	Unweighted Response Rate	Weighted Response Rate
1	94.7 (91,939)	94.6
2	96.2 (93,405)	96.2
3	96.5 (93,630)	96.4
4	96.4 (93,564)	96.4
5	94.5 (91,673)	94.4
6	95.6 (92,768)	95.6
7	95.7 (92,909)	95.7
8	96.1 (93,255)	96.1
1 – 2	93.1 (90,236)	92.9
3 – 8	87.9 (85,281)	87.7

[†] Basic CPS data from May 2001 – October 2003, excluding cases that were ever replacement households or deemed ineligible.

returning to the CPS sample for the first time in eight months. The overall response rates when combining response indicators from months one and two, and three through eight, fell to 92.9% and 87.7%, respectively.

Table 5 shows the correlation between response indicators across CPS waves, and between the main dependent measure (overall nonresponse at any time during months-in-sample three through eight) and nonresponse in either month-in-sample one or two. The coefficients shown in Table 5 reflect the strength of relationship between the response statuses across the life of the CPS panel. As expected, the largest correlations are between adjacent months-in-sample, and this is especially true for the last four CPS waves.

Table 5. Correlation (Phi Coefficient) of Response Status Across CPS Waves

	Month-in-Sample (MIS)								
MIS	1 - 2	1	2	3	4	5	6	7	8
1	.86								
2	.73	.45							
3	.40	.33	.51						
4	.35	.28	.43	.56					
5	.20	.14	.20	.26	.30				
6	.18	.11	.19	.24	.28	.54			
7	.16	.09	.17	.24	.27	.48	.62		
8	.14	.08	.14	.21	.25	.45	.56	.65	
3 – 8	.30	.24	.33	.52	.52	.65	.58	.57	.54

Note: Correlations reflect the extent to which two months-in-sample have the same response status (e.g., nonresponse). Each MIS was coded "1' if the household was a nonrespondent, "0" if it was a respondent. MIS "1 – 2" and "3 – 8" collapses response status across months-in-sample 1 and 2, and 3 through 8, respectively. MIS 1 – 2 was coded "1" if the household was a nonrespondent in either MIS 1 or 2, "0" otherwise. MIS 3 – 8 was coded "1" if the household was a nonrespondent in any of the remaining 6 MIS, "0" otherwise.

3.5.2 Predictors of CPS Response

The correlations in Table 5 suggest that nonresponse is a somewhat stable phenomenon across waves. I next examined four groups of variables for their potential impact on survey response—those associated with characteristics of the interview process; the demographic variables commonly found to influence survey response; busyness variables; and social capital variables. Table 6 displays estimates of the overall CPS completion rates in rounds three through eight as function of these variables. Both weighted and unweighted analyses were conducted and found to produce essentially the same pattern of results, so the results reported in this section are unweighted. Chi-square tests were conducted to determine whether rates of CPS participation varied by levels of these predictors.[9] Unless otherwise stated, all results reported in Table 6 were found to be significant ($p < .001$).

[9] For consistency of presentation in the table, continuous variables (e.g., population density) were transformed into quartiles and then analyzed using chi-square tests. T-tests of the untransformed continuous variables produced identical results.

Process Variables

The pattern of results for all the process variables was in the expected direction. The Northeast and West had lower response rates than other regions of the country. Response rates were lower in the winter and summer months than in spring and fall, though this finding failed to reach statistical significance. Strong effects were found for the each of the indicators of survey reluctance. Response rates in waves three through eight were substantially lower when the household was a nonrespondent in either wave one or two (52.1%) than when the household responded in both of the first two waves (90.5%). Lower nonresponse was associated with missing family income data and the need for additional personal contact attempts in the first CPS interview.

Demographic Controls

Estimates of response rates by the demographic characteristics were largely in the predicted direction. Response rates increased with age, were higher for females than males, and for White and non-Hispanic individuals than for minorities. The results for the family income variable were mixed, however. The finding that participation was highest in households with the lowest income was contrary to some previous research, though, as predicted, high income households were more likely to respond than those in the middle income category.

Busyness Variables

The relationship between CPS response status and alternative measures of labor force participation generally conformed to the busyness hypothesis. Response rates were lower for households in which all the adult members worked and when all

Table 6. Unweighted Percentage CPS Completion by Respondent Characteristics, Months 3 through 5 (and sample size) ‡*

Type of Predictor	Variable	Category	Complete	Total
Process	Region	Northeast	85.6%	22,366
		Midwest	90.3	24,480
		South	88.8	27,989
		West	86.4	22,218
	Season in which MIS 1 and 2 occurred	Winter	87.5	29,263
		Summer	87.9	32,818
		Spring or Fall	88.1	34,972
	NR in MIS 1 or 2	No	90.5	90,326
		Yes	52.1	6,727
	Family Income Missing	No	90.2	81,419
		Yes	75.6	15,634
	Personal Contacts MIS1	1	89.7	66,837
		2 – 3	84.9	20,147
		4 +	76.5	4,083
Demographic	Age	<15	73.3	30
		16-24	82.0	3,963
		25-44	85.4	33,716
		45-64	88.0	38,507
		65+	92.9	20,837
	Sex	Male	86.3	36,585
		Female	88.8	60,468
	Race	White	88.5	83,475
		Black	83.4	9,570
		Asian	84.4	1,035
		Other	85.0	2,973
	Hispanic	Hispanic	85.2	6,854
		Non-Hisp	88.1	90,199
	Family Income	Less than $30,000	90.7	18,360
		$30,000 – $75,000	86.3	41,279
		$75,000 +	88.2	37,414

‡ All results reported in this table were significant at the $p < .001$ level, with the exception of Season which was not significant ($p = .12$).

* Unless otherwise stated, the county-level data are for 1999. Data are from the U.S. Census Bureau and the Inter-University Consortium for Political and Social Research, University of Michigan.

Table 6 (continued).

Type of Predictor	Variable	Category	Complete	Total
	Hrs worked last week for main respondent	NILF or unemployed	91.7%	36,266
		Less than 35 hrs per week	84.7	17,161
		35 – 44 hrs per week	86.2	28,148
		Over 45 hrs per week	85.6	15,478
	Usual hours worked for main respondent	NILF or unemployed	91.9	35,025
		Less than 35 hrs per week	83.3	13,929
		35 – 44 hrs per week	86.8	33,352
		Over 45 hrs per week	85.1	14,747
	HH usual hours worked	All working HH members usually work over 40 hours	84.9	6,119
		Not all usually work over 40	88.1	90,934
	Adult HH workers	No adults working in HH	90.3	24,170
		At least 1 adult not working	90.1	27,880
Busyness		All adults in HH working	85.2	45,003
	Job type of main respondent	NILF or unemployed	91.8	28,177
		Executive/Profession	86.6	23,708
		Service	86.4	21,561
		Support/Production	85.8	23,607
	Urbanicity	Metropolitan	86.9	72,573
		Non-met, adjacent	90.1	10,205
		Non-met, not adjacent	90.1	10,555
		Rural	93.1	3,720
	Population Density	Less than 89	90.5	24,222
		89 – 360	90.0	24,225
		361 – 1400	87.7	24,125
		GT 1400	83.4	24,481
	Median travel time (min)†	Less than 15	90.6	21,122
		15 – 19.9	89.7	25,928
		20 – 21	88.1	24,455
		GT 21	83.5	25,548

† Data are from 2004, Bureau of Transportation Statistics

Table 6 (continued).

Type of Predictor	Variable	Category	Complete	Total
Social Capital	Marital Status	Married	89.4%	55,039
		Widowed	92.4	10,890
		Divorced	86.2	13,782
		Separated	84.8	2,292
		Never married	80.9	15,050
	Presence of Children	Yes	89.9	27,531
		No	87.1	69,522
	Presence of non-family	Only family members	87.3	66,529
		Relative in HH	90.5	3,970
		Non-family member in HH	88.9	26,553
	Household Size	One person HH	82.8	27,485
		2 – 3 person HH	89.5	46,712
		4 or more person HH	90.6	22,856
	Citizenship	Native US Citizen	88.2	88,298
		Naturalized US Citizen	84.7	4,616
		Not a US Citizen	84.2	4,139
	Employment Status	Employed	86.2	59,897
		Unemployed	83.6	2,809
		Not in Labor Force	91.1	34.347
	Home Ownership	Rent	84.4	22,057
		Own	88.9	74,996
	% change in businesses 1998-1999	Less than -.22%	88.1	23,295
		-.22 - .84%	88.2	24,744
		.85 – 1.9%	86.1	24,716
		GT 1.9%	89.2	24.298
	% Young Adult Change	-1.8% or more decline	88.1	23,995
		-1.7 – -.8% decline	87.3	24,470
		-.9 decline – .25% increase	87.3	23,880
		.26% or more increase	88.8	24,708
	Diversity Index‡	Less than 17%	90.8	23,568
		17 – 35%	89.5	24,208
		36 – 52%	87.5	24,955
		GT 52%	83.8	24,322

‡ The diversity index reports the percentage of times two randomly selected people would differ by race/ethnicity. Diversity indices between 0 and 14% reflect low diversity; indices above 60% reflect high diversity.

Table 6 (continued).

Type of Predictor	Variable	Category	Complete	Total
Social Capital	Urbanicity	Metropolitan	86.9%	72,573
		Non-met, adjacent	90.1	10,205
		Non-met, not adjacent	90.1	10,555
		Rural	93.1	3,720
	Population Density	Less than 89	90.5	24,222
		89 – 360	90.0	24,225
		361 – 1400	87.7	24,125
		GT 1400	83.4	24,481
	Per Capita Violent Crime	Less than 4.7	88.2	9,927
		4.8 – 35.9	89.0	38,704
		36 – 126	88.4	24,140
		127 +	85.9	24,282
	Unemployment Rate	Less than 3.8%	88.2	22,960
		3.8 – 4.8%	88.6	24,788
		4.9 – 6.1%	88.9	24.281
		6.2% +	85.8	25,024
	Median Family Income	Less than $27,000	88.1	24,263
		$27,000 – $34,999	88.2	24,252
		$35,000 - $44,999	86.1	24,256
		$45,000 +	89.2	24,282
	GINI Index	Less than 38.5	89.9	24,558
		38.5 – 41.1	89.2	24,307
		41.2 – 43.7	87.5	23,392
		43.8 +	84.9	24,796
	% HS Diploma, 25 yrs +	Less than 77%	87.4	24,143
		77 – 82%	88.0	24,404
		83 – 86%	88.7	24,271
		87% +	87.3	24,235

Note: The Gini Index is a commonly used measure of income inequality. It is a number between 0 and 100, where 0 corresponds to perfect income equality (i.e., everyone has the same income), and 100 corresponds to perfect inequality (i.e., where all the income is held by one person).

the employed adults in the household worked more than 40 hours per week. Measures of time stress due to commuting and traffic congestion also behaved as predicted. Median travel time, urbanicity, and density all had bivariate association with higher levels of nonresponse. However, the likelihood of CPS response actually was higher for full-time workers than part-time workers, contrary to the busyness notion.

Social Capital Variables

Consistent support was found for the effects of social capital variables. Response rates were relatively high for households in which the main respondent was employed (vs. unemployed), married, and a native-born US citizen.[10] Households with four or more people had response rates that were approximately 8% higher than single-person households. Response rates also were higher for households where there were children or other relatives present (though these results are confounded at the bivariate-level with household size), and home owners had a response rate in waves three through eight that was 4.5% higher than it was for renters.

At the community-level, growth in the percentage of young adults and business establishments in the county was associated with the highest levels of CPS participation, though the overall impact of these variables was relatively small. Larger effects were apparent for racial diversity, income inequality, and unemployment in the direction predicted by the social capital hypothesis. County-level measures of income and educational achievement showed little or no relationship to CPS response rates.

[10] Response rates for households in which the main respondent was not in the labor force (NILF) were 5% higher than those in which the main respondent was employed. However, this result appears to be driven largely by age—the majority (51%) of NILF respondents were over the age of 65, compared to 0.6% for employed persons and 14.7% for unemployed persons. Similarly, response rates for widowed individuals were about 3% higher than married individuals, but the vast majority of widows/widowers (77.2%) also were over the age of 65.

3.5.3 Multivariate Analyses of CPS Response Propensity

The crosstabulations just discussed identified patterns and potentially interesting relationships between CPS and community variables and CPS response propensity during waves three through eight. These relationships provided the baseline for subsequent multivariate analyses which control for possible confounding among the variables in the bivariate analyses. As noted earlier, predictors included in the multivariate model were selected based upon theoretical considerations discussed in Chapter 2, previous studies of CPS nonresponse reviewed in section 3.2, and the contingency tables reported in section 3.5.2 for which univariate tests indicated potential significance. Table 2 provides a list of the key process, demographic, busyness, and social capital variables examined. The model also included three interaction terms to capture the potential different effects of urbanicity, commute time, and racial diversity in different regions of the country. Preliminary multivariate analyses revealed that some variables were highly collinear with one another, and these were dropped from the multivariate model presented here.

To account for the complex stratified sampling design of CPS, the analyses were conducted in SAS-callable SUDAAN using the design PSU to calculate the appropriate standard errors for the logistic regression parameters. The data are reported for a weighted model using the survey base weights provided on the CPS dataset, and the goodness of fit of the model was assessed using the max-rescaled R-square coefficient (R^2) and Chi-square likelihood ratio test statistic.

The results of the multivariate model largely confirm those from the descriptive analyses. CPS nonresponse propensity was significantly and positively related to unit nonresponse in the first CPS interviews, item nonresponse for family income, and the number of personal contact attempts, and nonresponse propensity was higher for interviews administered in the winter. A significant interaction between urbanicity and region revealed the stronger effect of metropolitan status on nonresponse in the Midwest and South than in the Northeast and West. Among the demographic variables, being older, female, white, and non-Hispanic each were positively associated with the probability of CPS response.

One measure of busyness—household hours worked—was associated with a significant increase in nonresponse propensity. However, none of the remaining indicators of busyness were related to CPS participation. Neither the number of working adults in the household, commuting stress (commute time and population density), occupational demand, nor full household employment had significant effects on nonresponse, controlling for the other model variables. Moreover, the main respondent's hours worked variable was negatively associated with nonresponse propensity—as hours worked increased, CPS nonresponse in the last six months-in-sample actually decreased.

Substantially more support was found for the social capital hypothesis. As expected, CPS nonresponse was lower for married individuals, home owners, native-born citizens, and those living in larger households, households with young children or with non-family members than for those who were unmarried, renters, naturalized citizens, living in smaller households without children or non-family members. Nonresponse propensity also was lower in counties with greater income equality and more highly educated residents. No main effect was found for racial diversity, but the significant interaction between diversity and region revealed that nonresponse worsened as diversity increased in the Northeast, South, and West regions (with particularly negative effects of high diversity on nonresponse in the Northeast), whereas high levels of diversity in the Midwest actually were associated with lower nonresponse. Percent growth in businesses and young adult population in the county failed to reach significance in this model, but their effects were in the direction predicted by the social capital hypothesis. Two findings were contrary to social capital predictions, however: nonresponse increased with median family income and sample member employment, reversing the effects found at the bivariate level.

Table 7. CPS Propensity Model (predicting CPS nonresponse in waves 3 through 8)

Parameter	Odds Ratio	% Increase in Odds of NR	Signif.
CPS Nonresponse in MIS 1 or 2	5.133	413.3%	<.0001
# of Personal Contact Attempts in MIS1	1.071	7.1	<.0001
Family Income Missingness (No vs. Yes)	0.502	-50.2	<.0001
Region			.2096
NE vs. West	6.820	582.0	.7012
Midwest vs. West	1.792	79.2	.0233
South vs. West	2.789	178.9	.0009
Urbanicity			.6530
Metro vs. Rural	0.939	-6.1	.3920
Non-metro, adjacent vs. Rural	0.865	-13.5	.7323
Non-metro, non-adjacent vs. Rural	1.017	1.7	.7260
Urbanicity x Region			.0007
Season of CPS MIS1 interview			.0134
Winter vs. spring/fall	1.088	8.8	.0033
Summer vs. spring/fall	1.042	4.2	.1877
Respondent Age	0.973	-2.7	<.0001
Respondent Race			.0011
Black vs. White	1.152	15.2	.0014
Asian vs. White	1.297	29.7	.0311
Other vs. White	0.896	-10.4	.1989
Hispanic Origin (non-Hispanic vs. Hispanic)	0.896	-10.4	.0264
Respondent Sex (Female vs. Male)	0.903	9.7	.0012
Family Income			.1183
Low vs. High	.923	-7.7	.0091
Med. vs. High	.998	-0.2	.9114

Table 7 (continued).

Parameter	Odds Ratio	% Increase in Odds of NR	Signif.
Actual Hours Worked CPS MIS1, CPS Respondent	0.993	-0.7%	.0092
% Adults in HH Who Work	1.000	0.0	.4784
Do All Working HH Adults Work 40+ Hours? (Yes v. No)	1.161	16.1	.0010
CPS respondent job type			.3271
NILF vs. Executive/Professional	1.143	14.3	.4309
Service vs. Executive/Professional	0.856	-14.4	.1382
Support/Production vs. Executive/Professional	0.911	-8.9	.4858
Commute Time in County (min)	0.993	-0.7	.5171
Commute Time x Region			.2998
Marital Status (Married vs. Not Married)	0.917	-8.3	.0069
Presence of Young Child(ren)			<.0001
No vs. Children under 6	1.370	37.0	<.0001
Children over 6 vs. Children under 6	1.089	8.9	.0608
# of Non-family/Relatives Present	0.945	-5.5	.0114
Household size	0.973	-2.7	.0291
Citizenship			.0005
Naturalized vs. Native	1.184	18.4	.0010
Non-citizen vs. Native	0.899	-10.1	.3867
Employment Status			.0372
NILF vs. Employed	0.755	-24.5	.2151
Unemployed vs. Employed	0.628	-37.2	.0103
HH Ownership (Own vs. Rent)	0.892	-10.8	.0006
% Growth in Establishments	0.983	-1.7	.1348
% Growth in Young Adult Population	0.959	-4.1	.0621
% Diversity in County	0.999	-0.1	.6650
% Diversity x Region			.0103
% Adults with HS Diploma + in County	0.987	-1.3	.0066
Median Family Income (in tract)	1.00	0.0	.0125
Income inequality	1.046	4.6	.0132

3.5.4 Discussion of CPS Nonresponse Analyses

The findings presented in this section are consistent with previous literature on CPS nonresponse. Results of both the bivariate and multivariate analyses showed that older, white, female, and non-Hispanic individuals had the highest rates of CPS participation. There also was clear evidence that indicators of reluctance in the first two CPS waves (i.e., unit nonresponse, family income item nonresponse, number of personal contact attempts) were powerful predictors of nonresponse in subsequent waves.

The results offer relatively little support for the busyness account. Significant effects found at the bivariate level for the busyness measures largely disappeared after controlling for other factors. The multivariate analyses revealed that nonresponse did increase with overall household hours worked, but was unrelated to the other indicators of busyness—occupational demands, hours worked, and commuting times in the county. By contrast, most of the effects predicted by the social capital hypothesis were confirmed. Individual- and household-level indicators of social connectedness—marriage, presence of small children or non-family members, home ownership, and native US citizenship—produced gains in CPS participation, consistent with social capital predictions and prior CPS research. In addition, these analyses underscore the usefulness of including area-level characteristics in models of survey nonresponse. Several county-level variables not previously examined in the literature contributed significantly to CPS response propensities (cf., Johnson et al, 2006). Economic inequality, racial diversity, and lower educational attainment in the county were associated with lower rates of CPS participation, and the effects of business growth and increases in the young adult population in the county trended in this direction, as well. In the next section of this chapter, I will examine whether the pattern of results found in the CPS analyses generalize to the ATUS and discuss implications of these findings for the busyness and social capital hypotheses.

3.6 Results of ATUS Nonresponse Analyses

3.6.1 Relationship Between ATUS and CPS Nonresponse

Table 8 shows the overall response rate obtained from the 2003 ATUS dataset and the distribution of ATUS response rates based on the CPS households' month-in-sample response status. AAPOR response rate definition RR2 was used,

$RR2 = \frac{(\mathrm{I} + \mathrm{P})}{(\mathrm{I} + \mathrm{P}) + (\mathrm{R} + \mathrm{NC} + \mathrm{O}) + \mathrm{UE}}$, where the numerator is the number of complete interviews and the denominator is the number of completes plus the number of noninterviews and cases of unknown eligibility. The overall ATUS response rate was 59.9%. As with the CPS results, this rate is higher than the official BLS estimate because the rate reported here drops CPS replacement households and those that were ever CPS or ATUS ineligible. The bottom panel of Table 8 indicates that ATUS completion rates varied by CPS month-in-sample response status, with households that were never CPS nonrespondents also achieving the highest rate of ATUS participation.

Table 8. ATUS Response Rates—Overall and by CPS MIS Nonresponse

	Unweighted Response Rate	Weighted Response Rate
Overall ATUS Response Rate	59.9%	60.5%
CPS Nonrespondent (NR) Status by Month-in-Sample (MIS)		
Never a NR	62.2%	62.8%
NR in MIS 1	43.9	43.8
NR in MIS 2	41.0	40.9
NR in MIS 3	35.7	34.7
NR in MIS 4	38.8	37.1
NR in MIS 5	36.9	37.6
NR in MIS 6	28.9	29.5
NR in MIS 7	34.3	35.0
NR in MIS 8[†]	n/a	n/a
NR in MIS 1 or 2	44.1	43.8
NR in MIS 3 – 8	38.3	38.6

[†]CPS nonrespondents in MIS 8 are not eligible to be sampled for ATUS.

3.6.2 Bivariate Analyses of Predictors of ATUS Participation

Table 9 displays estimates of the overall ATUS completion rates as a function of the process, demographic, busyness, and social capital variables. The estimates are for unweighted data and were significant unless otherwise indicated.

<u>Process Variables</u>

As in the CPS analyses, the largest effects on ATUS nonresponse were found for the number of call attempts, item missingness on the CPS family income question, and unit nonresponse in CPS months-in-sample three through eight, all of which were negatively associated with ATUS participation. Additional significant effects were found for region, season of interview, and timing of interview attempt. Nonresponse was highest in the Northeast, in the summer months, and when the majority of calls were made in the afternoon, lowest in the Midwest, in the winter, and when call attempts were concentrated in the morning or equally distributed throughout the day.

In addition, ATUS participation improved significantly when the designated respondent also had been the CPS respondent in wave eight.

Demographic Controls

Consistent with findings from previous nonresponse studies, ATUS participation was significantly higher for females than males, whites than persons of other races, non-Hispanics than Hispanics, and for households with greater family income. Age had a curvilinear effect on ATUS participation: nonresponse was highest for young (< 34) and old (65+) sample members, while individuals age 35 – 64 exhibited participation levels that increased with age.

Busyness Variables

The busyness hypothesis received mixed support from the bivariate analyses of discretionary time indicators. As predicted, ATUS nonresponse was higher in households in which all employed adults worked more than forty hours per week than in households that included adults who worked forty hours or fewer. In addition, nonresponse increased as commuting times and population density increased and was lowest in large urban areas. Results for the remaining measures of discretionary time did not support the busyness hypothesis, however. For example, the hours worked variable did not have a consistent relationship to ATUS participation: full-time workers had nearly identical rates of ATUS participation as those who were not employed, and both groups had substantially lower levels of participation than individuals who worked forty-five hours or more per week. Similarly, participation rates were highest for households in which all adults worked and for individuals with the greatest occupational demands, contrary to busyness predictions.

Social Capital Variables

The bivariate analyses lend more consistent support to the social capital account of ATUS participation. Being married or a native-born US citizen was positively related to ATUS participation. Participation also was higher for ATUS sample members who were employed than for those not in the labor force (unemployed individuals had the highest rate of participation, but this result was based on a very small number of cases). At the household level, as predicted, nonresponse was lower for multiple-person than single-person households and for owners than renters. The presence of children variable failed to reach significance, but the distribution of response statuses took the expected shape: households with young children had lower nonresponse than households without children or with only older children. Finally, contrary to social capital predictions, households where non-family members were present had lower participation rates than households with family only.

Turning to the community variables, ATUS participation significantly increased as area income and educational attainment rose, and when business establishment growth was high; these findings are consistent with the social capital hypothesis. Also consistent were the findings that nonresponse increased significantly as levels of a community racial diversity, economic inequality, violent crime, unemployment, population density, and urbanicity increased. The one community factor that failed to conform to social capital predictions was the percentage change in young adults. I predicted that nonresponse would be higher in areas that had experienced declines in young adult population and lower in areas with stable or growing levels of young adults, but this difference failed to reach significance.

Table 9. Unweighted Percent ATUS Completion (and sample size) by Respondent Characteristics.

Type of Predictor	Variable	Category	Complete	Total
Process	Region	Northeast	61.3%	5,607
		Midwest	67.0	6,362
		South	62.2	8,837
		West	63.8	4,972
	Season in which last ATUS attempt was made	Winter	65.6	8,179
		Summer	61.3	6,869
		Spring or Fall	63.3	10,730
	NR in CPS MIS 3 - 8	No	65.1	23,882
		Yes	43.4	1,896
	Family Income Missing	No	65.3	22,553
		Yes	50.8	3,225
	Time of most frequent ATUS call attempt	Morning	67.0	13,366
		Afternoon	51.5	5,831
		Evening	60.5	3,482
		Equally distributed	74.4	3,099
	Is ATUS respondent same as CPS?	Yes, same person	65.4	15,068
		No	60.9	10,710
	ATUS call attempts	1	86.4	6,702
		2 – 4	76.4	6,338
		5 – 13	62.5	6,452
		14 or more	27.2	6,286
Demographic	Age	15-24	62.8	2,582
		25-34	61.2	3,712
		35-44	63.2	5,788
		45-54	65.2	5,060
		55-64	68.8	3,504
		65+	60.6	5,132
	Sex	Male	62.1	11,533
		Female	64.7	14,245
	Race	White	65.3	21,478
		Black	54.3	3,368
		Asian	62.8	207
		Other	55.3	725

Table 9 (continued).

Type of Predictor	Variable	Category	Complete	Total
Demographic	Hispanic Origin	Hispanic	60.5%	2,606
		Non-Hispanic	63.9	23,172
	Family Income	Less than $30,000	58.3	4,508
		$30,000 – $75,000	61.8	10,302
		$75,000 +	67.2	10,968
Busyness	ATUS respondent's hours worked	NILF or unemployed	61.9	10,716
		Less than 35 hrs per week	69.1	3,710
		35 – 44 hrs per week	61.2	7,668
		Over 45 hrs per week	67.5	3,684
	Adult HH workers	No adults working in HH	60.1	5,507
		At least 1 (but not all) adults working	63.4	8,728
		All adults in HH working	65.2	11,543
	HH usual hours worked[†]	All working HH members usually work over 40 hours	66.2	1,403
		Not all usually work over 40 hours	63.4	24,357
	Job Type	NILF or unemployed	61.1	9,399
		Executive/Profession	69.4	6,053
		Service	62.9	6,603
		Construction/Production	61.3	3,723
	Population Density	Less than 89	66.4	4,802
		89 – 360	66.2	6,545
		361 – 1400	63.7	7,720
		1401 +	58.9	7,160
	Urbanicity	Metropolitan	62.7	20,712
		Non-met, adjacent	67.5	2,625
		Non-met, not adjacent	66.0	1,799
		Rural	65.3	642
	Median travel time (min)	Less than 15	68.6	4,068
		15 – 19.9	64.5	6,603
		20 – 21	64.4	6,978
		GT 21	59.3	8,129

[†]Significant at the $p < .05$ level.

Table 9 (continued).

Type of Predictor	Variable	Category	Complete	Total
	Marital Status	Unmarried	61.8%	11,457
		Married	64.9	14,321
	Presence of Children*	No children	63.3	16,747
		Children under 6	64.7	3,785
		Children over 6	63.3	5,246
	Presence of non-family	Family members only	64.6	17,826
		Non-family in HH	61.1	7,952
	Household Size‡	One person HH	61.7	5,334
		2 – 3 person HH	64.2	12,322
		4 or more person HH	63.8	8,122
	Citizenship	Native US Citizen	64.2	22,991
		Naturalized US Citizen	56.1	1,341
		Not a US Citizen	59.5	1,446
	Employment Status	Employed	64.9	15,704
		Unemployed	66.4	675
		Not in Labor Force	61.1	9,399
Social Capital	Home Ownership	Rent	58.6	5,356
		Own	64.8	20,422
	% change in businesses 1998-1999†	Less than -.22%	64.3	6,051
		-.22 - .84%	62.5	6,720
		.85 – 1.9%	62.7	6,733
		GT 1.9%	64.7	6,274
	% Young Adult Change*	-1.8% or more decline	63.7	6,026
		-1.7 – -.8% decline	63.5	6,033
		-.9 decline – +.25% increase	62.5	6,835
		.26% or more increase	64.4	6,884
	Population Density	Less than 89	66.4	4,802
		89 – 360	66.2	6,545
		361 – 1400	63.7	7,720
		1401 +	58.9	7,160
	Urbanicity	Metropolitan	62.7	20,712
		Non-met, adjacent	67.5	2,625
		Non-met, not adjacent	66.0	1,799
		Rural	65.3	642

† Significant at the $p < .05$ level. ‡ Significant at the $p < .01$ level * Not significant.

Table 9 (continued).

Type of Predictor	Variable	Category	Complete	Total
Social Capital	Diversity Index	Less than 17%	68.2%	4,516
		17 – 35%	65.5	6,659
		36 – 52%	63.7	6,905
		GT 52%	58.9	7,698
	Per Capita Violent Crime‡	Less than 4.7	64.6	5,447
		4.8 – 35.9	64.5	5,871
		36 – 126	63.5	7,129
		127 +	61.9	7,331
	Median Family Income	Less than $27,000	61.5	6,504
		$27,000 – $34,999	63.2	6,273
		$35,000 - $44,999	62.9	6,336
		$45,000 +	66.3	6,665
	Unemployment Rate	Less than 3.8%	66.4	5,673
		3.8 – 4.8%	64.2	6,555
		4.9 – 6.1%	63.8	6,404
		6.2% +	60.4	7,146
	GINI Index	Less than 38.5	67.0	5,699
		38.5 – 41.1	65.8	5,612
		41.2 – 43.7	64.2	6,490
		43.8 +	58.8	7,977
	% HS Diploma in county	Less than 77%	60.3	7,552
		77 – 82%	62.9	6,618
		83 – 86%	64.6	6,275
		87% +	67.7	5,333

† Significant at the $p < .05$ level. ‡ Significant at the $p < .01$ level. * Not significant.

3.6.3 Multivariate Analyses of ATUS Response Propensity

As in the CPS analyses, a logistic regression model predicting overall nonresponse was fit to examine the influence of the previously specified ATUS survey process, demographic, busyness, and social capital factors. Weighted logistic regressions were performed using SAS-callable SUDAAN to allow for the different probabilities of selection and the effects of stratification and clustering, and model fit was assessed using max-rescaled R-square coefficients and Chi-square likelihood ratio test statistics.

Table 10 presents the results of this model. As before, I report effect sizes and their associated significance levels, and the results are grouped by panels separating process, demographic, busyness, and social capital variables. The results in the first panel confirm the effects of ATUS process variables found in the bivariate analyses and represent some of the largest effects in the model. Nonresponse was positively associated with the number of call attempts, a change in designated respondents between wave eight CPS and ATUS, frequent afternoon contact attempts, CPS unit nonresponse in rounds three through eight, and item missingness on the CPS family income question. Season of ATUS interview and region became non-significant at the multivariate level.

The second panel shows the effects of the demographic variables, including an age-squared variable added because the bivariate analysis suggested the effects of age were non-linear. The results are largely consistent with findings from the bivariate analyses. Nonresponse was negatively associated with family income, and was higher for Hispanics than non-Hispanics, and for individuals of "some other race" than whites (whites did not differ from blacks or Asians, however). Age was significantly related to the probability of nonresponse (older individuals had higher rates of participation, on average), but the significant effect for the age-squared variable revealed that nonresponse declined with age for sample members' age sixty-four or younger, but increased with age for individuals over the age of sixty-five. Gender did not reach significance after controlling for other model variables.

The results in the third panel offer no support for the busyness hypothesis. Contrary to busyness predictions, the designated respondents' hours worked, household hours worked, and county commute times were unrelated to ATUS participation, and participation actually increased as the percentage of adults in the household who worked increased. Similarly, participation rates were higher for those with the most demanding jobs (executives/professionals) than they were for those with service or support jobs and those not in the labor force, contrary to busyness predictions. Neither population density nor urbanicity was related to ATUS participation (fourth panel in Table 10) once the effects of the sample design and other model variables were taken into account, in contrast to both the busyness and social capital hypotheses.

In fact, as can be seen in the last panel of Table 10, the social capital hypothesis also only received weak support. Of the twelve predicted effects, seven were nonsignificant (household size, citizenship, home ownership, business growth, young adult population growth, county educational attainment, and income inequality) and two produced results in the opposite from the predicted direction (nonresponse increased with the presence of non-family members and children over the age of six in the household, respectively). Marital status was significantly related to probability of ATUS nonresponse in the direction predicted by the social capital hypothesis—sample members who were married were less likely to be nonrespondents than those who were unmarried. ATUS participation also was greater in communities with less racial diversity and higher median family income, consistent with social capital predictions.

Table 10. Final ATUS Propensity Model (predicting ATUS nonresponse)

Parameter	Odds Ratio	% Increase in Odds of NR	Signif.
# of call attempts	1.149	14.9%	<.0001
ATUS respondent same as CPS (Yes vs. No)	0.737	-36.3	<.0001
Time of most frequent contact			<.0001
AM vs. equally distributed	1.006	0.6	.9600
Afternoon vs. equally distributed	1.381	38.1	.0092
Evening vs. equally distributed	0.930	-7.0	.5823
Season of ATUS interview			.1659
Winter vs. spring/fall	0.975	-2.5	.5350
Summer vs. spring/fall	1.057	5.7	.1896
Region			.0750
Midwest vs. NE	0.557	-44.3	.0283
South vs. NE	0.550	-45.0	.0180
West vs. NE	.494	-50.6	.0201
CPS nonresponse MIS 3 – 8 (Yes vs. No)	1.864	86.4	<.0001
Item NR to CPS Family Income Q (Yes vs. No)	1.792	79.2	<.0001
Respondent Age	0.978	-2.2	.0410
Age squared	1.000	0.0	<.0001
Respondent Sex (Female vs. Male)	0.956	-4.4	.2331
Respondent Race			.0500
Black vs. White	1.011	1.1	.8459
Asian vs. White	0.857	-14.3	.4177
Other vs. White	1.333	33.3	.0089
Hispanic Origin (non-Hispanic vs. Hispanic)	.718	-28.2	<.0001
Family Income			<.0001
Low vs. High	1.335	33.5	<.0001
Med. vs. High	1.030	3.0	.5002

Table 10 (continued)

Parameter	Odds Ratio	% Increase in Odds of NR	Signif.
Hours Worked	1.002	0.2	.2664
% Adults in HH Who Work	0.998	-0.2	.0077
Do All Working HH Adults Work 40+ Hours? (Yes v. No)	1.152	15.2	.0988
Commute Time in County (min)	0.981	-1.9	.1624
Commute Time x Region			.0629
ATUS Respondent Job Type			<.0001
NILF vs. Executive/Professional	1.784	78.4	<.0001
Service vs. Executive/Professional	1.279	27.9	<.0001
Support/Production vs. Executive/Professional	1.228	22.8	.0010
Population Density	1.000	0.0	.2601
Urbanicity			.0757
Metro vs. Rural	0.864	-13.6	.1835
Non-metro, adjacent vs. Rural	0.759	-24.1	.0189
Non-metro, non-adjacent vs. Rural	0.864	-13.6	.1828
Household size	1.020	2.0	.2453
Marital Status (Married vs. Not Married)	.906	-9.4	.0150
Citizenship			.3329
Naturalized vs. Native	1.084	8.4	.3354
Non-citizen vs. Native	0.925	-7.5	.3586
# of non-family/relatives present	1.100	10.0	.0059
Presence of young child(ren)			.0253
No vs. Children over 6	.963	-3.7	.5325
Children under 6 vs. Children over 6	1.115	11.5	.0740
HH Ownership (Own vs. Rent)	0.966	-3.4	.4856
% Growth in Establishments	1.008	0.8	.4291
% Growth in Young Adult Population	1.004	0.4	.8010
% Diversity in County	1.006	0.6	.0287
% Diversity x Region			.0447
% Adults with HS Diploma + in County	0.997	-0.3	.4404
Median Family Income (in tract)	0.999	-0.1	.0275
Income inequality	1.000	0.0	.9992

3.6.4 Discussion of ATUS Nonresponse Analyses

The results reported in this section on the demographic correlates of ATUS nonresponse are consistent with previous studies by Abraham et al. (2006) and O'Neill and Dixon (2005). As in those studies, the results showed that being non-Hispanic, older, and having higher levels of family income is associated with increases in ATUS participation. The analyses also highlight the importance of several indicators of reluctance and survey process. Sample members were much more likely to be nonrespondents if they skipped the CPS family income question (odds ratio = 1.79), had been a CPS nonrespondent (odds ratio = 1.86), or were not the respondent in the last CPS interview (odds ratio = 1.36). ATUS nonresponse propensity increased as function of the number of call attempts and of the timing of those calls, as well.

The absence of findings supporting the busyness account of ATUS participation also is consistent with results reported in Abraham et al. (2006). Key busyness predictors of respondent hours worked, household hours worked, commuting durations, and population density all were unrelated to nonresponse, and ATUS participation actually increased as the percentage of adults in the household who worked increased and when the occupational demands of the sample members' job were high.

Despite strong indications at the bivariate level that ATUS nonresponse was related to social capital variables, the results of the multivariate social capital model largely failed to find the predicted effects. This is contrary to the findings of Abraham et al. (2006) who conducted a comparison of busyness and 'social integration' explanations of ATUS nonresponse and concluded that there was significant and consistent support for the role of social integration variables. It is worth noting that both this study and the one by Abraham and colleagues found a similar positive effect on nonresponse of marriage. Moreover, Abraham and colleagues and my study also found a negative effect of the presence of non-family members in the household, but Abraham et al. cited this as supportive of their 'social integration' hypothesis. As conceived in the present analyses, the presence of additional people in the household (family or non-family members) would increase social network opportunities and thereby produce gains in survey participation. This is exactly what was found in my earlier CPS analyses, but the finding was not replicated in ATUS models. Abraham et al. (2006) suggest that households that include non-family members may be more transient (and to have less well-integrated social networks). They attributed the negative effects of this variable to low contact rates in those households. Results from additional analyses on the present dataset (not presented here) support this conclusion—relatively high rates of non-contact (but not refusal) were found in the largest households and in households that include non-family members.

3.7 Effects of Excluding Cases That Have a High Probability of Nonresponse

Having examined the correlates of nonresponse propensity in the CPS and ATUS, I next assessed whether survey estimates would be affected had cases with high probabilities of nonresponse been excluded. Survey organizations often devote considerable resources to bringing hard-to-reach or reluctant sample members into the respondent pool under the assumption that incorporating them into estimates reduces nonresponse bias. One way to evaluate this assumption is to compare survey estimates obtained from the full sample to those derived from a truncated sample in which difficult (or high nonresponse propensity) cases are omitted (e.g., Curtin et al., 2000; Olson, 2006). Significant differences between the estimates indicate that bringing in the most difficult cases affects the nonresponse bias properties of the statistic.

Nonresponse propensity scores obtained from the multivariate models reported earlier served as the basis for the present analyses. For the CPS analyses, a single estimate of overall nonresponse propensity was obtained for each household member using the logistic regression model reported in section 3.5.3, and respondents then were grouped into propensity quintiles based on their predicted probabilities of nonresponse. ATUS respondents similarly were grouped into nonresponse propensity quintiles based on the results of the propensity model reported in section 3.6.3. I then conducted a variety of simulations comparing weighted survey results from the total sample to those based on a truncated sample in which respondents in the highest nonresponse propensity quintile were omitted. For the CPS, I examined means for the total and truncated samples for one hundred variables representing a range of demographic and labor force items.[11] For the ATUS, I investigated differences in mean activity duration between the full and truncated samples for forty different activities. Significant differences between the truncated and full samples were found in eighty-five percent of the CPS comparisons and forty-one percent of the ATUS comparisons.

[11] Several variables were more appropriately analyzed at the household-level than the person-level (e.g., home ownership, household hours worked, etc.), so for these variables nonresponse propensity quintiles were reformed based on the household's predicted probability of nonresponse.

Table 11. Effects of Excluding High CPS Nonresponse Propensity Cases from CPS Estimates Related to Busyness

Statistic	Estimate: Full Sample (1)	Estimate: Truncated Sample (2)	Diff (3)	Rel Diff (4)	High NR Prop. Group $\bar{X}$ (5)	T-value (6)
Retired	16.26%	18.14%	1.88%	11.56%	6.12%	51.41
No working adults	24.94	26.97	2.03	8.14	17.28	28.49
PT worker	18.53	19.49	0.96	5.18	14.31	18.16
Multiple job holder	6.84	7.08	0.24	3.51	5.78	6.80
Hours worked	38.44	38.21	-0.23	-0.60	39.50	-11.66
Exec/Professional	31.29	31.05	-0.24	-0.77	32.35	-3.81
Urban resident	80.13	77.42	-2.71	-3.38	90.39	-41.52
All HH adults work	45.36	41.10	-4.26	-9.39	61.46	-52.51
All work GT 40 hrs	6.13	4.96	-1.17	-19.09	10.53	-29.56

Tables 11 and 12 show the effects of removing the highest CPS nonresponse propensity quintile from CPS estimates related to busyness and social capital. The full and truncated sample means are given in columns 1 and 2, respectively. Column 3 shows the difference between the full and truncated sample means, and column 4 expresses this difference as the percent change in the full-sample estimate that occurred when the highest nonresponse propensity group was excluded. Results in the upper panel of both tables indicate statistics that would be over-estimated if the high nonresponse propensity quintile was excluded; results in the lower panel indicated statistics that would be underestimated using the truncated sample. The means for the high nonresponse propensity quintile are given in column 5; column 6 shows t-values resulting from tests examining whether this group differed significantly from the remainder of the sample on the statistic of interest. All of the effects reported in the tables were statistically significant ($p < .001$).

Table 12. Effects of Excluding High CPS Nonresponse Propensity Cases from CPS Estimates Related to Social Capital

Statistic	Estimate: Full Sample (1)	Estimate: Truncated Sample (2)	Diff (3)	Rel Diff (4)	High NR Prop. Group $\overline{X}$ (5)	T-value (6)
Rural resident	2.47	2.96	0.49	19.84	0.63	18.87
HH Size	2.50	2.64	0.14	5.60	1.99	57.89
Married	58.35	60.51	2.16	3.70	46.70	44.24
In school	51.19	53.06	1.87	3.65	44.03	11.57
Young child in HH	18.43	18.97	0.54	2.93	14.93	12.40
Native citizen	89.05	90.82	1.77	1.99	78.79	66.89
No child in HH	55.25	54.14	-1.11	-2.01	62.46	-19.77
Divorced	8.92	8.69	-0.23	-2.58	10.14	-7.98
Non-family in HH	3.21	2.98	-0.23	-7.17	4.59	-15.66
Unemployed (looking)	2.97	2.70	-0.27	-9.09	4.41	-15.81
Black	10.88	9.42	-1.46	-13.42	19.32	-55.01
Non-citizen	5.43	4.49	-0.94	-17.31	10.86	-48.61
Renter	23.04	18.82	-4.22	-18.32	39.42	-63.53

Both tables suggest that omitting high nonresponse propensity cases may introduce bias in CPS estimates, though the magnitude of the effects varied by statistic. For example, Table 11 reveals that estimates of *hours worked* and *executive/professionals* were only slightly lower in the truncated sample than the total sample, whereas the truncated sample produced fairly sizeable increases (as expressed in the relative difference measure) in estimates of *part-time workers* and the *retired*. Moreover, the results in Table 11 are largely consistent with expectations based on the busyness hypothesis, despite the fact that most busyness indicators failed to contribute to the underlying propensity model (section 3.5.3). Similarly, with the exception of the *non-family in household* estimate, the results in Table 12 are consistent with social capital expectations. CPS respondents in the high nonresponse propensity quintile appear to be busier and less socially integrated than the rest of the respondent pool, a result that echoes differences found between respondents and nonrespondents in the previous bivariate analyses (section 3.5.2).

The pattern was not so clear for the ATUS. Table 13 shows the effects of excluding high nonresponse propensity individuals from estimates of the average durations of activities reported in ATUS. If respondents in the high nonresponse propensity quintile were busier than other respondents, then we would expect that excluding these cases would lead to decreases in estimates of time at work and increases in estimates of time spent in leisure activities. In fact, the opposite occurred: high nonresponse propensity respondents reported significantly less time commuting to their job and working, and more time sleeping, relaxing, and watching TV. Thus, there was evidence of bias, but not in the direction predicted by the busyness hypothesis.

There were indications of bias stemming from low social capital, however. Respondents in the high nonresponse propensity quintile reported significantly less time socializing on the phone and caring for their own children than lower nonresponse propensity respondents; smaller, non-significant effects for reports of caring for non-household children, general socializing, and participation in religious activities also were in the direction expected by the social capital hypothesis. By contrast, ATUS respondents in the high nonresponse propensity quintile reported significantly more time spent in education, volunteer, and civic activities than other respondents, contrary to social capital predictions. It is worth noting that respondents who engage in these activities also may be busier or harder to reach, so these results highlight the potential relationship between busyness and social capital. In general, however, the effects reported in Table 13 are small—the full and truncated sample mean durations differ by less than a minute, on average—and indicate relatively little nonresponse bias due to the exclusion of high nonresponse propensity ATUS respondents.

Table 13. Effects of Excluding High ATUS Nonresponse Propensity Cases from ATUS Activity Duration Estimates (in minutes)

Activity	Estimate: Full Sample (1)	Estimate: Truncated Sample (2)	Diff (3)	Rel Diff (4)	High NR Prop. Group $\bar{X}$ (5)	T-value (6)	P-value (7)
Telephone call-social	5.24	5.87	0.63	12.02%	4.60	4.46	.0001
Care for HH child	23.73	24.88	1.15	4.85	19.74	3.98	.0001
Work (all)	198.34	203.34	5.00	2.52	181.05	4.65	.0001
Care for non-HH child	11.63	11.92	0.29	2.49	10.63	1.34	.1811
Housework	39.06	39.90	0.84	2.15	36.16	2.51	.0119
Commuting	16.85	17.18	0.33	1.96	15.71	2.20	.0276
Travel (all)	76.46	77.05	0.59	0.77	74.43	1.71	.0869
Socializing	38.90	39.06	0.16	0.41	38.34	0.48	.6341
Religion	8.45	8.47	0.02	0.24	8.36	0.17	.8629
Education	3.62	3.60	-0.02	-0.55	3.71	-4.20	.0001
Sleep	508.05	504.74	-3.31	-0.65	519.50	-6.16	.0001
Television	154.49	152.49	-2.00	-1.29	161.43	-2.97	.0029
Relaxing	223.06	219.66	-3.40	-1.52	234.83	-4.23	.0001
Volunteering	9.37	9.14	-0.23	-2.45	10.16	-1.10	.2728
Civic obligations	0.30	0.24	-0.06	-20.00	0.53	-2.28	.0228

To this point, the analyses have been based on respondent means. However, the relationship between nonresponse propensity and bias may manifest itself differently for different functions of a variable. The next analyses examine how the relationships between variables change if we exclude high nonresponse propensity cases. To illustrate these effects, I conducted a series of simple regression models, first using the entire sample and then comparing coefficients from these models to coefficients from the same models run on truncated samples. The first truncated sample excluded only the highest nonresponse propensity quintile (the "eighty percent sample”); the second excluded the two highest nonresponse propensity quintiles (the “sixty percent sample”). The predictors used in each model were respondents’ level of educational attainment, hours worked, and marital status. Regression models were run with the SAS *surveylogistic* procedure to account for the appropriate survey design features.

The upper panel of Table 14 shows the results of CPS analyses that regressed two different earnings measures on the model predictors. The lower panel shows similar results for ATUS work, leisure, and household activity durations. Column 1 gives regression coefficients from models run using the full sample. Column 2 gives model coefficients when omitting the highest nonresponse propensity group, and reports significance levels for tests of equality between the twenty percent group (i.e., the highest nonresponse propensity quintile) and the eighty percent group (the truncated sample). Similarly, column 3 reports coefficients from models run on the sixty-percent sample and significance levels for tests of equality between that group and the excluded sample (i.e., the two highest nonresponse propensity quintiles).

There are several observations that can be made about the results in Table 13. First, the three predictors are significantly related to the dependent measure in each model run using the full sample, and this continues to be true for all but one model run on the truncated samples. The one exception is the last model reported in the table, in which the *hours worked* variable becomes nonsignificant when high nonresponse propensity groups are omitted (in both columns 2 and 3). Second, significant differences were found between the truncated sample estimates and those of the excluded group in twenty-one of the thirty contrasts examined. The magnitude of the bias differs for different regression coefficients, but generally is not large. The largest effects were found for marital status on estimates CPS earnings, reflecting the lower earnings of unmarried respondents in the high nonresponse propensity groups.

Table 14. Effects of Excluding High Nonresponse Propensity Cases on the Relationship Between Variables

		Regression Coefficients		
		Full Sample (1)	Excluding the Highest NR Prop Group (80% sample) (2)	Excluding the 2 Highest NR Prop Groups (60% Sample) (3)
CPS				
Hourly Earnings	**Low education (Yes=1, No=0)**	-1.48	-1.37***	-1.05***
	Hours worked (for CPS ref week)	0.29	0.28**	0.27***
	Married (Yes=1, No=0)	3.91	4.22***	4.52***
Annual Earnings	**Low education (Yes=1, No=0)**	-19,977	-19,943	-20,559
	Hours worked (for CPS ref week)	1,200	1,158***	1,147**
	Married (Yes=1, No=0)	12,422	14,205***	14,297
ATUS				
Work	**Low education (Yes=1, No=0)**	14.8	12.6	16.7
	Hours worked (for CPS ref week)	6.9	6.8***	6.7***
	Married (Yes=1, No=0)	36.3	40.7***	44.8***
Leisure	**Low education (Yes=1, No=0)**	210.9	204.6***	184.5***
	Hours worked (for CPS ref week)	0.6	0.8***	1.2***
	Married (Yes=1, No=0)	157.7	155.8	142.0***
HH Activities	**Low education (Yes=1, No=0)**	59.7	59.4	52.2***
	Hours worked (for CPS ref week)	-0.2	-0.1**	0.0***
	Married (Yes=1, No=0)	114.7	116.2	117.0

***$p < .001$; **$p < .01$; *$p < .05$.

In terms of the relative change in coefficients, the smallest effects were found for ATUS durations, where the truncated samples typically produced estimates that differed from the full sample estimates by only a few seconds to a few minutes. Finally, not surprisingly, the estimates tend to get worse (i.e., become more biased) when more of the sample is excluded, but the pattern of changes apparent when omitting the highest nonresponse propensity cases remains the same when further restricting the sample.

3.8 Conclusions

The purpose of this chapter was to determine the person, household, area, and survey process characteristics that affect nonresponse in the CPS and ATUS. The analyses take advantage of the information available in these surveys about both respondents and nonrespondents. The results suggest that each causal component of nonresponse makes contributions to the final disposition of the sample unit.

A secondary purpose of this chapter was to assess the proper placement of these various causes within two alternative conceptual frameworks for survey participation—one that assumes that nonresponse arises mainly because of lack of discretionary time or a subjective feeling of time pressure (busyness), and one that assumes that nonresponse arises when individuals are poorly integrated in their social environment (social capital). The present analyses revealed little support for the busyness hypothesis. Results from both surveys largely failed to conform to predicted effects of busyness indicators, and often reflected a negative relationship between discretionary time and nonresponse. The findings offer mixed support for the social capital hypothesis. In the CPS analyses, the great majority of social capital variables had effects in the predicted direction. The finding that community-level indicators of social capital can have a measurable impact on CPS nonresponse marks a significant contribution to the existing literature on CPS nonresponse. By contrast, social capital variables in the ATUS analyses resulted in fewer significant findings in the expected direction. This finding may reflect the inherent difficulties of operationalizing social capital (particularly in selecting good indicators for area-level effects), but also may say something about the power of social capital predictors of nonresponse in surveys in which participation rates are already fairly low and likely dominated by a number of more proximal factors.

The discrepancy between the CPS and ATUS results may be accounted for by the different administration procedures of the two surveys. The CPS conducts at least two in-person interviews and the rest by telephone over the sixteen month cycle, whereas the ATUS is conducted only by telephone and is a one-time survey (though it may be viewed by respondents as a delayed extension of the CPS). The repeated and predictable interactions with CPS interviewers (sometimes face-to-face) may make more salient social norms related to social capital. The two surveys also have sharply different response rates, so it is possible that the effects of social capital variables are swamped by other sources of variance (e.g., the role of the interviewer and the nature of the interaction between interviewer and respondent). From an analytic perspective, it is also possible that I simply have more power to detect small effects in the CPS dataset (which has almost four times as many cases as the ATUS dataset). Whatever the reasons for the different pattern of results for the two surveys with respect to social capital, models for both surveys were able to account for a fairly substantial proportion of variance between respondents and nonrespondents, and in fact appear to fit the data significantly better than other models reported in the literature. In both instances, further analyses could be conducted to identify differences between refusals and noncontacts, and to include additional components of nonresponse (e.g., social psychological attributes of respondents, interviewer effects) in multivariate propensity models.

A third objective of this chapter was to examine how survey results changed when high nonresponse propensity cases were excluded from the respondent pool. Unlike previous research (e.g., Curtin et al., 2000; Keeter et al., 2000; 2006), I found evidence suggesting that omitting high nonresponse propensity cases may introduce bias in estimates. In both the CPS and ATUS, removing high nonresponse propensity cases produced significant changes in a variety of marginal (i.e., mean) estimates and estimates of the associations between variables (i.e., regression coefficients), though the frequency and relative magnitude of the effect was greater in the CPS than ATUS.

IV. Data Quality in the Current Population Survey and the American Time Use Survey

4.1 Introduction

Inaccurate or incomplete responses from survey participants reduce survey data quality. Differences between observed and true values (measurement error) can lead to bias and variance in a survey estimate and significantly affect statistical inference (Biemer et al., 1991; Lyberg et al., 1997). Literature on the causes of response error focuses on the cognitive processes of respondents as they attempt to answer survey questions, and how these processes interact with characteristics of the interviewer, questionnaire, and data collection method. Tourangeau (1984; 2000) identifies four major components in the response process—comprehension of the survey question, retrieval of relevant information from memory, integrating information to form a judgment or estimate, and reporting that response. Each component of the response process can be affected by characteristics of the survey design (e.g., question wording, reference period, interview mode) and the respondent (e.g., knowledge, ability, and motivation), and these attributes can interact to produce response errors (see, e.g., Sudman, Bradburn, and Schwarz, 1996, and Tourangeau, Rips, and Rasinski, 2000, for reviews).

For example, respondent rule decisions about whether to accept proxy reporting (i.e., reporting about other household members) can affect response error. Proxy respondents may not have the requisite knowledge to be able answer questions about other household members accurately (Tourangeau et al., 2000). Moreover, there is empirical evidence to suggest that proxy respondents are more likely to draw on general knowledge about the other person or their 'usual' behavior rather than on pertinent situational variation. This overreliance on dispositional information can make proxy reports more consistent within a given survey or across rounds of a panel survey, but not necessarily more accurate (Schwarz and Wellens, 1994).

Even when respondents have the necessary factual knowledge, they may shortcut the cognitive processes needed to generate an optimal answer and instead settle for a merely satisfactory response in order to reduce the effort of accurate reporting. This is the notion of survey *satisficing* developed by Krosnick and Alwin (1987) (see also Krosnick, 1999; and Cannell et al., 1981 for similar ideas). This satisficing tendency is strongest when certain predisposing features of the respondent (e.g., limited cognitive ability) and the survey design (e.g., a burdensome mode of survey administration) interact (Holbrook et al., 2003).

In longitudinal or panel surveys, random variation in respondents' answers to the same question across waves may produce response errors, as well. For example, Kalton, McMillen, and Kasprzyk (1986) found changes in occupation and industry status across interviews in the Survey of Income and Program Participation (SIPP) that reflected respondent misclassification errors, as well as changes in race and sex status that may have reflected interviewer (or possibly processing) errors. Panel surveys also are subject to the effects of time-in-sample bias, in which responses to questions repeated in later survey rounds may be influenced by those given in earlier rounds. Empirical studies of time-in-sample bias have found that reported frequency of criminal victimization, voter turnout, spending, and illicit drug use tend to decline with the number of previous rounds (e.g., Johnson et al., 1998; Kasprzyk, 2005; Neter and Waksberg, 1964; Pennell and Lepkowski, 1992).

4.1.1 Evaluating Measurement Error

A variety of techniques have been used to estimate measurement error. To study measurement bias, observed values from a survey can be compared to "true" values obtained from matched administrative records or aggregate estimates derived from independent, external sources (e.g., Cash and Moss, 1972; Coder and Scoon-Rogers, 1995; Marquis and Moore, 1990; Olson, 2006). Record-check and benchmark data are themselves subject to coverage, nonresponse, and measurement errors, however, and in practice it is often difficult or impossible to obtain the external data necessary for comparison. An alternative method for estimating response error is the reinterview study, in which questions from the original interview are re-administered in a second interview to a sample of the respondents. Response inconsistency between interviews is an indication of poor data quality and a number of authors have cited the usefulness of this approach for estimating response error and improving questionnaires (Feindt, Schreiner, and Bushery, 1997; Miller and Ennis, 2001; McGovern and Bushery, 1999; O'Muircheartaigh, 1991).

While record-check, benchmark, and reinterview approaches offer relatively direct means of quantifying measurement error, researchers have also utilized indirect indicators of measurement error. As noted above, one such approach is to examine misclassification errors reflecting spurious changes in respondents' answers across waves in panel surveys (Kalton et al., 1986). Another approach has been to examine item nonresponse as an indicator of potentially poor data quality (e.g., Atrostic and Kalenkoski, 2002; Dahlhamer *et al.,* 2003; Mason et al*.,* 2002). Participants' decisions about whether or not to respond to a survey question are affected both by their ability to comprehend the question and retrieve relevant information and their motivation to provide an answer. Failure to answer can be considered a form of satisficing that stems from these cognitive or motivational factors, as well as features of the question itself (e.g., difficulty, response complexity, question format; Beatty and Herrmann, 2002).

Another behavior that reflects data quality is the rounding of values. Respondents often are asked to provide information about continuous variables that is subject to rounding errors (e.g., age, time intervals, earnings, hours worked). As with other response errors, round value reports can occur because respondents misinterpret the question's intent (i.e., the level of precision required), lack precise knowledge of the characteristic of interest, or are not motivated to provide a fully accurate answer. In providing an imprecise yet 'plausible' answer, respondents have a systematic tendency to report prototypical values (often multiples of 5 or 10) (Sudman, Bradburn, and Schwarz, 1996).

The reporting of round values in surveys is a well-documented phenomenon (e.g., Hanisch and Rendtel, 2002; Myers, 1954), but its impact on measurement error largely has been neglected (Tourangeau et al., 2000). As Hanisch and Rendtel (2002) recently demonstrated, however, such neglect has rested on the erroneous assumption that the occurrence of rounding is a random event with little impact on statistical inference. Their study showed that rounding indicates a loss of data that affects the measurement scale and can bias distribution parameters (means and standard deviations) as well as statistics that focus on specific distribution points (medians and percentiles)

4.1.2 Chapter Overview

The purpose of this chapter is to investigate issues of data quality in the Current Population Survey (CPS) and the American Time Use Survey (ATUS). The next section of the chapter will report the results of previous studies that have applied these error detection techniques to the two surveys. In subsequent sections, I will present my own analyses of data quality indicators available in the CPS and ATUS datasets. I will begin by examining individual data quality indicators in each survey and exploring the causal roles that various household, respondent, and survey design attributes may have on the level of error. I then will present analyses of how these individual data quality indicators are related to one another within a given survey and between the two surveys. Finally, I will discuss the implications of these findings for evaluations of overall data quality in CPS and ATUS.

4.2 Prior Data Quality Studies of the CPS and ATUS

4.2.1 CPS

Record-check and benchmark studies of CPS data have been conducted on several occasions for the purposes of assessing CPS measurement error.[12] For example, Coder and Scoon-Rogers (1995) compared income estimates from the 1984 and 1990 CPS March Income Supplements to benchmark income data from Internal Revenue Service and Social Security Administration records. The results of their study showed that income estimates from the CPS were significantly lower than estimates reported by administrative sources. They found substantial disagreement

[12] A number of studies have also examined CPS-Census match data for the purposes of assessing measurement error *in the Census,* using the CPS as the standard of comparison (e.g., Palumbo and Siegel, 2004).

on questions of dividend income (48.2%) and worker's compensation income (51.8%), but only modest discrepancy between estimates for wages and salary income (2.7%). Similar or even lower rates of inaccurate reporting of wage and salary income were found in two studies that used matched tax records (Bound and Krueger, 1991; Herriot and Spiers, 1980; Moore, Stinson, and Welniak, 2000). Interestingly, however, Bound and Krueger (1991) reported that CPS income measurement error was negatively correlated with true earnings and was autocorrelated in adjacent waves of CPS, both violations of classical measurement error models.

CPS also regularly conducts analyses of its reinterview data to monitor and evaluate data quality. The index of inconsistency (IOI) is a common measure of response variance. It is the response variance over the total variance. Zero indicates that identical responses were given in the initial and follow-up interview; an IOI above 50 indicates a high degree of unreliability for a given characteristics. The most recent published IOIs for CPS labor force items indicate that classifications of "not in labor force" and "working full-time" have a relatively low response variance (8.1 and 9.4 IOIs, respectively), whereas "unemployed" and "with job, not at work" are relatively unreliable (30.8 and 30.1, respectively) (BLS, 2002, pg. 16.6). McGovern and Bushery (1999) report that only two-thirds of the people who say that they are "unemployed" in the original interview give the same answer during the reinterview. They demonstrate that unreliability in the estimates of the "unemployed" also causes bias, and found that estimates of the "unemployed" are understated by about thirteen percent in the CPS. In keeping with this finding, they found that inconsistent responses are more likely to be made by proxy respondents and by persons who are under the age of twenty-one, single, black, and poorly educated.

Analyses of CPS item nonresponse are regularly carried out as part of the CPS quality assessment efforts. The average item nonresponse in a given month is small for demographic and labor force items (1.5%), but higher for earnings (12.4%) suggesting the possibility for significant bias in earnings estimates (BLS, 2002). Independent studies of CPS item nonresponse corroborate these findings. Dahlhamer et al. (2003) report item missing rates below three percent for 2001 basic CPS questions on race, ethnicity, housing tenure, and hours worked last week, while Atrostic and Kalenkoski (2002) found nearly twenty-five percent item nonresponse for wage and salary income as reported on the CPS Annual Demographic (March) Supplement.

Only one study has analyzed the impact of round value reports on measurement error in the CPS. Schweitzer and Severance-Lossin (1996) found that over seventy percent of gross earnings reports in the CPS March Supplement were multiples of $1000, and that rounding was positively correlated with level of earnings. In addition, they found that round earnings reports significantly biased statistics that were sensitive to subtle changes in earnings distributions (e.g., measures of earnings inequality, quantiles, and wage rigidity). For example, they found that rounding changed inequality measures by up to 3 percent, an amount that was relatively small in an absolute sense, but was larger than the typical annual change in inequality or standard error estimates for these measures.

4.2.2 ATUS

To date no record-check, benchmark, or simple response variance reinterview studies have been conducted on ATUS data nor have any ATUS item nonresponse analyses been published. The only study in the literature dealing with potential ATUS response error was conducted by Bose and Sharp (2005). They examined reported durations of travel activities from the ATUS time diary and compared them to estimates from the National Household Travel Survey (NHTS). Although the authors were primarily interested in comparing duration estimates between two surveys with quite different methodologies, a figure in one of the appendices showed the extent and pattern of round values reported for travel durations by ATUS respondents. What is clear from the figure is that respondents are commonly providing round time data, with spikes at values divisible by 5 or 10, and the tendency appears to be stronger in ATUS than NHTS.

4.2.3 Summary of CPS and ATUS Data Quality Studies

Evidence from record-check and benchmark studies and analyses of reinterview data, item-nonresponse, and rounded value reports in the CPS demonstrate that survey measurement error is present in the survey, but that its magnitude and impact varies greatly depending upon the characteristic and statistic of interest. Measurement error appears to be small for estimates of many demographic variables in the CPS. On the other hand, estimates of income and earnings are susceptible to bias both from item nonresponse and rounding, and estimates of unemployment may be affected both by unreliability and bias. Very little empirical work has been done on ATUS data so far, but there is evidence that respondents report round time values for at least one activity. The remainder of this chapter is devoted to a further examination of data quality issues in these surveys.

4.3 Methodology

4.3.1 Analytic Procedures and Hypotheses

CPS

To assess CPS data quality, I investigated item nonresponse, round value reports, changes in classification between rounds of the survey (e.g., changes in race), and inconsistent reports between the basic CPS interview and the reinterview. Analyses first were carried out on the individual variables at the person-level for each CPS wave. Data then were aggregated across persons, variables, and waves of CPS in order to obtain an overall value for each data quality indicator for each household. I compared results at both levels of analyses with those found in previous studies of CPS data quality (when available).

I next examined the effects of respondent knowledge, ability, and motivation on CPS data quality. The variables used as indicators of these causal factors are listed in Table 1. Based on arguments outlined in section 4.1, if response errors are driven by gaps in respondent knowledge, then we would expect CPS data quality to be worse for proxy-reports than for self-reports and to decrease with household size (since it may be harder to have full and accurate information about other household members in larger households than in smaller households). If response errors stem from cognitive deficits, then we would expect elderly or poorly educated respondents to have the worst data quality.

Table 1. Causal Factors Examined for Their Effects on CPS Data Quality Indicators

Causal Factor	Indicator
Knowledge	Self vs. Proxy Household Size
Ability	Age Education
Motivation	Hours Worked Employment Status Marital Status Presence of Young Children Time in Sample Item Burden

Several indicators of respondent motivation were examined. To the extent that busyness reduces both the time that can be devoted to effortful processing of survey questions and the motivation to respond fully and accurately, we would expect that hours worked and employment would have a negative effect on data quality. If social isolation interferes with norms of cooperation and the motivation to engage in effortful processing, unmarried respondents and those without children should make more errors than those who are married or have young children in the household. The busyness and social capital hypotheses make opposite predictions for the effects of employment, marriage, and the presence of children. According to the busyness hypothesis, reductions in discretionary time associated with these factors will lead to reductions in data quality; the social capital hypothesis suggests that these factors are related to social integration and pro-social norms that will produce gains in data quality.

I also examined four measures of survey burden to assess its effect on data quality: interviewing period (i.e., first four rounds of CPS vs. the last four); extent of unit nonresponse (i.e., households that participated in all eight waves vs. those who did not); changes in household respondent across waves (i.e., same person responded in all waves vs. multiple people reported across waves); and item burden (i.e., total number of questions asked of the household relative to those asked of other households with the same time-in-sample and size). If the motivation to respond accurately is tied to perceptions of survey burden, data quality should be worst in the second half of CPS, and when the household responds all eight waves, the same person within the household reports across waves, and the number of questions asked of the household is large.

I regressed each CPS data quality indicator on these causal variables individually and in multivariate models. Finally, I looked to see if there were common causes across indicators, and then explored the relationships between the individual measures of CPS data quality.

ATUS

To evaluate data quality in the ATUS, I analyzed item nonresponse and round values in the ATUS labor force questions, 'don't know' responses on the time diary, round activity durations, and interviewer codes indicating poor case data quality. In addition, the time use literature suggested two other potentially useful measures of ATUS response quality: total number of time diary activities reported by respondents and the presence or absence of basic respondent activities (e.g., sleeping, eating, grooming) in which the vast majority of people engage on a given day.

As before, I examined the relationship between the ATUS data quality measures and potential causal variables, both individually and in multivariate regression models. Table 2 lists the causal factors examined. Because the ATUS is a one-time survey with a unique protocol (e.g., asking respondents to generate unstructured time-use activity reports, collecting very little proxy information), several causal variables related to respondent knowledge (e.g., household size, reporting for others) and survey burden (e.g., time in sample, extent of unit nonresponse) were dropped from these analyses. One indicator of survey burden that was available on the ATUS data file denoted ATUS respondents who also had been the respondent in the last round of CPS. The majority of CPS households have a single respondent for most if not all rounds, so if survey burden influenced respondents' motivations to report accurately, this variable would be negatively associated with ATUS data quality.

Table 2. Causal Factors Examined for Their Effects on ATUS Data Quality Indicators

Causal Factor	Indicator
Knowledge	n/a
Ability	Age Education
Motivation	Hours Worked Employment Status Marital Status Presence of Young Children ATUS Respondent Same as Last CPS Interview

Older respondents and those with lower levels of educational attainment would be expected to provide poorer data than younger, better educated respondents to the extent that response errors stemmed from cognitive limitations. Employment, marriage, and the presence of young children in the household were predicted to increase data quality under the social capital hypothesis, but these factors and hours worked were expected to reduce data quality under the busyness hypothesis.

4.3.2 The CPS Interview

Basic CPS Interview

During the first CPS interview, a roster of household members is collected from the respondent and information about each household member's relationship to the reference person, date of birth, race, ethnicity, sex, educational attainment, marital status, and armed forces status is recorded. Race, ethnicity, and educational attainment questions are asked again in the fifth CPS interview; educational attainment also is asked in February, July, and October. During the initial interview, respondents also are asked to report their household income, to indicate whether or not there is a working telephone in the household, and if the housing unit is owned, rented, or occupied without payment of cash rent. These three questions also are re-asked in the fifth interview.

Labor force questions are asked each interview about each household member 15 years of age or older. These questions include current employment status, actual hours worked during the reference period (the week of the month containing the 12th day), and usual hours worked for all eligible employed household members. Individuals who are classified as unemployed, on lay-off, or not in the labor force also are asked a series of questions related to these categories. In the fourth and eighth interviews, earnings data are collected for each eligible employed household member.

CPS Reinterview

Each month, CPS attempts to conduct a reinterview on a subsample of responding households. To minimize response burden, a given household only is reinterviewed once. Reinterviews are conducted only by telephone, and are carried out by senior interviewers between one and ten days after the original interview.

There are two components to the CPS reinterview program—a response error (RE) interview and an interviewer quality control (QC) interview. The latter does not yield estimates of response error, which are the focus here. About one percent of eligible CPS households are assigned to the RE component each month. Since 1994, all RE reinterviews are computer-assisted telephone interviews (CATI) that consist of the entire set of labor force questions; household membership is dependently verified, and no reconciliation is conducted. An effort is made to reinterview the person who responded to the original interview, but interviewers are allowed to conduct the reinterview with other knowledgeable household members (BLS, 2002).

Variables Used in CPS Analyses

Eight CPS demographic items were selected for analyses: sex, age, race, ethnicity, educational attainment, home ownership, telephone status, and family income. These items were chosen because they were asked at least once of every household member, are variables commonly of interest to researchers, and because they offered a range of potential response error. In addition, eight labor force items were examined. Column 1 of Table 3 presents each item and column 2 indicates how often and of whom the item is asked. Several of these items contribute to the monthly labor force variable used for determining the CPS's monthly unemployment rate. As with the demographic items, these labor force variables were selected because they are asked each round of most CPS respondents and offered a range of potential response error magnitudes.

Table 3. CPS Labor Force Questions Selected for Data Quality Analysis

Question	Item Protocol
Does anyone in this household have a business or a farm?	Once per household during each round
Last week, did *you/household member* do any work (either) for pay (or profit)?	All eligible household members each round
Last week, how many hours did *you/household member* ACTUALLY work at *your/their* main job?	All eligible household members each round
How many hours per week do (does) *you/household member* USUALLY work at your/their main job?	All eligible household members each round
What are *your/household member's* [periodicity] earnings on *your/this* MAIN job, before taxes or other deductions?	All eligible household members in rounds 4 and 8
Have *you/household member* been doing anything to find work during the last four weeks?	All eligible household members each round
What are all of the things *you/household member* have done to find work during the last four weeks?	All eligible household members each round
Last week, did *you*/*member* lose or take off any hours from *your/their* job, for ANY reason such as illness, slack work, vacation, or holiday?	All eligible household members each round

4.3.3 The ATUS Interview

Two months after the household's final CPS interview, a single household member fifteen years of age or older is randomly selected to participate in the ATUS interview. The designated person is assigned a specific reporting day of the week (e.g., Monday); if the interview cannot be completed on the designated day during the first week of the interviewing period, subsequent interview attempts are made on the designated day each week for up to eight weeks.

When the interviewer calls the ATUS respondent on their designated day, the time use interview begins by verifying the address and household roster information, and collecting labor force data about the respondent for the purposes of updating the CPS data. These questions are followed by the time diary which records information about the respondent's activities the day before the interview. After the diary has been completed, several summary questions are asked about the respondent's child care and paid work activities during the reference day, and about any absences from home the respondent may have had during the month preceding the first eligible interview date. Following the summary questions, additional labor force items are asked about the respondent's hours worked, industry, occupation, school enrollment, and earnings. ATUS interviews use standardized questions for the labor force items and summary questions, and less structured interviewing techniques for the time diary. The entire interview typically lasts about 20 – 25 minutes.

The core of the interview is the time diary. Respondents are asked to report their activities beginning at 4 a.m. the day before to 4 a.m. the day of the interview. They are asked to provide the type and duration of each activity, where they were when it occurred, and who they were with. Respondents are not asked or required to provide "who were you with" information for activities coded as work, school, or sleep, or for some other personal activities (e.g., grooming, getting dressed). During the interview interviewers can use thirteen codes for common activities. Time diary reports from completed cases then are entered into the ATUS coding application for processing by an independent coder (another interviewer). For quality assurance, all interviews are verified (i.e., recoded) by a second member of the coding staff, and the total time to code, verify, and adjudicate a case (if necessary) is recorded.

Variables Used in ATUS Analyses

The ATUS dataset was constructed from monthly data files from the first year of ATUS data collection (2003). It had 20,698 usable cases with respondents' time diary records, reports of who they were with during each activity, detailed call record information, labor force data collected in ATUS, and final outcome disposition (i.e., interview status) of the case. In addition, the file contained variables indicating the total time it took to code, verify, and adjudicate the interview, the total interview length, and the number of "don't know" reports given in the diary.

Table 4. ATUS Labor Force Questions Selected for Data Quality Analysis

Item
Do you own a business or a farm?
Did you have a job in the last seven days?
How many hours do you usually work at your main job?
How many hours do you usually work at your second job?
Did you do anything to find work during the last four weeks?
Are you currently enrolled in school?
What is your [periodicity] earnings for your MAIN job, before taxes or other deductions?
Did your spouse work in the last seven days?
Does your spouse usually work more than 35 hours per week?

Table 4 lists nine ATUS labor force items that were selected for analyses. Because ATUS does not re-ask most demographic items but confirms the data from the CPS, I selected two additional questions that are used to help calculate ATUS child care estimates: "What time did the first child awake?" and "What time did the last child go to bed?" Most ATUS respondents are asked at least a few of these eleven items, and there seemed to be sufficiently variation among the questions in their susceptibility to response errors.

4.4 CPS Results

4.4.1 Item Nonresponse

I first examined item nonresponse for eligible persons on each of the 16 variables in each of the eight CPS rounds, counting items that were supposed to be answered but got a “don’t know” or “refused” response. Consistent with previous findings (e.g., Dahlhamer et al. 2003), the levels of missingness for most items were quite low, generally less than three percent. Sex had the lowest item nonresponse rate, with less than one-tenth of a percent missing in each of the eight interviews. This is not surprising given that interviewers can easily infer the gender of most household members and only need to ask this question in rare circumstances. Questions about the existence of a household business, home ownership, telephone status, and employment status also had missing data rates that hovered around one percent or less. As anticipated, the items with the highest percent nonresponse were earnings and family income. In general, item nonresponse was worse in waves one and five than in the other waves, and slightly lower in the first four interviews than the last four.

A similar pattern of results emerged for the household-level measure of item nonresponse obtained for each variable by aggregating across all household members and rounds of CPS. Most households provided complete data for demographic variables over their time in sample. For example, the household mean missing data rates for the variables of sex, housing tenure, and “any work for pay or profit” were 0.1%, 0.7%, and 2.7%, respectively. Again, earnings and income produced the highest incidence of item nonresponse. Nearly half of the CPS households (47.6%)

failed to report earnings for one or more eligible household members and more than a quarter (28.5%) of households skipped the household income question at least once.

I next calculated the overall missing data rate for a household by summing the number of items skipped (across all variables and persons in the household), and dividing by the total number of items the household was eligible to be asked to over the eight rounds of CPS. Table 5 presents this aggregate household-level item missingness measure as function of some of the causal factors identified in Table 1 that were hypothesized to affect respondents' ability or motivation to respond accurately. The first column of the table lists these factors. The second column indicates the percent of households that had no item nonresponse during their time in sample, with superscripts denoting chi-square test significance levels. The third column presents the average percentage household item nonresponse, with superscripts denoting the factors' significance levels.

Item Nonresponse and Household Characteristics

As can be seen in the first row of the table, more than forty percent of households had no missing data across the eight CPS rounds and the average percent household item nonresponse was small (3.5%). In general, the sizes of effects presented in the table are small, though all but one reached statistical significance. Looking at the effects of indicators of respondent knowledge, item missingness rates were nearly twice as high when respondents were reporting about other household members than when they were reporting about themselves (*self- vs. proxy reporting*). Similarly, household size had fairly dramatic negative effects on the proportion of households that provided complete data (column 1), but there was a curvilinear effect for household size on item missing data rates (column 2). The average percent item nonresponse was higher in single person households than in multi-person households that consisted of fewer than seven individuals, but very large households had the highest missing data rates. The apparent discrepancy between these two findings is due in part to the relatively small effect of nonresponse on a few items or for a few people in large households compared to item nonresponse in smaller households. It likely also reflects significant social and demographic differences (uncontrolled for here)—e.g., in age, education, social trust, etc.—between respondents who live alone and those in multi-person households.

The lower panel in Table 5 examines the effects of several indicators of respondent motivation. Consistent with the notion that survey burden will reduce data quality, the percentage of households with complete data (i.e., no item nonresponse) was higher in the first four CPS waves than in the last four waves, and missing data rates were lower in households in which relatively few questions were asked. By contrast, missing data rates increased with increases in CPS unit nonresponse and the number of household respondents across rounds, contrary to expectations. Given these results, it seems likely that these two indicators may in fact be capturing the interaction between overall survey reluctance and item nonresponse rather than time-in-sample effects.

Table 5. Item Missingness in CPS Households by Household and Reporting Characteristics

		% of Households with No Item Nonresponse (1)	Average % Item Missing in HH (2)	Total
	Total	42.3%	3.5%	97,053
Knowledge	**Self- vs. proxy reporting**			
	Reporting for self	67.9	2.3	70,708
	Reporting for others	44.0***	4.1***	70,708
	HH size			
	1 person	55.1	3.8	26,354
	2 person	42.7	3.3	31,679
	3 -4 people	34.7	3.4	29,018
	5 – 6 people	30.7	3.7	8,713
	7 or more	18.9***	4.3***	1,298
Motivation	**Interviewing period**			
	Waves 1 – 4	59.0	3.5	96,068
	Waves 5 – 8	55.8***	3.4 ns	95,229
	Item Burden			
	Small	52.6	3.1	30,973
	Medium	41.5	3.3	33,374
	Large	33.4***	4.1***	32,706
	Extent of CPS unit nonresponse			
	None	43.8	2.9	81,777
	Waves 1 or 2 only	40.1	5.6	3,504
	Waves 3 – 8 only	31.4	5.9	8,549
	Waves 1 or 2 and 3 – 8	34.9***	11.2***	3,223
	# of changes in HH respondent			
	None	50.1	3.8	37,152
	One	40.9	3.1	48,856
	Two	23.5	4.0	9,299
	Three or more	16.2***	5.4***	1,746

*** Significant at the $p < .001$ level. ns Result is non-significant

Item Nonresponse and Respondent Characteristics

Table 6 presents additional descriptive statistics about how household item nonresponse varied by characteristics of the main CPS respondent (i.e., the person who served as the household respondent most often). Missing data rates were lower for females than for males and for whites and Asians than blacks or persons of some other race. Missing data rates also improved at higher levels of educational attainment, providing some evidence that CPS data quality was affected by respondents' cognitive abilities. However, age effects were in the opposite direction predicted—item nonresponse decreased with age.

Table 6. CPS Item Missingness by Main Respondent Characteristics

	Characteristics of the main respondent	% of Households with No Item Nonresponse	Average % Item Nonresponse Reported in HH	Total
Controls	**Sex**			
	Male	42.2%	3.7%	36,585
	Female	42.4 ns	3.4***	60,468
	Race			
	White	43.6	3.3	82,901
	Black	33.7	4.8	9,567
	Asian	42.4	3.3	1,027
	Other	33.1***	4.6***	2,955
Ability	**Age**			
	LT 15	6.7	9.8	30
	15 – 24	34.9	4.4	3,963
	25 – 64	40.6	3.5	72,223
	65 +	49.8***	3.4***	20,837
	Education			
	LT HS	41.2	3.7	13,252
	HS Only	40.6	3.6	31,395
	Some College	43.0	3.4	25,890
	BA/BS	44.1	3.4	17,012
	Advanced degree	46.3***	3.2***	8,858
Motivation	**Labor Force Status**			
	Employed	39.7	3.6	59,897
	Unemployed	40.8	3.3	2,809
	NILF	47.1***	3.4***	34,320
	Hours worked			
	0	46.6	3.6	38,262
	1 – 39	41.1	3.1	18,786
	40	35.1	4.1	19,700
	Over 40	43.3***	3.1***	16,942
	Marital Status			
	Married	45.2	3.2	55,039
	Unmarried	40.1***	3.9***	42,014
	Young Children in HH			
	Yes	45.9	2.6	11,631
	No	41.8***	3.7***	85,422

*** Significant at the p < .001 level. ns Result is non-significant

The results of the bivariate analyses were confirmed in a multivariate model that used the preceding causal factors to predict the presence of household-level item nonresponse (see Table 7). The model achieved a likelihood ratio chi-square of 6742.3 ($p < .001$) and a max-rescaled r-square of .10. The largest effects were found for unit nonresponse, household size, and relative item burden. Households with unit nonresponse in one or more CPS waves were significantly more likely to have item nonresponse than households that participated in every round. Large households and those asked a large number of questions were more likely to have item nonresponse than households with fewer members or fewer questions asked. In addition, households were least likely to provide complete data when there were no children present, and when the main respondent was male, black or 'some other race,' poorly educated, unmarried, employed, or working forty hours per week. The positive effect of age on item missingness found in the bivariate analyses was reversed here—the probability of having no missing data was highest in households in which the main respondent was under the age of twenty-five and lowest when the main respondent was over the age of sixty-five. The significant interaction between age and hours worked revealed the particularly negative effects of working forty hours per week or more for respondents over the age of sixty-five.

Table 7. Logistic Model Predicting Presence of CPS Item Nonresponse, HH-level

	Odds Ratio	Wald Chi-Square	Sig
Sex		8.0	.0046
Male vs. Female	1.02	8.0	.0046
Race		258.1	<.0001
Black vs. White	1.25	71.5	<.0001
Asian vs. White	0.75	31.1	<.0001
Other vs. White	1.21	28.8	<.0001
Age		93.7	<.0001
15 – 24 vs. 25 - 64	0.90	11.5	.0007
65+ vs. 25 – 64	1.27	68.2	<.0001
Education		171.9	<.0001
LT HS vs. BA/BS	1.17	77.7	<.0001
HS Only vs. BA/BS	1.1	62.7	<.0001
Some College vs. BA/BS	0.95	6.2	.0679
Advanced degree vs. BA/BS	0.88	46.2	<.0001
HH Size		1433.1	<.0001
2 person vs. 1	0.72	230.2	<.0001
3 – 4 person vs. 1	1.06	9.1	.0025
5 – 6 person vs. 1	1.38	142.1	<.0001
7 or more vs. 1	2.57	234.5	<.0001
Employment Status		19.5	<.0001
Unemployed vs. Employed	0.90	9.4	.0022
Not in Labor Force vs. Employed	0.96	3.5	.0601
Hours Worked		44.7	<.0001
0 vs. 40	1.02	0.2	.6286
1 – 39 vs. 40	0.88	19.1	<.0001
Over 40 vs. 40	0.92	4.7	.0301
Hours Worked x Age		72.7	<.0001
Any CPS Unit Nonresponse		1192.6	<.0001
Yes vs. No	1.48	1192.6	<.0001
Marital Status		228.5	<.0001
Not Married vs. Married	1.15	228.5	<.0001
Young Children in HH		838.8	<.0001
No vs. Yes	1.42	838.8	<.0001
Item Burden		220.9	<.0001
Small vs. Medium	0.77	217.7	<.0001
Large vs. Medium	1.24	167.3	<.0001

4.4.2 Changes in Classification Between Rounds in the Basic CPS Interview

The purpose of this analysis was to examine spurious changes that occurred in respondent's answers between rounds of the CPS. It was difficult to select variables for these analyses. Responses to CPS labor force items can naturally vary between months and many of the demographic items (e.g., age, sex, marital status) are asked only once or simply verified through dependent interviewing, so neither set was appropriate for an examination of changes in classification. I chose to examine five variables: race, educational attainment, educational attainment (restricted to individuals 30 years of age or older), housing tenure, and income. Since real changes in all of these variables except race are possible, I first examined between-round inconsistencies as a function of whether one respondent provided answers in both rounds or the answers came from two different people. These analyses are necessarily restricted to households with two or more members.

Table 8 shows the percent of cases (and totals) with any changes in classification between CPS rounds. The top half of the table shows person-level data and the bottom half shows household-level data. Column 1 gives the percent of cases with any classification changes when both answers were given by the same respondent. Column 2 presents change estimates stemming from two different reporters. Chi-square significance tests were run on data in these two columns (reporter type differences were significant for all variables). Column 4 gives the overall percent of cases with any changes in classification ignoring reporter type, and the last column indicates the average percent change for each variable.

Table 8. Changes in Classification Between CPS Rounds, by Type of Reporter

		Type of Reporter			
		Same reporter all rounds (1)	Different reporter between rounds (2)	Any Changes? (3)	Average % change (4)
Person-level	Race**	2.2% (38,952)	2.0% (178,637)	2.0% (217,589)	2.0% (217,589)
	Education***	12.7 (26,981)	17.3 (138,906)	16.6 (165,887)	10.9 (165,887)
	Education*** (Age 30+ only)	8.0 (21,396)	11.3 (105,743)	10.8 (127,139)	6.5 (127,139)
HH-level	Race**	3.3 (13,391)	2.8 (56,097)	2.9 (69,488)	1.7 (69,488)
	Education***	21.2 (13,403)	32.0 (55,968)	29.9 (69,371)	9.5 (69,371)
	Education*** (Age 30+ only)	12.2 (12,469)	18.7 (53,318)	17.5 (65,787)	6.5 (65,787)
	Tenure**	4.5 (14,088)	4.0 (56,620)	4.1 (70,708)	2.1 (70,708)
	Income**	38.2 (13,391)	39.4 (56,097)	39.2 (69,488)	19.6 (69,488)

*** Significant at the $p < .001$ level. ** Significant at the $p < .01$ level.

There are several features of these results worth noting. First, the amount of change varies a great deal by variable. For example, changes in race occurred only in two to three percent of the cases. However, this is the one variable in the group that truly should not ever change, so even this small amount of inconsistency is surprising.[13] Larger reporting inconsistencies occurred for income, and also for educational attainment even after restricting my analyses to persons 30 years of age or older (where actual change was less likely, though not impossible). Finally, classification changes in education and income were more likely to occur when

[13] CPS introduced a new race question beginning January, 2003 that allowed respondents to select multiple race categories. Since this likely would have had produced substantial rates of classification change for respondents who received this question, I excluded these cases from this analysis.

reports came from two different individuals than when they came from the same individual.

I aggregated these items (race, educational attainment for person 30 or older, housing tenure, and income) to obtain an overall measure of classification changes at the household level, and then constructed a logistic regression model to examine the effects of potential causal variables on this estimate. Table 9 presents the odds ratios and associated *p* values for the model.

The overall effect sizes in this model were relatively small compared to those found in the item nonresponse analyses. Classification changes were most likely when the main CPS respondent was under the age of 25, Asian, unemployed, and when young children were present in the household and the household had been some unit nonresponse. Each of these effects is the opposite found for measures of CPS missing data. Consistent with results from item nonresponse analyses, smaller households, married individuals, and college educated respondents produced fewer errors. The number of questions a household was eligible to answer across the eight rounds (i.e., *item burden*) again was negatively associated with this data quality indicator—changes in classification were more likely to occur when burden was high.

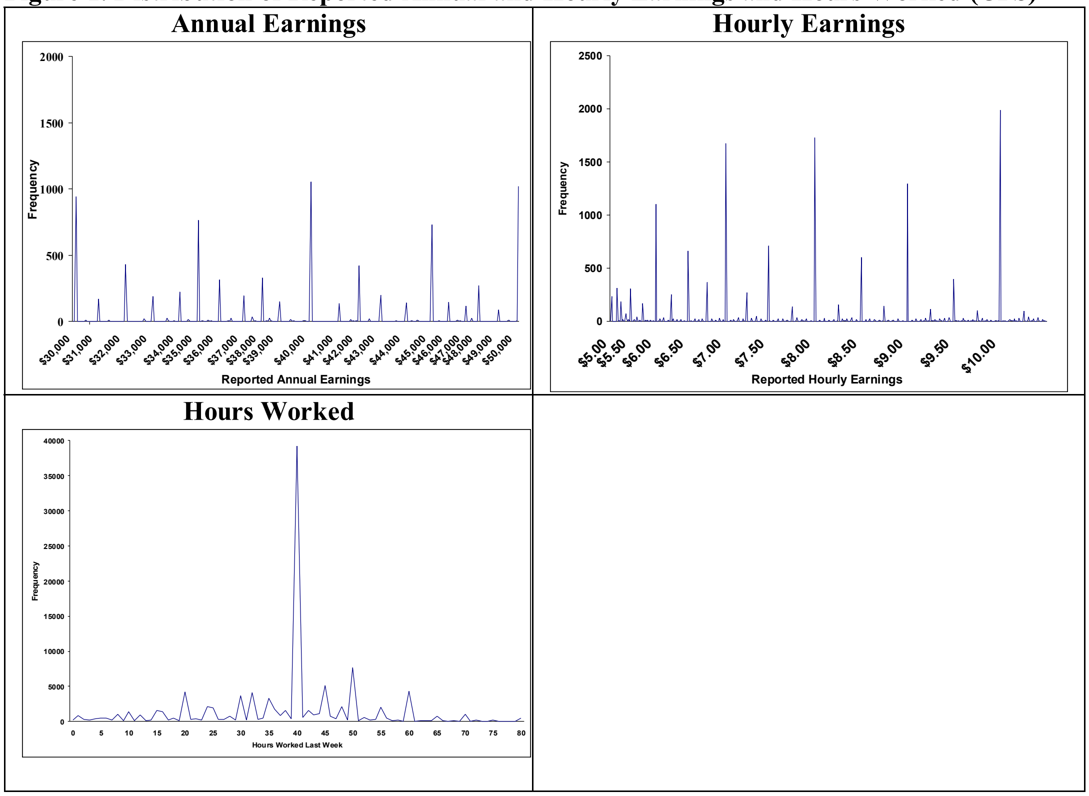
Table 9. Logistic Model Predicting Presence of Changes in Classification in Basic CPS

	Odds Ratio	Wald Chi-Square	Sig
Sex		9.0	.0027
Male vs. Female	1.03	9.0	.0027
Race		104.8	<.0001
Black vs. White	0.90	11.0	.0009
Asian vs. White	1.47	38.6	<.0001
Other vs. White	0.97	0.5	.4596
HH Size		26.9	<.0001
3 – 4 person vs. 2	0.92	20.6	<.0001
5 – 6 person vs. 2	1.00	0.0	.9894
7 or more vs. 2	1.19	14.1	.0002
Age		23.0	<.0001
15 – 24 vs. 25 - 64	1.17	21.0	<.0001
65+ vs. 25 – 64	0.91	7.6	.0058
Education		1269.3	<.0001
LT HS vs. BA/BS	1.59	502.9	<.0001
HS Only vs. BA/BS	1.19	152.4	<.0001
Some College vs. BA/BS	1.24	206.7	<.0001
Advanced degree vs. BA/BS	0.59	507.2	<.0001
Employment Status		39.9	<.0001
Unemployed vs. Employed	1.22	22.4	.0022
Not in Labor Force vs. Employed	0.86	23.7	.0601
Hours Worked		5.1	.1665
0 vs. 40	0.95	1.0	.3212
1 – 39 vs. 40	1.08	4.8	.0277
Over 40 vs. 40	1.00	0.0	.9544
Hours Worked x Age		11.94	.0633
Marital Status		283.6	<.0001
Not Married vs. Married	1.18	283.6	<.0001
Young Children in HH		52.3	<.0001
No vs. Yes	0.92	52.3	<.0001
Any CPS Unit Nonresponse		1063.2	<.0001
Yes vs. No	0.64	1063.2	<.0001
Item Burden		256.4	<.0001
Small vs. Medium	0.82	242.6	<.0001
Large vs. Medium	1.10	33.8	<.0001

4.4.3 Round Values

Four sets of CPS items were analyzed for round values: hours actually worked, hours usually worked, hours off from work, and earnings. I first examined the distribution of each variable to determine the extent to which round values occurred and to find the most appropriate/common multiple by which the item was divisible. I examined round values in the hours items for multiples of five and eight. Actual and usual hours were more commonly divisible by five, whereas hours away from the job were most commonly divisible by eight. At the household-level, 76.5 percent of usual hours reports, 70.3 percent of actual hours reports, and 51.4 percent of hours away from the job were round values. Earnings data in the CPS is reported in various ways depending upon respondents' preferred reporting periodicity. Annual earnings reports were most commonly divisible by $500 and $1000, whereas all other earnings were most commonly divisible by $50. The extent of round value earnings reports depended upon reporting periodicity, ranging from 67.2 percent for hourly earnings to 91.1 percent for annual earnings (multiples of $1000). Figure 1 shows the distributions of annual and hourly earnings reports for ranges that contain the median reported value.

Figure 1. Distribution of Reported Annual and Hourly Earnings and Hours Worked (CPS)



The results of bivariate analyses of household-level rounded value reports by household attributes are reported in Table 10. Round value reports increased as a function of time-in-sample, with higher average percent round values in the second half of CPS than the first. Again, unit nonresponse was associated with significantly higher rates of rounding than when the household was never a nonrespondent. Round reports also were higher for large (vs. small) households, for data reported about others (vs. data reported about oneself), and for households with a large amount of question burden. Analyses of round values by CPS main respondent characteristics (not presented here) revealed that persons over the age of 65, females, and Asians, and who were unmarried, without children, not in the labor force, or working fewer hours per week had relatively low amounts of round reports (though still 60 percent or more rounded).

Table 10. Household-level Round Value Reports by Household and Reporter Characteristics

		% of Households Reporting Only Non-round Values	Average % Round Value Reports in HH	Total
	Over all waves	1.6%	73.1%	74,464
Knowledge	**Self- vs. proxy reporting**			
	Reporting for self	2.7	71.7	73,129
	Reporting for others	2.4***	75.0***	58,855
	HH size			
	1 person	3.2	72.0	15,169
	2 person	2.3	71.8	23,622
	3 – 4 people	0.7	74.2	27,843
	5 – 6 people	0.4	74.2	8,554
	7 or more	0.2***	74.2***	1,276
Motivation	**Interviewing period**			
	Waves 1 – 4	2.3	71.5	74,264
	Waves 5 – 8	2.0 ns	75.5***	71,292
	Amount of Question Burden			
	Small	9.7	65.5	11,059
	Medium	0.5	73.8	33,519
	Large	0.0***	74.8***	31,886
	Extent of CPS unit nonresponse			
	None	1.4	72.8	63,873
	Waves 1 or 2 only	2.5	74.0	2,644
	Waves 3 – 8 only	1.8	73.9	7,407
	Waves 1 or 2 and 3 – 8	5.7***	75.0***	2,540
	# of changes in HH respondent			
	None	2.5	72.2	24,744
	One	1.3	73.3	41,329
	Two	0.6	73.7	8,808
	Three or more	0.7***	73.9***	1,583

*** Significant at the $p < .001$ level. ns Result is non-significant

Table 11. Logistic Model Predicting Presence of CPS Round Value Reports

	Odds Ratio	Wald Chi-Square	Sig
Sex		1.3	.2573
Male vs. Female	1.04	1.3	.2573
Race		4.0	.2595
Black vs. White	1.15	1.9	.1727
Asian vs. White	0.83	0.9	.3404
Other vs. White	0.90	0.6	.4554
HH Size		55.6	<.0001
2 person vs. 1	0.57	17.0	<.0001
3 – 4 person vs. 1	0.98	0.1	.9078
5 – 6 person vs. 1	1.39	2.8	.0960
7 or more vs. 1	2.57	3.9	.0463
Age		20.9	<.0001
15 – 24 vs. 25 - 64	8.86	0.1	.9458
65+ vs. 25 – 64	0.23	0.1	.9258
Education		8.9	.0515
LT HS vs. BA/BS	1.21	8.5	.0036
HS Only vs. BA/BS	1.01	0.1	.7586
Some College vs. BA/BS	0.98	0.1	.7889
Advanced degree vs. BA/BS	0.95	2.4	.1245
Employment Status		4.0	.1356
Unemployed vs. Employed	1.09	0.8	.3851
Not in Labor Force vs. Employed	0.87	3.9	.0458
Hours Worked		9.9	.0197
0 vs. 40	0.19	0.1	.9180
1 – 39 vs. 40	0.22	0.0	.9251
Over 40 vs. 40	0.49	0.0	.9648
Hours Worked x Age		4.5	.6087
Marital Status		4.1	.0423
Not Married vs. Married	1.09	4.1	.0423
Young Children in HH		0.2	.6267
No vs. Yes	0.96	0.2	.6267
Any CPS Unit Nonresponse		15.4	<.0001
Yes vs. No	1.18	15.4	<.0001
Item Burden		989.6	<.0001
Small vs. Medium	0.06	365.2	<.0001
Large vs. Medium	15.08	96.8	<.0001

Table 11 presents the results of the multivariate logistic model predicting the presence of household-level rounded value reports. The model achieved a likelihood ratio chi square of 3939.3 (p < .001) and a max-rescaled r-square of .33. The two effects driving the model clearly are age of the main respondent and relative item burden. Main respondents under the age of twenty-five are much more likely to provide round values than respondents over the age of sixty-five, and households with a large amount of question burden are much more likely to provide rounded values than households with average item burden. As before, being married and never a unit nonrespondent are significantly related to reductions in round value reports, whereas living in a large household is associated with more round values than living in a small household.

4.4.4 Basic CPS Interview-Reinterview Response Variance

After matching person-level data from the original interview to the reinterview data, 3,892 cases were analyzed for response consistency on sixteen labor force and demographic items. The overall response rate for the reinterview was 85.0 percent.

For each of the eight CPS rounds, I generated a single measure of person-level response inconsistency by summing across individual items for each person. I also created a single measure of household-level response inconsistency for each round by summing across items and persons within a household. (Unlike previous analyses, there was no reason to sum across waves because CPS households are only reinterviewed in one wave, if at all.) The rates of inconsistency across the eight rounds of CPS were very similar at both levels of analyses, ranging from about eight percent inconsistent up to eighteen percent. Response inconsistencies tended to be highest in the outgoing CPS rounds (i.e., the fourth and eighth months-in-sample), when earnings data is collected, and higher on average in rounds five through eight than in the first four waves of CPS.

Table 12. Response Inconsistency Between Basic CPS and Reinterview, by Characteristics of the Household and Reporter

		% of Households with No Inconsistency	Average % Inconsistency	Total
	Over all waves	36.6%	13.8%	3,892
Knowledge	**Self- vs. proxy reporting**			
	Self-report only	49.4	14.3	3,825
	Proxy-report only	46.3***	13.3^{ns}	2,991
	HH size			
	1 person	53.7	12.9	907
	2 person	35.0	14.3	1,347
	3 – 4 person	27.9	14.0	1,187
	5 – 6 person	27.6	13.5	355
	7 or more	15.6***	16.8*	45
Motivation	**Interviewing period**			
	Waves 1 – 4	36.3	13.2	2,752
	Waves 3 – 8	37.3^{ns}	15.1**	1,140
	Item Burden			
	Small	61.8	10.5	1,410
	Medium	29.1	15.5	1,263
	Large	15.2***	15.8***	1,219
	Extent of CPS unit nonresponse			
	None	35.3	13.9	3,389
	Waves 1 or 2 only	46.4	14.0	110
	Waves 3 – 8 only	41.6	11.9	291
	Waves 1 or 2 and 3 – 8	49.0**	18.8**	51
	# of changes in HH respondent			
	None	45.9	13.1	1,366
	One	31.7	13.9	2,016
	Two	27.3	15.6	388
	Three or more	31.0***	16.5*	71

*** Significant at the $p < .001$ level. ** Significant at the $p < .01$ level. * Significant at the $p < .05$ level.

Table 12 reports household-level response inconsistencies by characteristics of the household and reporter. The average total percent inconsistency for households was 13.8 percent, with a little more than a third of households reporting data free from response error. As noted, rounds one through four had significantly fewer inconsistent reports than rounds five through eight. The average percent inconsistency was highest for households that were nonrespondents in at least two previous CPS rounds, for the largest households, and for households with a relatively high degree of item burden. Response errors increased with the number of household respondents over rounds, and also were somewhat higher for self-reports than proxy reports, though this latter difference failed to reach significance. When similar analyses were conducted on this inconsistency measure by characteristics of the respondent in the reinterview, the average percent of inconsistent response in the household was significantly higher when the reinterview respondent was over the age of sixty-five, unmarried, not in the labor force, or working part-time.

The results of a multivariate model predicting the presence of response inconsistency between the basic interview and reinterview is presented in Table 13. Again, the strongest effects in the model were found for the relative question burden variable and household size. In addition, small but significant effects were found for main respondent age (respondents under the age of 25 reported more accurately than older respondents), education (respondents with at least some college had fewer inconsistencies than those with lower levels of education), employment status (persons not in the labor force reported more accurately than those employed or unemployed), and hours worked (those who reported working over forty hours in the preceding week were more accurate than those who working less than forty hours).

Table 13. Logistic Model Predicting Presence of Response Inconsistency Between Basic CPS and CPS Reinterview

	Odds Ratio	Wald Chi-Square	Sig
Sex		1.57	.2095
Male vs. Female	1.05	1.57	.2095
Race		1.70	.6362
Black vs. White	1.19	1.40	.2368
Asian vs. White	0.89	0.16	.6910
Other vs. White	0.93	0.13	.7217
HH Size		33.79	<.0001
2 person vs. 1	1.36	6.23	.0125
3 – 4 person vs. 1	0.79	3.93	.0474
5 – 6 person vs. 1	0.57	14.06	.0002
7 or more vs. 1	1.06	0.02	.8685
Age		6.64	.0360
15 – 24 vs. 25 - 64	0.48	2.26	.1323
65+ vs. 25 – 64	0.99	0.0	.9265
Education		9.01	.0421
LT HS vs. BA/BS	1.34	7.12	.0142
HS Only vs. BA/BS	1.02	0.01	.7746
Some College vs. BA/BS	0.84	5.69	.0170
Advanced degree vs. BA/BS	0.96	0.14	.7143
Employment Status		6.53	.0382
Unemployed vs. Employed	1.37	2.21	.1376
Not in Labor Force vs. Employed	0.70	6.04	.0140
Hours Worked		13.34	.0040
0 vs. 40	1.54	2.88	.0895
1 – 39 vs. 40	1.35	3.13	.0764
Over 40 vs. 40	0.88	0.34	.5587
Hours Worked x Age		7.26	.2977
Marital Status		0.34	.5602
Not Married vs. Married	1.03	0.34	.5602
Young Children in HH		1.84	.1747
No vs. Yes	1.10	1.84	.1747
Any CPS Unit Nonresponse		0.27	.6015
Yes vs. No	.097	0.27	.6015
Item Burden		392.60	<.0001
Small vs. Medium	0.22	384.38	<.0001
Large vs. Medium	3.89	268.07	<.0001

4.4.5 Summary of Variable Effects on CPS Data Quality Indicators

As can be seen in Table 14, a number of the variables examined had common effects across the four CPS data quality measures. The table indicates which variables had significant impact on the data quality indicators and the direction of the effect. Respondent sex and race were significantly related to item nonresponse and changes in classification between rounds, with fewer errors made by females and whites or Asians than by males or blacks or people of "some other race."

Table 14. Summary of Relationship Between CPS Data Quality Indicators and Covariates, Indicating Significant Effects and Direction of Effects

	Item NR	Changes in Classifications Between CPS Rounds	Round Value Reports	Basic CPS – Reinterview Inconsistency
Sex	**	**	n/s	n/s
Race	**	**	n/s	n/s
Proxy Reports (vs. Self)	** (+)	--	** (+)	n/s
HH Size	** (+)	** (+)	** (+)	** (-)
Age	** (+)	** (-)	** (-)	** (-)
Education	** (-)	** (-)	n/s	** (-)
Employment	** (+)	** (-)	n/s	** (-)
Hours Worked	** (-)	n/s	** (+)	** (-)
Marriage	** (-)	** (-)	** (-)	n/s
Children	** (-)	** (-)	n/s	n/s
Item Burden	** (+)	** (+)	** (+)	** (+)
2nd (vs. 1st) Half of CPS	** (+)	** (+)	** (+)	** (+)
CPS Unit NR	** (+)	** (-)	** (+)	n/s
# of Respondent Changes	** (+)	--	** (+)	** (+)

** Significant effect; n/s non-significant effect; (-) variable was negatively associated with measure; (+) variable was positive associated with measure; -- effect was not calculated for the measure

Factors that were likely to reduce respondent knowledge (i.e., reporting for others, larger households) generally were associated with poorer data quality, though household size effects were complicated by the unique properties of single person households and seemed to be driven largely by reporting deficiencies in the largest households. There was mixed support for the notion that response errors stemmed from respondents' cognitive limitations. Better educated respondents did have less item nonresponse and fewer inconsistencies between rounds and in the reinterview than those with lower levels of education, but education was not related to round value reports. Moreover, age was positively related to item nonresponse (older respondents had more missing data than younger respondents), as expected, but its effects on the three other quality indicators were in the opposite direction. Round values and response inconsistencies (both between CPS rounds and between the basic CPS interview and the reinterview) were most common in households in which the main respondent was under the age of twenty-five and least common among older respondents.

The effects of employment and hours worked generally were not supportive of the busyness hypothesis. Four of the six significant effects found for these variables were in the opposite direction predicted (with gains in data quality for respondents who were employed or who worked longer hours). In contrast, social integration factors (i.e., marriage and the presence of children in the household) generally were associated with better data quality. The largest and most consistent effects were found for two indicators of survey burden—item burden and CPS time-in-sample were positively related to all four data quality indicators. CPS unit nonresponse and the number of respondent changes across rounds generally had negative effects, as well.

4.4.6 Relationship of CPS Data Quality Indicators

Table 15. Correlations Between CPS Data Quality Indicators

	Round Value Reports	Changes in Classification	Reinterview Inconsistency
Item Missing	-0.11**	-0.12**	0.01
Round Value Reports		-0.01*	-0.13**
Changes in Classification			0.00

** Significant at $p < .001$. * Significant at $p < .05$.

The last step in the CPS analyses was to examine the relationship of the individual data quality measures. The zero-order correlations between the four CPS data quality indicators are presented in Table 15. In general, the indicators are only weakly and negatively correlated with one another. The negative association between item nonresponse and classification changes between rounds is due in part to the fact that the latter measure only examined pairs of valid responses. Items for which there was missing data in any given pair of CPS waves were not included in the classification change measure (i.e., I did not examine instances in which respondents gave a "don't know" or "refused" answer to a question in one month and then a valid response to that question in another month). Although I restricted my analyses of between round changes to items that were not particularly susceptible to missing data (e.g., race, education), item nonresponse on these variables would have attenuated the correlation between these two data quality indicators. The strongest relationship shown in Table 15 is between round value reports and interview-reinterview inconsistencies, suggesting that individuals who provide round values in the basic

CPS interview are unlikely to provide more specific responses for these items when they are reinterviewed.

4.5 ATUS Results

4.5.1 Missing or Poor Data

The item missing rate for the eleven labor force questions re-asked in ATUS ranged from zero percent (Do you own a business or a farm?) to seven percent for earnings. The two questions that asked respondents to report the time their first child awoke and their last child went to bed had the highest nonresponse rate (11.0% and 7.8%, respectively), though only about one-third of the sample were eligible for these questions. Aggregating across questions, the mean percent item missingness for ATUS respondents for these non-time-diary items was 6.5%. ATUS interviewers also record the number of times during the time-diary that respondents reported that they could not recall an activity. This occurred in 4% of ATUS cases.

In addition to these direct measures of item nonresponse, I examined time-diary records to determine if certain basic activities were missing. In a given day, most people sleep, eat, and perform personal care activities (e.g., grooming, dressing, going to the bathroom). When diaries do not contain one or more of these basic activities in the 24 hour period, it may be an indication that respondents intentionally omitted some behaviors or simply did not try to report their activities accurately. For each person, I coded the number of times these basic activities were reported, and flagged cases for which there was no data. A surprisingly large number of people (31.6% of ATUS respondents) failed to report at least one of these activities. Multivariate analyses revealed that missing activity reports increased with age and decreased with education, echoing earlier findings regarding role of cognitive limitations in data quality measures. Missing activity reports also were most likely to occur for males than females and for blacks than respondents of other races. Corroborating the idea that response errors are negatively related to social capital, unmarried respondents and those without children were less likely to report basic daily activities than those who were married or parents. In addition, I found that respondents who worked more than forty hours per week were more likely than those who did not to have missing activities, though employed respondents in general gave more complete reports than unemployed respondents. Finally, missing activity reports were more likely to occur when ATUS respondents also had been the respondent in the last CPS interview, providing some evidence for the harmful effects of survey burden.

Finally, there were a very small number of ATUS cases each month that were deemed of insufficient quality to include in the microdata files. In 2003, 668 cases, or 3.2% of the total number of respondents in that year, were discarded for this reason. These cases are of such poor quality that no substantive data analyses can be performed. However, ATUS interviewers are asked to identify cases during data collection in which they believe that the respondent is deliberately falsifying information, attempting to respond but unable to correctly recall activities, or deliberately reporting very long durations. 1.2 percent of 2003 ATUS cases were flagged with this code.

4.5.2 Round Values

Respondents' earnings, hours worked, and (when applicable) spouse's usual hours worked data were collected separately from the time-diary information. I analyzed these variables to determine how often the reports were round values. As was the case for similar items in the CPS, ATUS reports of earnings and hours worked tended to fall around round numbers in a consistent pattern. For example, nearly fifty percent of the earnings reports were round (i.e., in multiples of $50 for weekly and monthly earnings reports, and multiples of $500 for annual earnings reports), and the percent round increased as the reporting period became longer (hourly, weekly, monthly, etc.). The mean percentage of round values across all items and all ATUS respondents was 79.2%.

Another potential source of reporting error is round activity durations. I extracted duration measures for twenty-one different activity categories for each respondent. Activities lasting less that 30 minutes were coded as 'round' if they were divisible by 10; activities lasting between 30 and 60 minutes were coded as 'round' if they were divisible by 15; 61 to 180 minute activities were coded as 'round' if the were divisible by 30; and activities lasting longer than 180 minutes were coded as 'round' if they were divisible by 60. Segmenting durations in this way was done to accommodate changes in granularity (i.e., units of time) in respondents' reports over activities of different lengths.

Figure 2. Reported Durations of Select ATUS Diary Activites

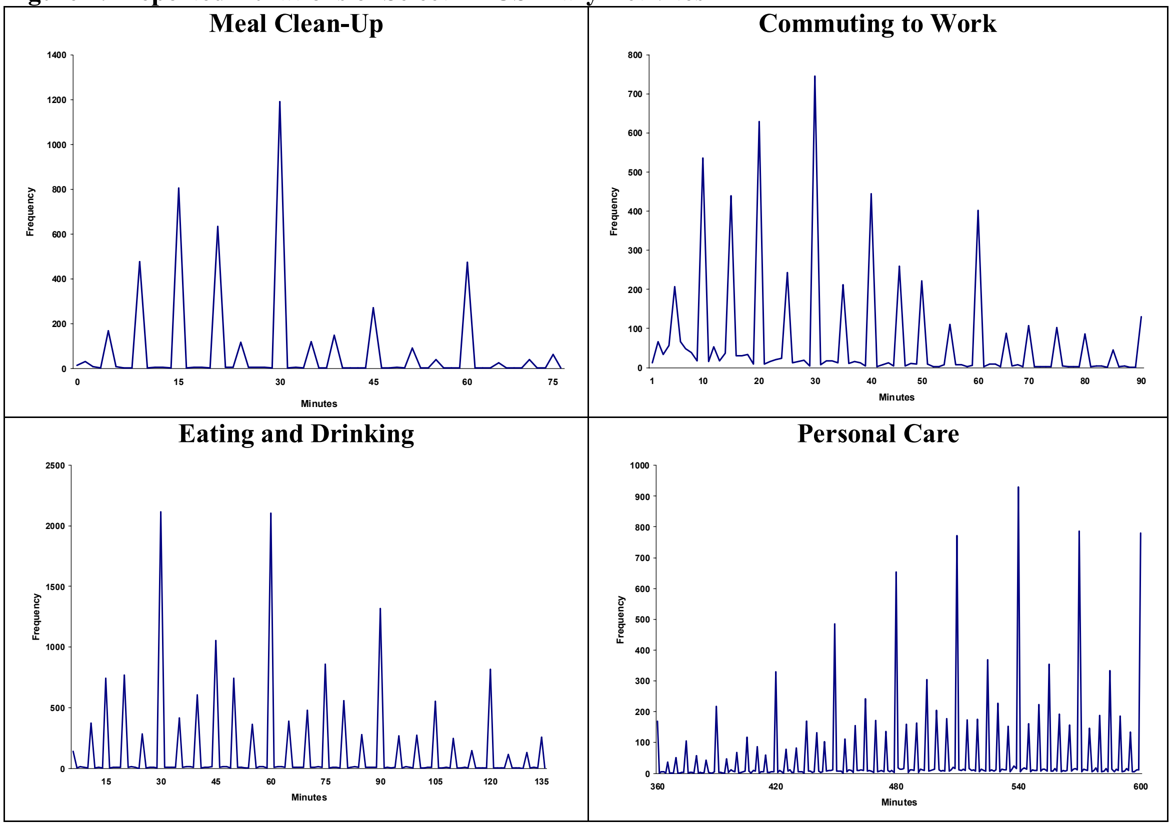


Figure 2 presents examples that illustrate the occurrence of round durations in four different activities. Table 16 presents the percent of reports that were round values for each of the ATUS diary activities examined. As can be seen in the table, there is significant variability in round reports across the different activities, ranging from about twenty percent round (e.g., personal care) to over sixty percent round (e.g., exercise). Aggregating across items, on average 59.2% of reported activity durations were round values. Results in the first and last several rows of the table indicate that in general activity duration was negatively related to round value reports. The distribution in the last panel of Figure 2 (i.e., reported durations for personal care activities), however, shows that respondents provided significant round value reports even for long activities, but that these included frequent reports using both multiples of sixty minutes (which contributed to the round value measure) and thirty minutes (which were not counted as "round" in the present analyses).

Table 16. Percent Round Durations, Mean Durations and Totals for Select ATUS Activities

Activity	% of Reports That Are Round	Mean Duration (min)	Total
Personal Care – any	20.2%	569.2	20,704
Work – at work only	21.2	430.0	7,639
Work – any	22.0	421.0	7,987
Leisure – any	22.7	301.9	19,714
Leisure - relaxing	23.9	288.0	19,540
Travel – any	27.4	87.5	17,954
Personal Care – private	28.1	22.3	783
Personal Care – sleep/rest	29.3	523.8	20,685
Caring for HH Child	29.9	117.3	5,497
Caring for non-HH child	34.2	86.3	1,196
Religious Activity	37.7	113.6	2,303
Leisure – TV	38.1	204.4	16,282
Volunteering	43.2	131.1	1,418
Travel – commuting	43.7	42.8	6,520
Eating/Drinking	48.7	73.0	18,911
Tobacco/Drug use	49.3	23.5	304
Telephone	51.1	43.6	3,658
Meal Preparation	55.2	52.6	10,328
Civic Participation/Activities	58.1	48.7	105
Exercise – active participation	61.5	103.4	3,475
Meal Clean-up	63.1	34.0	5,125

To a certain extent, the same issue also applies to shorter activities (e.g., for activities lasting between one and three hours, respondents frequently reported durations in 15 minute increments though these were not counted as "round"). This point underscores the difficulty of operationalizing measures of "round" reports in the absence of true values. Nevertheless, I selected criteria for this measure that I believed would minimize detection of 'false positives' (i.e., 'round' reports that also happened to be true reports) and at the same time allow me to identify probable reporting errors.

The next step in the analyses was to examine the effects of potential causal variables on rounding. I regressed respondents' percent round duration estimates on the variables listed in Table 2. The results of the model (F=54.7, df = 15, 20,681; $p < .001$) are presented in Table 17. There were clear effects of cognitive ability—older respondents and those with lower levels of educational attainment reported more round values than younger and better educated respondents. Contrary to the busyness account, respondents who were employed or worked more than forty hours per week reported fewer round values than the unemployed or those out of the labor force and those who worked less than forty hours per week. Being married was positively associated with percent round reports—married respondents had significantly more round durations than unmarried respondents—and the presence of children was not significantly related to round reports; both of these findings contradict social capital predictions. Finally, the percentage of reports that were round durations was not affected by whether the ATUS respondent was the same person who had responded in the outgoing CPS round.

Table 17. Results of Model Predicting ATUS Percent Round Value Reports by Possible Causal Variables

Effect	Estimate	t	F	Sig
Sex			0.76	.3842
Male vs. Female	0.25	0.87		.3842
Race			14.29	<.0001
Black vs. White	0.18	0.35		.7284
Asian vs. White	5.34	6.52		<.0001
Other vs. White	-0.40	-0.40		.6902
Age	0.15	14.62	213.64	<.0001
Education			2.60	.0344
HS only vs. LT HS	-0.87	-1.91		.0562
Some college vs. LT HS	-1.42	-3.11		.0019
BA/BS vs. LT HS	-1.22	-2.44		.0385
Advanced Degree vs. LT HS	-0.59	-0.98		.3258
Employment (ATUS LF item)			82.23	<.0001
Unemp vs. Employed	4.46	5.34		<.0001
NILF vs. Employed	4.22	12.29		<.0001
Hours Worked (ATUS LF item)			4.36	.0367
LE 40 vs. GT 40	1.13	2.09		.0367
Marital Status			33.31	<.0001
Unmarried vs. Married	-1.94	-5.77		<.0001
Presence of Young Child			0.27	.6057
No vs. Yes	-0.19	-0.52		.6057
Same R as CPS Wave 8			0.30	.5846
Yes vs. No	0.17	0.55		.5846

4.5.3 Diary Activity Reports

ATUS respondents reported an average of nineteen diary activities. The median activity duration was 75.8 minutes (mode = 80, mean = 86.7), and the entire ATUS interview took an average of 29.2 minutes to complete.

To assess whether the burden of completing the time-diary reduced the number of activities reported later in the interview, I examined the number of activities reported before and after noon (12 p.m.), before and after work (based on respondents' longest work duration or their first and last reported work activities), and between 6:00 and 9:00 a.m. versus between 6:00 and 9:00 p.m. In each instance, however, more activities were reported in the second block than in the first, suggesting that activity reports were not trailing off due to respondent fatigue.

I next examined the effects of variables related to respondent cognitive ability and motivation on the total number of reported activities. One additional variable on the ATUS data file was included in this analysis. On average, respondents indicated (through the "who were you with" probe) that they were by themselves during 42.0% of their reported activities. Since respondents are prompted to provide "who" information for each activity they report, reporting the presence of others incrementally adds to the burden of the survey, so fewer "who" reports (i.e., saying that one was alone during the activity) may indicate survey satisficing. Therefore, I recoded this variable into terciles, controlling for the number of activities reported and for household size, and examined its effect along with the other covariates listed in Table 2 in a multiple regression model predicting the total number of activities reported.

Some of the predictor variables in he regression model are directly related to the kinds and amounts of activities individuals engage in. For example, marriage and children likely will increase the number of reported activities irrespective of the effects of these factors on social capital. Similarly, working longer hours will reduce the number of reported activities since ATUS does not collect activity reports during work episodes. Thus, respondent knowledge and motivation as operationalized in this model may be confounded with the behavioral activities on which the dependent measure is based, and the results are less informative with respect to ATUS data quality than those from previous analyses.

Table 18 shows the effects of model variables on the total number of ATUS diary activities reported. The overall fit of the model was significant (F=150.98, df = 17, 20,680; p < .001), with an r-square value of .11. Consistent with the findings from analyses of ATUS round values, data quality (i.e., more reports) was negatively associated with age, positively associated with education, and unrelated to CPS respondent status. Also consistent were findings that significantly more diary activities were reported for females than males, and for whites and those of 'some other race" than blacks and Asians. In contrast to the round value results, a greater number of activities were reported for respondents who were out of the labor force than for those employed or unemployed, for individuals who worked fewer than forty hours per week, and for those who were married and had children. The hypothesized effects of level of "who" reporting also were confirmed: respondents who gave less detailed "who" reports (again, controlling for total number of activities reported and household size) also provided significantly fewer diary activities.

Table 18. Results of Regression Model Predicting Total Number of Reported Diary Activities

Effect	Estimate	t	F	Sig
Sex			841.82	<.0001
Male vs. Female	-3.36	-29.01		<.0001
Race			33.03	<.0001
Black vs. White	-1.68	-8.25		<.0001
Asian vs. White	-1.82	-5.63		<.0001
Other vs. White	0.52	1.33		.1836
Age	-0.02	-4.00	15.99	<.0001
Education			137.13	<.0001
HS only vs. LT HS	1.77	9.87		<.0001
Some college vs. LT HS	2.79	15.45		<.0001
BA/BS vs. LT HS	3.92	19.35		<.0001
Advanced Degree vs. LT HS	4.55	19.17		<.0001
Employment (ATUS LF item)			19.77	<.0001
Unemp vs. Employed	-0.86	-2.60		<.0001
NILF vs. Employed	0.74	5.25		<.0001
Hours Worked (ATUS LF item)			56.55	<.0001
LE 40 vs. GT 40	1.60	7.52		<.0001
Marital Status			20.45	<.0001
Unmarried vs. Married	-0.60	-4.52		<.0001
Presence of Young Child			247.01	<.0001
No vs. Yes	-2.31	-15.72		<.0001
Same R as CPS Wave 8			0.05	.8161
Yes vs. No	-0.03	-0.23		.8161
Number of "Who" Reports			56.64	<.0001
Low vs. Medium	-0.52	-3.77		.0002
High vs. Medium	1.02	7.02		<.0001

4.5.4 Relationship Between ATUS Data Quality Indicators

The final step in the analyses of ATUS data quality indicators was to assess their interrelationship. I first calculated the zero order correlations between the measures, and then examined the effects of partialing out the variables hypothesized to affect response error propensities. In addition to the four main dependent measures (item missing data on ATUS labor force questions, missing activity reports, round durations, and the total number of diary activities reported), I also included the following data quality indicators: round values in ATUS labor force items, “don’t know” responses reported in the ATUS diary, and interviewer codes for potentially bad diary reports.

Table 19. Correlations Between ATUS Data Quality Measures

	Missed Reports	Round Durations	LF Item NR	LF Round Values	“Don’t Know”	Bad Quality
Total Activities	-0.35**	-0.29**	-0.05**	-0.01	0.07**	-0.12**
Missed Reports		0.10**	0.03**	0.00	-0.01	0.10**
Round Durations			-0.04**	0.05**	-0.01	0.02*
LF Item NR				-0.01	0.00	0.02*
LF Round Values					0.00	0.00
“Don’t Know”						0.01
Bad Quality						
Total Activities	-0.34**	-0.31**	-0.06**	-0.02*	0.07**	-0.10**
Missed Reports		0.10**	0.03**	0.00	-0.01	0.10**
Round Durations			-0.02*	0.04**	-0.01*	0.02
LF Item NR				0.00	0.02*	0.03*
LF Round Values					0.00	0.00
“Don’t Know”						0.01
Bad Quality						

** Significant at $p < .001$. * Significant at $p < .05$.

The results of these analyses are presented in Table 19. The various data quality indicators had small to moderate correlations (top panel), and these relationships do not appear to change appreciably when controls are introduced (bottom panel). The total number of ATUS diary reports had a moderate, negative relationship with round diary durations, missed activity reports, and interviewer flags for potential poor data quality. Rather unexpected was the finding that the total number of diary activity reports was positively correlated with the use of "don't know" responses during the time diary. One possible explanation for this finding is that "don't know" responses actually are indicators of respondent effort to recall activities (and to admit when they cannot do so), whereas poor reporters cover up gaps in their memory simply by lengthening activities they can recall or choose to report. Finally, the findings that percent missingness in ATUS labor force items is negatively related to number of activities, and that round values in ATUS labor force items is positively correlated with round diary durations suggests that these may be a useful predictors for ATUS diary response quality, though the relations are quite weak.

4.6 Relationship Between CPS and ATUS Data Quality Indicators

One question is whether the CPS and ATUS data quality measures are related. Table 20, which reports the zero order correlations (top panel) and correlations partialing out the effects of covariates previously identified (bottom panel), shows that the two sets of measures are significantly related but only weakly so.[14] As before, controlling for potential common cause variables did not affect the size or direction of most associations. The strongest relationship was between CPS item nonresponse and ATUS labor force item nonresponse. In addition, there was a small, negative, and consistent relationship between the total number of ATUS diary reports and the three CPS data quality indicators. Two of the three CPS measures also were

[14] The indicator of CPS response inconsistency between the basic interview and CPS reinterview could not be included in these analyses because there were an insufficient number of cases.

positively associated with ATUS interviewer indicators of potential bad diary reports and missed diary activities. A final result worth noting it that the significant negative correlation between ATUS "don't know" responses and CPS round value reports offers some support for the idea that "don't know" responses are signs of respondent effort, though this indicator was not related to the other CPS data quality measures and the effect disappears when controlling for other variables.

Table 20. Correlations Between the CPS and ATUS Data Quality Measures

		Item NR	Changes in Classification	Round Value Reports
ATUS Data Quality Indicators	Total Activities	-0.05**	-0.06**	-0.03**
	Missed Reports	0.02*	0.04**	0.00
	Round Durations	0.00	0.02*	0.02*
	LF Item NR	0.10**	0.00	0.00
	"Don't Know"	0.00	-0.01	-0.02**
	Bad Quality	0.03**	0.02*	-0.01
	Total Activities	-0.04**	-0.06**	-0.02*
	Missed Reports	0.03**	0.04**	0.00
	Round Durations	0.00	0.02*	0.04*
	LF Item NR	0.09**	0.00	0.00
	"Don't Know"	0.00	-0.01	0.00
	Bad Quality	0.03**	0.01	-0.01

** Significant at $p < .001$. * Significant at $p < .05$.

4.7 Discussion

Much of what we know about response errors in surveys comes from studies that have examined respondents' cognitive processes and response strategies when producing survey reports. Models of response processes developed in the last twenty years inform survey researchers about when and how errors arise, and there is a large empirical literature that details these response problems (see, e.g., Tourangeau et al., 2000). Research on response errors in large, on-going surveys tends to be largely atheoretical with respect to these models of response processes, and to be heavily focused on measurable outcomes, in particular on direct measures of bias or response variance for specific items. The goal of the present study was to incorporate theoretical analyses of response processes in an examination of an expanded set of quantifiable measures of survey data quality in two national, household surveys.

The results of these analyses illustrate the value of examining multiple indicators of response quality while taking into account factors that drive respondents' propensity to report accurately (e.g., knowledge, ability, and motivation). Previous studies particularly have neglected to look at the effects of round value reports on measurement error, and only rarely have examined spurious changes in responses that occur between rounds of panel surveys. The typical approach of focusing on merely one aspect of response quality (e.g., item nonresponse) overlooks a large amount of useful information available in the data. The operational and analytic procedures employed in this study enabled me to gain more comprehensive understanding of overall survey quality and to better examine some the assumptions underlying response process models.

There are three main findings from this study. First, I found that data quality was often systematically affected by variables reflecting respondent knowledge, cognitive ability, and motivation, but that these effects differed across the different types of data quality measures. For example, errors in the CPS increased when respondents were less likely to have full and accurate information (e.g., proxy reports and in large households), and when they were poorly educated, less socially integrated, and when survey burden was high (e.g., as time-in-sample or the number of questions asked increased). Similar effects were found in the ATUS analyses for education and social integration, and for one measure of survey burden (respondents who were also the last CPS respondent were more likely to have missed activity reports than those who were not). There were two notable exceptions, however. The effects of busyness apparent in the ATUS analyses of data quality were absent in the CPS. In addition, age was negatively correlated with item missingness in the CPS but was consistently associated with poorer data quality in ATUS.

Second, the individual data quality indicators within each survey were significantly but weakly associated with one another. With the exception of several ATUS measures that had moderate negative correlations because of how they were constructed (e.g., round durations and total activities), the sizes of the associations typically were very small. The negative correlations among CPS measures (at least those that were significant) partially reflect their abilities to capture unique components of the response process (e.g., item nonresponse and round value reports). On the other hand, the indicator of ATUS labor force item nonresponse was positively associated with missed activity reports, and ATUS labor force round values was positively associated with round diary durations, and both may offer means of identifying potentially bad diary reporters. Although I did not examine the relationship of CPS indicators across rounds, I would expect that indicators of response errors from previous rounds could be used similarly.

Third, there were weak correlations between data quality measures in the two surveys. None of the interrelationships between data quality measures within a survey or between surveys went away when controlling for variables that might be common causes of both. Here again, the positive association between CPS item nonresponse and the ATUS labor force item nonresponse, and the significant correlations between all of the CPS data quality indicators and number of ATUS dairy reports may offer a promising tool for identifying potentially poor reporters.

V. The Relationship Between Response Propensity and Survey Data Quality in the Current Population Survey and the American Time Use Survey

5.1 Introduction

An important theoretical question in survey research over the past fifty years has been: How does bringing in late or reluctant respondents affect total survey error? Does the effort and expense of obtaining interviews from difficult to contact or reluctant respondents significantly decrease the nonresponse error of survey estimates? Or do these late respondents introduce enough measurement error to offset any reductions in nonresponse bias?

Evidence from some recent studies suggests that efforts to reduce nonresponse rates have little effect on nonresponse error (Curtin, Presser, and Singer, 2000; 2005; Groves, Presser, and Dipko, 2004; Keeter et al., 2000; Merkle and Edelman, 2002). For example, Curtin et al. (2000) found negligible differences between monthly estimates of consumer confidence derived from a full survey dataset and those derived from a dataset in which hard-to-interview respondents had been removed. Similarly, Keeter et al. (2000) and Merkle and Edelman (2002) found little correlation between low response rates and nonresponse bias.

Much less attention, however, has been given to the relation of response propensity and survey measurement error. In part, this neglect may reflect the assumption that the causes of nonresponse and measurement error are independent. Nonresponse typically is seen as a function of motivational variables (e.g., interest in the survey topic, time spent away from home), whereas measurement error is considered primarily a function of cognitive factors (such as ability). This assumption of independent causal factors may be untenable, however, because the same motivations that affect participation decisions also may affect performance. To the extent that individuals' response propensities are positively correlated with the level of effort that they give during the response process, bringing reluctant individuals into the respondent pool will increase measurement error and reduce the quality of estimates (Biemer, 2001; Groves, 2006).

Relatively few empirical studies have examined the relationship between nonresponse and data quality. Findings from these studies suggest that the relationship depends on the statistic of interest, how measurement error is operationalized, and the type of nonresponse (noncontact vs. noncooperation). Some studies that have examined of indirect data quality indicators (e.g., item nonresponse, response completeness) have found that late responders and initial refusers are more likely than early responders and those not requiring refusal conversion to skip items, give shorter, less informative answers to open ended questions, and provide DK, 'not applicable,' or 'no opinion' responses (e.g., Friedman and Clusen, 2003; Triplet et al., 1996; Willimack et al., 1995). By contrast, Yan et al. (2004) found that other indirect indicators of data quality (e.g., acquiescence, nondifferentiation) were unrelated to response propensity, or were negatively correlated with it—i.e., low propensity groups evinced better data quality than high propensity groups.

Several studies that have looked at direct estimates of measurement error (e.g., those based upon discrepancies between survey responses and administrative records) have found that low propensity respondents tend to provide worse data than high propensity respondents. For example, Cannell and Fowler (1963) found that individuals who responded at the end of the survey field period were 10 – 15% less accurate in their reports of the number and duration of their hospital stays than those who responded earlier. Similarly, Bollinger and David (2001) found that latent ‘cooperative’ respondents were less likely than ‘uncooperative’ respondents to drop out of a panel survey and to make reporting errors. More recently, Olson (2006) examined the separate impact of contact and cooperation propensity on several variables related to marital dissolution (e.g., time since divorce, length of marriage), partialing out the unique contributions of measurement error bias and nonresponse bias. She found that including reluctant respondents increased measurement error for some estimates, but that bringing in hard-to-contact respondents actually led to *decreases* in measurement error and overall error (see also, Voigt et al., 2005). For most of the estimates in her study, however, the resulting changes in measurement error were nonmonotonic across propensity strata and were very small relative to the size of the estimates.

The results of these studies suggest that there may be a relationship between response propensity and data quality, but the nature of that relationship and its causal mechanisms are not well understood. At the very least, these findings challenge the traditional assumption that nonresponse and measurement error are independent. One explanation for covariance between response propensity and data quality is that the relationship results from a cause (or vector of causes) common to both (Groves, 2006). The identification of appropriate common cause factors depends in part upon the particular survey protocol and respondent pool, but several candidates seem likely to apply to a broad range of surveys. For example, topic interest is one possibility. Interest in the survey topic may dispose individuals to agree to a survey request and also stimulate careful processing of the survey items. Alternatively, higher levels of social capital could activate stronger norms of cooperation (producing higher response propensities), and those same norms also could influence respondents' willingness to engage in more effortful response processes. Or, busyness or time-stress could produce a disinclination both to participate and to respond accurately if interviewed. Regardless, identifying and statistically controlling for the appropriate common cause(s) would eliminate the relationship between response propensity and data quality and provide a means for removing bias.

The purpose of this chapter is to explore the relationship between nonresponse and measurement error, and if there is evidence of covariation, to examine potential common causal factors. These issues will be investigated using data from two national, household surveys—the Current Population Survey (CPS), and the American Time Use Survey (ATUS). Response propensity scores and data quality indicators developed and presented for these surveys in Chapters 3 and 4, respectively, will serve as the basis for these analyses.

5.2 Data and Methods

5.2.1 CPS

The dataset used to create the CPS propensity scores and data quality indicators contained 97,053 households that were eligible for all eight CPS waves between May, 2001 and October, 2003. A single estimate of overall nonresponse propensity (not separating out noncontact and noncooperation) was obtained for each household using a logistic regression model predicting the probability that the unit would be a nonrespondent in any of the last six CPS rounds. Predictors in this model included level of effort (e.g., call attempts) and demographic control variables, as well as variables related to busyness and social capital constructs (see Table 7, Chapter 3 for the final CPS model specifications). On the basis of their predicted probabilities of nonresponse, households were divided into propensity quintiles which ranged in average nonresponse propensity from 1% for the low propensity group (Group 1) to 30% for the high propensity group (Group 5). Each CPS propensity quintile consisted of approximately 19,400 households.

Three household-level data quality indicators were derived for each CPS household: (1) percent item nonresponse, (2) percent round value reports, and (3) percent classification changes between CPS rounds. The percent household item nonresponse was calculated by as follows. $\mathrm{P}_{\mathrm{J\_INR}} = \sum_{k=1}^{W_k} \left(\sum_{i=1}^{P_i} \frac{m_{ijk}}{m_{ijk} + n_{ijk}} \right)$, where m_{ijk} is the total number of missing responses for person i in household j in wave k, and n_{ijk} is total number of non-missing responses for person i in household j in wave k, summing across all household members for all waves in which the household responded to the CPS. The percent round value reports and percent between-round classification changes were calculated in a similar manner, except classification changes were summed over wave-pairs rather than waves.[15] In addition, a fourth data quality indicator—percent inconsistent responses between the main CPS and the CPS reinterview—was created for the 3,851 households on the dataset that participated in the CPS reinterview program.

[15] A classification change indicates that a respondent provided different answers to the same question asked in adjacent waves (or between waves 1 and 4). The variables examined for this indicator were race, educational attainment (restricted to individuals 30 years of age or older), housing tenure, and family income.

Each of these data quality indicators was examined to see if it was related to likelihood of CPS nonresponse. I first analyzed the four indicators across propensity strata to assess the relative size and direction of the association. I then explored the extent to which controlling for potential common cause variables affected the association between indicators of data quality and nonresponse propensity. Finally, I repeated these analyses using CPS sample members' actual response status in rounds three through eight—i.e., whether they participated in all six rounds or were a nonrespondent in at least one of those rounds—to examine the relationship between observed CPS nonresponse (rather than respondents' nonresponse propensity) and CPS data quality.

5.2.2 ATUS

The dataset used to create the ATUS propensity scores had 25,778 records from individuals selected to participate in ATUS between January and December 2003. As in the CPS analyses, a logistic model was used to estimate a nonresponse propensity score for each ATUS sample member, and then ATUS respondents were grouped into quintiles based on these propensity scores (see Table 10, Chapter 3 for final model specifications). ATUS propensity groups were ordered from low nonresponse propensity (Group 1, $\overline{\hat{p}} = 10.6\%$) to high nonresponse propensity (Group 5, $\overline{\hat{p}} = 54.9\%$), with approximately 3,275 cases in each group.

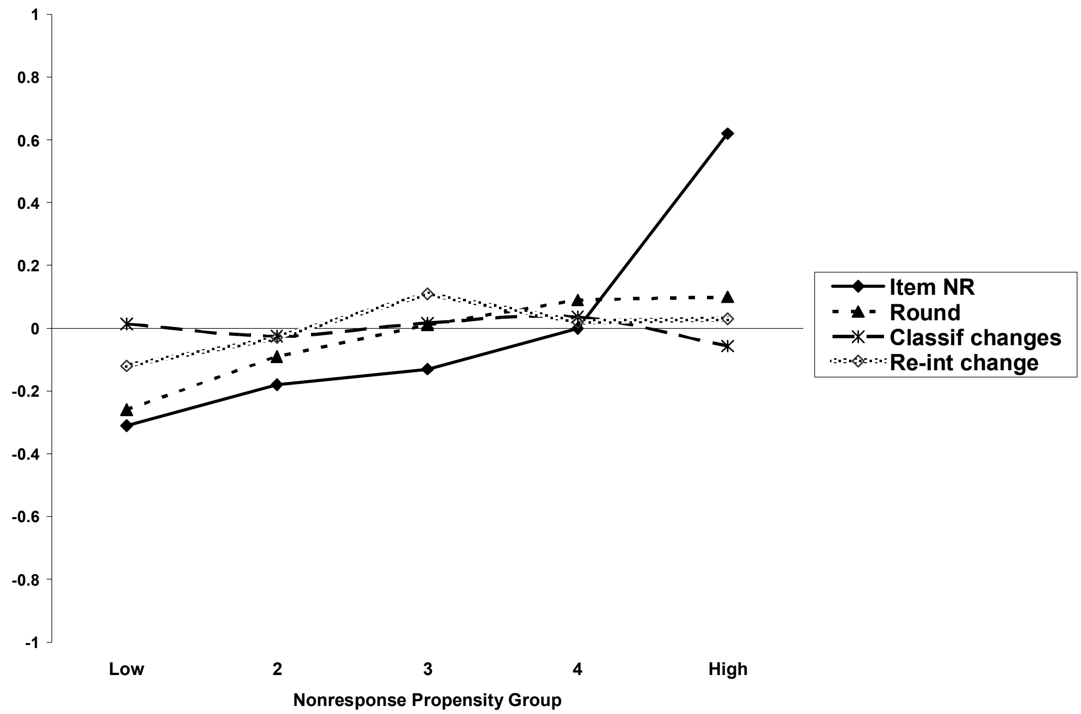
For each of the 20,698 individuals on the dataset who participated in ATUS, I created four data quality indicators: (1) total number of diary activities reported, (2) missing diary reports of basic daily activities, (3) round values for activity durations, and (4) item nonresponse on ATUS labor force questions. As in the CPS analyses, I began by examining the means for the four indicators across propensity strata, and then assessed the effects of controlling for potential common cause variables. I then conducted parallel analyses to examine how ATUS data quality varied as a function of nonresponse in the CPS and refusal conversion in the ATUS. I also analyzed the association between CPS data quality indicators and ATUS response status to see if poor response quality on the CPS was associated with ATUS nonresponse.

5.3 Results

5.3.1 CPS

Figure 1 presents the relationship between the CPS data quality indicators and CPS nonresponse propensity. The graph displays five nonresponse propensity strata, with likelihood of nonresponse increasing from left to right along the x-axis. In addition, the figure presents data quality indicators that have been standardized into standard deviation units in order to make it easier to compare the relative strength of each measure's association with propensity. Regression models were run (regressing the individual indicators on nonresponse propensity) to obtain slope estimates and significance tests (ANOVA models also were run to check for nonlinear trends).

Figure 1. Relationship of CPS Data Quality Indicators (in standard deviation units) to CPS Nonresponse Propensity



There are two main points to take away from this figure. First, the overall quality of CPS reports appears to decrease across nonresponse propensity strata. Taking the mean data quality score within each strata (i.e., averaging across the four indicators), we see that there is a monotonic increase in error as nonresponse propensity rises. Second, the strength of the covariance between propensity and error is highly dependent upon the type of data quality indicator. The relationship is strongest for item nonresponse ($\beta = .17$, $p < .001$): households with the highest probability of nonresponse had item missing rates that were almost a full standard deviation (or about six percentage points) higher than households with the lowest nonresponse propensity. Round value reports also were significantly related to nonresponse propensity, though the strength of the association was about two-thirds that of item nonresponse (β = .11, p < .001). The highest nonresponse propensity households provided about 10% more round value reports than the lowest propensity households. In contrast, nonresponse propensity was only weakly associated with the percent inconsistent reports between the basic CPS and reinterview (β = .07, p = .021), and in fact was slightly negatively correlated with the measure of between-wave classification changes (β = -.05, p < .001).[16]

Why might item nonresponse and round value reports be related to the level of nonresponse propensity? The common cause hypothesis suggests that this relationship may result from a shared explanatory factor (or factors). If the common cause model is correct, and the model is correctly specified with the appropriate variable(s), then the relationship between response propensity and data quality will be eliminated once the common causes are statistically controlled.

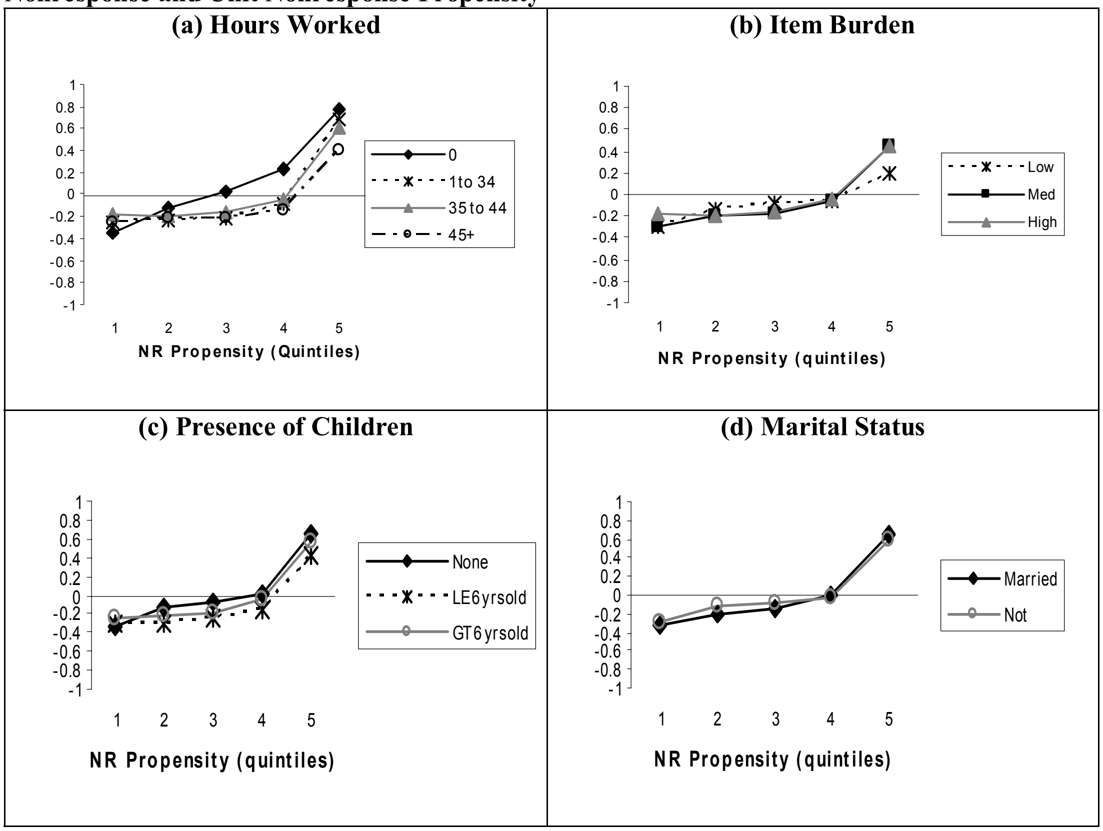
To test the common cause hypothesis, I examined several factors that potentially could contribute to both the likelihood of unit nonresponse and measurement error. Busyness and social capital, as discussed earlier, are two possible common cause candidates, and I included them in the present analyses. A third possibility is survey burden. In a panel survey like the CPS, the level of burden respondents' experience in one wave may affect both their likelihood of response in subsequent waves and their willingness to answer fully and accurately if they do participate. Since I did not have direct measures of these three factors, I examined a number of indicators for each construct. Hours worked and commute time served as

[16] Given the relatively small sample size of the reinterview dataset, and the fact that the between-wave classification change estimate itself likely had significant error (since some 'true' change could occur between waves for some of the variables used in this measure), it is not surprising that these two indicators proved less strongly and consistently related to nonresponse propensity.

indicators of busyness. For social capital, I examined marital status, home ownership, the presence of children in the household, and educational achievement in the community. Item burden (i.e., the number of items asked during the first two CPS waves) served as the measure of survey burden.

I began by looking at the effects of each of these variables individually on the association between nonresponse propensity and the two indicators of data quality that showed the strongest association with nonresponse propensity (item nonresponse and round value reports). If the covariance evidenced in Figure 1 is a direct effect of one of these common cause variables, then we would expect the covariance to diminish or go to zero after controlling for that variable. However, I found no evidence that busyness, social capital, or survey burden (at least as operationalized here) had any mediating effect on the relationship between propensity and data quality. This is illustrated in Figure 2 for the item nonresponse measure and in Figure 3 for round value reports. These figures reveal that the level of reporting error continued to covary with nonresponse propensity, even after taking into account measures of busyness (top-left panel), survey burden (top-right panel), and social capital (bottom two panels). The shapes of the curves in these figures are essentially the same as those found in Figure 1, and this finding also was true for the other common cause variables (not presented here) I examined.

Figure 2. Effects of Potential Common Cause Variables on the Relationship Between CPS Item Nonresponse and Unit Nonresponse Propensity



I next ran a simple regression model using nonresponse propensity to predict item nonresponse (or round value reports). I compared the results of this model to those from a series of models that included as a second predictor one of the common cause variables. The results of this analysis confirmed what is visually evident in Figures 2 and 3—that is, controlling for individual common cause variables had little effect on the size or direction of the relationship between nonresponse propensity and data quality. Moreover, this relationship was evident even when more complex, multivariate models were run that controlled for multiple common cause variables simultaneously (e.g., see Table 1).

Figure 3. Effects of Potential Common Cause Variables on the Relationship Between CPS Round Value Reports and Unit Nonresponse Propensity

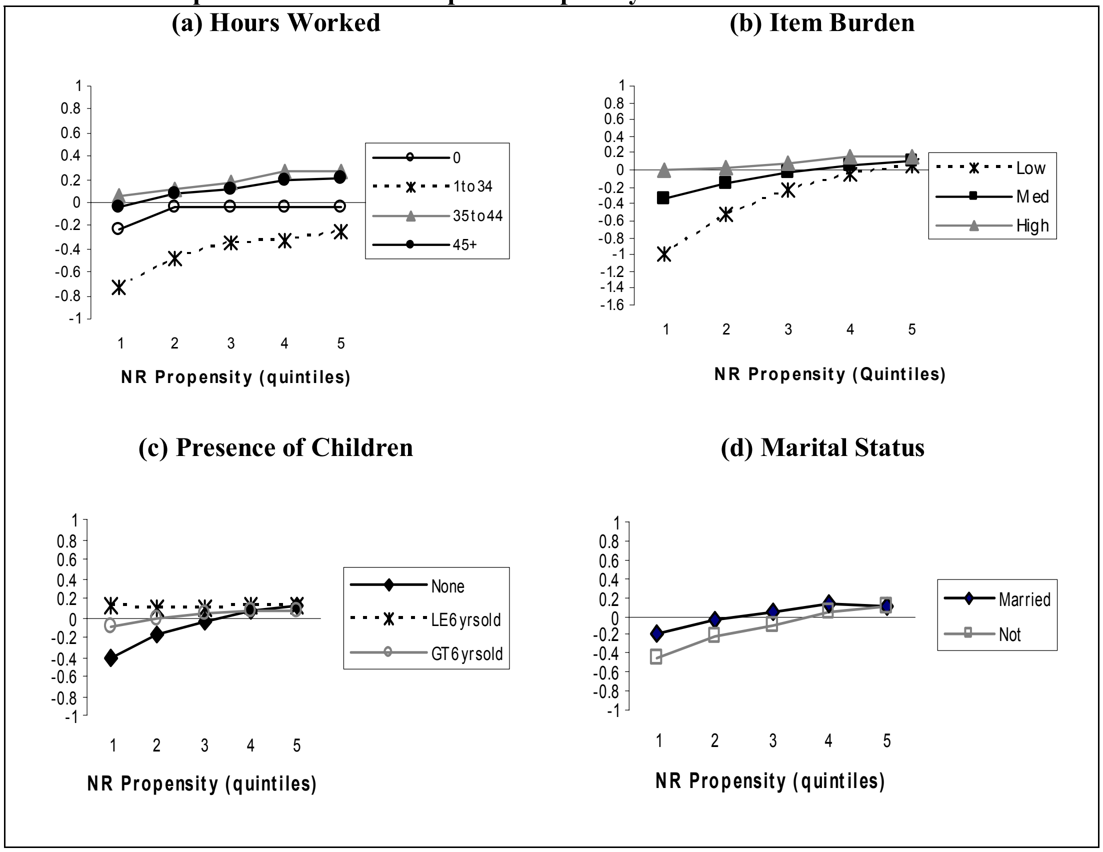
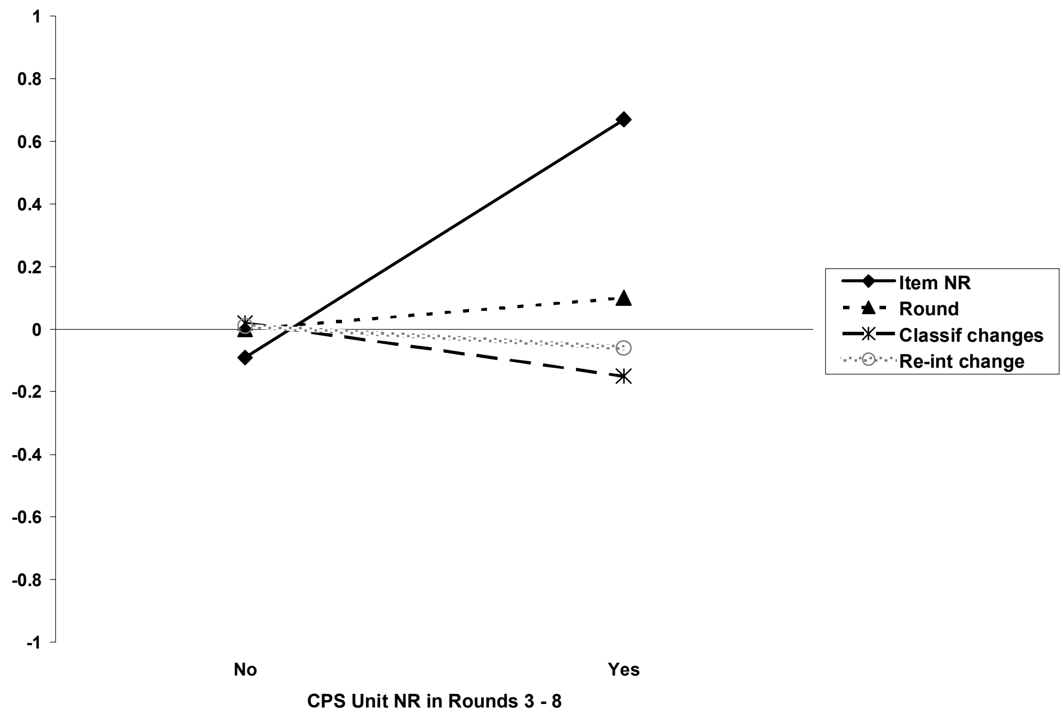


Table 1. Multivariate Regression Model Predicting CPS Item Nonresponse (standardized) from CPS Nonresponse Propensity Group and Potential Common Cause Variables

	Estimate	t	F	Sig
Age	0.01	32.86	1079.59	<.0001
Education			80.10	<.0001
LT HS vs. Advanced	0.18	14.02		<.0001
HS only vs. Advanced	0.16	14.43		<.0001
Some college vs. Advanced	0.10	9.06		<.0001
BA/BS vs. Advanced	0.07	5.42		<.0001
CPS NR Propensity Group (1 – 5)			3011.71	<.0001
2 vs. Lowest	0.29	27.88		<.0001
3 vs. Lowest	0.39	26.33		<.0001
4 vs. Lowest	0.56	49.68		<.0001
Highest (5) vs. Lowest	1.21	103.21		<.0001
Hours Worked	-0.01	-9.27	85.97	<.0001
Item Burden			30.76	<.0001
Low vs. High	-0.05	-5.28		<.0001
Medium vs. High	-0.06	-7.79		<.0001
Marital Status			6.34	.0118
Unmarried vs. Married	-0.02	-2.52		.0118
Presence of Young Child			10.79	<.0001
None vs. Older	-0.04	-4.59		<.0001
Young vs. Older	-0.03	-2.60		<.0001

The preceding analyses revealed a positive relationship between CPS nonresponse propensity and measurement error, but we can also look to see if measurement error varied as a function of *actual* CPS nonresponse. Figure 4 presents the relationship between the standardized measures of CPS data quality and an indicator of whether the household was ever a CPS nonrespondent during its last six months in sample. The effects of actual nonresponse mirror those from the nonresponse propensity analyses. Item nonresponse ($\beta = .75$, $p < .001$) and to a lesser extent round value reports ($\beta = .10$, $p < .001$) were significantly and positively related to actual nonresponse in rounds three through eight, whereas changes in classifications between CPS waves ($\beta = -.17$, $p < .001$) and basic interview-reinterview response inconsistencies (β = -.06, p = .244) were negatively related to nonresponse.

Figure 4. Relationship of CPS Data Quality Indicators (in standard deviation units) to CPS Nonresponse in Rounds 3 – 8.



Controlling for potential common cause variables—both individually and in multivariate analyses—had little effect on the associations apparent in Figure 4. Figure 5 shows the effects of potential common cause variables on the covariance between CPS unit nonresponse in waves three through eight and CPS item nonresponse across waves. There continued to be a significant, positive association between unit nonresponse and item nonresponse even in the presence of the common cause factors. Multivariate regression analyses revealed that this association remained when the common cause variables were statistically controlled (see Table 2), though this model accounted for less variance than the same model fit with CPS nonresponse propensities (r-square = .068 vs. .125, respectively).

Figure 5. Effects of Potential Common Cause Variables on the Relationship Between CPS Item NR and Actual CPS Unit NR

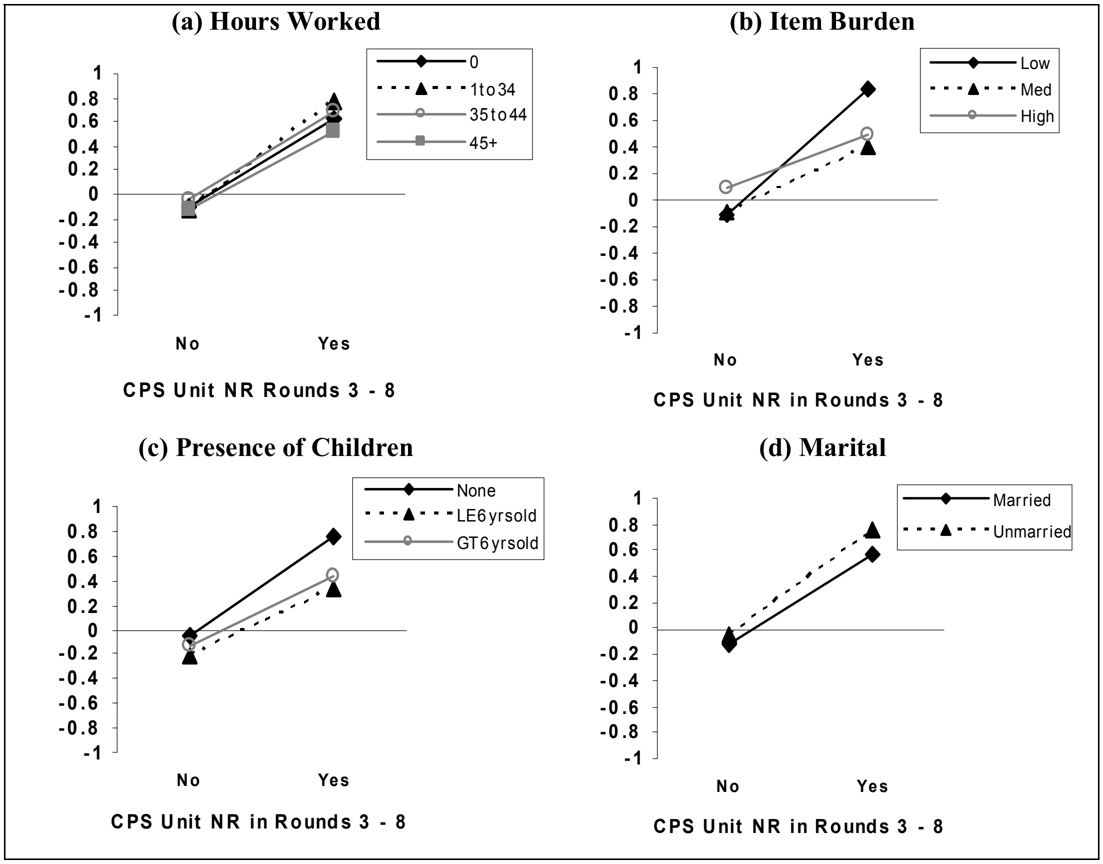


Table 2. Multivariate Regression Model Predicting CPS Item Nonresponse (standardized) from CPS Unit Nonresponse in Rounds 3 – 8 and Potential Common Cause Variables

Effect	Estimate	t	F	Sig
Age	0.01	6.00	36.02	<.0001
Education			29.96	<.0001
LT HS vs. Advanced	0.12	8.71		<.0001
HS only vs. Advanced	0.09	7.67		<.0001
Some college vs. Advanced	0.04	3.26		.0011
BA/BS vs. Advanced	0.04	3.51		.0005
CPS NR Waves 3 - 8			5425.18	<.0001
No vs. Yes	-0.74	-73.66		<.0001
Hours Worked	-0.0	-2.60	6.75	.0094
Item Burden			53.62	<.0001
Low vs. High	-0.08	-10.35		<.0001
Medium vs. High	-0.06	-5.94		<.0001
Marital Status			83.04	<.0001
Unmarried vs. Married	0.07	9.11		<.0001
Presence of Young Child			233.68	<.0001
None vs. Older	0.14	15.28		<.0001
Young vs. Older	-0.08	-6.96		<.0001

5.3.2 ATUS

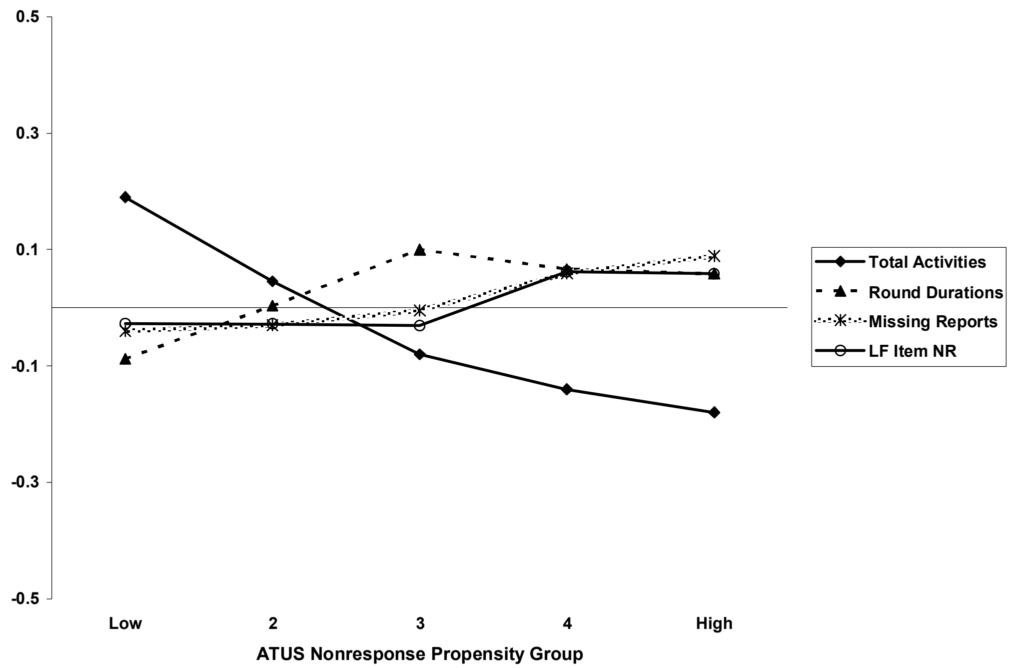
I began by looking at the relationship between CPS data quality indicators and ATUS response status to see if the CPS measures could be used as an indicator of potential ATUS unit nonresponse. Table 3 presents the weighted mean percents of the CPS data quality indicators for ATUS respondents and nonrespondents, the associated *t* and *p* values, and zero order correlations between ATUS response status and each of the CPS measures. The strongest effect was for CPS item nonresponse—ATUS nonrespondents had significantly higher CPS item missing data rates than ATUS respondents. ATUS nonrespondents also had significantly more round values in their CPS answers than ATUS respondents, but the relative difference between these groups in round reporting was quite small given the large amount of round reporting overall, and the correlation of round reporting with ATUS response status was considerably smaller than that for item nonresponse. In addition, there was a small, negative correlation between ATUS response status and CPS between-wave changes in classification—ATUS nonrespondents had fewer between-wave changes than ATUS respondents. When I examined the small number of ATUS cases that also had been in the CPS reinterview program, there were no differences between ATUS respondents and nonrespondents in the amount of CPS interview-reinterview response inconsistencies.

Table 3. Relation of CPS Data Quality Indicators to ATUS Outcome

CPS DQ measure	**ATUS Respondent**	**ATUS Nonrespondent**	**T-value**	**P-value**	**Correlation w/ ATUS nr**
Item missing rate	2.29%	4.05%	-27.83	0.0001	0.19
Round value reports	72.99	74.7	-5.62	0.0001	0.08
Change in classifications (basic CPS)	7.04	6.75	2.01	0.0445	-0.02
Inconsistent reports (CPS reinterview)	12.92	13.09	-0.18	0.8605	0.01

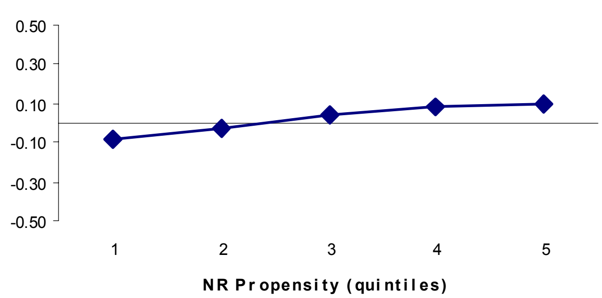
Figure 6 presents the relationship between ATUS data quality indicators and ATUS nonresponse propensity. As before, the graph shows nonresponse propensity increasing from left to right along the x-axis, and data quality indicators are presented in standard deviation units. For three of the measures—round activity durations, missing activity reports, and labor force item nonresponse—points above the zero deviation line indicate poorer data quality; points below the zero deviation line indicate better data quality. For the total number of diary activities reported, however, this is reversed—points above the zero deviation line indicate that respondents reported more than the average number of activities; points below the line indicate that they reported less than the average.

Figure 6. Relationship of ATUS Data Quality Indicators (in standard deviation units) to ATUS Nonresponse Propensity



As can be seen in the figure, there is a linear trend between nonresponse and overall data quality in ATUS. If we aggregate the standardized scores from the four data quality indicators within each nonresponse strata (after flipping the signs for the total activity measure), we see that error increases with nonresponse propensity (see Figure 7). However, the size of this effect is very small—less than .2 standard deviations separate the lowest and highest propensity groups.

Figure 7. Average Error by ATUS Nonresponse Propensity Group



This reflects the relatively weak correlations between nonresponse propensity and the individual data quality measures. Although each is positively (and significantly) related to nonresponse propensity, the only effect with any practical significance is for the total number of diary activity reports ($\beta = .05$, $p < .001$). Respondents in the highest nonresponse propensity group reported about three fewer diary activities than respondents in the lowest propensity group, which amounts to roughly 15 percent of the typical number of activities reported (20). When we couple this fact with the finding that respondents in the high nonresponse propensity group also are more likely than other sample members to report activities in round time blocks, neglect to report basic daily activities, and provide incomplete data on ATUS labor force items, it raises questions about the impact of including these individuals on ATUS estimates.

Having demonstrated a significant covariance between nonresponse and total activity reports, I next examined the effect on this relationship of controlling for a number of potential common cause variables. These analyses proceeded as they had for the CPS data, with one exception. Because ATUS is a one time survey in which only a very small number of standardized items are asked prior to the time diary, I could not use ATUS item burden as a measure of overall survey burden. However, ATUS respondents still may feel burden resulting from their experience in previous rounds of CPS. With this in mind, I looked at the effects of the number of items asked in the last CPS interview and the extent of CPS unit nonresponse as indicators for survey burden.[17]

[17] O'Neill and Sincavage (2004) report that a common reason given for ATUS nonresponse is survey fatigue related to CPS participation. I expected that fatigue would be greater for individuals who participated in all eight rounds of CPS, and that these individuals may be more likely to be ATUS nonrespondents and to provide poor ATUS data than those who participated in fewer CPS rounds.

Figure 8. Effects of Potential Common Cause Variables on the Relationship Between the Number of ATUS Diary Reports and ATUS Nonresponse Propensity

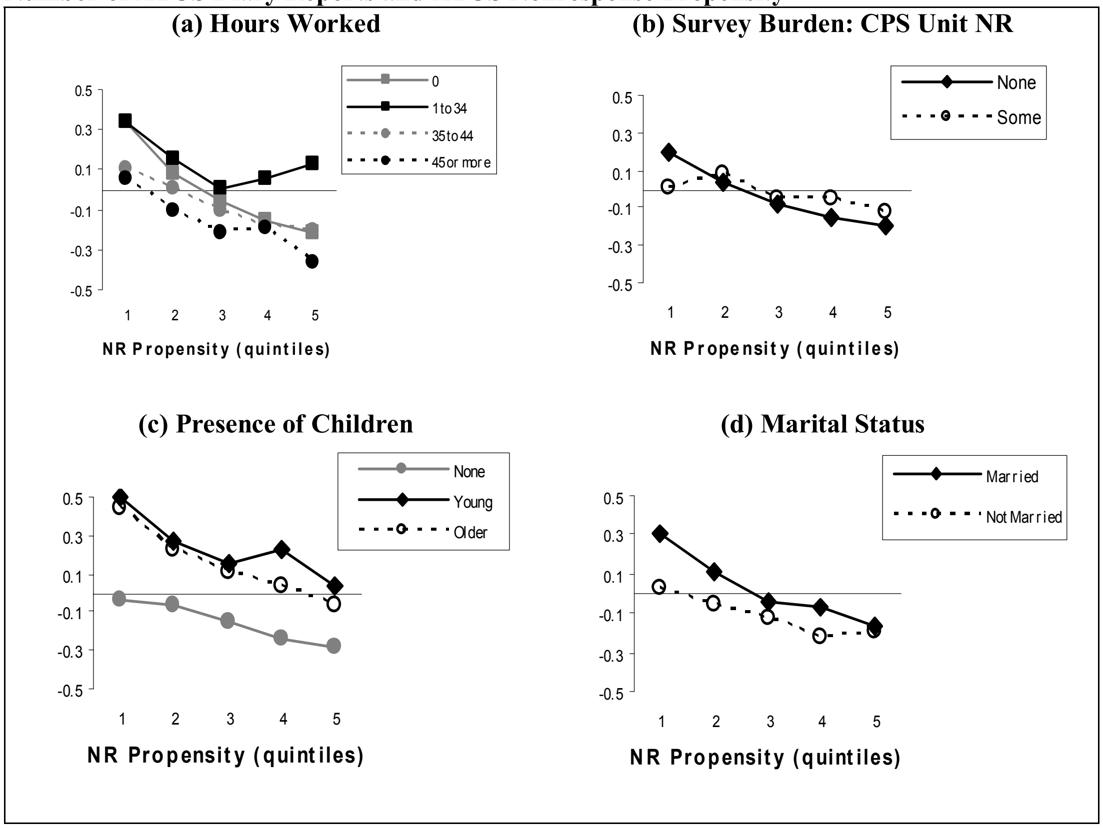
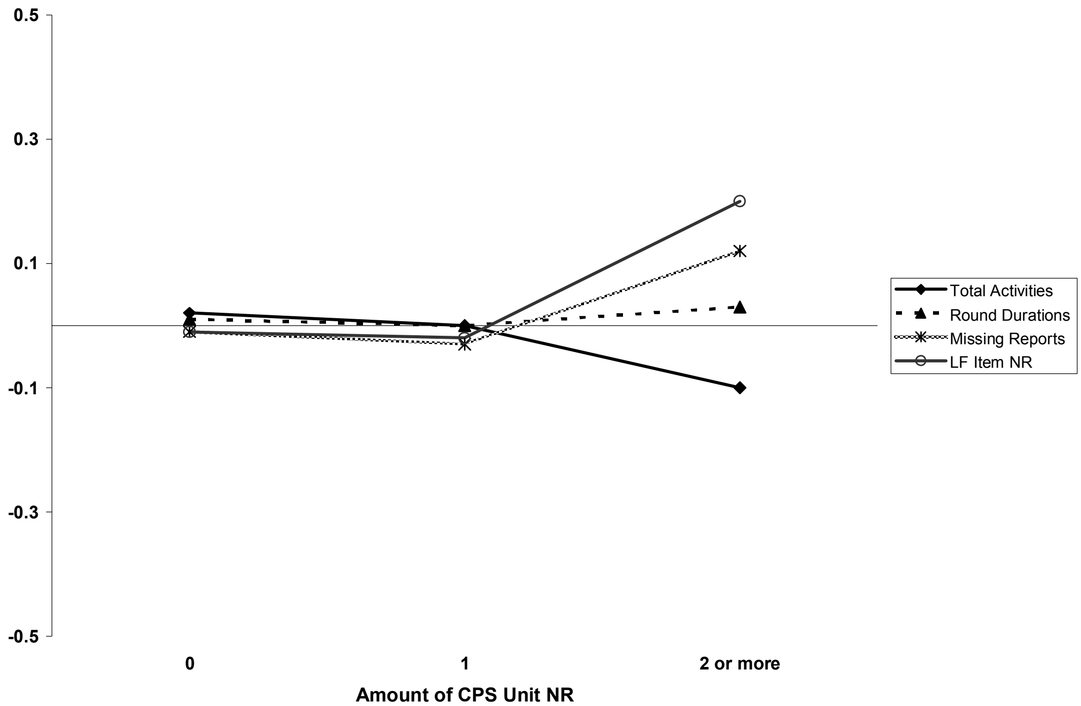


Figure 8 presents several examples that illustrate the effects of selected potential common cause variables on ATUS nonresponse propensity and number of reported diary activities. None of the common cause variables examined (including those not presented here) significantly weakened the covariance of these two factors. The overall magnitude and direction of the relationship between nonresponse and total activities was very similar to that shown in Figure 6. This finding was corroborated by results of regression analyses that controlled for the common cause variables individually and then multivariately. Nonresponse continued to be significantly related to the total number of items reported in the ATUS time diary even when busyness, social capital, and survey burden variables were taken into account. Fewer activities were reported by respondents with high nonresponse propensities, and for those without children; the number of activity reports also was negatively correlated with hours worked and positively correlated with educational attainment.

I next carried out parallel analyses that examined the association between ATUS data quality and two alternative indicators of ATUS response propensity. I first examined CPS unit nonresponse, which was shown to have a large, negative effect on ATUS response propensities in analyses presented in Chapter 3. To examine the relationship between CPS unit nonresponse and ATUS data quality, I created a variable to indicate if the ATUS respondent participated in all eight rounds of CPS (92.7%), failed to participate in one CPS round (5.5%), or failed to participate in at least two CPS rounds (1.8%).[18] Figure 9 shows the relation of this variable to the four ATUS data quality indicators. No difference was found in data quality between ATUS respondents who participated in every CPS interview and those who were nonrespondents in a single CPS round. However, ATUS respondents who failed to participate in two or more rounds of CPS provided poorer ATUS data—more round durations, missed diary activities, and item nonresponse on ATUS labor force questions, and fewer diary reports overall—than those who always participated in CPS or those who were only nonrespondents in one round. Regression analyses run on the individual data quality indicators revealed that only labor force item nonresponse and total reported activities were significantly related to CPS unit nonresponse.

[18] All ATUS respondents participated in wave eight CPS, by definition.

Figure 9. Relationship of ATUS Data Quality Indicators (in standard deviation units) to the Amount of CPS Unit Nonresponse



Figures 10 and 11 illustrate the effects of controlling for common cause variables on the relationship between CPS nonresponse and total reported diary activities (Figure 10) and item missing data rates for ATUS labor force questions (Figure 11). As before, data quality and nonresponse covaried even after controlling for potential common cause variables. In general, the pattern of results evident in Figure 9 also is apparent in Figures 10 and 11, though the effects of some of the variables are not in the expected direction (e.g., the effects of hours worked and the presence of children on the total number of diary activities; the effect of CPS item burden on ATUS labor force item nonresponse).

Figure 10. Effects of Common Cause Variables on the Relationship Between the Amount of CPS Unit Nonresponse and Number of Reported ATUS Diary Activities

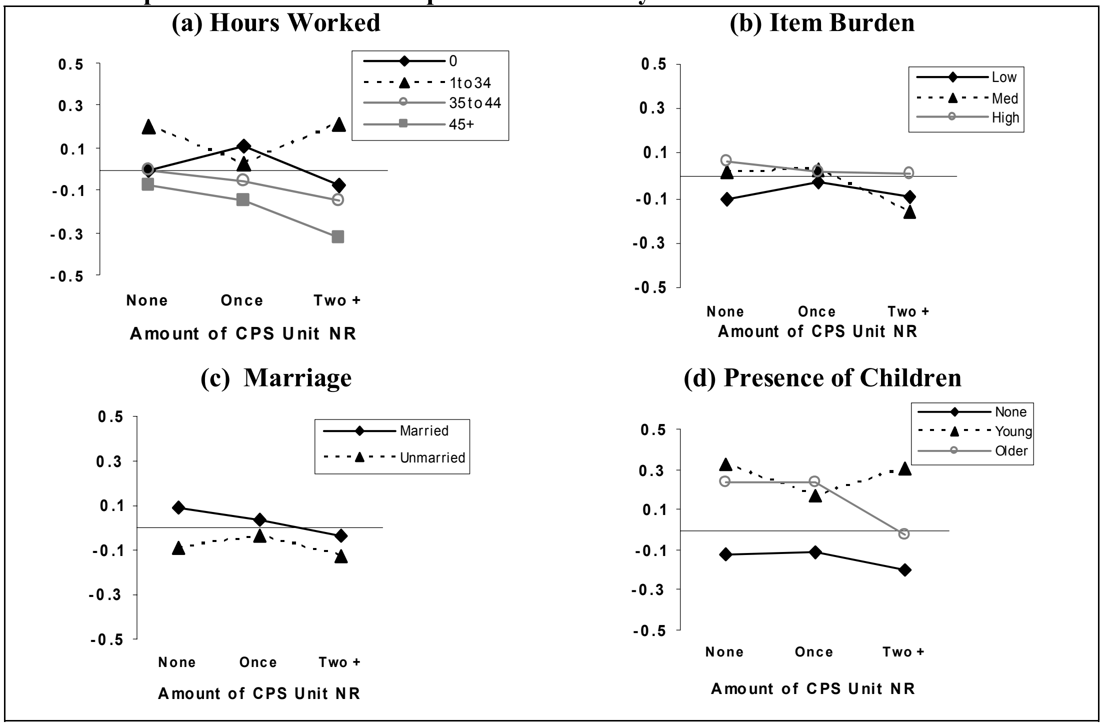
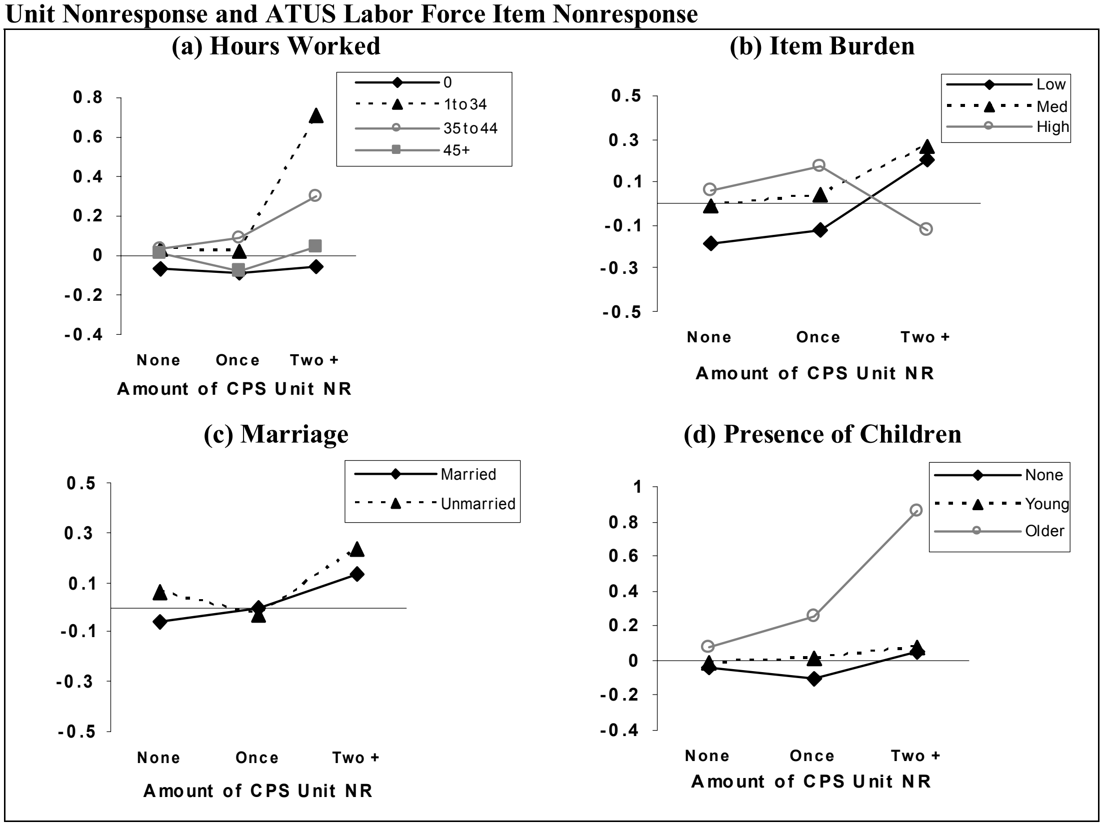


Figure 11. Effects of Common Cause Variables on the Relationship Between the Amount of CPS Unit Nonresponse and ATUS Labor Force Item Nonresponse



I also examined whether there were differences in data quality between ATUS cases that were refusal conversions and those that were not. Approximately twenty percent of ATUS sample members were flagged as a *refusal* at least once during the fielding period, and about five percent of ATUS respondents were refusal conversions. Consistent with previous findings, Figure 12 shows that data quality was worse for refusal conversion cases than for those that never refused, though regression analyses revealed significant effects only for the total number of diary activities and missing diary reports measures. Moreover, Figures 13 and 14 show that the associations between these two indicators of data quality and ATUS refusal conversion status do not disappear when controlling for potential common cause variables. At the multivariate level, total diary reports and missed activities continue to be related refusal conversion status, as well as to each of the covariates except marital status.

Figure 12. Relationship Between ATUS Refusal Conversion Status and ATUS Data Quality Indicators (in standard deviation units)

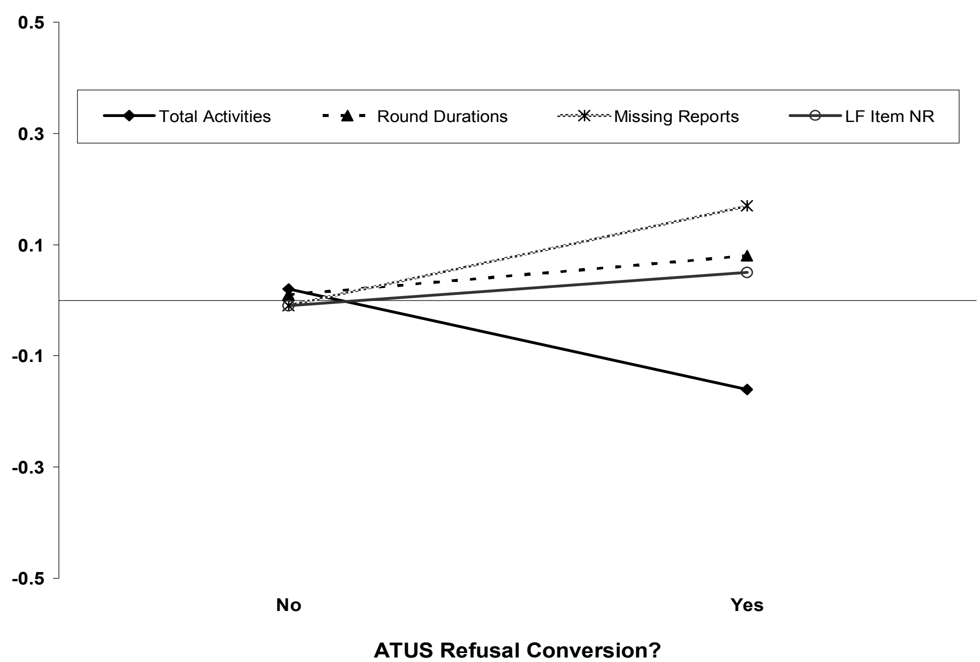


Figure 13. Effects of Common Cause Variables on the Relationship Between ATUS Refusal Conversion Status and Number of Reported ATUS Diary Activities

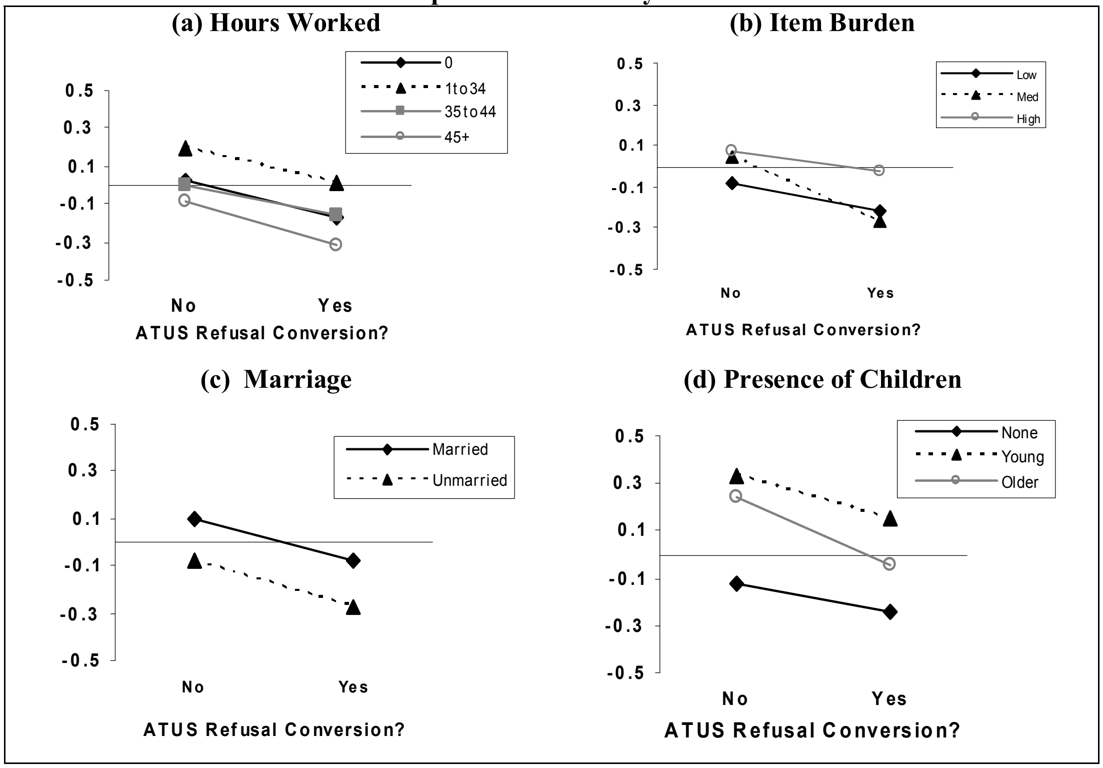
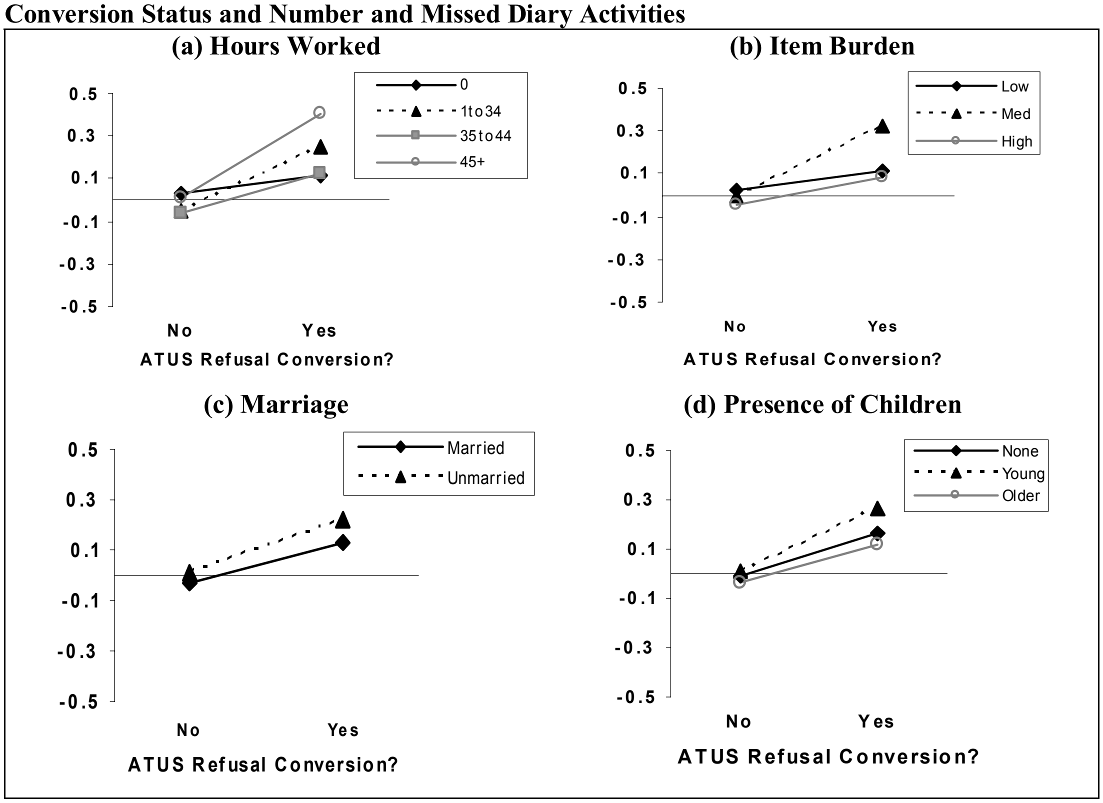


Figure 14. Effects of Common Cause Variables on the Relationship Between ATUS Refusal Conversion Status and Number and Missed Diary Activities



5.4 Discussion

The purpose of this chapter was to explore the relationship between response propensity and survey data quality. There are three main findings from the analyses presented in this chapter. First, data quality decreases as the probability of nonresponse increases. Second, the strength of this relationship varies by data quality indicator and by survey. The effects were stronger in the CPS, where nonresponse propensity was most strongly and positively related to item nonresponse and round values reports on continuous variables (e.g., hours and earnings). In ATUS, the relationship of nonresponse propensity to three of the four data quality indicators had essentially no practical significance. There was, however, a moderate, positive, and monotonic association between the total number of reported diary activities and likelihood of nonresponse. Third, when data quality and nonresponse did covary, controlling for potential common cause variables related to busyness, social capital, and survey burden did not weaken the relationship. Data quality continued to decline as nonresponse propensity rose, though there were main effects for some of the potential common cause variables.

These analyses have implications for survey organizations that strive for the highest response rates possible. Often, extraordinary persuasive efforts are made to bring difficult to contact or reluctant sample members into the respondent pool. The assumption is that these efforts are compensated by reductions in the total mean square error of survey statistics. Recent work by Curtin et al. (2000), Keeter et al. (2000), and others cast some doubt on this assumption, at least with respect to nonresponse error. The present analyses extend this work in two different and potentially opposing ways. On the one hand, it demonstrates that bringing in low propensity respondents also can produce significant increases in measurement error. If nonresponse error is not significantly increased by excluding low propensity cases (as these authors suggest), and these cases also are likely to be filled with measurement error (as we see here), then survey organizations may more comfortably divert resources away from recruitment of difficult respondents and focus instead on other error reduction techniques. On the other hand, evidence of covariation between measurement error and nonresponse may call into question previous investigations of nonresponse bias. The results of this study suggest that significant measurement error in late/difficult cases may in fact be concealing nonresponse bias undetected when examining respondent means (Groves, 2006). If higher nonpropensity individuals also are more likely to produce noisy data (increasing the variance of the statistic), then it becomes more difficult to detect if these individuals are different from low nonpropensity respondents; that is, it becomes more difficult to know the effects of excluding these individuals on nonresponse bias.

VI. Summary and Conclusions

There are five main findings that can be taken from the literature review and analyses presented in this paper. First, models of survey nonresponse were improved by adding community-level variables to the typical set of predictors (e.g., respondent, household, and survey process characteristics). Nonresponse in both the CPS and ATUS was more likely to occur when the amount of racial heterogeneity in the county increased. Additionally, county-level measures of income inequality and educational attainment were significantly related to CPS (but not ATUS) nonresponse propensity. Theories of survey participation point to the importance of social environmental influences of nonresponse (e.g., Groves and Couper, 1998), and the results of this study expand the list of potentially useful environmental variables and corroborate other research that has shown the beneficial effects of social integration on survey response (e.g., Abraham et al., 2006; Johnson et al., 2006).

Second, despite the intuitive appeal of busyness explanations of survey nonresponse, I found almost no evidence that busier people were less likely to participate in the CPS or ATUS. The indicators of busyness used in the present analyses were largely unrelated to nonresponse in both surveys, and in some instances actually were associated with higher levels of participation. These findings support the work of Abraham and colleagues (Abraham et al., 2006) and Groves and Couper (1998) which similarly failed to find a negative association between busyness and survey nonresponse. One explanation for this set of findings is that people who are busy may be engaged in activities that broaden and enrich their social networks and foster norms of trust and cooperation. This study provided some indirect evidence for this hypothesis. Two factors that have been shown in other work to reduce individuals' discretionary time and also to be associated with pro-social norms (e.g., Bellair, 1997; Vigdor, 2004; Robinson and Godbey, 1997)—marriage and parenthood—were found to be positively related to survey participation in both the CPS and ATUS.

Third, in contrast to recent studies by Curtin et al. (2000) and Keeter et al. (2000; 2006), analyses here that simulated lower response rates in the CPS and ATUS by removing high nonresponse propensity cases from the respondent pool resulted in significant changes (i.e., nonresponse bias) in CPS and ATUS estimates. These effects were largely confined to survey variables that were also related to the causes of nonresponse, as expected. Stronger effects were found in the CPS than ATUS, in part because more correlates of nonresponse were identified in the CPS propensity model than the ATUS model, and in part because the larger sample size improved the power of the CPS analyses. However, the fact that evidence of nonresponse bias was found in two surveys with such different designs and substantially different response rates, and for both estimates of survey means and associations between variables, is a significant contribution to the nonresponse literature.

Another contribution made by this study was its demonstration of the utility of a number of indirect measures of data quality. In particular, two data quality indicators that have been relatively neglected in the measurement error literature—round value reporting, and the presence of missed activity reports in time diaries—proved capable of capturing unique information about the kinds of errors respondents make in surveys, and were valuable additions to the set of quantifiable measures of data quality. In addition, examinations of the relationship between the data quality indicators and respondent and survey characteristics validated the assumptions underlying cognitive models of the survey response process. Respondents' cognitive abilities (e.g., age, education) and motivation (e.g., survey burden) had fairly consistent and predictable effects across each of the data quality measures.

Finally, I found evidence of a significant relationship between nonresponse propensity and survey data quality in both the CPS and ATUS. Respondents with higher nonresponse propensities consistently provided poorer data (i.e., more item nonresponse, round value reports, less complete responses, etc.) than those with lower nonresponse propensities. This result suggests that bringing in reluctant or hard to reach respondents may introduce significant measurement error in survey estimates. There is a clear tradeoff then in these surveys between the potential for nonresponse bias if late/reluctant respondents are excluded (as evidenced in Chapter 3) and the additional response error these same individuals likely introduce. The nonresponse models presented in this study can be used to improve post-survey nonresponse adjustments, but it is unclear from the present analyses what factors are driving the relationship between nonresponse and measurement error. None of potential common cause variables I examined had the expected moderating effect on this association.

In the absence of other commonly cited causal variables (e.g., topic interest) or more direct measures of respondents' social-psychological attributes (e.g., busyness, social integration, survey fatigue), we are faced with competing sources of error, and the survey researchers' familiar compromise between cost, timeliness, and accuracy. One solution would be to tie procedural decisions in these surveys to models of survey cost as a function of the total mean square error of key survey statistics. Even without such models, there is good evidence from the analyses presented in this paper that nonresponse bias in the ATUS is relatively small compared to the level of measurement error, so it may be appropriate for ATUS to devote fewer resources to nonresponse reduction and more to reductions in response error. Another potential solution lies in exploring alternative specifications of the various models of nonresponse and data quality examined in this study. For example, the correlates of the separate components of nonresponse (noncontact and noncooperation) and their relationship to data quality need to be investigated further, and more attention should be given to the effects of variable interactions (e.g., gender and marital status) in these models. Finally, additional measures of busyness and social capital (and others causal constructs) need to be examined in different survey contexts (e.g., non-governmental surveys, lower response rate surveys) to see if the findings of this study generalize and can be extended.

References

Abraham, K., Maitland, A., and Bianchi, S. (2006). "Nonresponse in the American Time Use Survey: Who is missing from the data and how much does it matter?" Public Opinion Quarterly, 70(5), 676-703.

Adler, P., and Kwon, S. (2002). 'Social Capital: Prospects For a New Concept.' Academy of Management. The Academy of Management Review, 27, pp.17-40.

Alesina, A. and La Ferrara, E. (2000). "Participation in Heterogeneous Communities". Quarterly Journal of Economics, 3, 847-904.

American Association for Public Opinion Research, (2000). "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys". Ann Arbor: AAPOR, 2000.

Apodaca, R., Lea, S., and Edwards, B. (1998). "The effect of longitudinal burden on survey participation." Paper presented at the annual meeting of the American Association of Public Opinion Research, St. Louis, MO.

Asch D, Jedrziewski K, Christakis N. (1997). "Response rates to mail surveys published in medical journals." Journal of Clinical Epidemiology,.50:1129–36.

Atrostic, B. and Kalenkoski, C. (2002) "Item Response Rates: One Indicator of How Well We Measure Income." Proceedings of the Annual Meeting of the American Statistical Association, August 2002.

Bates, N. and Buckley, S. (2000) "Effectiveness of the Paid Advertising Campaign: Reported Exposure to Advertising and Likelihood of Returning a Census Form." Census 2000 Dress Rehearsal Evaluation Results, Memorandum E1b. Planning, Research, and Evaluation Division, Bureau of the Census.

Bates, N., and Creighton, K. (2000). "The Last Five Percent: What We Can Learn From Difficult/Late Interviews." Proceedings of the American Statistical Association Section on Social Statistics. Alexandria, VA: American Statistical Association, pp. 120-125.

Beatty, P. and Herrmann, D. (2002). "To answer or not to answer: decision processes related to survey item nonresponse." In Groves *et al*. (Eds.), Survey Nonresponse, John Wiley & Sons, New York, pg. 71-85.

Bellair, P.E. 1997, "Social interaction and community crime: Examining the importance of neighbour networks." Criminology, 35, pp. 677–703.

Bethlehem, J (2002). “Weighting nonresponse adjustments based on auxiliary information.”In Survey Nonresponse, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, and RoderickJ. A. Little, pp. 275–88. New York: Wiley.

Biemer, P. (2001). “Nonresponse bias and measurement bias in a comparison of face to face and telephone interviewing.” Journal of Official Statistics, 17, 295-320.

Biemer, P., Groves, P., Lyberg, L., Mathiowetz, N., and Sudman, S. (1991) Measurement Errors in Surveys. New York: Wiley.

BLS (2002). Technical Paper 63 Current Population Survey: Design and Methodology. http://www.census.gov/prod/2002pubs/tp63rv.pdf.

Blau, P. (1964). Exchange and Power in Social Life, New York: Wiley & Sons.

Bogen, K. (1996). “The Effect of Questionnaire Length on Response Rates: A Review of the Literature.”.Washington, D.C.: U.S. Census Bureau.

Bollinger, C. and David, M. (2001). “Estimation with response error and nonresponse: Food Stamp participation in the SIPP.” Journal of Business and Economic Statistics, 19(2), 129-141.

Bose, J. and Sharp, J. (2005). “Measurement of Travel Behavior in a Trip-Based Survey Versus a Time Use Survey: A Comparative Analysis of Travel Estimates Using the 2001 National Household Travel Survey and the 2003 American Time Use Survey.” Paper presented at the 2005 ATUS Conference, December, Bethesda, Maryland.

Bound, J., and Krueger, A. (1991). “The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?” Journal of Labor Economics 9:1-24.

Brehm, J. and Rahn, W. (1997). “Individual-level evidence for the causes and consequences of social capital.” American Journal of Political Science, 41(3), pp. 999-1024.

Brooks, Arthur C. (2005). “Does social capital make you generous?” Social Science Quarterly 86(1): 1-15.

Burton, J., Laurie, H. and Lynn, P. (2004) “The long-term effectiveness of refusal conversion procedures on longitudinal surveys.” ISER Working Papers 2004-11.

Campanelli, P., Sturgis, P., and Purdon S. (eds.) (1997). “Can you hear me knocking: An investigation into the impact of interviewers on survey response rates.” London: Survey Methods Centre at SCPR.

Cannell, C.F. and Fowler, F.J. (1963). "Comparison of a Self-Enumerative Procedure and a Personal Interview: A validity Study" Public Opinion Quarterly, Summer, pp. 250-264

Cannell, C., Miller, P., and Oksenberg, L. (1981). "Research on interviewing techniques." In Leinhardt, S. (ed.) Sociological Methodology, pp. 389—437, San Francisco: Jossey-Bass.

Cash, W.S., and Moss, A. (1972). "Optimum recall period for reporting persons injured in motor vehicle accidents." Vital and Health Statistics, 2(50). Washington, D.C.: Public Health Service.

Chun, Y. and Scott, H. (2003). **"**A Theory-driven Approach to Reducing Nonparticipation of Twelfth Graders in National Surveys." Paper presented at the Annual Meeting of the American Association For Public Opinion Association, Fontainebleau Resort, Miami Beach, FL.

Coder, J. and Scoon-Rogers, L. (1995). "Evaluating the Quality of Income Data Collected in the Annual Supplement to the March CPS and the SIPP." SIPP Working Paper Series, No. 215, US Census Bureau, Washington, DC.

Conrad, F. and Schober, M. (2000). *"*Clarifying Question Meaning in a Household Telephone Survey.*"* Public Opinion Quarterly, 64:1-28.

Cook, K. (2000). "Charting Futures for Sociology: Structure and Action." Contemporary Sociology 29, 685-92.

Costa, D., and Kahn, M. (2003). "Civic engagement and community heterogeneity: An economist's perspective." Perspectives on Politics, 1, 103-111.

Couper, M. (1997). "A survey introductions and data quality." Public Opinion Quarterly, 61(2): 317-338.

Couper, M., Singer, E., and Kulka, R. (l997). "Participation in the Decennial Census: Politics, privacy, pressures." American Politics Quarterly, 26(1), 59-80.

Crowley, M. (2003). "Generation X Speaks Out on Civic Engagement and the Decennial Census: An Ethnographic Approach." Census 2000 Ethnographic Study. Washington, DC: U.S. Census Bureau.

Curtin, R., Presser, S. and Eleanor Singer (2000). "The Effects of Response Rate Changes on the Index of Consumer Sentiment." Public Opinon Quarterly, 64, Winter 2000 Number 4.

Curtin, R., Presser, S. and Singer, E. (2005). "Changes in telephone survey nonresponse over the past quarter century." Public Opinion Quarterly, 69(1), Spring, pp. 87-98.

Dahlhamer, J., Dixon, J., Doyle, P., Eargle, J., Griffin, D. and McGovern, P. (2003). "Quality at the Item Level: Terms, Methods, and Guidelines for Cross-Survey Comparisons." Presented at the Federal Committee on Survey Methodology Research Conference, November 17-19.

De Leeuw, E. and Heer, W. (2002). "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." In R. Groves, D. Dillman, J. Eltinge, & R. Little (Eds.), Survey Nonresponse (pp. 243-258). New York: John Wiley and Sons.

DeMaio, T. J. (1980), "Refusals: Who, where and why" Public Opinion Quarterly, 44, pp.223-233.

Deming, E. (1947). "Some criteria for judging the quality of surveys." Journal of Marketing, Vol. 12 (2), pp. 145-157.

DeNavas-Walt, C., Cleveland, R. and Roemer, M. (2001). "Money Income in the United States: 2000". U.S. Census Bureau, Current Population Reports, P60-213. U.S. Government Printing Office, Washington, DC. http://www.census.gov/prod/2001pubs/p60-213.pdf#search=%22%22gini%20index%22%20%22United%20States%22%20census%202000%22

Dillman, D., Gallegos, J. and Frey, J. (1976). "Reducing refusals rates for telephone interviews." Public Opinion Quarterly, 1, 66-78.

Dillman, D.A. (1978). Mail and Telephone Surveys: The Total Design Method. New York: Wiley & Sons.

Dillman, D. A. (2000). Mail and Internet surveys--The tailored design method. New York: John Wiley & Sons, Inc.

Dixon, J. (2000). "The Relationship Between Household Moving, Nonresponse and the Unemployment Rates in the Current Population Survey." In Proceedings of the Annual Meetings of the American Statistical Association, August 13-17, 2000.

Dixon, J. and Tucker, C. (2000). "Modeling household and interviewer nonresponse rates from household and regional characteristics." A paper presented at the International Workshop on Household SurveyNonresponse, Budapest, September, 2000. http://www.fcsm.gov/committees/ihsng/buda5.pdf

Drago, R., et al. (1999). "Time for surveys: Do busy people complete time diaries?" Society and Leisure 21(2):555–62.

Edgerton, M., Fisher, K., Gershuny, J., Pollman, A., and Torres, N. (2004). "US Historical Time Use Data File Production: Report on Activities, February to October 2004." Paper produced for the Institute for Social and Economic Research. http://www.timeuse.org/ahtus/reports/docs/ActivityReport.pdf (accessed December 18, 2006).

Feindt, P., Schreiner, I. and Bushery, J. (1997). "Reinterview: a tool for survey quality management." In Proceedings of the Section on Survey Research Methods. Alexandria, Virginia: American Statistical Association, pp. 105-110.

Fitzgerald, R., and Fuller, L., (1982). "I hear you knocking but you can't come in." Sociological Methods and Research, Vol. 11 No. 1, pp. 3-32.

Friedman, E. and Clusen, N. (2003). "Better Late? Characteristics of Late Respondents to a Health Care Survey." American Statistical Association, 2003, Proceedings of the Section on Survey Research Methods.

Gaertner, S. L. (1973). "Helping behavior and racial discrimination among liberals and conservatives." Journal of Personality and Social Psychology, 25, 335-341.

Glaeser, E., Laibson, D., Sacerdote, B. (2002). "An economic approach to social capital." The Economic Journal, 112, pp. 437-458.

Groves, R. (2006). "Nonresponse rates and nonresponse bias in household surveys." Public Opinion Quarterly, 70(5), pp. 346-675.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). "Understanding the decision to participate in a survey." Public Opinion Quarterly, 56, 475-495.

Groves, R. M., & Couper, M. P. (1998). Nonresponse in household interview surveys. New York: John Wiley.

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R. (2004). Survey Methodology. New York: Wiley.

Groves, R., Presser, S. and Dipko, S. (2004). "The role of topic interest in survey participation decisions." Public Opinion Quarterly,68 (1): 2-31.

Groves, R. M., Singer, E., and Corning, A. (2000). "Leverage-saliency theory of survey participation: Description and illustration." Public Opinion Quarterly, 64, 299-308.

Groves, R. M., Wissoker, D. A., Greene, L., McNeeley, M. E., Montemarano, D. A.. (2001). "Common influences on noncontact nonresponse across household surveys: Theory and data." Unpublished manuscript.

Guest, A., Cover, J., Kubrin, C., and Matsueda, R. (2006). "Neighborhood context and neighboring ties." City & Community, 5(4), pp. 363—385.

Hanisch, J. and Rendtel, U. (2002). "Quality of Income Data from Panel Surveys with Respect to Rounding." Chintex working paper No. 6, URL: http://www.destatis.de/chintex/download/paper6.pdf (accessed January 5, 2007)

Hansen, M. H. Hurwitz, W. N., and Pritzker, L. (1964). "The Estimation and Interpretation of Gross Differences and the Simple Response Variance." In C.R. Rao (Ed.) Contributions to Statistics, Pergamon Press Ltd, Calcutta, 111-136.

Harris-Kojetin, B. and Tucker, C. (1999). "Exploring the relation of economic and political conditions with refusal rates to a government survey." Journal of Official Statistics, 15(2), pp. 167-184

Herriot, R. and Spiers, E. (1980). "Measuring the Impact on Income Statistics on Reporting Differences Between Current Population Survey and Administrative Sources." Studies of Interagency Data Linkages, Report 11, US Department of Health, Education, and Welfare, Social Security Administration. 39 – 49.

Heyneman, S. P. (1998). "From the Party/State to Multi-Ethnic Democracy: Education and its Influence on Social Cohesion in Europe and Central Asia Region." Paper sponsored by International Child Development Centre, United Nations Children's Fund. Florence, Italy.

Holbrook, A., Green, M., and Krosnick, J. (2003). "Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias." Public Opinion Quarterly, 67: 79-125.

Hosmer, D. W., Jr., and Lemeshow, S. (1989). Applied Logistic Regression. John Wiley and Sons, New York.

Hyman, H., and Wright, C. (1971). "Trends in voluntary association memberships of American adults: Replication based on secondary analysis of national sample surveys." American Sociological Review, 36, 191-206.

Johnson, R., Gerstein, D., and Rasinski, K. (1998). "Adjusting Survey Estimates for Response Bias: An Application to Trends in Alcohol and Marijuana Use." *Public Opinion Quarterly* 62 (3):354-377.

Johnson, T. and Owens, L. (2003). "Survey Response Rate Reporting in the Professional Literature." Paper presented at the 58th Annual Meeting of the American Association for Public Opinion Research, Nashville, May 2003.

Johnson, T., Cho, Y., Campbell, R. and Holbrook, A. (2006). "Using community-level correlates to evaluate nonresponse effects in a telephone survey." Public Opinion Quarterly, 70(5), pp. 704—719.

Kalton, G., McMillen, D., & Kasprzyk, D. (1986). "Nonsampling error issues in the Survey of Income and Program Participation (SIPP)." Proceedings of the Bureau of the Census Annual Research Conference, *2*, 147-164.

Kasprzyk, D. (2005). "Measurement error in household surveys: sources and measurement." In The Handbook of Household Surveys, Studies in Methods, No. 31, United Nations publication No. E.83.XVII.13.

Keeter, S., Miller C., Kohut, A., Groves, R. and Presser; S. (2000). "Consequences of Reducing Nonresponse in a National Telephone Survey." Public Opinion Quarterly, 64(2).

Kellerman, S. E., and Herold, J. (2001). "Physician response to surveys: A review of the literature." American Journal of Preventive Medicine, 20, 61-67.

Kropf, Martha and Johnny Blair (2005). "Testing Theories of Survey Cooperation: Incentives, Self-Interest and Norms of Cooperation." Forthcoming. Evaluation Review.

Krosnick, J. (1999). "Survey research." Annual Review of Psychology, 50, pp. 537—567.

Krosnick, J. and Alwin, D. (1987). "An evaluation of a cognitive theory of response order effects in survey measurement." Public Opinion Quarterly, 51, pp. 201—219.

Krosnick, J. A., Narayan, S. S., & Smith, W. R. (1996). "Satisficing in surveys: Initial evidence." In M. T. Braverman & J. K. Slater (Eds.), Advances in survey research (pp. 29-44). San Francisco: Jossey-Bass.

Kubrin, C. and Weitzer,R. (2003). "New directions in social disorganization theory." Journal of Research in Crime and Delinquency, 40, 374-402.

La Due Lake, R. and Huckfeldt, R. (1998). "Social capital, social networks, and political participation." Political Psychology, 19, p. 567-584.

Letki, N. (2005). "Does diversity erode social cohesion? Social capital and race in British Neighborhoods." Nuffield College Working Papers in Politics.

Li, Y., Pickles, A., and Savage, M. (2005). "Social capital and social trust in Britain". European Sociological Review 21(2), 109-123.

Lin, I-Fen and Schaeffer, N.C. (1995). "Using Survey Participants to Estimate the Impact of Nonparticipation." Public Opinion Quarterly, 59(2), pp. 236-258.

Loosveldt, G. and Carton, A. (2002). "Utilitarian individualism and panel nonresponse." International Journal of Public Opinion Research, 14, pp. 428—438.

Lyberg, L., Biemer, P., Collins, M., De Leeuw, E., Dippo, C., Schwarz, N., and Trewin, D. (1997). Survey Measurement and Process Quality. Lars Lyberg et al. (Eds.) New York: Wiley-Interscience.

Lynn, P., Clarke, P., Martin, J. and Sturgis, P. (2002). "The effects of extended interviewer efforts on non-response bias." In Survey Nonresponse (Groves, R., Dillman, D., Eltinge, J. and Little, R., eds.), pp. 135-148. John Wiley, New York.

Marquis, K. and Moore, J. (1990). "Measurement error in SIPP Program Reports." Proceedings of the Sixth Annual Research Conference, Washington, DC: US Bureau of the Census, March 1990, 721—745.

Mason, R., Lesser, V., and Traugott, M. W. (2002). "Effect of Item Nonresponse on Nonresponse Error and Inference." In R. M . Groves, D. A. Dillman, J. L. Eltinge, and R . J. A. Little (eds.), Survey Nonresponse. New York: Wiley, Pp. 149-161.

McFarlane, E. (2006). "Nonresponse bias in a mail survey of physicians." Paper presented at the American Association for Public Opinion Research (AAPOR) Annual Conference, Montreal, Canada.

McGovern, P. and Bushery, J. (1999). "Data Mining the CPS Reinterview: Digging into ResponseError." Federal Committee on Statistical Methodology Research Conference Paper, available from: http://www.fcsm.gov/99papers/mcgovern.pdf accessed December 12, 2006.

Merkel, D., & Edelman, M. (2002). "Nonresponse in exit polls: A comprehensive analysis." In R. Groves, D. Dillman, J. Eltinge, & R. Little (Eds.), Survey Nonresponse (pp. 243-258). New York: John Wiley and Sons.

Mertler, C. (2003). "Patterns of response and nonresponse from teachers to traditional and web surveys." Practical Assessment, Research & Evaluation, 8(22).

Miller, D. and Ennis, S. (2001). "Response Error Reinterview of the 1999-2000 Schools and Staffing Survey." In Proceedings of Statistics Canada Symposium, 1- 9.

Moore, J., Stinson, L., and Welniak, E. (2000). "Income Measurement Error in Surveys: A Review," Journal of Official Statistics, (16) 4.

Morton-Williams, J. (1993). Interviewer approaches. Aldershot: Dartmouth Publishing.

Mulligan, C.B., Schneider, B., and Wolfe, R. (2000). "Time Use and Population Representation in the Sloan Study of Adolescents." Technical Working Paper 265. National Bureau of Economic Research, November 2000.

Myers, R. (1954). "Accuracy of age reporting in the 1950 United States census." Journal of the American Statistical Association, (44), 826-831.

Neter, J. and Waksberg, J. (1964). "A study of response errors in expenditure data from household surveys." Journal of the American Statistical Association, 59, 18-55.

O'Muircheartaigh, C. (1991). "Simple response variance: Estimation and determinants." In Biemer *et al*. (Eds.), Measurement Error in Surveys, John Wiley & Sons, New York, pgs. 551-574.

O'Neill, G. and Dixon, J. (2005). "Nonresponse bias in American Time Use Survey." Proceedings of the Annual Meeting of the American Statistical Association.

O'Neill, G. and Sincavage, J. (2004). "Response analysis survey: A qualitative look at response and nonresponse in the American Time Use Survey." BLS Working Paper.

Office of Management and Budget (2001). Statistical Policy Working Paper #31: Measuring and Reporting Sources of Error in Surveys. Statistical Policy Office, Office of Information and Regulatory Affairs, July 2001.

Oldendick, Robert W., Link, Michael W. 1999. "Call Screening: Is It Really a Problem for Survey Research?" Public Opinion Quarterly, 63, 577-589.

Olson, K. (2006). "Survey participation, nonresponse bias, measurement error bias, and total bias." Public Opinion Quarterly, 70(5), pp. 737-758.

Palumbo, T. and Siegel, P. (2004). "Accuracy of Data for Employment Status as Measured by the CPS-Census 2000 Match." Census 2000 Evaluation B.7, US Census Bureau.

Parten, M. (1950). Surveys, Polls, and Samples. New York: Harper & Brothers, pp. 624.

Pennell, S. and Lepkowski, J. (1992). "Panel Conditioning Effects in the Survey of Income and Program Participation." Journal of the American Statistical Association Survey Research Methods: 566-571.

Perkins, D., Brown, B., and Taylor, R. (1996). “The ecology of empowerment: Predicting participation in community organizations.” Journal of Social Issues, 52, 85-110.

Portes, A. and Landolt, P. (2000). "Social Capital: Promise and Pitfalls of its Role in Development." Journal of Latin American Studies, 32, 529.

Presser, S. and Singer, E. (forthcoming). “Privacy, confidentiality and respondent burden as factors in telephone survey nonresponse,” in J. Lepkowski et al. (eds.) Advances in Telephone Survey Methodology, New York: Wiley.

Price, J. H. (2000). “‘Show me the money’: Medical directors’ responses to a mail survey.” American Journal of Managed Care, 6, 1257-1259.

Putnam, R. (1995). "Bowling alone: America's declining social capital." Journal of Democracy, 6, 65-78.

Putnam, R. (2000). Bowling alone: the collapse and revival of American community. New York: Simon & Schuster.

Raghunathan T., Lepkowski J., Van Hoewyk J (2001). “A multivariate technique for multiply imputing missing values using a sequence of regression models.” Survey Methodology, 27, pg. 85–95.

Raghunathan T., Solenberger P., Van Hoewyk J. (1998). “IVEware: imputation and variance estimation software.” Ann Arbor, MI: Institute for Social Research, University of Michigan. (http://www.isr.umich.edu/src/smp/ive).

Read, D. (2004). “Utility theory from Jeremy Bentham to Daniel Kahneman.” Working Paper No: LSEOR 04-64, London School of Economics and Political Science.

Robinson, J. and Godbey, G. (1997). Time for life: The surprising ways Americans use their time. University Park, PA: The Pennsylvania State University Press.

Roemer, M. (2000). “Assessing the Quality of the March Current Population Survey and the Survey of Income and Program Participation Income Estimates, 1990 – 1996,” Housing and Household Economics Statistics Division, June, U.S. Census Bureau, Washington, D.C.

Rothstein, B., Uslaner, E.M. (2005). All for All: Equality and Social Trust. London School of Economics Health and Social Care Discussion Paper Number 15.

Sangster, Roberta L. (2003). "Calling effort and nonresponse for telephone panel surveys," Paper presented to the International Workshop on Household Survey Nonresponse,"Copenhagen, August, 2002.

Schieman, S. (2005) "Residential Stability and the Social Impact of Neighborhood Disadvantage: Gender- and Race-Contingent Patterns" Social Forces , 83, 3, 1031-1064

Schweitzer, M.. and Severance-Lossin, E. (1996): “Rounding in earnings data”, Federal Reserve Bank of Cleveland Working Papers, available through the www: http://www.clev.frb.org

Shaffer, H., Hall, M., and Bilt, J. (1997) “Program Evaluation: A Practical Guide to Discovering What Works.” Technical Report # EV-122297, Harvard Medical School, Division on Addictions.

Sharp, H. and Feldt, A. (1959). “Some factors in a probability sample survey of a metropolitan community.” American Sociological Review, 24(5), pp. 650-661.

Schwarz, N. (1999). “Self-reports: How the questions shape the answers.” American Psychologist, 54, 93-105.

Schwarz, N., Groves, R., and Schuman, H. (1995). Survey Methods. Survey Methodology Program Working Paper Series. Ann Arbor, Michigan, Institute for Survey Research, University of Michigan.

Schwarz, N., Park, D., Knäuper, B., & Sudman, S. (1999). Cognition, aging, and self-reports. Philadelphia, PA: Psychology Press.

Schwarz, N. and Wellens, T. (1997). “Cognitive Dynamics of Proxy Responding: The Diverging Perspectives of Actors and Observers.” Journal of Official Statistics 13(2), 159-179.

Singer, E., Mathiowetz, N. A., & Couper, M. P. (1993). The impact of privacy and confidentiality concerns on surveyparticipation. Public Opinion Quarterly, 57, 465-482.

Singer, E., Van Hoewyk, J., and Neugebauer, R. (2003). “Attitudes and Behavior: The Impact of Privacy and Confidentiality Concerns on Participation in the 2000 Census.” Public Opinion Quarterly, 67, 368-384.

Smith, T.W. (1983) ‘The hidden 25 percent: an analysis of nonresponse on the 1980 General Social Survey’, Public Opinion Quarterly 47: 386-404.

Smith, T.W. (1984) “Estimating nonresponse bias with temporary refusals.” Sociological Perspectives 27: 473-489.

"Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001.

Stinchcombe, A., Jones, C., and Sheatsley. P. (1981). "Non-Response Bias for Attitude Questions." Public Opinion Quarterly 45 (3): 359-75.

Stolle, D. (2001). "Clubs and congregations: The benefits of joining an association". In Karen Cook, ed., Trust in Society (New York: Russell Sage Foundation.

Stoop, I. (2005). The hunt for the last respondent: Nonresponse in sample surveys. The Hague: SCP.

Sudman, S., Bradburn, N. M., and Schwarz, N. (1996). Thinking about answers: The application of cognitive processes to survey methodology. San Francisco, CA: Jossey-Bass.

"*Technical Paper 63 Current Population Survey: Design and Methodology"*, March 2002, page G-2, on the Internet at http://www.census.gov/prod/2002pubs/tp63rv.pdf (visited Jan. 16, 2007).

Thibaut, J. and Kelly, H. (1959). The Social Psychology of Groups (New York: Wiley).

Tortora, B. (2004). "Response Trends in a National Random Digit Dial Survey." Metodoloski zvezki, 1(1), 21-32.

Tourangeau, R. (1984). "Cognitive science and survey methods." In T. Jabine et al. (Eds.), Cognitive Aspects of Survey Design: Building a Bridge Between Disciplines. Washington: National Academy Press, pp.73-100.

Tourangeau, R. (2004). "Survey research and societal change." Annual Review of Psychology, 55, 775-801.

Tourangeau, R., Rips, L.J., and Rasinski, K.A. (2000). The Psychology of Survey Response, Cambridge, New York: Cambridge University Press.

Triplett T., Blair J., Hamilton T., and Kang Y. (1996). "Initial cooperators vs. converted refusers: Are there response behavior differences?" American Statistical Association, Proceedings of the Section on Survey Research Methods, 1038- 1041.

Tuckel, P., and O'Neill, H. (2002). "The vanishing respondent in telephone surveys." Journal of Advertising Research, 42 (5), 26-48.

Triplett T., Blair J., Hamilton T., and Kang Y. (1996). "Initial cooperators vs. converted refusers: Are there response behavior differences?" American Statistical

Association, 1996, Proceedings of the Section on Survey Research Methods, 1038-1041.

U.S. Census Bureau, Current Population Reports, Series P23-205, *Population Profile of the United States: 2000 (Internet Release)*, "People on the Move: Geographic Mobility, 1999 – 2000."
http://www.census.gov/population/pop-profile/2000/chap03.pdf.

Uphoff, N. "Understanding Social Capital: Learning from the Analysis of Experience of Participation," in Ismail Serageldin and Partha Dasgupta, eds., Social Capital: A Multifaceted Perspective (Washington, DC: World Bank, 2000), pp. 215-249.

Voigt, L., Boudreau, D., Weiss, N., Malone, K., Li, C., and Daling, J. (2005). "Letter to the editor: RE: 'Studies with low response proportions may be less biased than studies with high response proportions." American Journal of Epidemiology, 161, pp. 401- 402.

Voogt, R. & Saris, W. (2003). "Political interest, the key for correcting to correction methods for non-response bias in election studies." Paper presented at the 14th International workshop on household non-response. September 22-24, Leuven, Belgium.

Willimack, D.K. , Schuman, H., and Lepkowski, J.M. (1995). "Effects of a Prepaid Nonmonetary Incentive on Response Rates and Response Quality in a Face-to-Face Survey, Public Opinion Quarterly, 59, 78-92.

Woltman, H. and Bushery, J. (1977). "Update of the National Crime Survey Panel Bias Study." Internal United States Bureau of the Census report, Washington, D.C.

Woolcock, M. and Narayan, D. (2000). "Social capital: Implications for development theory, research, and policy." The World Bank Research Observer, 15, 225-249.

Yan, T., Tourangeau, R., and Arens, Z. (2004). "When less is more: Are reluctant respondents poor reporters." American Statistical Association, Proceedings of the Section on Survey Research Methods, 4632 – 4651.