ABSTRACT

Title of dissertation: RESPONDENT CONSENT TO USE ADMINISTRATIVE DATA

Jenna A. Fulton
Doctor of Philosophy, 2012

Dissertation directed by: Professor Stanley Presser
Joint Program in Survey Methodology

Surveys increasingly request respondents' consent to link survey responses with administrative records. Such linked data can enhance the utility of both the survey and administrative data, yet in most cases, this linkage is contingent upon respondents' consent. With evidence of declining consent rates, there is a growing need to understand factors associated with consent to record linkage. This dissertation presents the results of three research studies that investigate factors associated with consenting. In the first study, we draw upon surveys conducted in the U.S. with consent requests to describe characteristics of surveys containing such requests, examine trends in consent rates over time, and evaluate the effects of several characteristics of the survey and consent request on consent rates. The results of this study suggest that consent rates are declining over time, and that some characteristics of the survey and consent request are associated with variations in consent rates, including survey mode, administrative record topic, personal identifier requested, and whether the consent request takes an explicit or opt-out approach. In the second study, we administered a telephone survey to examine the effect of administrative record topic on consent rates using experimental methods, and through non-experimental methods, investigated the influence of respondents' privacy, confidentiality, and trust attitudes and consent request salience on consent rates. The results of this study indicate that respondents' confidentiality attitudes are related to their consent decision; the other factors examined appear to have less of an impact on consent rates in this survey. The final study used data from the 2009 National Immunization Survey (NIS) to assess the effects of interviewers and interviewer characteristics on respondents' willingness to consent to vaccination provider contact. The results of this study suggest that interviewers vary in their ability to obtain respondents' consent, and that some interviewer characteristics are related to consent rates, including gender and amount of previous experience on the NIS.

RESPONDENT CONSENT TO USE ADMINISTRATIVE DATA

By

Jenna A. Fulton

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
2012

Advisory Committee:
Professor Stanley Presser, Chair
Professor Ernesto F. Calvo
Professor Mick P. Couper
Dr. Brian Harris-Kojetin, Office of Management and Budget
Professor Eleanor Singer
Professor Richard L. Valliant

Dedication

To my parents,

for their unconditional love, support,

and commitment to my education.

Acknowledgements

I am incredibly thankful for the help, kindness, and support of the many individuals who contributed in one way or another to this dissertation. I am very grateful that my dissertation advisor, Stanley Presser, generously adopted me as his student and patiently worked with me throughout this process. Working under someone with such great intellect and high standards has benefitted me tremendously and made me a better researcher. My research has significantly improved in his presence and I've grown an incredible amount in my time working with him.

I would also like to thank the other members of my dissertation committee who have each contributed a great deal of time and effort to my work. Mick Couper was very instrumental in the design of my dissertation research, even as early as during my Comprehensive Exams when I proposed the initial stages of this work. Ernesto Calvo and Rick Valliant always opened their doors to me and provided tremendous, invaluable help with analyses and seemingly intractable problems with my data. Brian Harris-Kojetin constantly stressed the importance of the statistical use of administrative records in a practical context and offered a real-world perspective on my research topic. Finally, I would like to thank Eleanor Singer for sharing her expertise in this area, and for giving her time to review and comment on my work.

I am also grateful for the advice and guidance of Katharine Abraham, Bob Groves, and Shelly Martinez who supported, shaped, and stressed the importance

of this project, especially during its initial stages, and encouraged me to pursue this line of research.

The research presented in this dissertation would not be possible without the generosity of numerous agencies who kindly shared their resources with me. I am very thankful that the U.S. Census Bureau awarded me a Census Doctoral Dissertation Fellowship which made much of this research possible. Several other individuals at the U.S. Census Bureau contributed their time to my project, including Nancy Bates, Patrick Benton, Cheryl Landman, Gregory Weyland, and Dann Doyle.

I truly would not have been able to conduct the majority of my dissertation research without the help of the National Center for Health Statistics (NCHS). Peter Meyer, Pat Barnes and Nat Schenker made my fourth chapter possible by solving many complex data problems. Jennifer Parker, Donna Miller and Jennifer Madans donated their time, effort, and resources so that I could examine consent rates in several NCHS surveys.

I also owe a great deal of thanks to JPSM for allowing me to use the 2011 Practicum survey as a vehicle for my dissertation research, and to Stanley Presser for guiding this project. Further, I am very thankful for the help of Rebecca Medway, as well as the Masters students who participated in the 2011 Practicum and conducted pretests, coded open-ended data, rightfully critiqued the and helped me to think about this research from a variety of new angles.

Sarah Gebremichael, Rupa Jethwa Eapen, Duane Gilbert, and Gina Hsu of JPSM have helped me tremendously over the years. Ben Skalland and Martin

Barron of the National Opinion Research Center were instrumental in my accessing and analyzing the National Immunization Survey, and answered all of my many questions about the data.

I'm also thankful for the generosity and thoughtfulness of the faculty, staff, and students in the Department of Government and Politics, especially those in American Politics. Outside of the University of Maryland, several individuals have acted as informal mentors to me – Sophia Tsakraklides, Skander Van den Heuvel, Brian and Lauren Harris-Kojetin, and Patrick Moynihan. I am very grateful for their advice, guidance, and the many, many times they have fed me.

Two professors have substantially contributed to my education while at the University of Maryland. The lessons learned from Wayne McIntosh and John Townshend cannot be contained within a dissertation. I will be forever grateful for their help.

This final spot is reserved for my most loving, wonderful, and supportive family. All of my grandparents have fostered and encouraged my educational goals throughout my life; most recently they've reinforced motivation to graduate through just four words, repeated on a weekly basis over these four years "Aren't you done yet?" My sister Marisa (who, Ph.D. or not, will always be one I.Q. point smarter than me (n.s.)) was a superb source of advice and constant provider of entertainment during this process. Finally, I cannot sufficiently express my gratitude to my parents who listened to me complain on a daily basis about anything and everything imaginable and provided me with everything I needed to complete my educational goals. As it turns out, having a psychologist for a father and a fiery Italian mother is the perfect combination of parenting that one needs to survive graduate school.

Table of Contents

List of Tables

List of Figures

1 CHAPTER 1: SUMMARY OF THE LITERATURE ON LINKAGE CONSENT

1.1 Introduction

A growing number of surveys ask respondents for permission to link their survey responses with administrative records. Such linked data enhance the utility of both surveys and administrative records by making possible studies that would be difficult or impossible to conduct using either source alone. Linking to administrative records can also reduce or eliminate the need to ask sensitive or difficult questions in surveys, thereby reducing respondent burden (Dahlhamer & Cox, 2007; Singer, Van Hoewyk, & Neugebauer, 2003). In addition, linked data can be used to validate survey responses and to replace data missing in surveys.

However, most linked datasets are contingent upon respondent consent to linkage (Bates, 2005; Singer, 2001). With evidence of declining consent rates (Bates, 2005; Dahlhamer et al., 2007), there is a growing need to understand factors associated with consent to record linkage. This chapter examines a variety of factors that may influence a respondents' decision to consent.

1.2 Auxiliary Data Requests

In addition to the initial participation decision, respondents are sometimes presented with additional requests for auxiliary information within the interview. Collecting auxiliary data within the survey capitalizes on the time and resources spent sampling, contacting, and interviewing each respondent and augments the data collected from the survey questionnaire. This auxiliary information can take various forms, for example, biological specimens collected during a physical examination with a respondent, locating information to recontact respondents for follow-up data collection, participation in multiple waves of data collection, or access to administrative records.

This dissertation is about a specific type of auxiliary information requested in a survey: for respondents' consent to grant access to their administrative records. A number of surveys such as the National Immunization Survey (NIS), the National Health Interview Survey (NHIS), and the Medical Expenditure Panel Survey (MEPS) request respondents' consent to access and append some of their administrative records (from sources such as healthcare providers or government agencies) to their survey responses. The consent question may be accompanied by a request for some personally-identifying information to facilitate linkage, such as vaccination provider contact information in NIS, or medical provider and pharmacy contact information in MEPS. Augmenting survey responses with this additional data increases the amount of information captured for each respondent beyond what can be collected during the survey interview. Unfortunately, the utility of the linked data can be compromised when respondents do not consent to access their administrative records.

In the remainder of this chapter we first provide an overview of administrative records, of linked data and their benefits, and of surveys that link responses with administrative records. Next, we discuss consent bias and characteristics of the consent request. Lastly, we discuss several hypothesized influences on consent to record linkage, including respondent-level and interviewer-level variables, and the salience of the consent request.

1.3 Administrative Records Defined

Administrative records are data created by program agencies to facilitate the operation of their programs (Davern, Roemer, & Thomas, 2006; Droitcour, 2001). Jones and Elias (2006:1) describe them as "information which arises via the operation of a transaction, registration or as a record of service delivery," and suggest that they "relate specifically to the administration of a system or process and are not primarily generated as research resources." Calderwood and Lessof (2009:2) similarly define them as "information that is routinely collected by organizations, institutions, companies, and other agencies in order that the organization can carry out, monitor, archive, or evaluate the function or service it provides." Others include additional types of information in the definition such as credit or debit card payments, stock prices, telephone and internet records, and other electronic transactions (Couper, 2002; Davern et al., 2006). At least one unifying feature of these various records is that none was initially collected for research purposes.

The earliest known use of U.S. administrative records for statistical purposes dates to 1890 when the Economic Census created an administrative record frame of mortgage holders by contacting real estate recorders' officers to identify all farm, home, and business owners with mortgages made between 1880 and 1890. The Census Office contacted all property owners by mail and requested additional information regarding their property and mortgage (U.S. Census Bureau, 2011). More recently, among other uses, the Census Bureau has relied on administrative records to improve within decade population estimates and decennial census operations (Potok, 2011). In 1976, § 6(c) was added to Title 13 calling for the statistical use of administrative records at the Census Bureau "To the maximum extent possible and consistent with the kind, timeliness, quality and scope of the statistics required" 13 USC §6(c).[1]

With limited financial budgets, time constraints, and declining response rates in surveys, researchers are investigating alternatives to primary data collection and ways to

---

[1] Retrieved from http://uscode.house.gov/download/pls/13C1.txt on August 11, 2011.

supplement traditional survey methods. This includes utilizing administrative records for statistical purposes and linking these data with survey responses. With administrative data, respondents can be linked to their own records; to administrative records of others, including their relatives, friends or coworkers; or to contextual information such as data about neighborhoods, schools, employers, states, or governments (GAO, 2001; Lane, 2010). The main focus of this dissertation is linkage between survey data and administrative records of the survey respondents, rather than records of other individuals or units.

1.4 Some Benefits of Linked Data

Though some speculate that efforts to link data extend back as far as humans have kept records (Scheuren, 1997), we know with certainty that linking survey responses with administrative data can be traced to at least the mid-twentieth century. British physicians Doll and Hill (1956) linked medical records of their co-physicians with survey responses evaluating their past and present smoking behavior in order to identify smoking as a cause of death. The growing use of data linkage since then is due to both technological change – in particular the growing capacities to process and store data over time (Calderwood & Lessof, 2009; GAO, 2001; Scheuren, 1997), and increased understanding of record linkage and its benefits (GAO, 2001). Nonetheless, the adoption of linkage methodology in survey research has been slow, especially in the U.S. (Martinez, 2010), possibly because the quality of records has yet to be thoroughly evaluated and verified (Davern et al., 2006).

Even so, the number of linked databases used for research purposes has increased substantially over the past few years (Lillard & Farmer, 1997). Bohensky et al. (2010)

estimated that the use of data linkage for medical research expanded nearly 6-fold over the past two decades: they identified 951 studies with the terms “medical record linkage” in the abstract or title published between 2002 and 2007 on Medline, compared to 161 between 1992 and 1997.[2] This increase is a testament to the perceived value of linked data and is reflected by the establishment of data linkage centers in Australia, North America and the UK (Bohensky et al., 2010).

The advantages to linking survey responses with administrative records are numerous: linked data offers potential for increased data quality, greater analytic potential, reduced interviewer and respondent burden, and cost savings to name a few. Record linkage capitalizes on the strengths and minimizes the weaknesses of each data source and produces a new enhanced source of information capable of addressing a broader set of research questions.

Surveys can contribute a broad set of socioeconomic and demographic variables to linked datasets; this information is typically representative of the target population and is often not available in the administrative records (McNabb, Timmons, Song, & Puckett, 2009; Drazga-Maxfield, 2008). Administrative records contribute information on program participation and administration generally excluded from surveys, and is often available in a more timely fashion (Calderwood et al., 2009; Drazga-Maxfield, 2008; McNabb et al., 2009). When combined, the linked data aggregates the breadth of variables included in the survey data with the accuracy and timeliness of administrative data. A single data source comprising this range of information is difficult to find (Lillard et al., 1997).

---

[2] Bohensky et al. (2010) do not provide information for 1998 through 2002.

There may also be quality benefits associated with linking to administrative records (Calderwood et al., 2009). Although administrative records undoubtedly contain errors[3], because they are used to determine outcomes such as program eligibility and benefit amounts, they are typically subject to quality control procedures and as a result, are thought to be of higher quality than survey data (Drazga-Maxfield, 2008). Consequently administrative records are sometimes used to correct for errors of observation and non-observation in surveys (McNabb et al., 2009).[4]

Linked data can also be used to evaluate methodological and reporting issues in surveys (Drazga-Maxfield, 2008). Linked data can assess respondent under or overreporting and correct for misreported or missing survey values (Jenkins, Lynn, Jackle, & Sala, 2008; Lillard et al., 1997; Yaffe, Shapiro, Fuchsberg, Rohde, & Corpeno, 1978), again assuming the accuracy of the administrative record. Information contained in the administrative records can increase the number of variables used for nonresponse adjustment (Sakshaug & Kreuter, 2011). Using linked data can reduce respondent and interviewer burden (Calderwood et al., 2009; Jenkins et al., 2008; Sala, Burton, & Knies., 2010) and permit collection of additional respondent data (Dahlhamer et al., 2007; Michaud, Dolson, Adams, & Renaud, 1995; Singer, 2001).

Linked data can reduce data collection costs if administrative data is used to replace primary data collection. By obtaining respondents' consent to access their tax records, the Survey of Labour and Income Dynamics (SLID) in Canada was able to reduce the number of interviews for consenting panel respondents. Eliminating a portion

---

[3] In a study on 2031 HRS respondents, including 441 diabetics, Sakshaug (2011) found that as compared to both self-reports and HA1c data, Medicare reimbursement data, on average, overestimated diabetes status.
[4] Davern et al. (2006) propose a research agenda to evaluate the coverage, nonresponse, sampling, and measurement error of linked data files.

of overall interviews resulted in about $160,000 savings in data collection costs (Michaud et al., 2005).

Linking survey and administrative data expands upon the research opportunities that would be available with only administrative or survey data (Drazga-Maxfield, 2008). With linked data, researchers can more easily investigate complex problems that would be difficult or impossible to study with only survey or administrative records. Linked data provides a more extensive and comprehensive range of available data. For example, the utility of survey measures of health status and behaviors can be enhanced by appending Medicare records which include administrative measures of health care costs and utilization of services (Lillard et al., 1997).

Nevertheless, the statistical use of linked data is not without disadvantages. The U.S. lacks a national registry, or complete system of records, from which other countries such as Denmark, Finland, Norway and Sweden benefit (Redfern, 1986) and American's willingness to employ a purely administrative record census is decreasing over time.[5] Public opinion is likely influenced by privacy concerns associated with record linkage, as linked data are often more personally-identifiable than survey data alone, given the amount of detail they can provide about a respondent and the opportunities for re-identification of a specific individual (Lane, 2010). Because of this, most linked data will never be made available in the public domain (Davern et al., 2006).

---

[5] The percent of respondents selecting "favor" to the following question from the Census Survey of Privacy Attitudes has declined over time: "Another proposal is to do away with Census forms entirely. No one would be asked to fill out a form. Instead, the Census Bureau would count the entire population by getting information from other government agencies. Would you favor or oppose the Census Bureau getting everyone's sex, age, date of birth, Hispanic origin, and race from the records of other government agencies, so no one would have to fill out a census form?" *%"Favor": 1995 - 59%; 2000 - 42%; 2010 - 37%.* See, for example, Singer, Bates, &Van Hoewyk (2011), and Singer et al. (2003).

There are also technical limitations associated with linkage. Even if respondents do grant consent, constraints associated with the linkage procedure can produce unsuccessful matches potentially introducing systematic biases stemming from the non-matched cases (Calderwood et al., 2009). Further, if estimates differ between the administrative record and survey data, it is not always clear which source is more accurate.

Perhaps most significantly, in many cases accessing and linking administrative data with survey responses requires respondents' consent (GAO, 2001). Without universal consent from survey respondents, the subsequent linked data can lack the representativeness and generalizability of the survey data. There is potential for any differential willingness to consent to introduce a systematic bias into the linked data.

1.5 Examples of Linked Survey and Administrative Data

Two examples of surveys that request respondents' consent to link answers with administrative records are the Residential Energy Consumption Survey (RECS) and MEPS. These surveys differ in many ways, not limited to overall survey design, sponsor and topic. Further, they also vary in the manner consent to record access is requested from respondents as well as the type of records that are requested.

1.5.1 The Residential Energy Consumption Survey

RECS, conducted by the Energy Information Administration (EIA), is an example of a survey that links responses with administrative records. Since its first administration in 1978, RECS has been conducted a total of 13 times; the most recent administration was in 2009. The RECS household interview is conducted in person, and nonresponding cases are followed up by mail or phone (EIA, 2011).

Respondents are asked several questions during the 45 minute household interview for which they must likely consult records. The interview consists of 13 sections, including questions on housing unit and household characteristics, household appliances, heating, and fuel usage. Toward the end of the interview, the interviewer takes various measurements of the respondent's dwelling unit.

During the second half of the interview, after a series of questions on fuel usage, respondents are asked to provide the interviewer with a copy of their most recent statement from each of their utility providers (for example, electric, gas, propane, kerosene, etc.).

> In this interview you have told me how your household uses energy. In addition, we would like to find out how much [fuel] you actually used in the past year. Getting that information directly from your energy suppliers would add to the data you've given me and improve our forecasts of energy consumption. At the end of the interview, I will ask for your authorization to contact your energy suppliers to retrieve this additional information. First, however, I'd like to collect some information about each of your energy accounts. You'll probably want to get any recent bills that were sent to you by your suppliers to help with these questions. If it is alright with you, at the end of the interview I will also scan copies of these bills into my computer. (EIA, 2009)

If respondents refuse to provide their utility statements, starting in 2009, interviewers note whether they believe this was due to a confidentiality or privacy concern, or for another reason, for example because the household's bill is an electronic one. Regardless of whether or not the respondent provides a bill, the interviewer asks for the name and address of each supplier.

The consent question was asked at the very end of the survey in 2009, but was asked earlier in prior administrations (see Appendix for the text of the consent form).

> Thank you for this information about your energy suppliers. So we can collect additional information from your fuel suppliers about the actual amounts of energy you use, would you please sign this authorization form that gives them your permission to give us that information? (DOE, 2009)

A second, related survey, the Energy Supplier Survey (ESS), is conducted among utility providers for whom authorization to contact was granted. The ESS is a required survey that collects data on household energy expenditures through mail and, beginning in 2009, internet as well. EIA produces national household energy use estimates by jointly modeling data from both the household survey and the ESS (EIA, 2011).

1.5.2 The Medical Expenditure Panel Survey

Another example of a survey that requests respondents' consent to link personal records with their survey responses is MEPS Household Component, conducted by the Agency for Healthcare Research and Quality (AHRQ). Since 1996, the MEPS Household Component has collected data from individuals, families, medical providers, and employers. A related survey, the MEPS Insurance Component, surveys employers on matters relating to employer-based health insurance.

The MEPS Household Component collects data from a nationally representative subsample of households who participated in NHIS the prior year. The interview includes questions on respondents' health status, health conditions, use of medical services, health expenditures, health insurance, and demographic and socioeconomic characteristics. Each panel of MEPS includes five interviews spanning two calendar years, permitting detailed measurement of change in respondents' health status and income, and the interplay between the two.

In addition to survey data, MEPS collects records from pharmacies, medical providers, and hospitals for respondents who report using these services. The collection of this data varies by panel round. Permission to contact the respondents' medical providers is requested during most rounds of data collection. Respondents are asked to sign an authorization form for *each* of these providers, authorizing MEPS to contact them and release the respondent's records.

> As I mentioned during the last interview, it is important for us to get accurate names and addresses for medical providers so that we can contact them for more information about the services they provide. To do this, we must have written authorization from the family members receiving these services. I would like to get authorization from the following people: [Lists providers.] These materials explain more about why we contact medical providers and answer questions people sometimes ask about this part of the study. Please take a minute to review this information while I prepare the forms. (AHRQ, 2009)

Respondents are asked separately for permission to contact their pharmacy provider if they mentioned using any prescription medication. This consent request is not included in all MEPS rounds.

> As you know, the Department of Health and Human Services is very interested in obtaining the most complete and accurate information about health care use and expenditures, including prescription medicines. Many pharmacies now offer their customers a summary of their prescription medicine charges. People sometimes request these summaries to help in preparing their taxes or insurance claims. To help us get the best information about the family's prescriptions, we would like to obtain a printed summary from each pharmacy used by this family during the past year. To do this, we must have written authorization. From the information I have, I would like to get a signed authorization form for: (person)'s prescriptions filled at (pharmacy). These materials explain more about why we contact pharmacies and answer questions people sometimes ask about this part of the study. Please take a minute to review this information while I gather the forms. (AHRQ, 2009)

Once the provider receives the authorization form, the Medical Provider Component is conducted over the telephone and information is collected regarding the respondent's diagnoses, payments, and services provided. Pharmacies contribute information on prescriptions filled, their quantity, dosage, and strength, and payment information – their data is either provided over the phone or uploaded on a disk and sent to MEPS.

The information collected from the Medical Provider Component and Pharmacy Component is not used to provide national estimates, but rather to edit and impute data collected during the Household Component. AHRQ considers provider-contributed information to be more complete and less prone to reporting errors than information collected in the household survey (Machlin & Taylor, 2000).

1.6 Consent Bias

In situations for which respondents are required to consent in order for record linkage to occur, not everyone usually consents (Jenkins et al., 2008). If consent is not universal, there is potential for bias in analyses using the linked data (Huang, Shih, Chang, & Chou, 2007; Sala et al., 2010). Harris, Cook, Victor, Beighton, DeWilde and Carey (2005:336) provide an illustration: if respondents in poorer health are more likely to consent, than research based on the linked data would overestimate disease prevalence.

Analogous to survey response rates, maximizing consent rates can potentially reduce, but not necessarily eliminate, consent bias (Harris et al., 2005), which is affected by the difference between the consenting and non-consenting respondents on the estimated variables. As with nonresponse bias, consent bias is a multiplicative function of the consent rate and the difference between consenters and nonconsenters. As consent to record linkage is conditional on survey response, differences between consenting and nonconsenting respondents can increase the disparity between survey respondents and nonrespondents (Harris et al., 1995; Tate, Calderwood, Dezateux, & Josh, 2005). Theoretically, this could minimize differences between survey respondents and nonrespondents as well.

Consent is typically high among respondents (Dunn, Jordan, Lacey, Shapley, & Jinks, 2004) although consent rates vary widely among surveys (Klassen, Lee, Barer, & Raina, 2005) and some surveys demonstrate consent rates lower than response rates reported by the same survey (Sakshaug & Kreuter, 2011). Because not all respondents consent to record linkage, researchers may need to increase the initial survey sample size in order to compensate for those who do not consent, although this would not alleviate any bias caused by consent refusal (Angus, Entwistle, Emslie, Walker, & Andrew, 2003; Dunn et al. 2004; Harris et al., 2005; Huang et al., 2007; Jenkins, Cappellari, Lynn, Jackle, & Sala, 2006; Sakshaug et al., 2011). There is also concern that merely requesting consent to record linkage affects survey response rates (Angus et al., 2003; Jenkins et al., 2006; Korkeila, et al. 2001; Nelson et al. 2002). However, some research indicates that including such a request has no effect (Shah et al., 2001).

It is typically infeasible to access the administrative records of the respondents who do not consent to linkage, and therefore it is rarely possible to estimate bias due to consent refusal. In the German Labour Market and Social Security survey (PASS), a study of German benefit recipients, for which differences between consenting and non-consenting respondents were available at the aggregate level, Sakshaug (2011) identified consent biases for some variables (age and foreign status). These biases were smaller than the survey's nonresponse and measurement biases and the author concluded that data linkage was advantageous from a Total Survey Error perspective. However, in addition to the narrowness of the target population, other limitations of this study may limit its generalizability, including its response rate (26.7%), the unknown quality of the administrative data, and the fact that the administrative data were merged across various sources.

Evidence from several national surveys demonstrates that the public is becoming less willing to consent to linkage requests (Bates, 2005; Dahlhamer et al., 2007). For example, consent refusals in the Survey of Income and Program Participation (SIPP) increased from 12% to 35% between the 1996 and 2004 panels; in the Current Population Survey (CPS) from 10% to 24% between 1994 and 2003 (Bates, 2005); and in the NHIS from 15% to 50% between 1993 and 2005 (Dahlhamer et al., 2007).

1.7 Variations in the Consent Request

Little is known about what consent approaches are best at maximizing respondents' willingness to consent to record linkage (Partin et al., 2008) and there do not appear to be any widely accepted "best practices" for soliciting permission to access respondent records. However, it is important to mention that at least some of the variability in consent procedures stems from differences in institutional review board and legal requirements specific to individual institutions.

1.7.1 Informed Consent Requirements

Requirements for obtaining respondent consent as a prerequisite in conducting record linkage can vary by country (Sala et al., 2010) and by type of record, (Tu et al., 2004). Within the U.S., requirements vary by state (McCarthy, Shatin, Drinkard, Kleinman, & Gardner, 1999) and there is potential for variation by sponsoring research organization as well (GAO, 2001). According to GAO (2001:57), "Perceptions about the need for consent may vary according to type of linkage." This suggests that at least within the U.S., consent requirements can vary depending on whether survey responses are being linked with person-level records or contextual records such as characteristics of a geographic area. Regarding only person-by-person linkage, the focus of this dissertation, a variety of viewpoints have been proposed as to whether or not consent should be obtained before linkage is attempted.[6]

Gastwirth (1986) and Scheuren (1997) among others reason that respondents should be given the opportunity to evaluate any potential benefits of linkage as well as any confidentiality or security risks in order to make an informed decision. Explicitly asking respondents for their consent provides them with a degree of control over the use of their personal records because linkage is not performed for those who do not consent (GAO, 2001). If the consent request is not proposed, individuals are likely unaware that any linkage is taking place (GAO, 2001). To further maximize control, researchers should follow up with consenting respondents periodically in case respondents' consent preferences or researchers' objectives change (Calderwood et al., 2009).

Others argue that obtaining consent from all respondents, or at least some, is unnecessary. For example, Wallman and Coffey (1997) suggest that securing consent may not be necessary in certain situations if proper confidentiality and security measures are in place. Melton (1997) suggests that if the vast majority of those who are asked consent, than consent does not need to be requested from all respondents. Requesting consent could be burdensome, biasing, and impractical, potentially hindering researchers from executing valuable research (Melton, 1997). Others speculate that requesting

---

[6] For surveys that link responses with administrative records including the Medicare Current Beneficiary Survey, New Beneficiary Survey, and National Long Term Care Survey, the sampling frame is drawn directly from the administrative record file which greatly facilitates record linkage (Lillard et al., 1997).

consent has a detrimental effect on survey recruitment and representativeness (Partin et al., 2008:1033).

However, another point worthy of consideration is that while most surveys are voluntary, participation in administrative records is typically not optional (GAO, 2001). Individuals sampled to be in a survey can choose whether to respond. However, given the mandatory nature of the administrative records, if they are not explicitly asked to consent to record linkage, they have little control over the use of these records. Some have advocated for obtaining consent due to the nonvoluntary nature of filing these records, as well as for linkages that are not substantively related to the original data collection, for vulnerable populations, for high risk linkages, or for linkages conducted within longitudinal surveys where a greater amount of data is accumulated over time relative to cross-sectional surveys (Calderwood et al., 2009; GAO, 2001).

1.7.2 U.S. Policy Relating to Record Linkage and Privacy

Arguably, individuals should be able to control the way that personal information about them is used (GAO, 2001). Yet a tension exists between fully explaining the details of the proposed record linkage to respondents and the likelihood that they will agree to have their records linked (Lane, 2010). For many, an individual's right to maintain their personal privacy is at odds with the greater good of expanding research and knowledge (Melton, 1997). As concerns about linkages grow, ethical and regulatory policies have expanded to address them (Calderwood et al., 2009). In the U.S. government, this is addressed through several policies, notably the Privacy Act, specific agency statutes, and the Common Rule (GAO, 2001).

The Privacy Act of 1974 is a government-wide statute that "governs the responsibility of federal agencies concerning the content, access, and disclosure of records concerning individuals" (GAO, 2001:22). According to the Privacy Act, an individual's consent is required before their information can be disclosed to third parties:

> No agency shall disclose any record which is contained in a system of records by any means of communication to any person, or to another agency, except pursuant to a written request by, or with the prior written consent of, the individual to whom the record pertains… (The Privacy Act of 1974 5 U.S.C. §552a)[7]

However, there are twelve exceptions to accommodate legitimate needs for identifiable information. This includes conducting research and statistical activities involving record linkage (GAO, 2001).

> (b) Conditions of disclosure[8]
>
> No agency shall disclose any record which is contained in a system of records by any means of communication to any person, or to another agency, except pursuant to a written request by, or with the prior written consent of, the individual to whom the record pertains, unless disclosure of the record would be-- …
>
> > (5) to a recipient who has provided the agency with advance adequate written assurance that the record will be used solely as a statistical research or reporting record, and the record is to be transferred in a form

---

[7] Retrieved from http://www.justice.gov/opcl/privstat.htm on August 11, 2011.

[8] Numbers and labels are from original document.

that is not individually identifiable (The Privacy Act of 1974 5 U.S.C. §552a)[9] [10]

In addition to the Privacy Act, some agency-specific statutes apply to record linkage. For example, Section 9 of Title 13 requires that information at the Census Bureau is kept confidential and only used for the statistical purposes for which it was intended (GAO, 2001). The Public Health Service Act of the National Center for Health Statistics (NCHS) ensures that, “no information obtained in the course of NCHS' activities may be used for any purpose other than for which it is supplied unless there has been consent” (GAO, 2001:24).

In addition to the Privacy Act and agency-specific statutes, the Federal Policy for the Protection of Human Subjects, frequently known as the Common Rule, is a federal regulation governing research that involves human subjects; this may include record linkage. Under the Common Rule, Institutional Review Boards evaluate and approve research projects using criteria such as whether subjects' informed consent is obtained (GAO, 2001).

1.7.3 Agencies Sponsoring Record Linkage

Record linkage is sponsored by a variety of government agencies. Research agencies, such as the National Cancer Institute and the National Institute for Occupational Safety and Health, conduct record linkage in order to study various aspects of health (GAO, 2001). The statistical and research offices of program agencies with access to large datasets routinely conduct linkage, for example, Statistics of Income at the

---

[9] Retrieved from http://www.justice.gov/opcl/privstat.htm on August 11, 2011.

[10] Although the records must be de-identified during transfer, they must be personally-identifiable to be useful for linkage.

Internal Revenue Service (IRS) and the Office of Research Evaluation and Statistics at SSA. Agencies such as the Substance Abuse and Mental Abuse and Mental Health Services Administration and others administering block grants also conduct record linkage, for example, linking the records of individuals treated for drug addiction with their employment and treatment records (GAO, 2001).

The Centers for Medicare and Medicaid Services also provide enrollment and claims administrative data that is linked with surveys such as the Panel Survey of Income Dynamics (PSID), HRS, and several NCHS surveys. For respondents who provide their Medicare number and for whom a Medicare record match can be made, several files (for example, claims filed by outpatient providers or hospice providers) can be linked with survey responses. Linkages can also be made with the Denominator file which provides data on all enrolled Medicare beneficiaries and demographic information on those enrolled during that calendar year. Beginning in 2006, Medicare Part D files were also linked with survey responses for NCHS survey participants (NCHS, 2010).

The Social Security Administration (SSA) probably holds the largest collection of administrative records used for research. SSA data is collected with the primary purpose of administering SSA programs. Over time, this data demonstrated valuable use for research and policy purposes (Haines & Greenberg, 2005). SSA first entered into a Memorandum of Agreement with the Census Bureau in 1967 which effectively permitted linkages of data from Census Bureau surveys with SSA administrative data (Haines et al., 2005).

Through the SSA, respondents can be matched with a variety of administrative data sources. For example, survey data can be matched to the Master Beneficiary Record which records an individual’s beneficiary and payment history and is created when an individual applies for benefits. Another source, the Supplemental Security Income (SSI) Record, records an individual’s Federal and State SSI Benefits, as well as disabled individuals’ disability diagnoses (Koenig, 2003). In addition, SSA distributes SSNs to nearly all legal residents of the US. The record of these SSNs is kept in the NUMIDENT file, which also contains applications of changes to SSNs as well as the holder’s name, date, and place of birth, their parents’ name, and eventually, their date of death. The coverage of these records is one of its most obvious benefits, considering that the vast majority of Americans have a SSN, and nearly 7.2 million Americans received SSI benefits in 2006 (Drazga-Maxfield, 2008).

1.8 Potential Influences on Consent

Relatively little is known about who gives consent and who refuses to give consent to record linkage or why they do so. In the next section, we review relevant theory and findings from the prior work on this issue.

1.8.1 Cost Benefit Analysis

According to Morton-Williams (1993), people’s decision to participate in a survey may be influenced by their perceptions of the interviewer, a sense of personal responsibility, confidence in their willingness to help, empathy, their mood, perceived legitimacy of the request, or a cost-benefit comparison of perceived gains and losses in participating. Similarly, in an application of Lewin’s (1942) Field Theory of Motivation, Kahn and Cannell (1957) suggest that in deciding whether or not to cooperate, individuals assess their relevant goals, needs, and desires, which may be in opposition to one another, in order to determine whether or not they will ultimately cooperate.

As they do in deciding whether or not to cooperate (Singer & Ye, 2010), survey respondents presumably evaluate the potential gains and losses of their decision of whether to consent to record linkage. Respondents may base this decision on their perception of the attractiveness (or unattractiveness) of the basic features of the request (Groves, Cialdini, & Couper, 1992). According to the Leverage Salience Theory of survey participation (Groves, Singer, & Corning, 2000), sampled individuals vary in the characteristics they perceive relevant to their response decision and the importance that they assign to these characteristics, or their "leverage." The propensity to respond depends on both the leverage they assign to survey features and how salient those characteristics are made in the request. The valence of these features is also critical, that is, whether its associated emotion is positive or negative. The valence and perceived importance of the various features is a function of an individual's background and prior experience.

This theory can be extended to the consent request. Respondents might note various features of the consent request (for example, the sponsor, records or personally-identifying information requested, or the degree of confidentiality protection offered), and weight this information differentially in deciding whether they will consent to record linkage. Further, the weight applied to different elements and the valence associated with the elements can vary by respondent.

The Subjective Expected Utility Theory framework, initially proposed by Savage (1954), describes how individuals make decisions by considering their potential losses and gains. In an application of Utility Theory to survey response, Rasinski, Baldwin, Willis, and Jobe (1994) and Willis, Sirkin, and Nathan (1994) found that in deciding whether or not to respond truthfully to a survey question, respondents will consider both potential gains and losses associated with their response.

Extending the Subjective Expected Utility Theory and Leverage Salience Theory to the consent request, we expect respondents to consider the various features of the consent request– both beneficial and detrimental, and weigh any potential gains and losses accordingly in making their decision. Respondents will consider both costs and benefits of the request. As described by Morton-Williams regarding survey interviews (1993), costs include preserving themselves from danger or stress, limiting any disruption or intrusiveness and avoiding an unpleasant or embarrassing experience. Benefits include taking part in a pleasant, interesting, or novel experience, and participating in a worthwhile endeavor. One could also add to these benefits contributing to research or the sheer act of helping.

The potentially serious consequences of disclosure of sensitive information and identity theft associated with record linkage could more substantially influence respondents' consent decision than positively-valenced aspects such as fulfilling an altruistic motive, increasing rapport with the interviewer and gaining the interviewer's approval upon making their decision. When faced with the consent decision, respondents may be prone to weight these negative features more heavily. Other features, such as the source of the request or record topic may be positively, negatively, or neutrally-valenced, depending on an individuals' background and personal experience.

1.8.2 Respondent Level Influences

1.8.2.1 Respondent Personal Characteristics

Research examining predictors of linkage consent is largely confined to requests for medical records, and focuses primarily on respondent demographic characteristics such as age, gender, education, race, and ethnicity. While these studies often identify demographic differences between consenters and refusers, the differences are not consistent across studies. Indeed, according to Sala et al. (2010: 2), “characteristics that are associated with higher consent in one study are negatively associated with consent in another.”

For example, males are more likely to consent to record linkage in some surveys (Dunn et al., 2004; Koenig, 2003; Woolf et al., 2000), although others find no significant gender differences (Baker, Shiels, Stevenson, Fraser, & Stone, 2000; Harris et al., 2005; Huang et al., 2007; Jenkins et al., 2008; Korkeila et al., 2001). Older respondents are more likely to consent in some surveys (Harris et al., 2005; Partin et al., 2008; Woolf et al., 2000), while younger respondents are more likely to consent in others (Cleary, Mechanic, & Weiss, 1981; Dunn et al., 2004; Huang et al., 2007; Jay, Belli, & Lepkowski, 1994; Sala et al., 2010). Jenkins et al. (2008) found that middle-aged respondents were least likely to consent and several other studies found no age differences (Baker et al., 2000; Finkelstein, 2001; Korkeila et al., 2001).

Similar inconsistencies exist in regards to respondents’ education, with more educated respondents consenting more often in research by Cleary et al. (1981) and Huang et al. (2007), and less educated respondents consenting more often in research by Jay et al. (1994). Tate et al. (2005) concluded that the respondents with the greatest and least amounts of education were least likely to consent, and Korkeila et al. (2001) found no differences by education.

Minority respondents consented less often in research conducted by Haider and Solon (2000); Jay et al. (1994); Koenig (2003); Partin et al. (2008); Tate et al. (2005); and Woolf et al. (2000), yet more often in research by Huang et al. (2007). Korkeila et al. (2001) did not identify any significant difference in consent rates by minority status. Research by Cleary et al. (1981) and Klassen et al. (2005) concluded that lower income respondents were less likely to consent, though Huang et al. (2007) identified that lower income respondents were more likely to consent, and Jay et al. (1994) and Korkeila et al. (2001) found no difference by income status.

Some research shows consent rates higher for married respondents (Koenig, 2003; Partin et al., 2008), while another study shows no difference by marital status (Jay et al., 1994). Other research associates consent with area (Huang et al., 1997; Yaffe et al., 1978); employment status (Haider et al., 2000; Klassen et al., 2005; Tate et al., 2005), and item nonresponse on the income question (Jenkins et al., 2006; Olson, 1999; Sala et al., 2010; Woolf et al., 2000).

The Joint Program in Survey Methodology (JPSM) 2010 Practicum survey asked respondents about their (hypothetical) willingness to consent to administrative record linkage for seven different types of records: medical records, credit history, tax returns, government benefits, health insurance information, employment history, and Medicare records if the respondent was 65 years or older. Demographic characteristics such as age, race, gender or education did not demonstrate any consistent or meaningful differences.

Together, this evidence suggests that respondents' demographics and socioeconomic characteristics are not strong or consistent predictors of respondents' willingness to consent to record linkage. However, it is also possible that these relationships vary across studies because the role of respondents' personal characteristics varies in conjunction with more direct influences in the survey and consent request.

Some research demonstrates that data collection features such as the survey topic and record subject are associated with consent rates (Jenkins et al., 2006; Singer et al., 2003). Respondents may give consideration to these characteristics if pertinent features of the consent request are made salient. Further, respondents who find the request to record linkage more salient may be more likely to consent (Sala et al., 2010).

In an analysis of the 2005 PSID, respondents with a chronic illness were significantly more likely to grant access to their Medicare records and provide their Medicare number than respondents without a chronic illness (60.1% vs. 48.7%). Combining consent requests across the 2005, 2007, and 2009 panels demonstrates the same trend, with chronically-ill respondents being more likely to grant Medicare record access (74.2% vs. 60.5%; Fulton, Schoeni, & Freedman, 2011).

Dunn et al. (2004) demonstrate that respondents with the symptom under investigation in the survey were more likely to consent to medical record linkage than respondents without the symptom, even after controlling for age and gender. Similarly, Harris et al., (2005); Klassen et al., (2005); Korkeila et al., (2001); Partin et al., (2008); and Woolf et al., (2000) reported that less healthy respondents were more likely to consent to linkage of their medical records. Petty et al. (2001) found that respondents with more repeat prescriptions were more likely to consent to a medication record review.

Sala et al. (2010) determined that saliency indicators specific to the consent request, such as number of hospital stays or receiving government benefits were positively associated with consent. However, Baker et al. (2000); Huang et al. (2007); and Jay et al. (1994) found no such overrepresentation of less healthy persons.

In an analysis of the representativeness of the HRS Social Security Earnings Sample, Haider et al. (2000) determined that consent to SSA records (and willingness to provide SSN) varied by characteristics relating to their records. For example, respondents who reported that they never worked were less likely to consent as were respondents born outside the U.S.

Although the scenarios were hypothetical, respondents in the 2010 JPSM Practicum survey differed in their perceived sensitivity of income-related records and willingness to provide access to income and employment-related records. Higher income respondents generally found this information more personal and expressed less willingness to grant access to it (see Tables 1.01 and 1.02). Respondents who refused to answer the income question, an indicator of privacy, also rated this information more personal and indicated that they would be less willing to grant access.

Table 1.01

Perceived Sensitivity of Administrative Records, by Respondent Income from the 2010 JPSM Practicum Survey[11]

	2009 Household Income (n)			
	<30k (460)	30-75k (627)	>75k (485)	REF (166)
Tax Return**	2.97	3.15	3.12	3.18
Credit History**	2.91	3.29	3.13	3.26
Employment History	2.43	2.55	2.50	2.61
Medical Records***	3.18	3.32	3.60	3.36
Health Insurance Info*	2.66	2.92	3.04	3.07
Gov. Benefits	2.93	2.87	2.87	2.98

Table 1.02

Willingness to Grant Record Access, by Respondent Income from the 2010 JPSM Practicum Survey[12]

	2009 Household Income (n)			
	<30k (460)	30-75k (627)	>75k (485)	REF (166)
Tax Return*	2.05	1.84	1.80	1.49
Credit History ***	2.02	1.73	1.51	1.45
Employment History	2.29	2.08	2.19	2.08
Medical Records ***	1.71	1.51	1.31	1.34
Health Insurance Info *	2.10	1.91	1.86	1.41
Gov. Benefits*	1.96	2.06	2.05	1.43

---

[11] Q38: I am going to read a list of different types of records you might have. For each one of them, please tell me how personal you think it is. Use a scale where 1 is 'not at all personal' and 4 is 'extremely personal.' *$p<0.05$, **$p<0.01$, ***$p<0.001$

[12] Q39: How likely would you be to give your consent to the Census Bureau to obtain each of the following types of records? Use a scale where 1 is 'very unlikely' and 4 is 'very likely.' *$p<0.05$, **$p<0.01$, ***$p<0.001$

As with survey response, this suggests that a respondent's decision to consent may be specific to each individual request, and that respondents consider their own personal circumstances in conjunction with the specific identifier or records requested. If an individual's propensity to consent does vary within and across studies, their consent is a stochastic or probabilistic rather than deterministic phenomenon based on fixed attributes. As a result, the bias due to consent nonresponse will depend on the covariation between the consent decision and the survey variables (Bethlehem, 2002):

Average consent propensity among those who consent

$$bias(\bar{y}_{consent}) = \frac{\sigma_{yp}}{\bar{p}}$$

where σ_{yp} is the covariance between the variable of interest, y, and the consent propensity, p, and $\bar{p}$ is the mean consent propensity over the responding sample.

Davern et al. (2006) suggest that if the linking information is missing in a clear, understandable pattern, the difference between linked and unlinkable cases can be controlled at least to some extent by increasing the weight of the linked cases and treating the unlinkable cases like survey nonrespondents. Unfortunately, as demonstrated above, consent is typically not refused in any predictable or consistent pattern.

Current research on consent does not identify the conditions where consent refusal leads to consent bias. Rubin (1987) and Groves, Presser, & Dipko (2004) suggest that nonresponse error can result when topic interest motivates survey participation as those less interested in the topic may differ on key survey variables. Further, these key variables pertaining to the survey topic are likely to be most affected. Those more involved or interested in a survey topic may respond at a higher rate than those less involved (Goyder, 1987). Regarding consent it is unclear if respondents for whom the request is especially salient are consenting at a higher rate because of topic interest, perceived relevance, or some other reason.

1.8.2.2 Privacy and Confidentiality Attitudes

There are several features of the consent request that are indisputably negatively-valenced. For example, if requested, respondents may consider the burden involved in providing personally-identifying information if it is not accessible from memory (as when respondents have to look up the requested identifier from records (Jenkins et al., 2008)). Furthermore, respondents generally receive no incentive for consenting to record linkage and the absence of incentives may affect their consent decision. However, the most significant impediment is likely respondents' privacy concerns and the threat of disclosure of personal information. The extent to which these concerns influence a respondent's decision may depend on how salient they are made in the request (Singer, 2011).

Privacy and confidentiality concerns impact willingness to consent to record linkage. According to Singer, Mathiowetz & Couper (1993), privacy concerns reflect one's hesitation to disclose any information to others, including interviewers, whereas confidentiality concerns involve how the information is handled (or mishandled) once respondents disclose it. Respondents want to ensure that their personal information is kept safe from potentially dangerous third parties (Tourangeau et al., 2000). The role that privacy concerns play in survey participation has been explored in the survey methodology literature. Research investigating the influence of privacy and confidentiality concerns on participation in the decennial census has found that such concerns have a small but significant negative effect on the probability of returning the decennial questionnaire (Singer et al., 1993; Singer et al., 2003).

Census Bureau-sponsored research including focus groups conducted in 1992 and CATI surveys conducted from 1995 to 2010 gauged the public's attitudes toward the statistical use of administrative records, trust, and privacy over the past two decades. This research demonstrates that respondents' concern about their personal privacy is increasing, with 32% of respondents saying they were "very worried" about their personal privacy in 2010, up from 25% in 2000 and 24% in 1995. As another indication of increasing privacy concerns, in 2010, 31% said that the Census questions are an invasion of privacy, 10 percentage points higher than in 2000. However, in 2010, 40% strongly agreed with the statement "People have lost all control over personal information," somewhat of a decrease from 44% in 2000 (it was 40% in 1995). Further, 42% in 2010 strongly agreed with the statement "The government knows too much about me." It was 52% in 1995 and 43% in 2000 (Singer et al., 2011).

Existing research demonstrates respondents' discomfort in providing personally-identifying linking information. Both the 1992 Test Census and an experiment embedded in the 2000 Census found that including a request for the respondent's SSN increased unit and item nonresponse (Dillman, Sinclair & Clark, 1993; Guarino, Hill, & Woltman, 2001). The percent of respondents in the Survey of Privacy Attitudes who said that they would provide their SSN if the census form requested it fell from 68% in 1996 to 56% in 2000 (Singer, 2001). Singer (2001) found that among respondents who were opposed to providing their SSNs, the most frequently cited reason (22%) was privacy and confidentiality fears. However, when asked hypothetically, many more people say that they would refuse to provide their SSN than actually refuse the request (Guarino, et al., 2001; Singer, Bates, & Miller 1992).

Respondents' privacy and confidentiality concerns are likely to be magnified when individual databases are linked with others, given the amount of detail that linked records can provide on an individual respondent (Fienberg, 2006) and the increased risk of re-identification with linked data (GAO, 2001). When linked, any risks associated with sensitive data are exacerbated, and even non-sensitive data can become more sensitive when linked (GAO, 2001).

Trust concerns also may be influential. Those who refused consent to medical record linkage in research by Cleary and Jette (1984) were less trusting of their physician. Sala and colleagues (2010) found that generally trusting others was positively associated with consent in the BHPS.

Respondents may be especially cautious in disclosing personal identifiers to facilitate record linkage or grant record access due to fears of identity theft, or the unlawful access to personal information, which they may perceive as a risk associated with record linkage.[13] To date, no study has demonstrated this link. Even though survey estimates of the prevalence of identity theft are quite low (see Table 1.03), the fear of occurrence is quite high. According to a 2009 Gallup Poll, 66% of respondents worry frequently or occasionally about being the victim of identity theft, the highest level for any crime respondents rated in the survey.

---

[13] In fact, the Federal Trade Commission (FTC) includes SSN in its definition of identity theft, defining it as, "when someone uses your personally identifying information, like your name, Social Security number, or credit card number, without your permission, to commit fraud or other crimes" retrieved from http://www.ftc.gov/bcp/edu/microsites/idtheft/consumers/about-identity-theft.html on August 10, 2011.

Table 1.03

Survey Estimates of Identity Theft Prevalence[14]

% Victim of ID Theft	Reference Period	Sponsor	Year	n	Mode
8	12 mos	Gallup	2010	1025	CATI
10	12 mos	Gallup	2009	1013	CATI
10	2 years	AARP	2008	1007	CATI
4	12 mos	FTC	2006	4,917	CATI
5	12 mos	FTC	2003	4,057	CATI

Simply requesting consent to access additional data may bring privacy and confidentiality concerns to mind (Jenkins et al., 2006), although others suggest that giving respondents control over their information increases trust (McCarthy et al., 1999). Respondents' refusal may simply be a byproduct of unfamiliarity with record linkage. GAO (2001:17) suggests that "much record linkage likely remains invisible to the general public and some policy makers as well." In research utilizing an online web panel that manipulated the length of record linkage consent statements, Das (2011) determined that respondents who received the longer consent statement were more likely to consent, understood more about the linkage, and had a lower perceived risk.

The income question is sometimes used as a proxy for privacy concerns (Hurd, Juster & Smith, 2003; Juster & Smith, 1997), and those who refuse the income question are more likely to refuse consent (Jenkins et al., 2006; Olson, 1999; Sala et al., 2010; Woolf et al., 2000). Respondents who believed that their linked data could be used to detect fraud were less likely to consent (Gray, 2008) as were those who did not believe their personal information would be kept confidential (Armstrong et al., 2008).

---

[14] Gallup/AARP: "Please tell me which, if any, of these incidents have happened to you or your household within the last (twelve months/two years)…. *You or another household member was the victim of identity theft…?"*

FTC: "Have you ever been notified by a company, government agency, or other organization that it had lost your personal information, such as an account number or your social security number, or that the information had been stolen or hacked?

The request to disclose an SSN or some other identifier to an interviewer to facilitate record linkage might be especially alarming to respondents. Fifty-three percent of respondents to a survey conducted by Shell Oil (1999) believed that using their SSN as a standard form of identification is a major invasion of privacy and another 29% felt that it was a minor invasion of privacy. In another poll conducted that year, 92% said that they were "very concerned" or "somewhat concerned" about the protection of their SSN (National Consumers League Survey, 1999). Furthermore, 36% say that they have received an unsolicited email requesting personal information such as a bank account number or their SSN (Pew Internet and American Life Project Poll, February 2007).

Some surveys do attempt to justify these intrusive requests to respondents. The Survey of Privacy Attitudes found that respondents were more likely to say they would allow their administrative data to be linked if the purpose was to increase the accuracy of the information than if the purpose was to reduce costs (Singer, 2001; Tourangeau, Singer, & Presser, 2003).

Other surveys "sell" the request as reducing the burden on respondents (Bates, 2005; Dahlhamer et al., 2007). However, with the exception of the SLID (Michaud et al., 1995), respondents themselves do not directly benefit with a survey that is reduced in length or burden if they consent to record linkage. The survey questions are not any shorter or less difficult.

Respondents' reluctance to provide personal identifiers to an interviewer or grant consent to record linkage may reflect the absence of any obvious benefit to the respondent in doing so. It is likely that many respondents focus on the potential disadvantages of consent (potential for disclosure and fear of prosecution) rather than the advantages to the agency (reduced data collection costs, increased accuracy, and broader analytic possibilities).

1.8.2.3 Respondent Altruistic Motives

Respondents may consider fulfilling an altruistic motive or sense of civic obligation as benefits to consenting to record linkage, especially if this is highlighted in the consent request. For example, if respondents withhold consent in the NIS, interviewers are instructed to tell them that the success of the NIS "depends on the voluntary cooperation of thousands of concerned households like yours" (CDC, 2011). Similar factors are thought to motivate survey response (Couper, Singer, & Kulka, 1998; Goyder, 1987; Groves et al., 1992; Morton-Williams, 1993; Singer, 2011), though respondents are likely more familiar with and less threatened by using surveys to conduct research. They may understand how surveys benefit society and research more clearly than how providing an SSN in a survey context achieves these same goals.

Some empirical findings support the idea that an altruistic motivation may lead respondents to consent. Sala et al. (2010) found that community-minded respondents (those more liberal, more likely to trust others, and who participated in volunteer work) were more likely to consent to record linkage in the BHPS. Similarly, consent was higher among respondents who believed that society would benefit from data linkage (Dunn et al., 2004; Jenkins et al., 2006).

1.8.3 Interviewer Level Influences

The existing research on consent focuses almost exclusively on respondent characteristics, largely ignoring the influence of interviewers. Although a substantial amount of work evaluates the impact of interviewers on survey response and response rates, to date, far less work has examined how these interviewers affect respondent consent. This dissertation extends prior work by investigating interviewer effects on consent to record access.

Research demonstrates that, on average, interviewers tend to produce responses that are more alike than in the sample overall (Groves et al., 2004; Hox, de Leeuw, & Kreft, 1991). Recent research yields evidence that some interviewer variance may be nonresponse variance as interviewers may differ in their success at gaining cooperation from different types of respondents (West & Olson, 2010). Measures of interviewer variance, such as the intraclass correlation, indicate the amount of variance attributable to the interviewer, but do not indicate which interviewer characteristics are responsible for these effects. Most research investigating the effects of specific interviewer characteristics suggest that observable traits influence estimates only when they are perceived by the respondent as relevant to the survey questions (Groves et al., 2004).

Interviewers vary in the response rates they achieve with some interviewers exhibiting greater effectiveness in reducing refusals and noncontacts than others (Campanelli & O’Muircheartaigh, 1999; Lyberg & Dean, 1992; Lyberg & Lyberg, 1991; O’Muircheartaigh & Campanelli, 1998; Singer, Frankel, & Glassman, 1983; Snijkers, Hox, & de Leeuw, 1999). In their review of past studies of interviewer effects on survey responses, O’Muircheartaigh et al. (1998) demonstrated that for each of the 12 studies reviewed, the average ρ_i value, or intra-interviewer correlation, was no greater than 0.02. Still, ρ_i is a property of a specific estimate and each of the 12 studies reviewed by O’Muircheartaigh et al. (1998) have ρ_i values greater than 0.02, including one value of 0.2 for a study of mental disabilities conducted by Freeman and Butler (1976).

In a review of the literature on interviewer effects, Hox et al. (1991:440) conclude that “The only consistent findings concerning the sociodemographic characteristics were for interviewer race, especially when race connected questions were asked.” Additional findings demonstrate that effects of interviewers’ race are only evident or strongest when the questions concern race (Anderson, Silver, & Abramson, 1988; Hatchet & Schuman, 1975; Kane & Macaulay, 1993; Schaeffer, 1980; Schuman & Converse, 1971).

Similarly, the effect of interviewers’ gender on responses seems more marked for gender-related questions, though interviewer gender has demonstrated effects on a broader range of questions. Kane et al. (1993) note gender-of-interviewer effects on responses for gender-sensitive questions, and Nealon (1983) found that female respondents reported higher farming activity to male interviewers. Groves and Fultz (1985) obtained more optimistic responses regarding economic outlook when respondents were interviewed by a male, but otherwise found no differences in item missingness or responses to factual questions by interviewer gender.

Interviewers’ age appears to affect responses but the direction of this effect is inconsistent. Berk and Bernstein (1988) and Hanson and Marks (1958) found less item nonresponse when the interviewers were older and younger, respectively. Sudman and Bradburn (1974) and Singer et al. (1983) concluded that older interviewers caused less bias and response variance; however, Freeman et al. (1976) reported that older interviewers (in combination with older respondents) contributed the highest amounts of interviewer variance. Singer et al. (1983) found that older interviewers obtained higher response rates, yet Collins (1980) concluded that interviewers’ age had no clear effects.

Interviewer experience, although likely less visible to respondents, can still influence responses and response rates. Groves and Couper (1998) suggest that with increased experience, interviewers become more effective at gaining cooperation by learning how to tailor their behavior to meet the needs of different situations. However, in practice, interviewer experience demonstrates mixed effects on responses and response rates. There is evidence that interviewers perform better with experience, with some research indicating higher response rates and better data quality (Feldman, Hyman, & Hart, 1951; Singer et al., 1983). Yet other research reports poorer data quality with increased experience (Bailar, Bailey & Stevens, 1977) or no identifiable effect of experience on responses (Berk et al., 1988; Boyd & Westfall, 1955; Collins, 1980).

A logical extension of the research on interviewer effects is the application of these findings to the request for respondent consent to record linkage. Just as some interviewers are more successful at obtaining unit and item response, it is possible that certain interviewers are more effective at gaining respondent consent to record linkage, due to their differing levels of experience, their ability to tailor the request to respondents' concerns, or other characteristics. The consent request can be considered a sensitive and challenging question for both interviewers and respondents and may require more interviewer skill and tact on the part of the interviewer than most items.

Some previous research identifies significant interviewer-contributed variance on requests for consent to record access. Cleary (1981) identified strong interviewer effects for a request for respondents' written consent granting access to medical records in a state-level mental health survey. After controlling for respondent age, income and education, Sakshaug et al. (2010) identified a significant interviewer variance component for the consent request in the HRS. Sakshaug et al. (2011) similarly identified a significant interviewer variance component in PASS. Albeit not directly comparable to traditional survey interviewers, in research requesting consent to access medical records, consent rates varied by whether office staff or a physician made the request, suggesting that respondents are sensitive to who initiates this request (Armstrong et al., 2008; Baker et al., 2000).

Further, interviewers may be uncomfortable requesting consent. Focus groups conducted in conjunction with the NHIS confirmed that interviewers are reluctant to request sensitive information, like income, or personal identifiers, such as an SSN (Meyer, Dahlhamer, & Pleis, 2006); this suggests that greater interviewer experience may facilitate higher rates of consent. Olson (1999) hypothesized that greater interviewer experience on average contributed to higher rates of successful SSN collection in SIPP compared to other surveys. Cleary (1981) determined that experienced interviewers were more effective at gaining respondents' written consent to their medical records. Sala et al. (2010) concluded that interviewers with more experience on a particular wave of the BHPS achieved higher consent rates on that wave, yet interviewer job perception measures, personality, and attitudinal measures had no effect on their ability to gain consent.

In the BHPS, likelihood of consent was weakly associated with having been interviewed by the same interviewer the prior year (Sala et al., 2010). Jenkins et al. (2006) determined that consent was positively related to the quality of the interviewer-respondent interaction, or the interview "smoothness," suggesting that consent likelihood may be influenced by rapport.

It is possible that similar to survey response, interviewer sociodemograhic characteristics affect their ability to gain consent as well. Sakshaug et al. (2010) found that black interviewers were significantly less successful than white interviewers in obtaining SSNs in the 2004 HRS. However, research by Sala and colleagues (2010) found that interviewer sociodemographic characteristics, including gender, age, and education, were not related to consent rates in the BHPS. Interviewer gender was not a significant factor in determining whether or not a respondent consented to link their pregnancy and birth records with survey responses in research conducted by Tate et al. (2005).

1.8.4 Characteristics of the Request

1.8.4.1 Sponsorship

Surveys vary in the wording of the consent request, in what records are to be accessed, and in whether consent is expressed orally to the interviewer or in writing with a signature. In addition, some surveys request identifying information to facilitate the record linkage. The surveys that request access to records differ in their mode, topics, sponsorship, and response rates. It is not known whether or how these factors affect respondents' consent likelihood. For example, the source of the request may affect the consent decision. Authority is clearly an influence on other requests (Cialdini & Goldstein, 2004) and respondents are more likely to consent to a request made by a sponsor that they trust and consider to be legitimate (Groves et al, 1992).

In a meta-analysis of mail surveys, Heberlein and Baumgartner (1978) identified a 12.4% gain in response rates associated with government sponsorship, controlling for survey topic and number of contacts. Groves et al. (2004) speculate that government surveys attain higher response rates than academic or private sector-sponsored surveys because sampled individuals believe that the government needs and will use survey information to benefit them. Just as survey sponsors can influence response rates (Heberlein et al., 1978), more trusted survey sponsors may also gain higher rates of consent. It is more likely that respondents will consent to record linkage if the request comes from a sponsor viewed as having legitimate reason to make such a request, such as a government agency (Groves et al., 1992; Morton-Williams, 1993). A legitimate source may also be seen as more likely to safeguard information.

In the Census Survey of Privacy Attitudes, favorability towards the statistical use of administrative data (specifically, sex, age, DOB, race and Hispanic origin information) varied by the agency that was sharing the information with the Census Bureau. In 2010, the most recent year of the survey's administration, 60% favored the SSA while only 52% favored the IRS; a minuscule 23% favored the Census Bureau receiving administrative data from a private credit agency (Singer et al., 2011).

During the typically lengthy and potentially hard-to-understand request statements (Willis, 2006) respondents may decide to agree to these requests, not because they understand them, but because they originate from a trusted authority figure. For example, consider this request for Medicare number from the 2009-2010 NHANES survey:

> May I please see your Medicare card to determine the type of coverage and to record the Health Insurance Claim Number? This number is needed to allow Medicare records of the Center for Medicare and Medicaid Services to be easily and accurately located and identified for statistical or research purposes. We may also need to link it with other

> records in order to re-contact you. Except for these purposes, the Department of Health and Human Services will not release your Health Insurance Claim Number to anyone, including any other government agency. Providing the Health Insurance Claim Number is voluntary and collected under the authority of the Public Health Service Act. Whether the number is given or not, there will be no effect on your benefits. This number will be held in strict confidence. (NCHS, 2010)

A similarly long example is the request for respondents SSN made in the 2009 NHIS:

> Finally, we would like the last four digits of your Social Security Number. This information will help us link your survey data with health-related records of other government agencies, and allow us to conduct additional research without taking up your time with more questions. The National Center for Health Statistics uses this information for research purposes only. Providing this information is voluntary. Federal laws authorize us to ask for this information and require us to keep it strictly private. There will be no effect on your benefits if you do not provide this information. (NCHS, 2009)

This mechanism may also work at the interviewer level: respondents may consent because they view the interviewer as an authority figure, or because they seek the interviewer's acceptance or social approval. Similar motives have been proposed to explain survey response (Cialdini et al., 2004; Groves et al., 1992).

1.8.4.2 Opt-Out Requests

Of surveys that request consent, the ambiguity of the request varies. Some surveys such as HRS are very explicit. The request explains in clear, comprehensible language which records will be transferred and linked, why linkage is required and the conditions of release. Further, respondent's written consent is required to conduct the linkage. The unambiguous request made in the HRS is a clear contrast to opt-out consent requests.

Until 2006, the Census Bureau requested respondents' SSNs to facilitate record linkage in the SIPP and CPS. However, because respondents' became less willing to provide this information over time and because the Office of Management and Budget (OMB) asked agencies to limit their collection of SSNs in surveys (OMB, 2007), the Census Bureau changed their consent requirements from an active procedure to a passive "opt-out" request (McNabb et al., 2009).

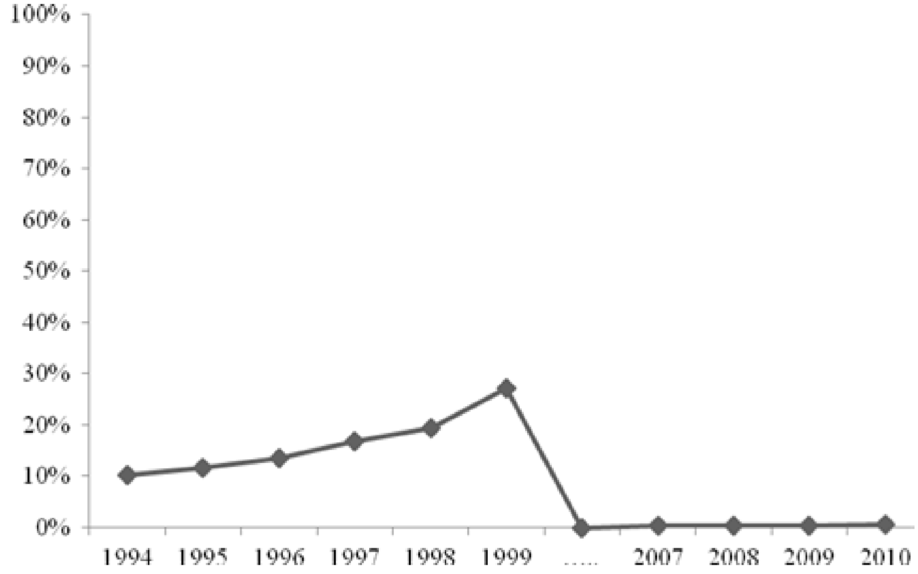
Reacting to this increasing SSN nonresponse, the Census Bureau stopped directly requesting SSN as an identifier. Instead of asking respondents to provide personally-identifying information to indicate consent, respondents under the "opt-out" consent procedure are assumed to have given their consent unless they explicitly state otherwise. Unless a respondent objects, or "opts-out", the respondent's survey data will be linked with other personal information for research purposes (McNabb et al., 2009). In absence of their SSN, the Census Bureau aggregates information for each consenting respondent from the SSA NUMIDENT file which contains information from the respondents' SSN application with their address records from the IRS, SSA, and other sources in order to identify the respondent's SSN. When a match is found, the SSN is used to link survey data with administrative record data for that respondent (McNabb et al., 2009). Here is an example of the opt-out procedure used in the 2008 CPS prenotification letter:

> Occasionally, we may combine data from the CPS with data we obtain from other government agencies to provide a comprehensive set of summary information about employment, income, and participation in various government programs. The same confidentiality laws that protect your survey answers also protect any additional information we collect (Title 13, United States Code, Section 9.) To ensure your protection, the laptops used for the data collection are password protected and all survey responses are encrypted. *If you wish to request that your information not be combined with information we obtain from other agencies, we ask that you notify the Field Representative at the time of the interview* [emphasis added]. (U.S. Census Bureau, 2008)

The passive, opt-out consent approach improved consent rates and now only a very small percentage of their survey respondents explicitly refuses consent to record linkage (see Figure 1.01). The opt-out procedure shifts the burden from consenting (which entails recalling and reporting an SSN or other personal identifier), to refusing to consent (which requires reading the pre-notification letter, identifying and understanding the opt-out statement, and notifying the field representative). Census initially conducted research in 2005, including cognitive interviews and a field test, which examined the effectiveness of opt-in as compared to opt-out wordings. Although the opt-out wording tested in that research is quite different from the current wording used, it did obtain higher rates of consent than more traditional “opt-in” consent requests that include a request for personally-identifying information (Bates, 2005).

Figure 1.01

Rates of Consent Refusal to Link Survey and Administrative Data in the CPS



1.8.4.3 Personally-Identifiable Linking Information

Some surveys request personally-identifying information to facilitate record linkage, for example, respondents’ medical provider contact information, as in MEPS and NIS; their SSN, as in HRS; or Medicare number, as in PSID. If linking information is requested and provided by a respondent, some surveys assume consent to linkage (GAO, 2001), although this is not the case in all surveys, such as NIS, RECS, and MEPS, which request personally-identifying linkage information and consent separately. If personally-identifying information is not requested, linkage may otherwise be facilitated through information collected as part of the survey (e.g., in the CPS and SIPP). This is typically demographic information.

The SSN is the most common identifier in the U.S. and suggested by some to be the most important linking variable (Jabine & Scheuren, 1986). As a nearly universal identifier in the U.S., it has high discriminatory power (Jabine et al., 1986) making it an ideal candidate to link data. However, because of its widespread use to identify individuals, in the wrong hands it can enable identity theft (McNabb et al., 2009). It is perhaps for this reason that many respondents are reluctant to provide their SSN in a survey context (Jabine et al., 1986).

Even if respondents are willing to provide their SSN when requested in a survey it may be reported with error (Jabine et al., 1986). Like the National Insurance Number (or NINO), the UK's analog to the SSN, one disadvantage of the SSN is that it lacks an internal check digit, or a single digit which can be computed from the other digits in the number (Jenkins et al., 2008; Sala et al., 2010; Jabine et al., 1986). An internal check digit could quickly and easily identify reporting errors.

Both interviewers and respondents may contribute to SSN reporting error. Interviewers may record the number with error, even if respondents report it correctly. Respondents may not know their SSN, or they may misreport their SSN either intentionally or unintentionally, and report another's SSN (Jabine et al., 1986). They may also refuse to answer or otherwise skip the question. Further, as of about 1975, about 6 million individuals had two or more SSNs (Jabine et al., 1986). Fortunately, the SSA enumeration verification system applies certain tolerances during the record linkage process. This includes checking for transposed digits of the SSN and variations of the number (McNabb et al., 2009).

With linking technology increasing in sophistication over time and respondents' willingness to provide linking information decreasing, some surveys request only the serial portion of the SSN (the last four digits) and perform record linkage based on this information, along with other identifying information collected in the survey interview, such as consenting respondents' name and date of birth (DOB). The NHIS, for example, began requesting this reduced information in 2007; prior to that they requested the full 9-digit SSN (Miller, Gindi, & Parker, 2011). Requesting this reduced information produces considerably higher consent rates than requesting the full 9-digit SSN (Dahlhamer et al., 2007). Conducting linkage without explicitly requesting personally-identifying information from respondents results in the highest rates of consent (Bates, 2005).

Other information that is typically already collected during the survey interview can be used to facilitate record linkage, such as sex, DOB, and address. Conducting record linkage based on this information does not require any additional respondent or interviewer burden, and likely circumvents the privacy concerns associated with giving a SSN to an interviewer. However, this information may not be as discriminating as a SSN, even after combining several variables (Jenkins et al., 2008). Other issues arise as well: for example, if the survey and administrative data code this information differently, it will be harder to match information across sources (Jenkins et al., 2008).

Although respondents are often unwilling to provide it, SSNs do effectively facilitate linkage between survey responses and administrative records. As demonstrated in research by Abreu, Daniel, Iwig, and Hoge (2009), attempting linkage with less than a full SSN increases the burden on agencies in linking survey and administrative data and results in fewer successful matches. Sayer and Cox (2003) estimate that without SSN, 7% of correct matches would be missed in linking survey responses with National Death Index records, but demonstrated via simulation that the last six digits of the SSN are an equally effective replacement for the full 9-digit SSN in matching. Abreu et al. (2009) varied SSN lengths and found that the last four digits in conjunction with survey predictors resulted in 1-2% missed matches. Using survey predictors only (no SSN) resulted in approximately 4-6% missed matches. Abreu et al. (2009:1) also noted that the "percentage of missed matches will increase as the size of the datasets being matched increase," for example, in larger states.

Jabine et al. (1986) comment on other types of demographic variables commonly used to facilitate record linkage when collection of SSN is not feasible. Typically respondent names are readily available in both survey and administrative data but the consistency of this data is problematic. In one or both sources, nicknames may be reported, the use of middle initial varies by situation, last name may be changed through marriage or divorce, and the order of names varies by culture. Address is also often available in both sources but respondents may vary in whether they report their mailing or home address; the variability across these sources can prevent the likelihood of a match. Sex and DOB are well-reported but may be excluded from administrative records and are less discriminating.

Jenkins et al. (2008) compared match rates of five variations of linkage criteria in the UK: respondent provided NINO[15] and four combinations of survey collected information including sex, DOB, zip code, first name, last name, and address. They found

---

[15] A National Insurance Number, or NINO, consists of two letters, six numbers, and a final letter, for example: QQ 12 34 56 A (Her Majesty's Revenue and Customs, retrieved from **http://www.hmrc.gov.uk/manuals/nimmanual/nim39110.htm on August 21,** 2011. Jenkins et al. (2008) report that only 1.5% of respondents refused to provide their NINO to the interviewer (another 9.9% of respondents did not know it), suggesting that the NINO is much less sensitive than SSN in the US.

that the combination of sex, DOB, and first and last name resulted in the highest rate of matches: 61.7% of consenters. NINO alone resulted in the second highest number of matches – 62.1%. Nearly 75% of consenters matched on at least one of the five criteria. Overall linkage rates using survey information were lower for the respondents who did not provide a NINO, suggesting that these respondents provided lower quality survey data than those willing to provide it.

1.8.4.4 Additional Influences

Empirical findings suggest that the placement of the consent request can influence a respondent's consent decision. Requesting consent within the survey produces equal or better survey response rates (Partin et al., 2008; Shah et al., 2001) and consent rates (Partin et al 2008) as compared to requesting consent as part of a separate communication. In the British Household Panel Survey (BHPS), Sala et al. (2010) concluded that within a household, a respondents' likelihood of consenting was associated with the number of previous respondents interviewed in that household who had already consented. Later survey respondents were more likely to consent if earlier respondents in the household consented as well. Sala et al. (2010) speculated that this could be a household contagion effect or an indicator of interviewer burden, with increasing pressure to complete later interviews as briefly as possible.

The presence of multiple consent questions within a single survey is increasingly common. Within the U.S., existing surveys such as MEPS, NHIS, HRS, and NHANES include more than one request for record linkage. Multiple requests are typically for various types of health, income, and employment-related records (Jenkins et al., 2006) or multiple types of medical records (Klassen et al., 2005). Later consent requests may only be made if consent is given in earlier requests; this can potentially introduce additional selection biases in later requests (Jenkins et al., 2006).

Several panel surveys in the U.S. including the PSID, HRS and MEPS benefit from linking survey responses with administrative records. Using research conducted in the UK on the BHPS, Sala et al. (2010) discovered that respondents' propensity to consent decreased with the number of years they had been in the panel. The researchers hypothesize that panel respondents may be suspicious of a new innovation introduced after so many years, or they may feel as though they have provided enough data and do not see why they should contribute more.

1.9 Remainder of Dissertation

The following three chapters present research that investigates the impact of respondent, interviewer, and consent request characteristics on the decision to consent to record linkage and consent rates. The descriptive analysis in Chapter 2 examines the relationship between characteristics of the survey and consent request and consent rates from surveys conducted in the U.S. Chapter 3 uses primary data collection to assess the effect of consent request topic on consent rates, as well as the influence of respondents' privacy, confidentiality, and trust attitudes and consent request salience. The effects of interviewers and interviewer characteristics on consent to record linkage are examined using data from the 2009 NIS in Chapter 4.

2 CHAPTER 2: DESCRIPTIVE ANALYSIS OF INFLUENCES ON CONSENT

2.1 Introduction

In this chapter we describe characteristics of requests for consent to administrative record linkage and the surveys that contain these requests. Through qualitative and descriptive research methods, we evaluate the effects of various features of the survey and consent request on consent rates, and examine trends in consent rates over time.

2.2 Methods

2.2.1 Overview

There are three components to this research. We first assess whether rates of consent to record linkage have declined overall using all available consent rates. The second and third objectives of this research overlap: we describe several characteristics of surveys that request consent to record linkage, and examine these characteristics as potential sources of variation in consent rates. We selected attributes of the survey and consent request that vary across surveys in the target population, for which sufficient information was available in the methodological documentation, and for which we predicted an influence on consent rates. This includes survey mode, sponsor and response rate; whether consent is requested orally or in writing, whether the request takes an explicit or opt-out approach, the topic of the records requested, and any personally-identifying information requested to facilitate record linkage.

2.2.2 Hypotheses

This section describes the seven noted characteristics of the survey and consent request that may affect consent rates and their predicted influence. In areas where there is limited existing research to inform hypotheses, we rely on the relevant literature from survey methodology more broadly.

2.2.2.1 Survey Response Rate

Similar to the decline in survey response rates over time (Curtin, Presser & Singer, 2000, 2005; Steeh, Kirgis, Cannon, & DeWitt, 2001; Tourangeau et al., 2000), there is evidence that consent rates within some surveys are declining (Bates, 2005; Dahlhamer & Cox, 2007). To the extent that the same factors contribute to both unit response (effective interviewers, refusal conversion efforts, and advance letters (Groves et al., 2004)) and willingness to consent to administrative record linkage, we predict that rates of survey response and consent will be positively related.[16]

In recent years, some surveys began employing measures to improve consent rates such as reducing or eliminating the amount of personally-identifying information requested. Where such procedures are used, we predict that the resulting correlation between response rates and consent rates will be negative because such procedures should boost consent rates, even as response rates are likely to decline over time.

In addition, we also predict there will be no relationship between consent rates and response rates in telephone panel surveys specifically. This is because in some waves of these surveys, the request is presented to respondents for the first time, and in other waves, it is only asked of respondents who previously refused the request. Therefore we do not expect there to be any relationship between response rates and consent rates over time in these surveys.

---

[16] It is worth noting that consent is many times only requested after a person has cooperated to the survey request and responded to some items. Given this, at the individual respondent level, the motivation for consenting may be different from survey cooperation.

2.2.2.2 Survey Mode

In-person surveys, on average, achieve higher response rates than mail or telephone surveys (Cannell, Groves, Magilavy, Mathiowetz, & Miller, 1987; Goyder 1985; Groves & Kahn, 1979; Sykes & Collins 1988). Compared to surveys conducted in other modes, they can more easily establish legitimacy through display of official badges and materials (Biemer & Lyberg, 2003; Groves et al., 2004) and often involve more probing and rapport than telephone surveys (Groves, 1989).

These advantages of in-person surveys could benefit consent rates. Respondents' questions and concerns regarding the request may be most easily and effectively addressed in person. Increased respondent-interviewer rapport (Jenkins et al., 2006) and perceived legitimacy of the interview may improve consent rates as well.

2.2.2.3 Survey Sponsor

The survey sponsor can affect survey response rates and may also affect consent rates. Government-sponsored surveys can achieve higher response rates than comparable surveys with academic or private sponsors (de Leeuw & de Heer, 2002; Goyder, 1985; Heberlein et al., 1978; Linsky, 1975). Sample members may assume that responding to government-sponsored surveys is mandatory (National Academy of Sciences, 1979) or perceive greater importance of surveys with government sponsorship (Heberlein et al., 1978). For the same reasons, government sponsorship could positively influence consent rates as well. We hypothesize that consent rates will be greater in surveys with government sponsorship as compared to surveys with another type of sponsor.

2.2.2.4 Topic of Administrative Records Requested

In this chapter, we examine if consent rates vary as a function of the topic of the records requested.[17] In developing a hypothesis for which record topic respondents are more likely to grant access we consider results to hypothetical questions on the 2010 JPSM Practicum survey. On several related questions, respondents indicated that they would be less likely to consent to medical record access as compared to income and employment-related records.[18] In characterizing records as related to either respondents' health or income and employment, we hypothesize that rates of consent to health-related records will be lower than records related to income and employment.

One survey, the Residential Energy Consumption Survey, or RECS, requests consent to access utility records. We hypothesize that records regarding how much electricity one uses are less sensitive than their income and employment or health-related records and these consent rates will therefore be higher.

2.2.2.5 Consent Mode

Singer (1978:152) finds that the request for a signature negatively affects survey response rates and "appears to function simply as another sensitive question." Respondents may be more suspicious of participation when asked to provide a signature. It is hypothesized that requiring that respondents provide written consent to authorize record linkage is similarly detrimental to consent rates as compared to oral consent.

---

[17] We conduct this analysis at the respondent level in Chapter 3.

[18] The BHPS, which requests consent to health and income and employment-related records, obtained higher consent rates to the health-related consent request. As the research in this chapter is limited to surveys conducted in the U.S., we do not consider the BHPS findings in formulating hypotheses. Further, the HRS requests consent for both types of records but the subset of respondents asked to provide access to their income and employment records is not identical to those asked to provide access to their health records, limiting the generalizability of these consent rates.

2.2.2.6 Personal Identifier Requested

Requests for consent to record linkage often include a request for a personal identifier to facilitate this linkage, such as a Medicare number, SSN, or the name and address of the respondents' healthcare provider. When a personal identifier is requested, consent is often assumed if respondents provide it. Respondents' perceived sensitivity of different identifiers likely varies, and thus their willingness to consent to a linkage request could vary depending on the associated identifier.

A battery of hypothetical questions evaluated respondents' perceived sensitivity of personal identifiers on the 2010 JPSM Practicum Survey. Using a 4-point scale where four indicated "Extremely Personal" and one indicated "Not at all Personal", respondents rated their 9-digit SSN as the most personal identifier (3.6), followed by Medicare number (3.3; rated only by those over the age of 65), 4-digit SSN (3.0), and doctor's contact information (2.6).

In lieu of any prior research, we expect respondents to be slightly more willing to provide their Medicaid number as compared to their Medicare number. Both grant access to government-provided health records, yet Medicare number comprises an individual's SSN.[19] [20]

Accordingly, we predict consent rates to be lowest for requests for 9-digit SSN, followed by Medicare, Medicaid number, 4-digit SSN, and healthcare provider information. We predict that consent rates will be higher when utility provider information is requested (in RECS) as this information is likely less sensitive than other

---

[19] Retrieved from http://ssa-custhelp.ssa.gov/app/answers/detail/a_id/1366/~/meaning-of-the-letters-after-a-social-security-or-medicare-number

[20] It is unclear if respondents are aware of the similarity between their SSN and Medicare numbers.

identifiers. We expect consent requests unaccompanied by a request for a personal identifier to have the highest rates of consent.

2.2.2.7 Explicit vs. Implicit Consent

Some record linkage is authorized as a function of respondents' implicit consent in which they must overtly opt out of record linkage if they do not want their administrative records to be accessed and linked with survey responses. As compared to explicit consent requests, the passive opt-out procedure shifts the burden from consenting (which may require recalling and reporting a personal identifier, or responding to a direct consent request) to withholding consent (which, for example in the Census surveys, requires identifying and understanding the opt-out procedure in the prenotification letter or other materials and notifying the field representative that they wish to opt out of the record linkage). It is hypothesized that the added steps required to refuse consent will lead to greater consent rates when an opt-out procedure is used as compared to more direct requests.

2.2.3 Eligibility Criteria

At the most basic level, we considered surveys eligible for inclusion in the descriptive analysis if they presented respondents with the choice to participate in record linkage. That is, respondents were given the opportunity to exclude themselves from record linkage without affecting their ability to participate in the survey. Eligibility was unaffected by the directness of the consent request: we included surveys that used explicit consent requests and surveys that used implicit consent requests. Eligibility was not contingent on whether or not identifying information was also requested, or the type of identifying information requested.

Some surveys use administrative records to draw a sampling frame, such as the New Beneficiary Survey (drawn from SSA records). Surveys such as this one which do not require respondents' consent to access records are excluded from this research. Also excluded are surveys for which sample members are unable to participate unless they consent to record linkage, and therefore survey participation and consent are intertwined. This includes some surveys conducted by the National Center for Education Statistics.[21]

Finally, we limit our scope to only those surveys conducted within the U.S. The sensitivity of such a request, and the guidelines for requesting consent and using administrative records vary across countries (Baker et al., 2000). For example, written consent is required to access health records in the U.K. (Tate et al., 1995) and Finland (Korkeila et al., 2001) and is typically required to access personal records in Australia (Silva et al, 2002). This country-level variation limits the comparability of requests made across countries.

To identify eligible surveys, we searched the websites of the federal statistical agencies as well as other organizations that conduct large, nationally representative surveys (for example, the University of Michigan Survey Research Center, National Opinion Research Center at the University of Chicago, and RAND), and consulted Diane Rourke's list of survey organizations produced through the Survey Research Laboratory at the University of Illinois at Chicago. Personal communication was also made with relevant, knowledgeable individuals at many of these agencies and organizations, and a message was posted to the American Association for Public Opinion Research listserv

---

[21] For example, by agreeing to participate in the Beginning Postsecondary Students Study and the Baccalaureate and Beyond Study, sample members must consent to the following statement: "Your responses, combined with student record information (such as transcripts and financial aid data), may be used only for statistical purposes and may not be disclosed, or used, in personally identifiable form for any other purpose, unless otherwise compelled by law."

requesting that members notify us of surveys that request respondents' consent to record linkage.

We also sought to identify additional surveys through research published in relevant journals. We used Boolean searches to identify such research, including combinations of the following terms: record, linkage, consent, survey, SSN, Medicare number, personal records, personal identifier, and administrative records. We searched journals relevant to survey methodology such as Public Opinion Quarterly, Survey Methodology, Survey Practice, Survey Research Methods, Social Security Bulletin, and the Journal of Official Statistics, as well as the proceedings from the American Association for Public Opinion Research and the Joint Statistical Meetings. We also looked for eligible surveys in journals in related fields that use linked databases in a practical setting, for example, medicine, public health, and epidemiology, including Biomedical Health Services Research, American Journal of Epidemiology, Journal of Epidemiology and Community Health, Annals of Internal Medicine, Health Services Research, New England Journal of Medicine, Medical Care, and the Journal of Clinical Epidemiology. We consulted the bibliographies to identify other publications containing additional relevant surveys.

For all surveys we identified, we examined available publications, methodological documentation, and questionnaires to determine eligibility. In several cases, we contacted the researchers for additional information.

2.2.4 Coding

For each survey administration, we coded a series of characteristics pertaining to the survey and consent request from relevant publications, methodological documentation, and questionnaires.[22] We recorded the surveys' mode of administration and sponsor. In determining survey sponsorship, we consider respondents' perception of the sponsor.[23] We also documented a series of characteristics related to the consent request including consent mode, whether the survey utilized an explicit consent request or an opt-out approach, if any personal identifier was requested from the respondent to facilitate consent, and the topic of the administrative records requested.

We also recorded the overall response rate for each survey administration. For all studies, we used the response rate provided in the survey documentation or publication. Considering the mix of survey designs in this population (for example, panel and cross sectional surveys), the calculation of response rates varies across surveys.

Finally, the percent of respondents granting consent was recorded for each survey administration. If a survey contained multiple consent requests, the rate of consent was recorded for each request. Broadly, we define the rate of consent as the percent of respondents permitting linkage between survey responses and administrative records. However, there is variation in how consent is requested, and so we further specify our definition of consent by the nature of the request. We categorize consent requests into three categories. (The final column in Table 2.01 indicates the type of each consent request included in this research.)

First, some consent requests are straightforward appeals for respondents' consent to record linkage either orally or a request for a signed authorization form. If an identifier

---

[22] Please see the Appendix for references indicating the source of all data included in this chapter.

[23] Although surveys such as AHEAD, HRS, and PSID have some government funding, respondents likely view the University of Michigan as the survey's sponsor. When contacting sample members, interviewers address themselves as calling from the University of Michigan, advance letters have the University of Michigan return address and the University of Michigan footer, and provide IRB information for the University of Michigan.

is requested, it is in addition to this consent request, and respondents must explicitly agree to the direct consent request in order for record access and linkage to occur. For these requests, consent is calculated as the proportion of respondents who grant consent to the direct consent request.

Others request personally-identifying information to facilitate linkage, and respondents are presumed to have consented if they provide this requested linking information. Conversely, respondents refusing to provide the requested identifier are implicitly considered to have refused record linkage (Miller et al., 2011). Thus, providing the requested identifier and granting consent are intertwined. For these types of requests, consenting respondents are defined as those who provide the identifying information.

Opt-out consent requests comprise a third subset. Here, consent is assumed to be granted if the respondent does not explicitly opt-out of the record linkage. In this research, we define consenting respondents as those who did not explicitly opt-out.

2.2.5 Analyses

In the first portion of this research, we estimate changes in consent rates over time through linear regression. Next, we use correlations to describe the relationship between survey response rates and consent rates. Then, descriptive statistics including frequencies, mean consent rates, and 95% confidence intervals, are presented for surveys and consent requests by each design characteristic of interest: survey mode, survey sponsor, consent mode, identifier requested, whether consent is requested explicitly or implicitly, and administrative record topic. Line graphs illustrate trends in mean consent rates over time for each of these characteristics. We then use multivariate regression to describe the relationship between these characteristics of interest and consent while controlling for other factors.

All regression analyses, correlations, and descriptive statistics were calculated using Stata 11 and account for clustering by survey using Stata SVY procedures. All analyses are unweighted. Graphics were produced using Stata and Microsoft Excel.

2.3 Results

2.3.1 Eligible Surveys

Twenty-two surveys met the eligibility criteria, and at least one consent rate was available for all 22 surveys.[24] [25] Collectively, the 22 surveys included in the descriptive analysis contribute 110 survey administrations and 162 consent rates. Table 2.01 provides a list of these surveys and consent rates.

Some surveys are missing consent rates for certain administrations or requests. A list of missing data is provided in Table 2.02. Some agencies were unable to retrieve consent rates from older surveys due to storage complications, for example, if the data is housed on mainframe computers. Other consent rates are missing because the agency never processed the data.

Although consent rates from all available administrations are included in this research, several surveys are conducted on an ongoing basis and have newer, additional

---

[24] Olson (1999) and Zell et al. (2000) discuss consent rates included in research published elsewhere. These are excluded from our research. Dahlhamer & Cox (2007) provides consent rates for the first two quarters of the NHIS. As consent rates from the full 2007 administration are included in the analyses, consent rates from Dahlhamer & Cox (2007) are excluded.

[25] Project Talent is an additional survey that would be eligible for inclusion in this research, however consent rates have not been released yet. The survey was fielded from January to May of 2012, conducted by the American Institutes of Research, and sponsored by the National Institute on Aging via the Health and Retirement Survey. Respondents who participated in the 1960 baseline Project Talent (currently ages 65-70) were mailed surveys (a selection of items from the HRS) as well as SSN and signature request forms to authorize linkage of responses with SSA records. (Retrieved from http://www.projecttalent.org/)

administrations yet to be released. This includes CPS, HRS, MEPS, NHANES, NHIS, NIS, NIS Teen, PSID, RECS, and SIPP. We do not consider data missing if the survey administration has not yet been released.

2.3.2 Descriptive Statistics

The 162 consent rates included in this research ranged from 9.9% to 100.0%, with a mean of 70.2% and a median of 73.1%. Response rates ranged from 52.1% to 96%[26] with a mean of 76.6% and a median of 76.9%. All available consent rates were from surveys conducted between 1982 and 2010 (see Table 2.01 for additional characteristics).

---

[26] For consent and response rates where a decimal place was not available, one is not added.

Table 2.01

Characteristics of Data Included in Descriptive Analysis

Survey	Consent Requests	Years Included	Comp-letes (Ave.)[27]	Response Rate (Ave.) %	Consent Rate (Ave.) %	Consent Request Type[28]
1. Assets and Health Dynamics Among the Oldest-Old (AHEAD)[29]	SSN	1993, 1995	8,032	84.8	37.0	2
	Medicare Number	1993, 1995			64.5	2
	Medicaid Number	1993, 1995			64.0	2
2. Cleary & Jette (1984)	Medical/Pharmacy Record Access (1 request; No ID)	1984	1,026	88.0	88.9	1
3. Current Population Survey (CPS) Annual Social and Economic Supplement	SSN	1994-2005	282,083	84.2	82.2	2
	Opt-Out	2006-2010			99.7	3
4. Health Interview Evaluation Survey (HIES)	Medical Record Access (No ID)	1990	1,077	76	94	1
5. Health and Retirement Study (HRS)	SSN	1992-2010 (biennial)	16,665	87.4	48.0	2
	Medicare Number	1996-2010 (biennial)			54.1	2
	Medicaid Number	1996; 2000-2010 (biennial)			54.4	2
6. Health Field Study (HFS)	Medical Record Access (No ID)	1993	2,006	66.6	94.7	1

---

[27] For surveys with multiple administrations included, the number of completed cases, response rates, and consent rate is the unweighted average across all included administrations.

[28] 1= Separate consent question; 2=providing identifier is considered implicit consent; 3=opt-out consent request. See page 11 for the full description of each consent request type.

[29] AHEAD is the precursor to the HRS. We consider them individual surveys because of the differences in sample design and target population.

7. *Longitudinal Study of Aging* II (LSOA II)	SSN	1994	8,265	mis.	94.0	2
	Medicare Number	1994			96.8	2
8. McCarthy, Shatin, Drinkard, Kleinman, & Gardner (1999)	Medical Record Access (No ID)	1997-1998 (1 administration)	73	52.1	35.6	1
9. Medical Expenditure Panel Survey (MEPS)	Healthcare Provider Contact Info.	1996-2010	29,998	68.4	68.5	1
	Pharmacy Contact Info.	1996-2010		68.4	75.8	1
10. Murdoch, Pietila, & Partin (2010)[30]	Medical Record Access (No ID) (Check Form) (a)	2010	367	67.5	59.1	1
	Medical Record Access (No ID) (Sign Form) (b)	2010	367	71.9	52.3	1
	SSN (c)	2010	365	74.2	48.7	1
11. National Health and Nutrition Examination Survey (NHANES)	SSN	1999/2000, 2001/2002, 2003/2004, 2005/2006, 2007/2008[31]	10,325	80.8	75.0	2
	Medicare Number	1999/2000, 2001/2002, 2003/2004		81.7	80.2	2
12. National Health and Nutrition Examination Survey III (NHANES III)	SSN	1988-1994[32]	33,994	85.6	96.7	2
	Medicare Number	1988-1994			100.0	2
13. NHANES I Epidemiologic Follow-up Study (NHEFS)[33]	SSN	1982-1984 (1 administration)	14,407	96	98.9	2
	Medicare Number			96	98.2	

---

[30] Murdoch et al. (2010) included three administrations that varied consent procedures. The first administration required respondents' authorize consent by checking a box, the second required that they provide their written signature, and the third required that they provide their SSN.

[31] NHANES public use files are released in 2-year increments.

[32] NHANES III interviews cover 6 calendar years (1988-1994) and the data is generally analyzed as one dataset.

[33] NHEFS is a series of follow-up studies of respondents who completed a medical exam at NHANES I 1971-1975 and is generally analyzed as one dataset.

14. National Health Interview Survey (NHIS)	SSN (9-digits)	1997-2006	95,502	73.1	51.3	2
	SSN (last 4-digits or permission to link)	2007-2010	79,482	65.3	76.3	2
	Medicare Number	1997-2006	95,502	73.1	46.7	2
	Medicare Number (last 4-digits or permission to link)	2007-2010	79,482	65.3	70.1	2
15. National Immunization Survey (NIS)	Healthcare Provider Data	1995-2010	30,223	73.6	82.8	1
16. National Immunization Survey Teen (NIS Teen)	Healthcare Provider Data	2008-2010	33,084	58.3	75.0	1
17. Partin et al. (2008)	Medical Record Access (No ID)	2006	686	77.1	50.3	1
18. Panel Study for Income Dynamics (PSID)	Medicare Number	2005-2009 (biennial)	5,590	93.8	42.5	2
19. Residential Energy Consumption Survey (RECS)	Utility Provider Contact Info.	1978-2009	5,376	81.7	89.0	1
20. Survey of Health Insurance and Program Participation (SHIPP)	Opt-Out	2010	mis.	53.5	90.0	3
21. Survey of Income and Program Participation (SIPP)	SSN	1984, 1990, 1992, 1993, 1996, 2001, 2004	mis.	81.4	89.1	2
	Opt-Out	2008		80.1	100.0	3
22. Woolf, Rothemich, Johnson, & Marsland (2000)	Medical Record Access (No ID)	2000	1,106	94	67	1

Notes: Data collection of LSOA II, NHEFS, NHANES III may not be as accurate as later NCHS surveys because of both the age of the data and changes in collection procedures over time as the NCHS record linkage program evolved. Multi-year surveys (NHANES, NHANES III, NHEFS) are included in all analyses as the final year they were administered.

Table 2.02

Consent Rates Excluded from Analyses

Survey	Consent Requests	Missing Consent Rates
CPS	SSN	2000
HRS	Industry and occupation data (linked to pension plan information)	1992-2010 (biennial)
NHANES	SSN	2009/2010
	Medicare Number	2005/2006, 2007/2008, 2009/2010
NHIS	SSN	1984-1996, 2010
	Medicare Number	1984-1996, 2010
NIS	Healthcare Provider Contact Info	1994
RECS	Utility Provider Data	1978, 1979, 1980, 1981, 1982, 1984
SIPP	SSN	1986-1988[34]

2.3.3 Changes in Consent Rates over Time

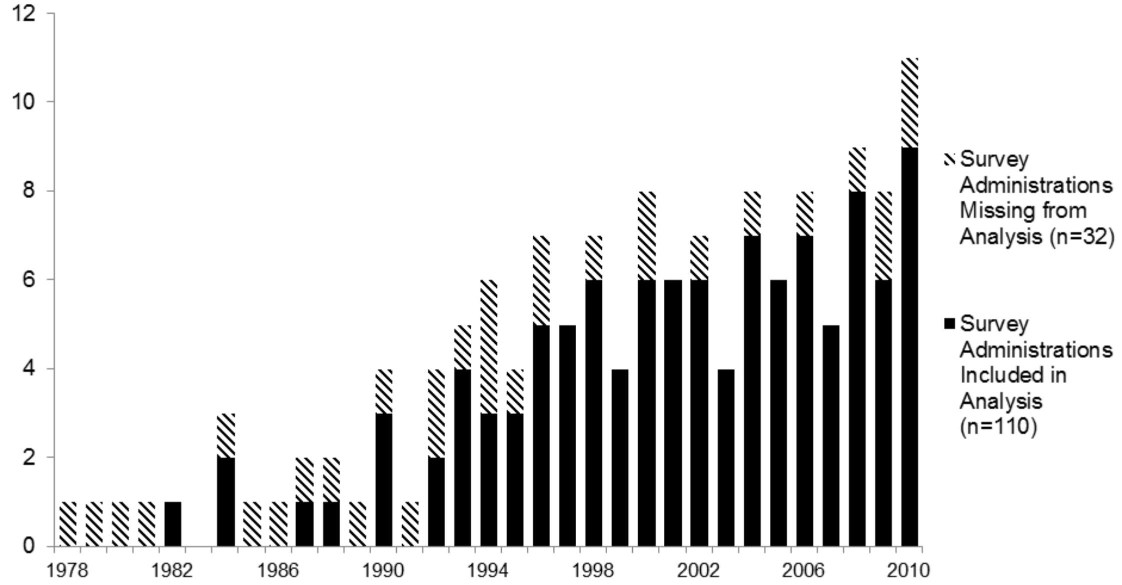
Surveys have requested consent to administrative record linkage for several decades. RECS requested access to respondents' utility provider information as early as 1978, and surveys conducted by the Census Bureau and NCHS began requesting consent to record linkage shortly thereafter. Figure 2.01 illustrates the number of surveys requesting consent to record linkage by year of survey administration. As shown, more surveys are requesting respondents' consent to record linkage over time, with a marked increase in the number of surveys requesting consent in the early 1990s.

---

[34] Please refer to Appendix for details about missing data in SIPP.

Figure 2.01

Number of Survey Administrations with Consent Requests, by Year

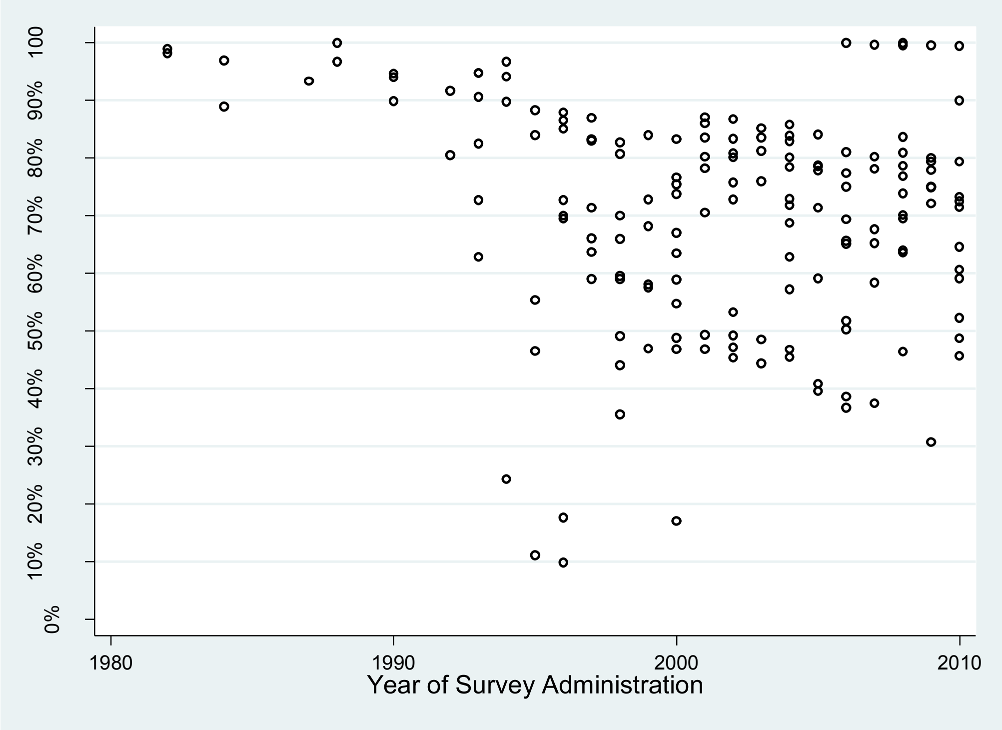


Notes: All surveys with consent requests included in Table 2.01 and Table 2.02 are included in Figure 2.01, regardless of whether the consent rate was available for this research.

The 162 consent rates included in this research are plotted by year of survey administration in Figure 2.02. The distribution of consent rates suggests a slight decline in consent likelihood over time, with the exception of consent rates produced through opt-out consent requests (observed in the far upper right hand corner of the figure).

Figure 2.02

Scatterplot of Consent Rates Included in Descriptive Analysis, Plotted by Year of Survey Administration



Notes: N=162 consent rates.

We estimate a simple linear regression predicting consent rate only as a function of survey year, and find that willingness to consent is slightly declining with time (*B* = -0.58; *p* = 0.096; $r^2 = 0.04$). However, if we hold survey constant by including an indicator for each survey in the dataset, we find no significant relationship between survey year and consent rate (*B* = 0.40; *p* = 0.116; $r^2 = 0.69$; see Table 2.03). In this model, the individual surveys are significantly associated with consent, likely because there is a great deal of variability in consent rates across surveys.

Table 2.03

Linear Regression Predicting Consent Rate as a Function of Survey and Survey Year

	Coef.	SE	*p*
Constant	49.96	3.18	0.000
Year	0.40	0.24	0.116
Survey			
AHEAD *(ref)*			
Cleary et al. (1984)	37.74	2.45	0.000
CPS	29.22	1.99	0.000
HIES	40.43	0.98	0.000
HRS	-6.66	2.06	0.004
HFS	39.93	0.24	0.000
LSOA II	40.24	0.00	0.000
McCarthy et al. (1999)	-21.17	0.98	0.000
MEPS	13.38	2.20	0.000
Murdoch et al. (2010)	-8.21	3.92	0.048
NHANES	18.08	2.26	0.000
NHANES III	45.59	1.47	0.000
NHEFS	48.20	2.94	0.000
NHIS	-4.20	2.20	0.070
NIS	24.24	2.08	0.000
NIS Teen	13.82	3.67	0.001
Partin et al. (2008)	-9.68	2.94	0.003
PSID	-17.92	3.18	0.000
RECS	33.20	0.37	0.000
SHIPP	28.42	3.92	0.000
SIPP	34.48	0.49	0.000
Woolf et al. (2000)	9.43	1.47	0.000

Notes: N=162; r^2=0.69; adjusted Wald test for all parameters: $F(1,21) = 4.40$, $p<0.048$. Year is coded as the number of years between 1982 and the year of administration for a particular survey.

Figure 2.03 illustrates the trend in mean consent rates by year of survey administration. Consent rates were highest during the first decade of administration but began to fall during the early 1990s.[35] Although we cannot identify the cause from the available data, this decline may be related to the introduction of surveys with a wider variety of characteristics requesting consent during this time as compared to the previous decade. Surveys with consent requests conducted during the 1980s were limited to in-person interviewing and had primarily government sponsorship. In later years, surveys

---

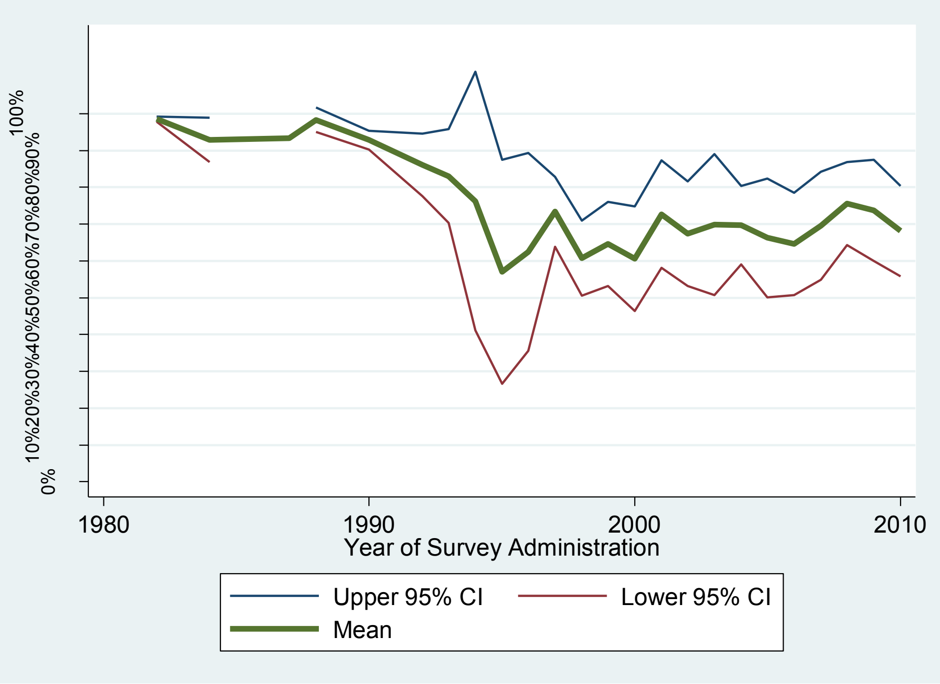
[35] Four of the seven consent rates from the 1980s are from NCHS surveys, which, as noted in Table 2.01, are to be interpreted with caution as changes in data collection procedures over time may reduce the precision of this data. Thus, the particularly high consent rates depicted during the 1980s in this research may not accurately reflect respondents' true willingness to consent during that period.

containing consent requests had more diverse characteristics, including mode and survey sponsorship.

Since around 1995, consent rates have leveled off and fluctuated between approximately 60-70%. Rates are slightly higher in more recent years, perhaps due to various measures taken to prevent further decline, e.g., opt-out consent requests and reducing the amount of personally-identifying information requested.

Figure 2.03

Mean Consent Rates and 95% Confidence Intervals by Year of Survey Administration



Notes: N=162 consent rates. Confidence intervals could not be calculated for 1987 because only one data point was available. (No data points were available for 1985 and 1986.)

2.3.4 Current Consent Practices and Impact on Consent Rates

2.3.4.1 Overall Survey Response Rate

The overall correlation between survey response rates and consent rates is $r = -0.024$ ($p = 0.762$; n = 156).[36] This suggests that overall, there is essentially no relationship between consent rates and response rates among the surveys included in this research. However, this correlation does not account for differences in the survey design or type of consent request procedure which may positively or negatively affect consent rates, thus affecting the correlation among individual surveys. The identifier requested, whether an explicit or opt-out procedure is used, or if respondents who refuse consent are presented with the request again in later waves, as in the telephone panel surveys, may impact consent rates. The overall correlation may mask different relationships across surveys between consent rates and response rates.

While survey response rates may be declining over time, some consent procedures may boost consent rates, for example, requesting only the last four digits of an SSN or Medicare number instead of the full identifier, or using an opt-out consent approach. With increases in consent rates over time and potentially declining response rates, the relationship between consent rates and response rates may be negative.

Other consent procedures produce a more complicated relationship between consent rates and response rates. This specifically pertains to the panel surveys conducted by telephone at the University of Michigan (HRS, PSID, and AHEAD). In these surveys, eligible respondents are typically asked for consent during their initial wave in the survey. Any respondents who refuse the initial request may be presented with

---

[36] This correlation does not include the following survey administrations for which response rates were unavailable: HRS: 2010; CPS: 1994-1996; LSOA II.

the consent request again in subsequent waves. For example, in the PSID, respondents receiving Medicare benefits were asked to provide their Medicare number for the first time in 2005. Any respondents who refused this request in 2005 were asked again in 2007, along with respondents who began receiving Medicare benefits between the 2005 and 2007 waves. In 2009, consent refusers from prior waves were asked to provide Medicare number, in addition to any respondents who began receiving Medicare benefits between 2007 and 2009 (Fulton et al., 2011).

AHEAD, conducted in the mid-1990s, uses a similar design. Regarding the collection of SSNs in the 1993 administration, Soldo et al. (1997:16), explain that "another attempt will be made in wave 2 [1995] to secure SSA linkage permission from those unwilling to provide written consent in 1993." This consent procedure may account for the decline in consent rates in these surveys in some waves if the majority of those asked the consent request refused consent in prior waves.

Other panel surveys, including CPS, MEPS, and SIPP do not use the same approach to requesting consent. Consent is requested only once from CPS respondents in the March Annual Social and Economic Supplement. In most panels, all SIPP respondents are asked to consent at each wave. Eligible MEPS respondents are asked for consent at each wave, though the administrative records to which access is requested by wave varies.[37]

Table 2.04 provides correlations between survey response rate and consent rates by consent request type for the 11 surveys with three or more available consent rates.

Table 2.04

---

[37] Consent to contact health providers is requested in all waves of MEPS; consent to contact pharmacy providers is generally requested during waves three and five.

Correlations between Consent Rates and Response Rates by Survey and Consent Request Type

Survey	Consent Request	Number of Comparisons	Correlation (*p*-value)
AHEAD	Overall	6	-0.767 (0.075)*
CPS	Overall	16	0.014 (0.960)[38]
	SSN (9-digit)	11	-0.656 (0.029)**
	None (Opt-Out)	5	-0.769 (0.128)
HRS	Overall	21	-0.241 (0.292)
	SSN (9-digit)	8	-0.432 (0.285)
	Medicare Number	7	-0.678 (0.094)*
	Medicaid Number	6	0.611 (0.198)
MEPS	Overall	30	-0.368 (0.045)**
	Healthcare Provider Contact Info.	15	-0.489 (0.065)*
	Pharmacy Contact Info.	15	-0.452 (0.091)*
NHANES	Overall	8	0.574 (0.137)
	SSN (9-digit)	5	0.978 (0.004)***
	Medicare Number	3	-0.522 (0.650)
NHIS	Overall	26	-0.452 (0.020)**
	SSN – Overall	13	-0.435 (0.138)
	SSN (9-digit)	10	-0.614 (0.059)*
	SSN (4-digit + link)	3	-0.931 (0.239)
	Medicare Number – Overall	13	-0.493 (0.087)*
	Medicare Number (full)	10	0.831 (0.003)***
	Medicare Number (4-digit + link)	3	-0.724 (0.488)
NIS	Healthcare Provider Contact Info.	16	0.252 (0.346)
NIS Teen	Healthcare Provider Contact Info.	3	0.542 (0.635)
PSID	Medicare Number	3	0.563 (0.620)
RECS	Utility Provider Contact Info.	6	-0.843 (0.035)**
SIPP	Overall	8	0.511 (0.196)
	SSN (9-digit)	7	0.365 (0.421)

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.10$; surveys and consent requests excluded had less than three consent rates.

Contrary to our hypothesis, we do not observe a strong or consistent association between consent rate and survey response rate. Of the 27 correlations examined, twelve

---

[38] During the years that CPS requested SSN (and for which data was available), both consent rates and response rates fluctuated (1997-2005), resulting in an increase in consent rates and a slight decrease in response rates. The remaining years during which an opt-out request was used (2006-2010), response rates increased and consent rates very slightly decreased. When these years are examined in aggregate, the resulting correlation is zero.

are significant and two of these are positively related. Where relationships exist, the directionality is inconsistent and the association between consent rate and response rate does not appear to be related to the identifier or survey design.

We find both positive and negative correlations between consent rates and SSN, Medicare number, and other identifying information. An examination of the underlying consent rates and response rates shows that consent rates are not declining in a linear fashion over time (nor are response rates in most surveys) and so the resulting relationship between the two is less straightforward than predicted.

For example, NIS exhibits a positive non-significant correlation between consent and response rates. Although both consent rates and response rates are declining with time in this survey, year to year, there are some small positive fluctuations in consent rates that appear to skew the correlation. In contrast, Although MEPS demonstrates significant negative correlations between response rates and consent rates, there is very little net change in consent rates over time, but response rates have declined nearly 20 percentage points from 1996-2010, the years data was available.

As predicted, we do find that reducing the amount of identifying information requested produces a negative (but not significant) correlation in the NHIS and CPS.

Table 2.05

Additional Characteristics of Surveys Included

Survey	Identifying Information	Records	Consent Mode	Consent Request Type	Record Topic	Survey Mode	Survey Target Population	Sponsor
AHEAD	SSN Medicare Number Medicaid Number	SSA Records Medicare Records Medicare Records	Written Oral Oral	Explicit Explicit Explicit	Inc/Emp Health Health	Phone: Panel	Older Adults	Other
Cleary & Jette (1984)	None	Medical Records	Written	Explicit	Health	In Person	Household	Other
CPS	SSN None	SSA Records SSA Records	Oral Oral	Explicit Opt-Out	Inc/Emp Inc/Emp	In Person	Household	Government
HIES	None	Medical Records	Written	Explicit	Health	In Person	Household	Government
HRS	SSN Medicare Number Medicaid Number	SSA Records Medicare Records Medicaid Records	Written Oral Oral	Explicit Explicit Explicit	Inc/Emp Health Health	Phone: Panel	Older adults	Other
HFS	None	Medical Records	Written	Explicit	Health	In Person	HMO Members	Other
LSOA II	SSN Medicare Number	SSA Records Medicare Records	Oral Oral	Explicit Explicit	Inc/Emp Health	In Person	Older Adults	Government
McCarthy et al. (1999)	None	Medical Records	Written	Explicit	Health	Mail	Patients	Other

MEPS	Healthcare Provider Contact Information	Healthcare Provider Records	Written	Explicit	Health	In Person	Household	Government
	Pharmacy Contact Information	Pharmacy Records	Written	Explicit	Health			
Murdoch et al. (2010)	No ID (a)	Medical Records	Written	Explicit	Health	Mail	Patients	Other
	No ID (b)	Medical Records	Written	Explicit	Health	Mail	Patients	Other
	SSN (c)	Medical Records	Written	Explicit	Health	Mail	Patients	Other
NHANES	SSN	SSA Records	Oral	Explicit	Inc/Emp	In Person	Household	Government
	Medicare Number	Medicare Records			Health			
NHANES III	SSN	SSA Records	Oral	Explicit	Inc/Emp	In Person	Household	Government
	Medicare Number	Medicare Records	Oral	Explicit	Health			
NHEFS	SSN	SSA Records	Oral	Explicit	Inc/Emp	In Person	Household	Government
	Medicare Number	Medicare Records	Oral	Explicit	Health			
NHIS	SSN (9-digit)	SSA Records	Oral	Explicit	Inc/Emp	In Person	Household	Government
	SSN (4-digit)	SSA Records	Oral	Explicit	Inc/Emp			
	Medicare Number	Medicare Records	Oral	Explicit	Health			
NIS	Healthcare Provider Contact Information	Vaccination Records	Oral	Explicit	Health	Phone: Cross-Sectional	Infants	Government
NIS Teen	Healthcare Provider Contact Information	Vaccination Records	Oral	Explicit	Health	Phone: Cross-Sectional	Teens	Government
Partin et al. (2008)	None	Medical Records	Written	Explicit	Health	Mail	Older Patients (50-75)	Other

PSID	Medicare Number	Medicare Records	Oral	Explicit	Health	Phone: Panel	Household	Government
RECS	Utility Provider Contact Information	Utility Records	Written	Explicit	Utility	In Person	Household	Government
SHIPP	None	SSA Records	Oral	Opt-Out	Inc/Emp	Phone: Cross-Sectional	Household	Government
SIPP	SSN	SSA Records	Oral	Explicit	Inc/Emp	In Person	Household	Government
	None	SSA Records	Oral	Opt-Out	Inc/Emp			
Woolf et al. (2000)	None	Medical Records	Written	Explicit	Health	SAQ	Patients	Other

2.3.4.2 Survey Mode

All 22 eligible surveys were administered in person or by telephone, mail, or as part of an in-person self-administered questionnaire (SAQ).[39] Most surveys use in-person administration. Some in-person surveys have cross-sectional design such as HFS, HIES, NHANES, NHANES III, NHIS, RECS and the survey described by Cleary et al. (1984). Others have panel designs such as CPS, MEPS, and SIPP. Finally, some in-person surveys incorporate follow-up waves although all waves are typically combined during analysis; this includes LSOA II and NHEFS.[40] As shown in Table 2.05, the target population of many of these in-person surveys is the household, and most have government sponsorship.

Six of the 22 surveys making requests for consent to record linkage are conducted by telephone. This includes cross-sectional surveys with government sponsorship such as NIS, NIS Teen and SHIPP[41] that rely on RDD samples, and panel surveys such as PSID, HRS, and AHEAD. The target populations for these surveys vary. Young people are of primary interest in NIS and NIS Teen, while older individuals are the target population of HRS and AHEAD. Both SHIPP and PSID conduct interviews with household samples.

We identified three mail surveys that request consent to record linkage. These mail surveys are conducted on smaller samples than the surveys conducted in other modes and have narrower target populations. McCarthy et al. (1999), Murdoch et al. (2010), and Partin et al. (2008), describe mail studies that include requests to link

---

[39] Several surveys including AHEAD, HRS, LSOA II, and SIPP and those mentioned in McCarthy et al. (1999) and Partin et al. (2008) utilized multiple modes of administration including for nonresponse follow-up. We considered the mode of administration to be that in which the most interviews were conducted.

[40] NHEFS is a follow-up study to NHANES I and LSOA II is a follow-up study to SOA II.

[41] SHIPP supplemented the RDD sample with a list of Medicare enrollees (Pascale, 2011).

respondents' medical records with survey responses.[42] Target populations of these studies include patients with PTSD (Murdoch et al., 2010), patients with colorectal cancer (Partin et al., 2008), and health plan members (McCarthy et al., 1999). A final survey requesting consent does so via an in-person SAQ administered to patients at a family medical clinic (Woolf et al. 2000).

The information requested from respondents varies slightly by interview mode. Seven of the ten surveys requesting SSN are conducted in person, and six of the eight surveys requesting Medicare number are conducted in person. Mail surveys tend to request consent without any personally-identifying information; this may help boost mail response rates; only one mail survey identified requests personally-identifying information (Murdoch et al., 2010).

Table 2.06

Descriptive Statistics by Survey Mode

	# Surveys	# Consent Requests	Mean Consent Rate	95% Conf. Int.
			%	%
Survey Mode				
In-Person	12	103	75.0	62.6, 87.3
Phone	6	53	63.1	46.1, 80.1
Phone: Panel	3	33	51.6	49.3, 53.9
Phone: Cross-Sectional	3	20	82.0	79.2, 84.8
Mail	3	5	49.2	41.3, 57.1
SAQ	1	1	67	--

Table 2.06 presents descriptive statistics by mode of survey administration. As predicted, consent rates are highest in in-person surveys (75%). The difference in

---

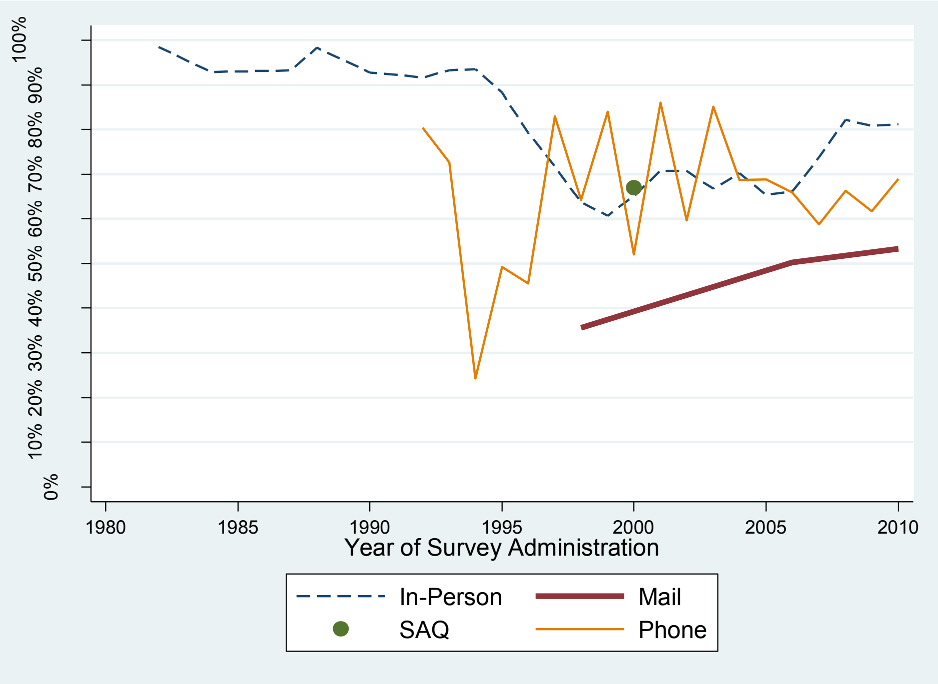
[42] As of 1997, a statute requires that patients in Minnesota consent before researchers examine their medical records (Woolf et al., 2000). Three surveys included in this research (McCarthy et al. (1999), Murdoch et al. (2010), and Partin et al. (2000)) were conducted in Minnesota, possibly to identify methods that maximize consent rates for this purpose.

consent rates between in-person surveys and mail surveys (49%) is significant ($p<0.05$); but in-person surveys do not achieve significantly higher consent rates as compared to telephone surveys (63%).

As previously discussed, the three telephone panel surveys included in this analysis are distinct from the others in that they continue to request consent from refusing respondents in later waves of the survey. This distinction in consent request procedures between cross-sectional and panel phone surveys warrants examining consent rates by phone survey design.[43] On average, consent rates in telephone cross-sectional surveys are significantly higher than in telephone panel surveys (82% vs. 52%; $p<0.001$).

Figure 2.04

Mean Consent Rates by Survey Mode and Year of Survey Administration



---

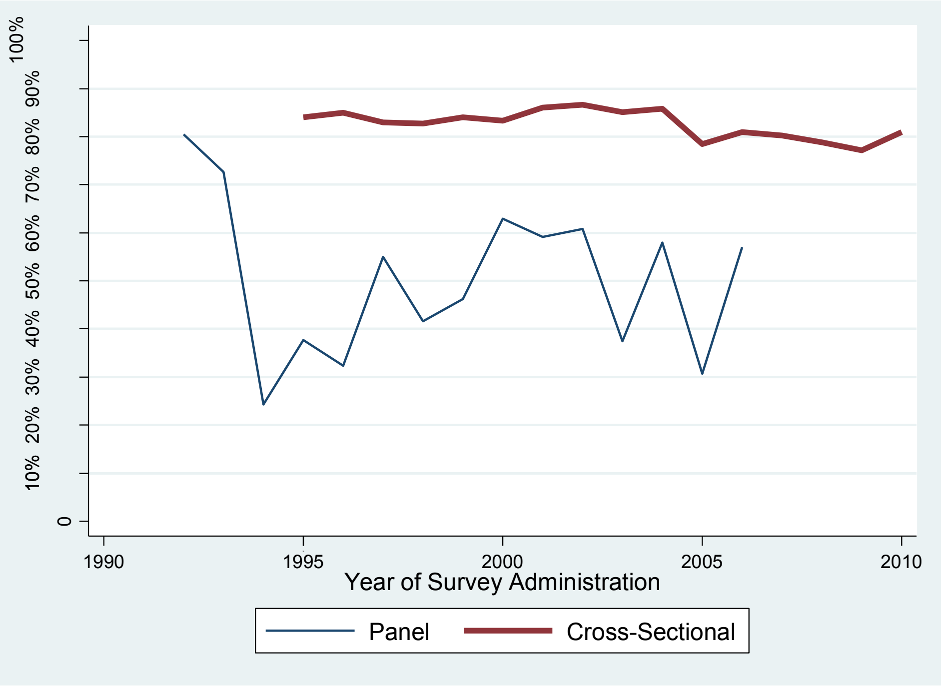
[43] Consent procedures do not vary the same way in in-person panel surveys. In-person panel surveys vary in terms of whether respondents are asked for consent at one point in time (e.g., CPS) or at each wave (e.g., SIPP). Given this variation we do not further segment in-person panel surveys in our analysis.

Figure 2.04 shows mean consent rates over time, by survey mode. Mean consent rates in in-person surveys peak in the early years of administration but dip in the late 1990s and early 2000s. More recently, rates of consent to surveys in this mode appear to be increasing, perhaps because of measures taken to boost consent rates including opt-out procedures and requesting minimal identifying information. The lack of a similar gain in consent rates in telephone surveys during this time may be because they have not adopted similar measures to increase consent rates. Mean consent rates in mail surveys are lower across years, but appear to be increasing with time. The single consent rate available from an SAQ is near the middle of consent rates (67%).

Consent rates in telephone surveys do not appear to demonstrate any stable pattern or trend over time in Figure 2.04. Separating consent rates by telephone survey design – cross-sectional or panel – provides a clearer picture. Consent rates in telephone panel surveys can fluctuate by survey administration depending on whether respondents refused consent in a previous wave (and thus may be less likely to consent in the current wave). This fluctuation is clearly illustrated in Figure 2.05. Consent rates in cross-sectional surveys conducted by telephone are higher and more stable over time.

Figure 2.05

Mean Consent Rates by Telephone Survey Design and Year of Survey Administration



2.3.4.3 Survey Sponsor

Federal statistical agencies administer a large portion of the surveys that request consent to record access. NCHS conducts several such ongoing surveys (NHIS, NIS, and NIS Teen) as well as several past surveys (HIES, LSOA II, NHANES, NHANES III, and NHEFS). The Census Bureau conducts two ongoing surveys that link responses with administrative records - SIPP and CPS. AHRQ conducts MEPS, an ongoing survey, and EIA conducts RECS, which is also ongoing.

We identified a smaller number of surveys conducted outside of the federal government that request respondent’s consent to record linkage. This includes surveys sponsored by or associated with universities (AHEAD, HFS, HRS, PSID, and the survey described by Cleary et al., 1984). Other surveys are conducted in medical settings like hospitals and medical centers, and in conjunction with health and insurance plans (McCarthy et al., 1999; Murdoch et al., 2010; Partin et al., 2008; Woolf et al., 2000).

As predicted, we find in the bivariate analyses that consent rates in surveys with government sponsorship are significantly higher than in surveys with other types of sponsors (76% vs. 54%; p<0.01; see Table 2.07).

Table 2.07

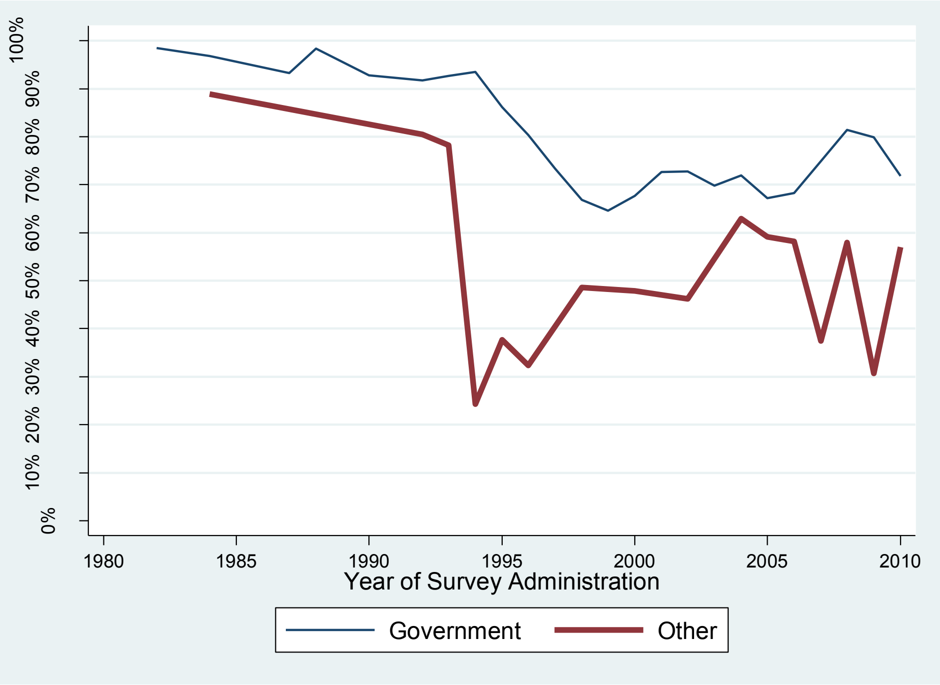
Descriptive Statistics by Survey Sponsor

	# Surveys	# Consent Requests	Mean Consent Rate	95% Conf. Int.
			%	%
Survey Sponsor				
Government	13	121	75.8	64.8, 86.7
Other	9	41	53.7	49.5, 57.8

As illustrated in Figure 2.06, consent rates both in surveys with government and other sponsorship peaked in the first decade of administration followed by a sharp decline. After plateauing, government surveys experienced an increase in consent rates upon nearing 2010. This increase may be related to the introduction in opt-out procedures and requests for reduced personally-identifying information in government-sponsored surveys. Variation in consent rates in surveys with other types of sponsorship may be attributed to volatility of consent rates in telephone panel surveys; telephone panel surveys contributed the majority of non-government consent rates. Of the 41 consent rates from surveys with non-government sponsorship, 33 of these were from panel surveys.

Figure 2.06

Mean Consent Rates by Survey Sponsor and Year of Survey Administration



2.3.4.4 Topic of Administrative Records Requested

Administrative records linked with survey responses typically relate to either respondents' health or income and employment (GAO, 2001). Health-related records include Medicare and Medicaid enrollment and claims records, or records maintained by the respondents' healthcare providers or pharmacist. Social Security Benefit histories are generally accessed to provide information about an individual's earnings and other benefits (NCHS, 2011). One survey requests records related to neither health nor income and employment: RECS focuses on household energy usage and requests access to household utility provider records.

The topic of the administrative records requested is generally related to the survey's content, which could make the request seem more germane to the interview and potentially more acceptable to the respondent. Consent requests in health surveys that contain questions on health conditions or behaviors are generally for access to respondents' health-related records. For example, the survey described by Cleary et al. (1984) includes questions about respondents' use of health care services and access to care, and consent to access respondents' medical records is requested. The HFS contains questions about healthcare utilization, insurance coverage, injuries, health conditions, behavior, and mental health, and requests access to respondents' medical records.

Similarly, surveys that focus primarily on respondents' earnings, finances, sources of income, and employment status generally request access to records containing related information. This includes the CPS and SIPP which both link responses with SSA records.

Some surveys, such HRS and AHEAD, inquire about respondents' health and economic circumstances, and request access to records related to both of these types of information. For example, AHEAD describes its purpose as examining the "interrelationships of changes and transitions for older Americans in three major domains – health, financial, and family" (AHEAD, 1993: 3) and requested access to respondent's Medicare, Medicaid, and SSA records.

We predicted that consent rates would be highest when records were related to respondents' utility records ("other"), followed by income and employment. As predicted, we found that consent rates are highest when administrative records are related to an "other" topic (89%); the difference in consent rates compared to health-related records is significant ($p<0.05$). Consent rates to income and employment records are not significantly higher than consent rates to health-related records (72% vs. 68%; see Table 2.08).

Table 2.08

Descriptive Statistics by Record Topic

	# Surveys†	# Consent Requests	Mean Consent Rate	95% Conf. Int.
Record Topic			%	%
Health	18	99	67.8	58.9, 76.8
Income/Employment	10	57	72.4	56.8, 88.1
Other	1	6	89.0	80.6, 97.4

Notes: †Seven surveys contain requests to both health and income and employment-related records. “Other” includes RECS which requests access to utility provider records.

Figure 2.07 presents mean rates of consent by record topic and year. Consent rates to records related to another topic have remained fairly stable and only minimally declined over time. However, we observe a decline in rates of consent to both health and income and employment-related records in the mid-1990s. Since then, on average, willingness to consent to requests for health-related records has ranged between 60-70%.

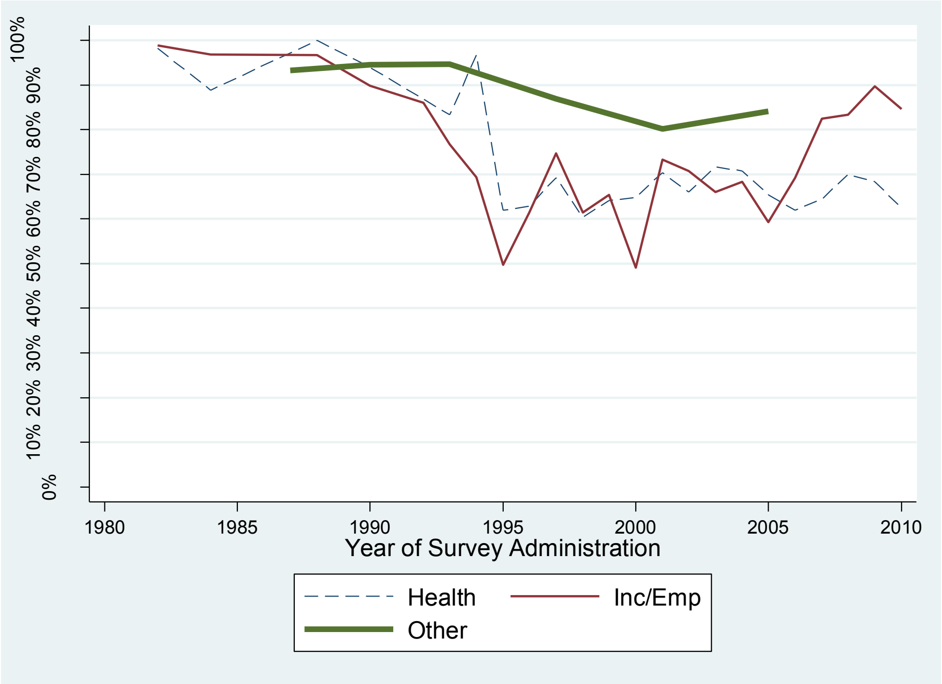
Rates of consent to requests for income and employment-related records began to climb in the mid 2000’s. This is most likely due to the introduction of opt-out consent requests around this time which obtained very high consent rates. To date, opt-out requests have not yet been used to obtain records related to other topics.[44]

---

[44] Excluding the income and employment opt-out consent rates demonstrates a trend in mean consent rates for income and employment-related records that is similar to those for health-related records (not shown).

Figure 2.07

Mean Consent Rates by Record Topic and Year of Survey Administration



2.3.4.5 Consent Mode

Consent mode – whether respondents must consent to record linkage orally or in writing – is dictated by both survey mode and institutional requirements. Regarding mode, mail surveys appear to only request consent in writing, and telephone and in-person surveys tend to request consent orally. Consent mode is also heavily influenced by institutional requirements of the agencies disclosing the records. The procedures for obtaining respondents' consent to certain records, specifically, records maintained by SSA, vary depending on whether a government or non-government agency is requesting access. In addition, states vary in their regulation concerning the release of certain types of records, with at least one state, Minnesota, requiring written permission to authorize the release of health-related records (Woolf et al., 2000).

If the survey is conducted by a government agency, such as the Census Bureau or NCHS, oral consent permits access and linkage of SSA records with survey responses, and thus respondents can provide their SSN orally to the survey interviewer.

Respondents' written consent is required if SSA records are to be linked with a survey conducted outside of the government. For example, in the HRS, conducted at the University of Michigan, respondents must provide their SSN in writing to authorize the release of SSA records. However, phone administration of the HRS somewhat complicates distribution and collection of written authorization forms. During the telephone interview, interviewers read respondents the SSA consent request. If respondents agree to this request, HRS mails a consent form to the respondent with a pre-paid mailer and cover letter. Respondents are to indicate their SSN on the consent form, authorize consent for their earnings and benefits information separately, and sign the form permitting record linkage. Forms mailed back to HRS can be submitted to SSA for record linkage (Olson, 1999; Weir, 2007). This procedure mirrors the one used to obtain access to SSA records in AHEAD**.**

The three mail surveys included in this research (McCarthy et al., 1999; Murdoch et al., 2010, Partin et al., 2008), all conducted in Minnesota, and the SAQ (Woolf et al., 2000) request respondents' written consent. Only one of these surveys - a single administration discussed by Murdoch et al. (2010) – requests a personal identifier (SSN).

Regardless of survey sponsor and mode, respondents consent to Medicare record linkage and provide their Medicare number orally, in surveys including AHEAD, HRS, LSOA II, NHANES, NHANES III, NHEFS, NHIS, and PSID. As respondents may not be able to recall their Medicare information from memory, in-person interviewers are often instructed to ask to copy the information from respondents' Medicare card. Respondents are also asked to provide their Medicaid number and permission to Medicaid records orally in HRS and AHEAD.

Consent mode for other types of information requested varies. MEPS respondents, who are interviewed in-person, are asked to endorse their consent in writing to authorize access to pharmacy and healthcare provider records. RECS, also an in-person survey, requests respondents' written authorization indicating their consent to utility provider information. In the NIS and NIS Teen surveys, both conducted by telephone, respondents are asked for oral consent to vaccination provider records. Two in-person surveys that request consent to health-related records without identifiers, HIES and HFS, both requested respondents' written consent.

Surveys containing multiple consent requests often make both requests in the same mode, such as MEPS which requests access to both pharmacy and healthcare records in writing. LSOA II, NHANES, NHANES III, NHEFS, and NHIS all request that respondents provide their Medicare number and SSN orally. HRS and AHEAD are unique in that they request consent both in writing, for access to SSA records, and orally, for access to Medicare and Medicaid records.

Unlike oral consent, which is typically requested within the survey, if respondents must provide written consent, the request for respondents' signature is typically made after the interview. HFS, HIES, MEPS, RECS and the survey described by Cleary et al. (1984), request written consent at the very end of the interview, AHEAD and HRS mail consent forms to respondents after the interview's completion.[45] Therefore, in these surveys, consent mode is confounded with the timing of the consent request.

Table 2.09 shows that, contrary to our predictions, there is little difference in mean consent rates by consent mode.

Table 2.09

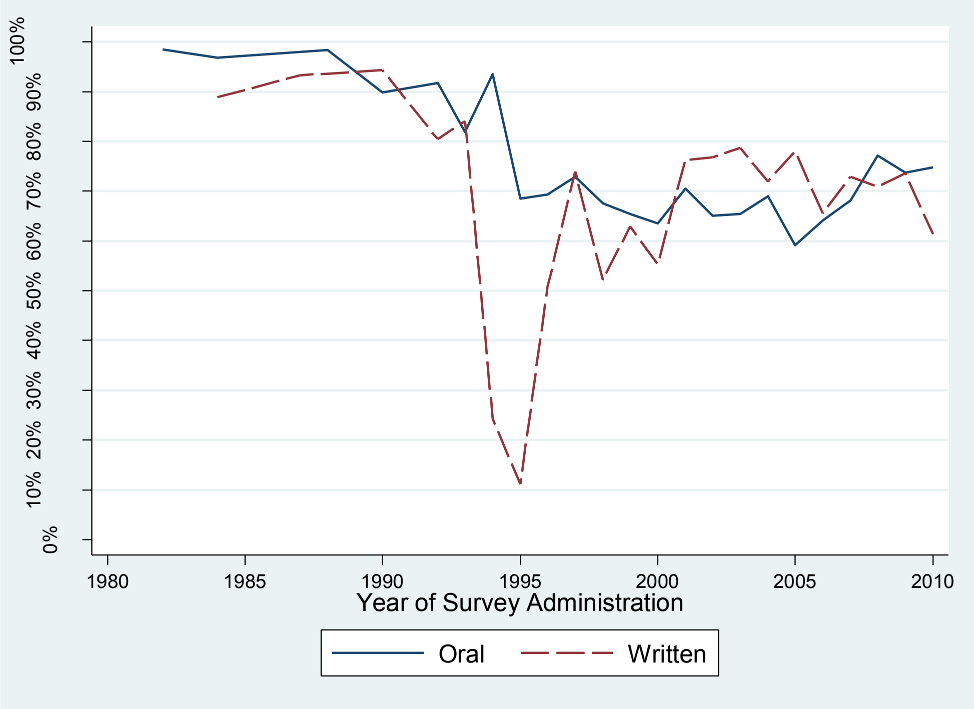
Descriptive Statistics by Consent Mode

	# Surveys†	# Consent Requests	Mean Consent Rate	95% Conf. Int.
Consent Mode			%	%
Oral	13	112	70.7	46.6, 94.8
Written	11	56	67.8	43.4, 63.9

Notes: †Two surveys contain oral and written consent requests.

Figure 2.08

Mean Consent Rates by Consent Mode and Year of Survey Administration



---

[45] Two-thirds of sample members in Murdoch et al. (2010) were mailed the consent form after the survey and half of sample members in Partin et al. (2008) were mailed the consent form after the survey. Information regarding the timing of the consent request was not available for McCarthy et al. (1999).

With the exception of one very low consent rate in a telephone panel survey in 1995, mean rates of consent to oral and written requests follow approximately the same trend over time, with variation, and tend to stay within 10% of one another (see Figure 2.08).[46] We observe an increase in average consent rates to oral consent requests from about 2005 onward, to a high of nearly 75%, likely because of the use of opt-out techniques and reductions in the amount of personally-identifying information requested. Such features were only incorporated in surveys that request consent orally.

2.3.4.6 Personal Identifier Requested

Surveys request different kinds of personally-identifying information to link responses with various types of administrative records. Government-issued identifiers requested in some surveys, including SSN, Medicare number, and Medicaid number, enable access to government benefit records. Other surveys ask respondents to provide the contact information of an individual or organization who maintains records about the respondent, for example, a doctor, pharmacist, or utility provider. Finally, some surveys request respondents' consent only, and do not ask for identifying information to facilitate linkage.

SSN is the identifier requested most frequently to link records, requested in 10 of the 22 surveys. All surveys request the full 9-digit SSN with the exception of newer administrations of the NHIS. As respondents became less willing to provide this information, NHIS started requesting only the last four digits of respondents' SSN (and Medicare number) in 2007. Those who refused to provide the last four digits were asked

---

[46] Removing consent rates from telephone panel surveys does not change the direction or significance of the findings presented in Table 8.

for permission to link using their name and date of birth (Miller et al., 2011). Also experiencing declining consent rates, the Census transitioned from a direct SSN request in the CPS and SIPP to an opt-out approach.

Requests for Medicare number are almost nearly as prevalent as requests for SSN and were made in eight surveys. As the majority of Medicare recipients are over the age of 65, the request may not apply to all survey respondents, and so questions to determine Medicare status typically precede this request (for example, in NHANES and NHIS). HRS, which requests Medicaid number, similarly identifies respondents who receive Medicaid benefits before asking for their Medicaid number.

Three health surveys, MEPS, NIS, and NIS Teen, request the contact information of respondents' health service providers. MEPS respondents are asked for the contact information of any medical provider or pharmacy if respondents report using their service during the reference period of that interview. (This is in addition to signing a permission form authorizing the release of these records.) All NIS and NIS Teen respondents are asked for the contact information of their child's healthcare provider to request vaccination record information, even if respondents say their child has not been vaccinated. Respondents who say that their child has not received any vaccinations may have forgotten about them or be unaware they occurred (Lee et al., 1994).

To link survey responses with records from utility providers, RECS respondents are asked to provide the contact information for their energy suppliers, found on their utility bill, for each type of fuel they use, in addition to signing authorization forms granting EIA permission to contact each of these companies, as well as providing scanned copies of their fuel bills. The specific types of energy the respondent uses is determined earlier in the survey.

Several surveys do not request any personally-identifying information from respondents. If no identifier is used to assist linkage, surveys must match responses with records in another way. For example, to facilitate record identification and linkage for opt-out consent procedures (where consent is not directly requested), the Census Bureau extracts information from respondents' SSA NUMIDENT file which contains their SSN. The SSN is then used to link survey and administrative record data (McNabb et al., 2009).

Several healthcare surveys did not request any personal identifiers. The smaller size of these surveys permitted more manual matching methods. Surveys such as the HFS, HIES and those conducted by McCarthy et al. (1999); Murdoch et al. (2010), Partin et al. (2008), Woolf et al. (2000) had rich frame information available from which consenting patients could be identified. Researchers in the survey described by Cleary et al. (1984) conducted record linkage by auditing records from all area physicians and pharmacies and searching them by consenting respondents' name to match records with survey responses.

Table 2.10 shows the range in consent rates by identifier requested to facilitate consent - from 57% of respondents who provided their Medicaid number to 89% of respondents who consented when another type of identifier was requested (RECS). When no identifier was requested to facilitate linkage, 82% of respondents consented.

Table 2.10

Descriptive Statistics by Identifier Requested

	# Surveys	# Consent Requests	Mean Consent Rate	95% Conf. Int.
Identifier Requested			%	%
SSN (all)	10	51	68.4	54.7, 82.1
SSN (9-digit)	10	48	67.9	51.1, 84.7
SSN (4-digit or link)	1	3	76.3	61.9, 90.8
Health-Related Identifiers	11	90	68.1	58.4, 77.7
Medicare Number (all)	8	32	59.4	49.6, 69.3
Medicare Number (9^+-digit)	8	29	58.3	46.5, 70.2
Medicare Number (4-digit or link)	1	3	70.1	54.8, 85.3
Medicaid Number	2	9	56.5	51.5, 61.5
Healthcare Provider Info.	3	34	75.8	66.0, 85.6
Pharmacy Info.	1	15	75.8	73.7, 77.9
Other Identifier	1	6	89.0	84.0, 93.9
None	10	15	82.0	64.8, 99.2

Notes: “Other Identifier” refers only to utility provider contact information requested in RECS.

Although not a significant difference, it appears that requesting only the last four digits of respondents’ SSN and allowing respondents to participate in linkage without providing an identifier benefits consent rates as compared to requesting the full SSN (76% vs. 68%, n.s.). However, as indicated in Table 2.10, reduced SSN information was requested in only three survey administrations, and all from NHIS.

Figure 2.09 presents mean consent rates by type of personally-identifying information requested and year of survey administration. Across surveys overall, willingness to provide SSN (all variations) and health-related identifiers (Medicare number, Medicaid number, and healthcare provider and pharmacy contact information) declined in the mid-1990s and from then on, fluctuates between 60-70%. Rates of consent when no identifier is requested fell in the late nineties and peaked in the late 2000s, likely with the introduction of opt-out consent requests. Consent rates when another identifier is requested (RECS) are higher overall and more consistent.

Figure 2.09

Mean Consent Rates by Identifier Requested and Year of Survey Administration

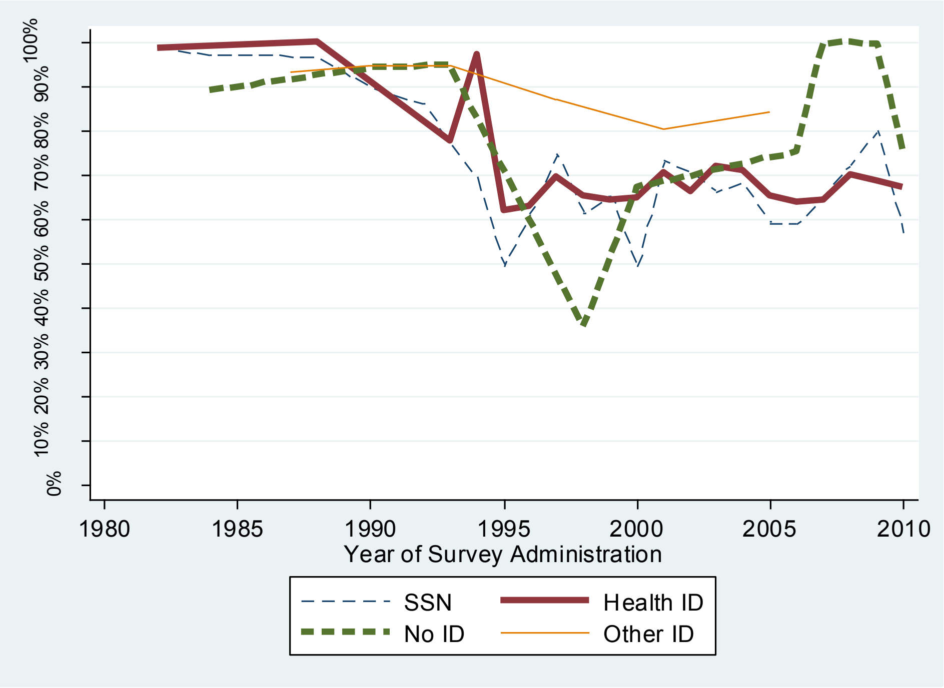
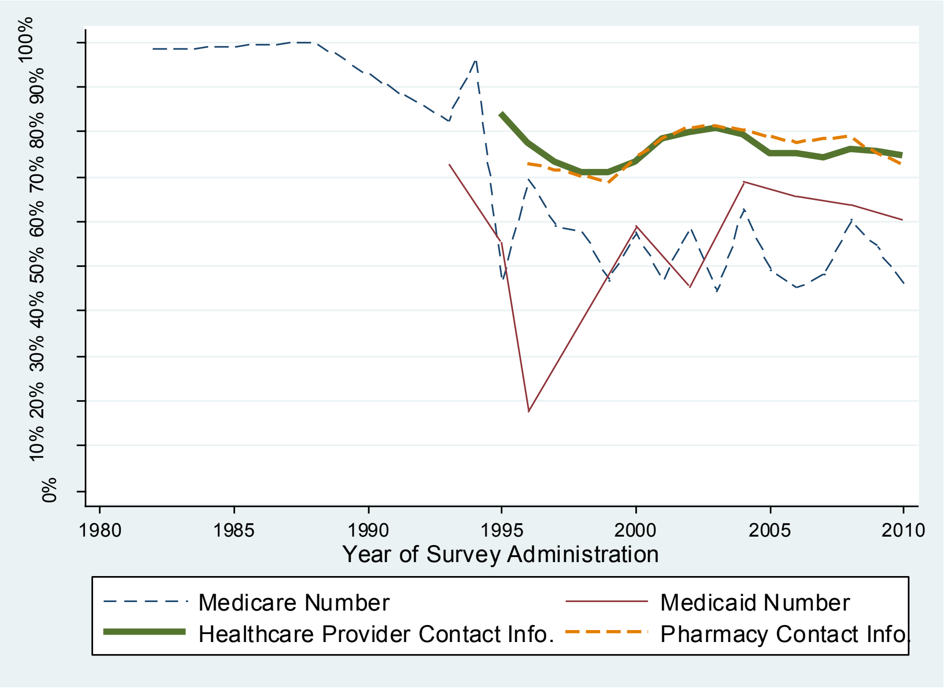


Figure 2.10 examines the effect of the different types of health-related identifiers on consent rates. Mean consent rates to requests for Medicare number fell in the mid-1990s, since then, they have fluctuated between about 50% and 60%. Some of the variation in Medicare and Medicaid consent rates over time is likely related to variability in the consent rates in telephone panel surveys. Mean consent rates for healthcare-related contact information (pharmacy or healthcare provider) to enable record linkage is greater in years this data is available and more stable.

Figure 2.10

Mean Consent Rates by Health Identifier and Year of Survey Administration



2.3.4.7 Explicit vs. Implicit Consent

A final distinction in consent procedures we examine is whether respondents' consent to record linkage is granted explicitly, as in the majority of survey administrations, or implicitly, as in CPS and SIPP administrations including and after 2006, and in SHIPP, a third Census-sponsored survey.[47] Prior to 2006, CPS and SIPP requested consent explicitly.

As predicted, consent rates obtained through implicit procedures are significantly higher than those obtained from explicit procedures (p<0.001; see Table 2.11).

---

[47] Implicit consent procedures varied slightly in SHIPP compared to CPS and SIPP. Households were mailed an advance letter informing them of the record linkage. If respondents acknowledged receipt of the letter and did not object during the interview, this was considered implicit consent. If respondents did not recall receiving the letter or were not mailed a letter, they were asked for explicit consent. Thirty-eight percent of respondents gave implicit consent, and overall, 90% of respondents consented (Pascale, 2011).

Table 2.11

Descriptive Statistics by Consent Request Type

	# Surveys[†]	# Consent Requests	Mean Consent Rate	95% Conf. Int.
Request Type			%	%
Explicit	21	155	69.0	59.7, 78.2
Implicit	3	7	98.3	95.1, 101.6

Notes: [†]Two surveys use both explicit and implicit procedures in different years.

Figure 2.11

Mean Consent Rates by Consent Request Type and Year of Survey Administration

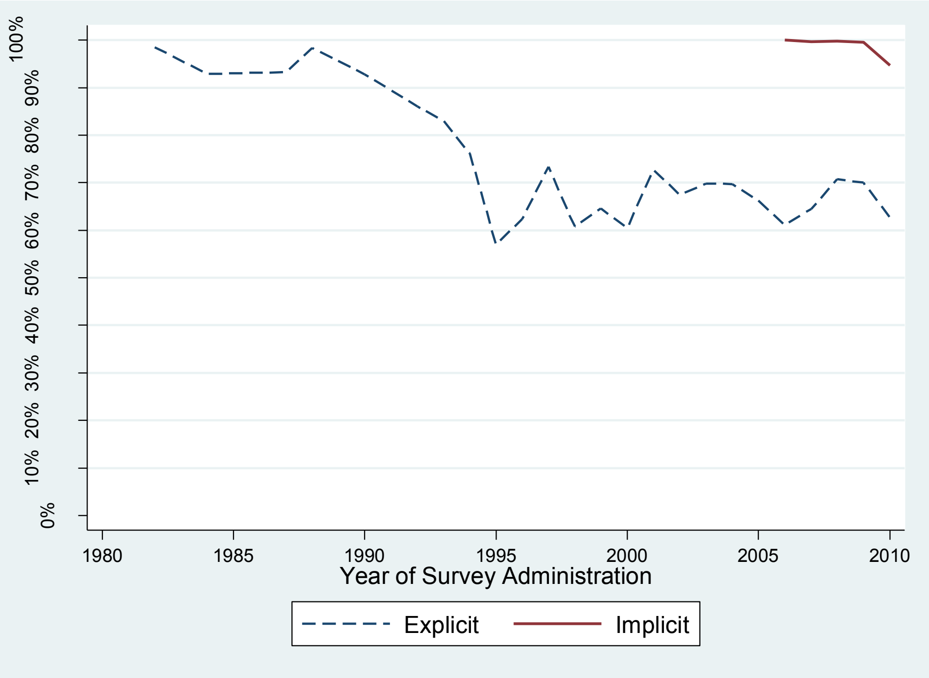


Figure 2.11 shows the trend in implicit and explicit consent rates over time. After declining in the mid-1990s, mean rates of consent to explicit requests range between approximately 60-70% from 1995 onward. In contrast, mean rates of consent obtained through opt-out requests are drastically higher for the years these consent rates are available (2006-2010), with rates of consent not falling below 90%.

2.3.5 Multivariate Analysis

The bivariate analyses identify that several characteristics are related to consent, including survey mode and sponsorship, identifier and topic of records requested, and the use of opt-out consent procedures. However, the results produced through the bivariate analyses are affected by confounds in the data. For example, no mail or telephone panel surveys had government sponsorship which is positively associated with consent. Additionally, no mail surveys used opt-out consent procedures, another positive predictor of consent. Government-sponsorship is confounded with other characteristics including the use of opt-out consent requests, which is a strong positive predictor of consent, and survey mode.

There are also confounds that are not significantly associated with consent rates: all mail surveys required respondents' written consent and the majority of in-person surveys requested oral consent, most mail surveys requested consent without an identifier, and all mail surveys requested access to health records.

To begin to address these confounds and sort out the extent to which the bivariate effects are spurious, we estimate a multivariate regression model predicting consent rate as a function of the seven characteristics of interest in this chapter and year. This model is presented in Table 2.12.

Table 2.12

Linear Regression Predicting Consent as a Function of Survey and Consent Request Characteristics[48]

		Coef.	SE	*p*
Constant		99.53	11.03	0.000
Year		-0.66	0.32	0.052
Survey Mode†	In-Person (ref.)			
	Mail	-29.62	10.30	0.010
	Phone: Panel	-18.34	7.09	0.018
	Phone: Cross-Sectional	10.97	7.89	0.181
Survey Sponsor	Government (ref.)			
	Other	-3.46	5.12	0.507
Record Topic	Health (ref.)			
	Income/Employment	-16.73	8.49	0.064
	Other (*utility*)	-2.83	6.54	0.670
Consent Mode	Oral (ref.)			
	Written	-0.16	7.09	0.982
Identifier Requested†	SSN (ref.)			
	Health-Related[49]	-16.29	7.26	0.037
	None	-1.09	6.58	0.870
Request Type	Explicit (ref.)			
	Implicit	29.94	9.10	0.004
Survey Response Rate		0.04	0.05	0.479

Notes: N=154; r^2=0.47; adjusted Wald test for all parameters: $F(10,19) = 2648.41$, $p<0.001$. † SAQ (Survey Mode) was dropped from the model; Other (Identifier Requested) was dropped from the model. Year is coded as the number of years between 1982 and the year of administration for a particular survey.

When controlling for other sources of variation using multivariate regression, we find that some characteristics of the survey and consent request are related to consent rates in the same manner as the bivariate analyses. Similar to the bivariate analyses, there is a significant effect of survey mode on consent rates, where mail and telephone panel surveys obtain significantly lower consent rates than in-person surveys (p<0.05). We

---

[48] Only one interaction term accurately portrayed the characteristics of the surveys in this analysis; other combinations of characteristics were not observed. The term - *Survey Sponsor x Record Topic* - was not significantly associated with consent and was excluded from the model.

[49] An analogous multivariate model predicting consent rates that includes specific healthcare identifiers is included in the Appendix. Briefly, in that model we find that there is a negative effect of Medicare number (p<0.01), Medicaid number (p<0.001), and healthcare provider information (p<0.05) on consent rates, but no effect of pharmacy information on consent rates.

also continue to find that, as in the bivariate analyses, request type is a significant predictor of consent rates with consent rates obtained through implicit procedures significantly higher than those obtained through explicit procedures ($p<0.01$). Also, as in the bivariate analyses, we find no relationship between consent mode or survey response rates and consent rates, even after controlling for other sources of variation in the multivariate model.

In the multivariate model, we find that consent rates are significantly lower when a health-related identifier is requested compared to requesting SSN ($p<0.05$), a finding not observed in the bivariate analyses. Because the model also controls for record type, this finding may be a function of increased burden of providing this information which is less likely to be accessible from memory than SSN.

In the bivariate analyses, survey sponsorship was a significant predictor of consent rates, with government-sponsored surveys achieving significantly higher consent rates than those with another sponsor ($p<0.01$). Yet, after controlling for other factors in the multivariate model, we see no significant effect of sponsorship on consent rates, suggesting that the effects in the bivariate analyses may have been spurious and related to confounds in the data, for example survey mode or the use of opt-out consent requests.

Regarding record topic, consent rates were highest when records related to another topic (utility records) in the bivariate analyses, but the difference between health-related records and income and employment-related records was not significant. When controlling for other factors in the model, we find that rates of consent to income and employment-related records are significantly lower than those relating to health ($p<0.10$), but we see no difference in consent rates for records relating to another topic.

Additionally in this model we find a negative effect of survey year on consent rates ($p<0.10$). This contrasts with results that we presented from an earlier model in which we controlled for the individual survey.

2.4 Overall Conclusions

Over time, more surveys are linking responses with administrative records and subsequently include requests for consent to such linkage. This increase in surveys conducting record linkage is likely due to greater technological capabilities over time, including improvements in the ability to process and store data (Calderwood et al., 2009; Scheuren, 1997), as well as a more comprehensive understanding of the benefits of linked data (GAO, 2001).

In this chapter, we sought to describe the current consent environment in the U.S. using all available consent rates from such current and past surveys, as well as information characterizing each of these surveys and requests. The mean consent rate across the 162 available consent rates is 70.2%; after falling in the early 1990s, consent rates have fluctuated between approximately 60-70%. We used linear regression to investigate the relationship between survey year and consent rates, and found that when controlling for the individual survey, we do not observe any trend in consent rates over time. However, a separate model that excludes predictors for the individual survey but controls for survey and consent request characteristics suggests that consent rates are declining across surveys ($p<0.10$)

We also investigated the relationship between seven characteristics of the survey and consent request on consent rates including survey mode, sponsor, and response rate; whether consent is requested orally or in writing, whether the request takes an explicit or opt-out approach, the topic of the records requested, and any personally-identifying information requested to facilitate the linkage. Yet, the number of confounded characteristics across the surveys included in this research increased the difficulty in drawing conclusive findings through bivariate analyses, and so multivariate analyses were then used to address these confounds.

Contrary to our hypothesis, we did not observe a strong or consistent relationship between survey response rate and consent rate in both the bivariate and multivariate analyses, even when examining this relationship by consent request within surveys. This suggests that, counter to our hypotheses, the mechanisms contributing to unit response may differ from those contributing to consent. We also note that some consent requests are not asked of all respondents, such as for Medicare and Medicaid number. Thus, as the survey response rate applies to all respondents and some consent requests apply to only a subgroup of respondents, correlating these numbers does not provide a perfect index of the relationship between consent and response rate for all requests.

Mean consent rates varied by survey mode. As predicted, we observed the highest mean consent rates for surveys conducted in person (75%) as compared to those conducted by telephone (63%) or mail (49%); the consent rate from the SAQ was 67%. Consent rates from in-person surveys were significantly greater than those in mail surveys in both the bivariate and multivariate analyses. This may be related to benefits of interviewer administration, such as probing or rapport, or increased perceived legitimacy of the interview when conducted in person, as initially hypothesized. Average consent rates in telephone surveys fell between in-person and mail surveys in the bivariate analyses; this directionality may suggest that the consent request benefits from interviewer administration.

We also investigated differences in consent rates by telephone survey design. Average consent rates in telephone surveys with panel designs were significantly lower compared to those collected through telephone surveys with cross-sectional designs (52% vs. 82%; $p<0.001$). This is at least partly due to differences in the sample composition of respondents asked for consent, with prior refusers asked to consent again in the telephone panel surveys.

Consent rates varied by sponsorship in the hypothesized direction in the bivariate analyses: surveys with government sponsorship achieved significantly higher consent rates on average (75%) compared to surveys with another type of sponsor (54%). However, when controlling for other factors in the multivariate analyses, we no longer observe an effect of sponsorship, suggesting that the findings in the bivariate analyses were a result of confounds, such as the use of opt-out requests and survey mode.

We predicted that consent rates would be highest for utility records, and higher to requests for income and employment-related records as compared to health-related records. In the bivariate analyses, RECS respondents (categorized as “other”), had the highest overall average consent rates (89%); the difference in consent rates compared to records pertaining to health is significant. Though, as only one survey requests access to records of another topic, it is unclear if the especially high consent rates observed in RECS can be attributed to RECS specifically or the topic of the records requested (utility records). Consent rates did not significantly vary between health and income and employment-related records in the bivariate analyses (health: 72%; income and employment: 68%).

However, in the multivariate analysis, we find that rates of consent to income and employment-related records are significantly lower than those relating to health ($p<0.10$), but we no longer see any difference regarding records of an “other” topic. This suggests that when controlling for other factors – specifically the identifying information requested – respondents may find the content of income and employment-related records more sensitive than those relating to health or utility.

We hypothesized that requesting written consent would be detrimental to consent rates because it would heighten respondents’ suspicion of the request. However, we did not observe any difference in mean oral and written consent rates in the bivariate analysis (71% vs. 68%) or multivariate analysis.

We propose that the lack of differences in consent rates is due to the placement of the written request rather than the mode of consent. As discussed earlier, requests for written consent are often made at the end of the survey interview or after the interview’s completion. (In contrast, requests for oral consent are typically presented within the survey.) Respondents may feel more comfortable consenting after gaining a better understanding of the purpose of the research. In self-administered surveys, sample members willing to complete the survey after reviewing it may be the same individuals who are willing to sign written consent statements. In interviewer-administered surveys, the placement of the request could positively influence willingness to consent as respondents may feel obligated to cooperate after developing rapport with the interviewer.

Average consent rates also varied by the identifier requested to facilitate consent; however, most of these differences were not significant, and several were not in the hypothesized direction. Although we hypothesized that consent rates to 9-digit SSN request would be lowest, we found that Medicaid and Medicare number consent rates were lowest, respectively in the bivariate analyses. In the multivariate analysis, consent rates were significantly lower when a health-related identifier was requested as compared to a SSN ($p<0.05$).

Medicare and Medicaid numbers are less frequently used and requested identifiers, and thus, respondents may not be able to recall them from memory. The additional burden of looking up these identifiers may prevent some respondents from consenting. Further, for all surveys included, only a subset of respondents are eligible for the Medicare and Medicaid number requests, and these respondents' willingness to provide this information may differ from the broader population of respondents.

As noted, Medicare number is an individual's SSN followed by a letter and potentially a number.[50] Given the similarity of the two identifiers, it is surprising that overall consent rates to 9-digit SSN are higher than Medicare number. This could suggest that the added burden of retrieving the additional Medicare information (letter and/or number) is detrimental. We also note that other characteristics can vary across surveys that contain SSN and Medicare number requests (such as survey topic) that could influence willingness to provide these identifiers.

---

[50] For example, if an individual's SSN is 123-45-6789, their Medicare claim number would be 123-45-6789A if they are the both the wage earner and the individual receiving the Medicare benefits. Retrieved from http://ssa-custhelp.ssa.gov/app/answers/detail/a_id/1366/~/meaning-of-the-letters-after-a-social-security-or-medicare-number

We find that average consent rate for 4-digit SSN is higher than 9-digit SSN, in line with our predictions, though few data points for this request were available and the difference is not significant in the bivariate analysis (76% vs. 68%). Mean consent rates for healthcare and pharmacy provider contact information are quite high (76% and 76%). The mean consent rate for another type of identifier (contact information for energy providers) is even higher – 89% - suggesting that respondents don't find this information prohibitively sensitive. Overall, not requesting an identifier results in greater consent rates (82%) than requesting an identifier (69%) though the difference is not significant in the bivariate or multivariate analyses.

The characteristic that appears to have the most significant influence on consent rates is whether the survey uses an implicit or explicit consent request procedure. Consent rates obtained through implicit procedures are significantly higher than those obtained through explicit procedures in the bivariate analyses (98% vs. 69%; $p<0.001$), and after controlling for other factors in the multivariate analyses ($p<0.01$). We initially hypothesized that consent rates to explicit requests would be lower due to the added burden of these requests, but an alternative explanation for these findings is that some respondents may be unaware that they are consenting to record linkage when an opt-out procedure is used.

Opt-out consent requests are incredibly effective in eliciting high consent rates. Yet, if a secondary purpose of this procedure is to inform respondents that their records are being linked, further research needs to determine the success of that goal. Eliminating both a request for an identifier and a direct consent question makes the record linkage less transparent.

2.5 Limitations

There are several limitations of the research presented in this chapter, including confounding characteristics across studies, missing data, a small number of eligible consent rates, variability in data quality, and an inability to determine causality given the methods used.

As noted, characteristics of the survey and consent request are not evenly distributed across the surveys in this data which substantially limits our ability to determine the effect of specific characteristics on consent rates. True experimental research that randomizes the assignment of each of these features (for example, survey sponsor or consent mode) is needed to accurately assess their effect on consent rates.

We identified 22 eligible surveys, producing a dataset with 162 consent rates across 28 years. Further subsetting our dataset by various characteristics of interest (e.g., any identifier requested) limited the power of the analyses and our ability to draw robust conclusions. In addition, several surveys are missing consent rates for certain administrations or consent requests.

Although we conducted a thorough and extensive search to identify all surveys requesting consent to record linkage, it is possible that we failed to locate some eligible surveys. Specifically, we acknowledge a risk of missing smaller surveys or surveys conducted by mail, similar to those we identified in the published literature included in this research. These surveys were more difficult to locate than the larger surveys conducted in person or by telephone that were mentioned on multiple web pages and more frequently cited in the literature.

The research presented in this chapter is also limited by the accuracy of the data examined. As noted, the quality of the earlier NCHS surveys may be somewhat compromised. In addition, as we collected much of the data by hand, there is potential that we introduced error through the coding process.[51] Further, the definition of consent varied across surveys, contingent on the nature of the request, and some surveys, such as those described by Partin et al. (2008) and Murdoch et al. (2010), incorporated experiments into consent procedures which introduce additional variation into our dataset. Response rate calculations varied as well, depending on the survey design.

Lastly, in using largely qualitative and descriptive methods, we are unable to determine causal relationships between the characteristics examined and variation in consent rates. If the number of surveys with consent requests continues to increase over time, and there is more variation in the characteristics of surveys requesting consent, there will be opportunities to quantitatively describe these relationships. However, only experimental research can overcome confounds in this data that result as a function of institutional requirements (i.e., consent mode and survey sponsorship).

---

[51] Data was largely collected by a single individual (the author) and so no measure of coding reliability was calculated.

3 CHAPTER 3: INFLUENCES ON CONSENT IN THE 2011 JPSM PRACTICUM SURVEY

3.1 Introduction

This chapter investigates three potential determinants of consent to record linkage: the type of administrative records requested; respondents' feelings about privacy, confidentiality, and trust; and the salience of the consent request.

3.2 Research Questions

3.2.1 The Effect of Type of Record for Which Consent to Link Is Requested

As demonstrated in Chapter 2, record linkage consent procedures vary by survey. Such differences may contribute to variations in consent rates (Fulton & Tourangeau, 2011; Jenkins et al., 2006; Sala et al., 2010; Singer et al., 2003) although much of the research identifying such effects relies on hypothetical scenarios.

Experimental research is needed to assess the impact of variations of the consent request on consent rates; e.g., the personally-identifying linking information and administrative records requested, the amount of detail contained in the consent request, and whether consent is obtained written or orally. The first section of this chapter begins to fill this need by reporting the results of an experiment that randomly varied the type of administrative records requested from respondents (either health or income and employment-related records).

3.2.2 The Effect of Concerns about Privacy, Confidentiality, and Trust

This chapter also investigates whether privacy, confidentiality, and trust concerns are related to cooperating with the request for record linkage (Singer et al., 1993; Singer et al., 2003) through non-experimental research methods. In addition, we examine the reasons respondents offer in explaining why they decided to consent or withhold consent.

This should identify other factors that respondents see as impediments to consenting, as well as shed light on how respondents see the potential benefits of consenting.

3.2.3 The Effect of Consent Request Salience

Finally, this chapter examines the effect of the salience of the request on consent likelihood through non-experimental research methods. Respondents who find the consent request for record linkage to be more salient may be more likely to grant consent (Sala et al., 2010).

Like survey nonresponse, consent refusal has the potential to produce bias (Groves et al., 2004; Sakshaug et al., 2011). The present research does not involve access to any administrative records, regardless of whether or not respondents consent, and so it will not be possible to directly determine whether bias results from refusals to consent. However, analysis of survey items that are associated with the information in the records will provide an indirect indicator of the likelihood of bias.

3.3 Methods

3.3.1 Practicum

We carried out this research in the 2011 Joint Program in Survey Methodology (JPSM) Practicum survey, which is part of a two-semester JPSM course for master's students. The 2011 Practicum provided data for this dissertation and for another dissertation on the impact of an advance monetary incentive on data quality.

3.3.2 Survey Development and Pretesting

3.3.2.1 Focus Groups

Two 90-minute focus groups, conducted the evening of March 10, 2011 at Shugoll Research in Bethesda, MD, informed the design of the Practicum survey. Both groups explored attitudes toward surveys; more specifically, types of information respondents would be willing to provide in a survey, privacy concerns, and the impact that incentives would have on their responses. Students assisted in the development of the moderator's guide.

Twenty-one individuals participated and received an honorarium of $85. Groups were balanced on gender and age. Participants tended to be more educated and white. About one-quarter of group members were black and no group members were Asian or any other race. No group members were Hispanic.

As an introduction to the privacy section, the moderator asked group members about the kind of information they would be uncomfortable or unwilling to provide in a survey. In response, members of both groups mainly included identifying information such as name, address, and SSN. One respondent offered an explanation for his opinion:

> "Certain questions… for surveys, they are just trying to get a broad group of the population which is why they ask age, race… questions of that nature. But when they start getting too personal, like with Social Security Numbers, or addresses, any type of thing that they can… particularly pinpoint you, that is when it gets awkward."

To explore respondents' reactions to a consent request for record linkage, the moderator asked respondents to read an abbreviated version of the consent request from the NHIS, which stimulated further discussion.

> To help us link your survey data with vital statistics and health-related records of other government agencies, we would like the last four digits of your Social Security Number. The National Center for Health Statistics, a government agency, uses this information for research purposes only. Providing this information is voluntary. Federal laws authorize us to ask for this information and require us to keep it strictly private. There will be no effect on your benefits if you do not provide this information.
>
> The specific federal laws are the Public Health Service Act (Title 42, United States Code, Section 242K) and the Confidential Information Protection and Statistical Efficiency Act (Title V of Public Law 107-347).

When asked, the overwhelming majority of respondents said they would not provide their SSN in a survey context and many noted that they would not grant record access even if SSN was not needed to facilitate the linkage. (As noted above, respondents may overestimate their unwillingness to consent (Singer et al., 1992).) Respondents expressed concern about what they perceived to be a lack of detail regarding the use of their records in the request statement, and about the potential for confidentiality breach.

> "I don't think it would remain confidential, that is my concern. It's linking with other data, I don't know where it's going to end up eventually. I don't know who is going to see it. There is nothing specific that says who else is going to look at it. A lot of this information now is accessible through the internet."
>
> "What is hanging me up is truly the SSN part of it. You could trust them in that moment but there is no protection then. What if they do release it? Then you're screwed."

The process of linking survey responses with administrative records baffled some respondents, as did the choice of SSN as an identifying variable.

> "They'd do a lot better doing other identifiers... maybe gender, year of birth, left handed versus right handed, what census tract do you live in. There are other identifiers that are a lot better than the last four digits of your Social Security Number. If you surveyed a million people, you'd have a lot of duplication!"
>
> "I don't think it would make any sense on how they would match it."
>
> "I don't think of four [digits] as being personally-identifiable."

Even though some respondents remained unconvinced that such a request would be part of a federally-funded survey, the content of these discussions suggested that respondents understood the request and the linkage process, at least in a broad sense.[52]

> "I don't believe that the government would be asking you for your SSN, because they warn you every day, they tell you every step of the way – 'never give out your Social Security Number!'"

---

[52] Through cognitive testing, Bates (2005) concluded that respondents did not equally comprehend the material contained in the consent requests for record linkage.

> "I do not think that there is this much cooperation within the federal government, among federal agencies."

Before conducting the focus groups, we considered incorporating a request for the last four digits of respondents' SSN as part of the consent request in the Practicum survey. Considering the overwhelmingly negative responses to the four-digit SSN request in the focus groups, the decision was made to exclude any request for identifying-information in further pretests and the final version of the survey to prevent any breakoff. As several surveys conduct linkage without explicitly asking for identifiers, including surveys sponsored by the Census Bureau, this should not limit the usefulness of our results.

3.3.2.2 Other Pretesting

The Practicum students conducted a total of twenty cognitive interviews during the week of April 7, 2011 using an early iteration of the survey. Students employed various cognitive interviewing protocols (mainly concurrent and retrospective think-aloud procedures) and developed specific probes for the draft survey questions. As groups, they identified problematic questions and proposed solutions.

Two small-scale field tests further evaluated the survey instrument at different stages of progress. In the first, students conducted 29 interviews at Westat on May 1, 2011 after they received an hour of standardized interviewing training. After further revisions, Princeton Survey Research Associates International (PSRAI) conducted the second pretest during the evenings of July 6-7, 2011. This pretest resulted in 42 completed interviews averaging 22 minutes. PSRAI prepared a memo based on interviewer feedback identifying problems that surfaced during the pretest and suggested improvements. No Practicum students were directly involved in the final pretest, and neither pretest utilized incentives or an advance letter.

3.3.3 Final Instrument

The final Practicum survey instrument spanned a range of topics to address the needs of both this consent dissertation and those of the incentives dissertation (see Table 3.01). The questionnaire included items on health and health attitudes; attitudes toward the country, current events, privacy, confidentiality, and trust; employment; income; the "Big 5" personality trait of conscientiousness; a request for consent to record linkage; and demographics. (See Appendix for the full instrument.) The majority of items were either taken from surveys such as the Census Surveys of Privacy Attitudes, the American Community Survey (ACS), GSS, NHANES, and NHIS, or were modified from items used in other surveys.

Table 3.01

Topics Included in the Practicum Survey

Question Topic	# of Questions
Health Attitudes	5
General Health	11-19
Health Insurance and Expenditures	2-5
Attitudes Toward America and American Institutions	15
General Public Opinion	3
Income and Employment	5-7
Consent Request Module	2
Measures of "Big 5" Conscientiousness Items	10
Privacy, Confidentiality, and Trust	9
Television, Telephone, and Computer Usage	2
Demographics	6
Prenotification Letter Recall	1-2

Notes: The number of questions a respondent was asked depended on the responses they endorsed.

3.3.3.1 Consent Request

Respondents were randomly assigned to a request for their consent to record linkage for either health-related administrative records or income and employment-related administrative records.

Health-Related Administrative Record Consent Request

> We would like to understand how the use of health care may change as people age. To do that, we need to obtain information about vital statistics, health care costs and diagnoses from your health-related records. In order for us to retrieve these records, we need your consent. This will allow us to conduct more research without asking additional questions. Your consent is voluntary and the information that you provide will be kept completely confidential. May I have your consent to access these records?

Income and Employment-Related Administrative Record Consent Request

> We would like to understand how people's income changes as they age. To do that, we need to obtain information about income and employment from your income and employment-related records. In order for us to retrieve these records, we need your consent. This will allow us to conduct more research without asking additional questions. Your consent is voluntary and the information that you provide will be kept completely confidential. May I have your consent to access these records?

Although the consent request came approximately three-quarters through the survey, after the interviewer and respondent had sufficient time to build rapport (Morton-Williams, 1993), a series of questions relating to earnings and sources of income directly preceding the request may have heightened respondents' privacy concerns and affected their consent decision.

The consent statements in the Practicum survey were modeled on the requests for Medicare number in the PSID and for SSN in the NHIS. Although no actual record linkage took place, PSRAI and interviewers were unaware of the inauthenticity of the record linkage request. To further promote cooperation, we developed and provided interviewers with a one-page document of what we anticipated would be commonly-asked questions regarding the consent request and how interviewers should address them (see Appendix). The need for materials that addressed these concerns was identified during the student-administered pretest.

Directly after the consent request, interviewers asked respondents to elaborate on the reasons for their consent decision, regardless of whether they did or did not consent to record linkage. The wording of this question varied depending on whether or not the respondent granted consent. We attempted to soften the language of this question for respondents who did not consent to prevent any break-off in case respondents were especially bothered by the follow-up request.

If respondents consented:

> "Can you tell me why you decided to consent to this request to access your records?"

If respondents did not consent:

> "I appreciate your patience and I indicated that you do NOT consent. Before we move on to the next section, can you tell me why you decided not to consent to this request to access your records?"

We chose to capture respondents' rationale for their decision through an open-ended question rather than a forced choice question. With little prior insight into the type of responses that might be offered, it was important not to limit the range and diversity of responses allowed. Further, open-ended questions are useful for allowing respondents to describe their reasoning behind a conclusion or behavior.

3.3.3.2 Sample

A directory-listed telephone and mail sample of non-institutionalized persons age 18 and over in the 48 contiguous states and Washington, DC was selected for the Practicum survey by Survey Sampling International (SSI). This was chosen over a more comprehensive frame (for example, RDD) in order to meet the goals of the incentives dissertation, which required addresses to mail an advance incentive.

SSI's database, the frame from which the Practicum sample was drawn, was generated by merging various public and private databases, such as residential telephone numbers from InfoUSA, birth records, voter registration records, real estate transactions, credit sources, vehicle registrations, and proprietary sources owned by SSI. SSI stratified the Practicum sample on FIPS codes, ZIP code, and telephone number.

3.3.3.3 Incentive and Advance Letter

PSRAI mailed an advance letter to sample members shortly before attempting to obtain an interview (see Appendix). Letters were sent in two batches: the first batch was sent to 3,400 sampled individuals on July 14th and 15th; and the second was sent to 3,800 individuals on July 28th and 29th.

The letter explained that the sample member would be contacted by PSRAI for a phone survey conducted by researchers at the University of Maryland on "Americans' health and their views on various social issues." Both the letterhead and envelope contained the University of Maryland logo. For a randomly-selected 40% of the sample, a $5 bill was included with the letter. This fraction was selected in order to achieve fairly equal portions of incentive and non-incentive receiving respondents (as suggested by results from Tourangeau, Groves, & Redline, 2010).

3.4 Survey Results

PSRAI fielded the final version of the survey from July 18, 2001 to August 17, 2011. All interviews were conducted in English. Up to six call attempts were made to gain cooperation with each sample person, and additional calls were made to complete partial interviews. The average interview length was 21 minutes (range 13 – 62 minutes). The AAPOR Response Rate 1 was 15.7%. Nearly two thirds (63%) of the 900 completed cases received an incentive.

Compared to estimates from the CPS 2010 March Supplement, the unweighted Practicum sample was considerably older and underrepresented females and minorities (see Table 3.02).[53] JPSM created replicate weights through raking using the jackknife procedure in WesVar. Data was weighted to match age, gender, region, and education control totals from the 2010 CPS March Supplement. Missing data on these variables in the Practicum survey were first imputed using hot deck imputation prior to weighting. Individual weights greater than 2,000,000 were trimmed to prevent any individual respondent from contributing too much influence. Raking was performed a second time

---

[53] Females are often overrepresented in telephone surveys (Salmon & Nichols, 1983). Here, we observe the opposite potentially because many of the names available for contact information in the list sample were for the heads of household who are often male.

on the trimmed weights. Estimates produced with the replicate weights closely aligned to CPS estimates.

Table 3.02

Sample Demographics

	Unweighted Practicum Estimates	2010 CPS March Supplement	Weighted Practicum Estimates
	%	%	%
Gender			
Male	67.2	48.3	48.3
Female	32.8	51.7	51.7
Age			
18-44	11.3	48.2	48.2
45-54	15.6	19.4	19.4
55-64	23.7	15.5	15.5
65+	49.5	16.9	16.9
Education			
<HS Grad/HS Grad	32.2	44.8	44.8
Some Col/Assoc Deg	29.7	27.9	27.9
Bachelors Deg	21.7	18.0	18.0
Graduate Deg	16.4	9.3	9.3
Region			
Northeast	15.4	18.5	18.5
Midwest	24.3	22.0	22.0
South	35.8	36.9	36.9
West	24.4	22.6	22.6
Race/Ethnicity			
White Non-Hispanic	84.6	68.3	68.3
Black Non-Hispanic	5.5	11.5	11.5
Hispanic	3.5	14.0	14.0
Other Non-Hispanic	6.3	6.2	6.2

Note: Practicum estimates are calculated on 900 cases.

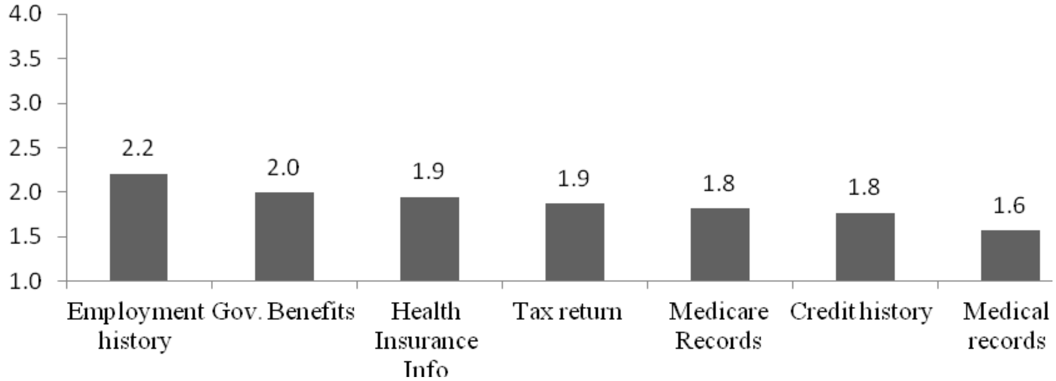
3.5 Research Design and Hypotheses

3.6 Research Question 1: The Effect of Type of Record for Which Consent to Link is Requested

To examine whether the content and topic of the administrative records requested affected respondents' willingness to consent, equal numbers of respondents were randomly assigned to one of two consent requests. Respondents were asked for permission to link their survey responses with either their health-related records or their income and employment-related records. These requests came at the same point in the survey, after questions on health, employment, and income.

Figure 3.01

Respondent Willingness to Provide Access to Administrative Records from the 2010 JPSM Practicum Survey[54]



Findings from hypothetical consent scenarios included in the previous year's 2010 JPSM Practicum survey motivated the hypothesis for which request respondents would be more willing to consent. Respondents to that survey rated their willingness to grant the Census Bureau access to seven different types of personal and administrative records (see Figure 3.01). Albeit hypothetically, respondents indicated that they would be less likely to grant access to their medical records as compared to income and employment-related records such as their tax return, employment history, or program benefit history ($p<0.05$). In a separate question in the same survey, respondents rated their medical records as somewhat more personal than their credit history or tax return (3.3 vs. 3.1 and

---

[54] How likely would you be to give your consent to the Census Bureau to obtain each of the following types of records? Use a scale where 1 is 'very unlikely' and 4 is 'very likely.' (Sample size is approximately 1200-1250 respondents per item; only respondents ages 65 and older were asked about the Medicare record item (n=272).)

3.1 respectively, on a 4-point scale).[55] These results indicate that respondents view health-related records as more personal and sensitive than income and employment-related records. We therefore hypothesize that consent rates to the health request will be lower than consent rates to the income and employment request.

In formulating this hypothesis, we also consider results from surveys that request consent to both health and income-related administrative records from the same respondents within the same survey administration. This includes the BHPS which requests consent to both of these types of records from respondents aged 16 and older (Sala et al., 2010). [56] In the most recent wave of this survey for which data is available (Wave 18, 2008-2009), more BHPS respondents consented to the request for health records (41%) than economic records (32%), counter to what would be suggested by the 2010 JPSM Practicum findings. This implies that at least some BHPS respondents consider their economic records to be more sensitive than their health records.

However, differences in policy surrounding the access and use of administrative records abroad (Baker et al., 2000) as well as potential differences in respondents' data privacy concerns limit the generalizability of the BHPS results to the U.S. In the U.K., respondents' informed consent must be obtained in order to link individual-level administrative record data with survey responses (Sala et al., 2010). Further, for respondents in the U.K., written consent is necessary to acquire health records (Tate et al., 2005).

---

[55] I am going to read you a list of different types of records you might have. For each one of them, please tell me how personal you think it is. Use a scale where 1 is 'not at all personal' and 4 is 'very personal.'

[56] The HRS also requests consent to access both health and income and employment-related administrative records from respondents, however, the sample of respondents from which these records are requested is not comparable: Medicare information is requested primarily from those over the age of 65, and Social Security information is requested from all sample members.

Both studies reviewed have caveats. Although the reliance on hypothetical scenarios is a limiting factor of the findings from the 2010 JPSM Practicum survey, this research was conducted on a nationwide sample of respondents within the U.S., and provides multiple data points that express respondents' hesitancy to grant access to health-related records. Findings from the BHPS were not produced under research conditions comparable or generalizable to those in the U.S. Thus, even considering the BHPS findings, it is hypothesized that rates of consent to the health record request in the current survey will be lower than the consent request for income and employment records as suggested by the 2010 JPSM Practicum survey.

We also evaluate the effect of the $5 prepaid cash incentive on consent. Although the incentive was randomly assigned to sample members, this assignment was independent of assignment to consent conditions. To date, no research has examined the influence of cash incentives on consent to administrative record linkage. It is hypothesized that the incentive will counteract what respondents perceive as the more negative features of the consent request, and increase consent rates, analogous to its influence on survey response (Groves et al., 2000).[57]

3.6.1 Consent Request Variation: Analyses

To check the success of the randomization to the two consent conditions, the sample balance across conditions was compared across a variety of survey items using unweighted data. Consent rates for the full sample and for each of the two consent request conditions were then calculated using weighted and unweighted data.

---

[57] An alternative framework in considering the effect of incentives on consent is described in research by Singer and Couper (2008) and related biomedical research which finds that study participants are not willing to accept greater risk in exchange for a larger monetary incentive.

Bivariate analyses were used to describe demographic and socioeconomic characteristics of consenting respondents. Odds ratios were calculated to express the effect of consent request type on consent rate. Multivariate logistic regression was used to examine the presence of a significant main effect of consent request type, and of the effect of the incentive on consent, while controlling for socioeconomic and demographic characteristics. Wald Tests evaluated the contribution of categorical predictors in the logistic regression model and Archer & Lemeshow's Goodness of Fit test for Survey Data (Archer & Lemeshow, 2006) assessed model fit. All of the following analyses in this chapter were estimated using Stata 12 software and examined on both weighted and unweighted data, and we note any discrepancies between the weighted and unweighted findings. All unweighted results are in the Appendix. Jackknife replicate weights were incorporated using the Stata SVY procedure.

Slightly more respondents who completed the interview were assigned to the income and employment consent request (n=455) than to the health consent request (n=445).[58] Overall, characteristics of respondents assigned to each of the two requests are generally equivalent; the random assignment to each of the two consent request conditions appears to be successful with regards to the variables evaluated in Table 3.03. Chi-square tests conducted on unweighted data indicate that none of the differences reaches significance at the 0.10 level. Although it is possible that variation in consent rates can stem from variation in characteristics beyond those evaluated in this table, considering the success of the random assignment on the variables evaluated, it is

---

[58] An analysis of respondents who did not complete the interview shows that 8 respondents broke off during the consent request. This includes 3 respondents assigned to the health request and 5 respondents assigned to the income and employment request.

reasonable to attribute any differences in consent rates to the experimental treatment effects.

Table 3.03

Demographic Characteristics by Consent Request Condition

	Consent Request	
	Health (n=445)	Income/Employment (n=455)
	% (n)	% (n)
Male	67.2 (299)	67.3 (306)
18-44	11.5 (51)	11.0 (50)
45-54	14.4 (64)	16.5 (75)
55-64	24.5 (109)	23.5 (107)
65+	49.7 (221)	49.0 (223)
<HS Grad/HS Grad	33.0 (147)	31.2 (142)
Some Col/Assoc Deg	28.1 (125)	31.7 (144)
Bachelors Deg	23.2 (103)	20.0 (91)
Graduate Deg	15.7 (70)	17.1 (78)
White Non-Hispanic[59]	86.3 (384)	83.5 (380)
Black Non-Hispanic	4.7 (21)	6.2 (28)
Hispanic	3.2 (14)	3.7 (17)
Other Non-Hispanic	5.8 (26)	6.6 (30)
<25K	19.1 (85)	17.6 (80)
25- <75K	36.9 (164)	38.9 (177)
>75K	26.1 (116)	26.6(121)
Income DK[60]	6.3 (28)	4.6 (21)
Income REF	11.7 (52)	12.3 (56)
Incentive	61.6 (274)	64.2 (292)

Note: Unweighted estimates.

---

[59] Due to the small number of Non-White respondents and because respondent race and ethnicity distinctions are not central to this analysis, respondents are categorized as White and Non-White for the remainder or this chapter.

[60] Interviewers entered several responses to the income question incorrectly; this includes 16 “don’t know” and 25 “refused” responses. These responses are appropriately coded in the dataset, however the 25 respondents who initially refused were not directed to the bracketed income questions like others who initially refused the income question. See question Q28A in questionnaire (Appendix) for details.

3.6.2 Consent Request Variation: Results

3.6.2.1 Consent Rates

Across the two requests, 269 of the 900 respondents consented (see Table 3.04), resulting in an unweighted consent rate of 30% and a weighted consent rate of 33%. The unweighted consent rate for the health request was 34%[61] and the weighted consent rate was 31%. For respondents assigned to the income and employment consent request, the unweighted consent rate is 26% and the weighted consent rate is 36%. Thus, when examining the weighted data, the direction of the difference reverses, suggesting that more respondents consented to the income and employment consent request. Given this reversal, we continue to examine both the weighted and unweighted findings in the remainder of the chapter and note any discrepancies.

Substantial weighting due to the sample composition contributed to this reversal of direction in consent rates. Although this reversal is likely due to cumulative effects of weighting across respondents rather than one specific subgroup, the youngest respondents in the sample, those between the ages of 18 and 44, appear to have specifically contributed to the consent rate reversal. Compared to the average weight of all respondents (252,105), the youngest respondents were assigned, on average, a weight of 2,101,961 due to low response among this age group. When excluding the youngest respondents from consent rate calculations, the weighted consent rates are in the same direction as the unweighted consent rates (health consent rate: 37%; income and employment consent rate: 29%).

---

[61] Three respondents assigned to the income and employment consent request and 5 respondents assigned to the health consent request responded “don’t know” to the consent request. These respondents are considered consent refusals in all analyses.

Table 3.04

Consent Rates by Consent Request Condition

	Consent Request					
	Overall (n=900)		Health (n=445)		Income/Employment (n=455)	
	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	%	%	%	%	%	%
Consent	33.4	29.9	30.9	33.5	35.9	26.4
Nonconsent	66.6	70.1	69.1	66.5	64.1	73.6

Note: Weighted estimates.

3.6.2.2 Characteristics of Consenting Respondents

Bivariate analyses were used to describe characteristics of consenting respondents overall, and respondents by consent request type (see Table 3.05), and we conducted Chi-square tests of independence for respondent characteristics and consent status. As discussed in Chapter 1, existing literature on consent to record linkage suggests that in general, most respondent demographic and socioeconomic characteristics are not consistently predictive of their decision to consent, with the same characteristic positively associated with consent in one study and negatively associated in another (Bohensky et al., 2010; Sala et al., 2010). Considering the existing literature, we briefly comment on the results from these bivariate analyses.

Overall, gender is not related to consent, but rates of consent to the health request were significantly greater among women compared to men ($p<0.10$). This finding contrasts with past research examining consent rates to health-related record linkage which generally finds that more males consent (Dunn et al., 2004; Koenig, 2003; Woolf et al., 2000), or that no significant differences in consent rates by gender exist (Baker et al., 2000; Harris et al., 2005; Huang et al., 2007).

Table 3.05

Percent of All Respondents Who Consent to Linkage, and Percent Consenting to Each Request, by Demographic Category

		Consent Request	
	All Respondents (n=900)	Health (n=445)	Income/ Employment (n=455)
	%	%	%
Male *(605)*	29.2	23.5	35.0
Female *(295)*	37.4	38.0	36.7
18-44 *(101)*	36.4	26.5	**47.2**
45-54 *(139)*	21.0	16.2	**24.4**
55-64 *(216)*	34.3	45.1	**22.0**
65+ *(444)*	38.2	43.4	**32.4**
<HS Grad/HS Grad (289)	40.8	**40.0**	41.7
Some Col/Assoc Deg (269)	30.6	**17.5**	40.1
Bachelors Deg (194)	31.1	**35.0**	26.6
Graduate Deg (148)	10.8	**6.0**	14.6
White (764)	32.0	**47.7**	**17.6**
Non-White (136)	34.1	**23.7**	**45.0**
<25K (165)	*40.6*	**49.6**	*29.4*
25- <75K (341)	*41.8*	**28.9**	*50.4*
>75K (237)	*21.2*	**25.6**	*16.0*
Income DK (49)	*46.0*	**43.9**	*48.7*
Income REF (108)	*5.2*	**0.8**[62]	*13.5*
No Incentive (566)	30.3	29.5	31.4
Incentive (334)	35.1	31.9	37.8

Notes: Weighted estimates. *Italics* indicates differences by demographic category $\chi 2$ p<0.01; **bold** indicate differences by demographic category $\chi 2$ p<0.05; underline indicates differences by demographic category $\chi 2$ p<0.10

Consent status does not vary by age overall, but consent rates to the income and employment request were greater among younger respondents; consent rates to the health

---

[62] Of respondents who refused the income question, 2 out of 52 assigned to the health consent request granted consent compared to 7 out of 56 assigned to the income and employment consent request.

request were greater among older respondents. The overall consent rate is not related to race, but race is related to each of the separate requests.

The Chi-square test of independence of education level and consent status indicated that consent also varies by educational attainment ($p<0.10$), with more educated respondents generally less likely to grant consent overall and for both request types. Existing research on consent by educational attainment is largely mixed, with more educated respondents consenting more often in some research (Cleary et al., 1981; Huang et al., 2007), and less educated respondents consenting more often in other research (Jay et al., 1994).

Consent varies as a function of income and willingness to provide a substantive response to the income question ($p<0.01$). Refusal to the consent request is associated with higher income, and for respondents assigned to the income and employment request, refusal is associated with lower earnings. Respondents who refused the income question are highly unlikely to consent: this result is consistent with past findings (Jenkins et al., 2006; Olson, 1999; Sala et al., 2010; Woolf, 2000). Refusal to the income question is sometimes used as a proxy for privacy concerns (Hurd et al., 2003; Juster et al., 1997) and so refusing both requests may stem from similar objections. This will be examined in more depth in the next section of this chapter.

The bivariate analyses do not demonstrate a significant effect of the incentive on consent rates. Of respondents who received an incentive, 35% consented, compared with 30% of those who did not receive an incentive (n.s.). The incentive appears to have a smaller effect on the health consent request compared to the income and employment consent request: 32% of respondents who received an incentive consented to the health consent request as compared to 38% of respondents assigned to the income and employment consent request condition, though again, these differences are not statistically significant.

In examining the unweighted estimates analogous to those presented in Table 3.05, we find a small number of additional differences in characteristics of consenting respondents and by request type not observed in the weighted data. In the unweighted data, consent rates to the income and employment request are significantly lower among men as compared to women (24% vs. 32%; $p<0.10$), yet we see no difference in consent rates to the health request by gender as we did in the weighted data.

Also in the unweighted data, consent rates vary as a function of age overall ($p<0.01$); a finding not observed in the weighted data. A final difference not evident in the weighted data is that consent rates are significantly greater among respondent who were mailed an incentive. This is true in the overall sample (32% vs. 26%, $p<0.05$) as well as for the health consent request (38% vs. 26%, $p<0.01$).

We next used logistic regression to examine the main effect of consent request type on consent rates. When accounting for sample composition and nonresponse, the unadjusted weighted odds ratio is 0.80 CI [0.42, 1.51] (the corresponding log odds ratio is -0.23). This suggests that the odds of consenting to the health record consent request are about 20% lower than the income and employment consent request, though this difference does not achieve statistical significance.

The direction of this effect is reversed when the analysis is calculated without weights: the unweighted and unadjusted odds ratio evaluating the effect of consent request type on consent rates is 1.41, CI [1.06, 1.87]. (The corresponding log odds ratio is 0.34.)

Next we examine the effect of consent request type after controlling for demographic and socioeconomic predictors.

3.6.2.3 Logistic Regression Analyses

The primary goal of this initial research question was to determine if the content and topic of the records mentioned in the consent request influenced respondents' decision to grant consent. The unadjusted logistic model did not identify a significant effect of consent request type; however, we next examine the effect of this predictor after controlling for other potential sources of variation examined in the bivariate analyses. The baseline multivariate logistic regression models consent and incorporates predictors for consent request type, demographic and socioeconomic variables, and an indicator for whether or not a respondent was mailed a prepaid incentive. We then conduct separate regression analyses for the health and income and employment consent conditions to evaluate whether predictors of consent vary as a function of consent request type.

As illustrated in Table 3.06, the coefficient for the predictor of interest, consent request type, demonstrates that the log-odds of consenting are lower for the health request condition as compared to the income and employment condition – consistent with the unadjusted model – even after accounting for respondents' characteristics and whether or not they received an incentive, though this effect is not statistically significant. This suggests that in the Practicum survey, respondents' decision to consent to record linkage was largely unaffected by the topic and content of records requested after controlling for other factors.

In the model that includes all respondents, age, education, and income are negatively associated with consent (see Table 3.07). Gender and race are not related to consent, nor is the incentive.

In the unweighted baseline logistic regression model (Appendix) we find that, like the unweighted and unadjusted logistic regression, there is a significant positive effect of the health consent request on consent rates ($p<0.05$) when controlling for other factors. Other differences between the weighted and unweighted baseline models include a significant positive effect of the incentive on consent rates ($p<0.05$).

Table 3.06

Consent Request Variation Baseline Logistic Regression Model Predicting Consent to Record Linkage

		Coef.	SE	p-value
Constant		0.04	0.64	0.954
Consent Request	Income/Employment (ref.)			
	Health	-0.22	0.32	0.498
Gender	Female (ref.)			
	Male	-0.54	0.34	0.113
Age	18-44 (ref.)			
	45-54	-1.15	0.40	0.007
	55-64	-0.17	0.46	0.712
	65+	-0.35	0.43	0.411
Education	<HS Grad/HS Grad (ref.)			
	Some Col/Assoc Deg	-0.73	0.39	0.070
	Bachelors Deg	-0.36	0.51	0.478
	Graduate Deg	-1.82	0.65	0.007
Race	Non-White (ref.)			
	White	0.48	0.36	0.191
Income	<25K (ref.)			
	25- <75K	0.12	0.48	0.796
	>75K	-0.64	0.64	0.321
	Income DK	0.18	0.66	0.791
	Income REF	-2.52	1.06	0.021
Incentive	No Incentive (ref.)			

Incentive	0.30	0.40	0.458

Notes: Weighted estimates. Model based on 900 cases; Archer and Lemeshow Goodness of Fit test for survey data suggests a slight lack of fit [F(9,41)=1.80; Prob>F=0.097].

Table 3.07

Wald Tests for Categorical Parameters in Baseline Logistic Regression Model

Categorical Predictor	F-Test Statistic	p-value
Age	$F_{(3, 47)} = 2.96$	0.042
Education	$F_{(3, 47)} = 2.84$	0.048
Income	$F_{(4, 46)} = 2.22$	0.082

Table 3.08

Consent Request Variation Baseline Logistic Regression Model Predicting Consent to Record Linkage, by Request Type

		Health (n=445)			Income and Employment (n=455)		
		Coef.	SE	p-value	Coef.	SE	p-value
Constant		-0.90	1.19	0.454	-1.16	1.08	0.287
Gender	Female (ref.)						
	Male	0.64	0.54	0.236	0.19	0.42	0.644
Age	18-44 (ref.)						
	45-54	-0.78	0.82	0.344	-1.51	0.56	0.009
	55-64	0.70	0.68	0.306	-1.24	0.63	0.053
	65+	0.35	1.00	0.725	-1.12	0.53	0.041
Education	<HS Grad/HS Grad (ref.)						
	Some Col/Assoc Deg	-1.20	0.57	0.042	-0.33	0.40	0.417
	Bachelors Deg	-0.41	0.85	0.627	-0.52	0.83	0.539
	Graduate Deg	-2.46	1.08	0.027	-1.18	1.01	0.252
Race	Non-White (ref.)						
	White	-0.65	0.71	0.362	1.60	0.76	0.041
Income	<25K (ref.)						
	25- <75K	-0.35	0.76	0.649	0.62	0.57	0.282
	>75K	-0.09	0.85	0.916	-0.71	0.88	0.426
	Income DK	-0.11	1.39	0.937	1.14	0.97	0.246
	Income REF	-4.25	2.07	0.045	-0.81	1.18	0.495
Incentive	No Incentive (ref.)						
	Incentive	0.39	0.69	0.577	-0.11	0.52	0.831

Notes: Weighted estimates. Archer and Lemeshow Goodness of Fit test for survey data suggests a lack of fit for both models: Health [F(9,41)=3.80; Prob>F=0.002]; Income/Employment [F(9,41)=9.21; Prob>F=0.000].

Table 3.09

Wald Tests for Categorical Parameters in Baseline Logistic Regression Model Predicting Consent to Record Linkage, by Request Type

	Health		Income and Employment	
Categorical Predictor	F-Test Statistic	p-value	F-Test Statistic	p-value
Age	$F_{(3, 47)} = 0.93$	0.435	$F_{(3, 47)} = 2.95$	0.042
Education	$F_{(3, 47)} = 2.58$	0.064	$F_{(3, 47)} = 0.51$	0.680
Income	$F_{(4, 46)} = 1.03$	0.403	$F_{(4, 46)} = 2.11$	0.094

We next examine the baseline multivariate logistic regression model by request type, and find that predictors of consent somewhat vary depending on the records requested (see Table 3.08). We note that the smaller number of significant predictors of consent in these models may be related to the reduction in sample size compared to the full baseline model presented in Table 3.06.

Here, age is associated with consent for the income and employment request, with older respondents significantly less likely to consent than younger respondents, but is not associated with the health request. Design-adjusted Wald tests justify the inclusion of the categorical predictors for age in the income and employment model only ($p<0.05$; see Table 3.09).

Respondents with a graduate degree or some college are significantly less likely to consent to the health request compared to those with a high school or less than a high school education, however, there is no association between education and consent to the income and employment consent request. The categorical predictor for education is significant only in the health model ($p<0.10$)

There is no effect of race on consent to the health request, but white respondents are significantly more likely to consent to the income and employment request ($p<0.05$). Finally, refusal to the income question is negatively associated with consenting to the health request ($p<0.05$), but not associated with the income and employment request. The categorical predictors for income is significant only in the income and employment model ($p<0.10$).

We also compare findings from the unweighted models and find some minor differences in predictors of consent. In contrast to the weighted data, the oldest respondents (65+) are more likely to consent to the health request in the unweighted data ($p<0.10$), and in addition, age is no longer a significant predictor of consent to the income and employment request. Income is significantly associated with both requests in the unweighted analyses, with refusal to provide a substantive income response significantly associated with consent refusal ($p<0.01$).

Finally, the incentive has a positive effect on consent to the health request ($p<0.05$). It appears that the incentive had a greater effect among older individuals assigned to this request. Forty-six percent of respondents ages 65 and older who received an incentive consented to the health request, compared to 33% of respondents in this age group who did not receive an incentive. As a comparison, 31% of respondents in this age group who received an incentive consented to the income and employment request. The overrepresentation of older respondents in the sample may have influenced the positive incentive effect.

3.6.3 Consent Request Variation: Conclusions

After controlling for respondents' demographic and socioeconomic characteristics, we find that there is no significant difference between respondents' willingness to consent to link their health or income and employment records in the weighted analyses. Several explanations could account for the lack of significant differences between the two consent request conditions. First, the lack of prominence of the consent request in the survey may have contributed to these null findings. It is typical for surveys requesting consent to record linkage to ask a series of questions relating to consent, as in MEPS, RECS, and NIS, or mention the consent request in an advance letter, as in HRS and surveys sponsored by the Census Bureau. With greater time and attention given to the consent request in the survey and survey materials, respondents would have had more time to process this request and potentially have responded with a more carefully thought out response.

A second reason lies within the design of the Practicum questionnaire. The diverse content of the survey, lack of a well-known, non-government sponsor, and reliance on telephone administration may have compromised the legitimacy of the survey. This is suggested by the overall consent rate which is lower than in most surveys. If the design of this survey increased its legitimacy -- through government sponsorship or in-person administration, for example -- rates of consent to one request could have been disproportionately affected. If respondents considered one request to be slightly more sensitive than the other, increased legitimacy could positively benefit rates of consent to this request. Or, with government rather than academic sponsorship, respondents could be more willing to grant access to records that appear to be government-related. For example, respondents might be more likely to associate information requested in the income and employment request with government records than the health consent request.

In addition, with such a low response rate (15.7%) it is possible that only the most compliant sample members participated in our survey. Further, those willing to consent may not be concerned with whether the records to which they were granting access were related to health or income and employment.

The analyses do identify that certain demographic and socioeconomic characteristics are significantly associated with consenting overall in the bivariate analyses; this primarily includes education and responses to the income question. A greater number of differences appear when examining these characteristics by consent request type.

In the baseline logistic regression models aggregated across request type, we find that age, education, and income predict consent, but that predictors of consent vary somewhat depending on the records requested. Age, race, and overall income are significantly associated with consent to the income and employment request, and education and income refusal are significantly associated with consent to the health request. The effect of the incentive was not statistically significant in any of the models.

It is important to consider that the sample distribution was quite skewed from overall population estimates and required substantial weighting. This contributed a large design effect and sizable variance to all of the analyses included in this chapter and could have masked actual significant findings. To begin to investigate the extent to which weighting affected estimates, we compared the weighted results presented in this chapter with unweighted findings, though we note that the unweighted estimates are not representative of the target population.

The most notable difference in the unweighted analyses is the reversal and significance of the main effect of consent request type. In the unweighted data, respondents assigned to the health request were more likely to consent than those assigned to the income and employment request. We observe this in the bivariate analyses as well as the multivariate analyses that control for other respondent characteristics. Older respondents consented at higher rates to the health request, and were overrepresented in the unweighted analyses, potentially influencing these findings. We also note that, in the unweighted analyses, the incentive had a significant effect on consent rates in the overall sample and on consent rates for the health request specifically.

3.7 Research Question 2: The Effect of Privacy, Confidentiality, and Trust Attitudes

We next evaluated the relationship between respondents' privacy, confidentiality, and trust attitudes and their willingness to consent through a battery of items measuring these constructs. We hypothesized that respondents with greater privacy, confidentiality, and trust concerns would be less willing to consent to linkage. Some existing research demonstrates this. Respondents who refuse to answer a survey question on income, a proxy for privacy concerns (Hurd et al., 2003; Juster et al., 1997), were less likely to consent to record linkage (Jenkins et al., 2006; Olson, 1999; Sala et al., 2010; Woolf, 2000), as were those concerned with the confidentiality of their records (Armstrong et al., 2008), or who believed their linked data could be used to detect fraud (Grey, 2008). Socially detached (Cleary et al., 1984; Sala et al., 2010) and less trusting individuals also consented at lower rates (Cleary et al., 1984).

Privacy, confidentiality, and trust were each measured by three items initially developed for use in other surveys. Some of these items were modified from their original wording to more appropriately address the present research questions (see Table

3.10). The privacy items were used previously in the Census Surveys of Privacy Attitudes, the Census Survey of Participation, as well as a survey by Gallup.

The three confidentiality items were based on ones from the Census Surveys of Privacy Attitudes. Two of these items were modified from an item asking respondents how bothered they would be if their Census responses were not kept confidential. Instead of Census responses, the Practicum items asked respondents how bothered they would be if their medical and income tax records were not kept confidential. The third item -- unmodified -- asked whether people have lost all control over how personal information about them is used.

The three trust items are from the GSS. They ask respondents whether others can be trusted, try to be helpful, and try to be fair.

All the privacy, confidentiality, and trust items were asked after the consent request, separated by a series of unrelated items. Thus it is possible the answers to the privacy, confidentiality and trust items could have been affected by either the consent request or the income questions that directly preceded the consent request. But asking the privacy, confidentiality and trust items after the consent request seemed preferable to asking them before the consent request, where they might have affected the response to the consent request.

Table 3.10

Privacy, Confidentiality, and Trust Items Included in the Practicum Survey and Sources

Question	**Response**	**Source**	**Original Wording**	**Measures**	**Q #**
Every ten years, including 2010, most households are sent a Census questionnaire that includes a few questions about everyone living there. Would you agree or disagree that the Census is an invasion of privacy?	Agree/Disagree	1993: Survey of Census Participation 2010: Census Surveys of Privacy Attitudes	1993: The Census is an invasion of privacy. (Agree/Disagree) 2010: Do you feel it is an invasion of privacy for the Census Bureau to ask these questions? (Yes/No)	Privacy	46
Do you think the government bothers you too much with requests for information?	Yes/No	Survey of Census Participation		Privacy	47
How much would it bother you if your medical records were not kept confidential? Would it bother you a lot, some, a little, or not at all?	A lot/Some/A little/Not at all	Census Surveys of Privacy Attitudes	How much would it bother you if your answers to the Census were not kept confidential? Would it bother you a lot, some, a little, or not at all?	Confidentiality	48/49 (random -ized)
How much would it bother you if your income tax records were not kept confidential? Would it bother you a lot, some, a little, or not at all?					
Please tell me if you strongly agree, somewhat agree, somewhat disagree, or strongly disagree: People have lost all control over how personal information about them is used.	Strongly agree/ Somewhat agree/ Somewhat disagree/Strongly disagree	Census Surveys of Privacy Attitudes	Please tell me if you strongly agree… People have lost all control over how personal information about them is used."	Confidentiality	50
Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?	Most people can be trusted/ You can't be too careful	GSS		Trust	51
Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?	Try to be helpful/ Just look out for themselves	GSS		Trust	52
Do you think most people would try to take advantage of you if they got a chance, or that they would try to be fair?	Try to take advantage of you/ Try to be fair	GSS		Trust	53
How often do you worry about being a victim of identity theft – frequently, occasionally, rarely, or never?	Frequently/ Occasionally/ Rarely/ Never	Gallup		Privacy	54

3.7.1 Privacy, Confidentiality, and Trust Attitudes: Analyses

Descriptive and bivariate analyses for the privacy, confidentiality, and trust items were first produced for all respondents, and compared with the original and modified versions of the existing items (for example, from the Census Surveys of Privacy Attitudes). Although substantial differences between the Practicum survey and the earlier surveys complicate the inference, these comparisons help determine if respondents' concerns were heightened by the consent request or the income questions preceding it. Responses to the privacy, confidentiality, and trust items are then related to the answers to the consent request by consent condition.

3.7.2 Privacy, Confidentiality, and Trust Attitudes: Results

3.7.2.1 Benchmark Comparisons

As presented in Table 3.11, results from the first privacy item from the Census Surveys of Privacy attitudes (whether the respondent agrees or disagrees that the Census is an invasion of privacy) suggests less privacy concerns in the Practicum survey than the most recent administration of the Census Survey of Privacy Attitudes (2010). However, and as with the other comparisons to the items in this battery, this may not represent true differences in attitudes across the surveys. The Census Surveys of Privacy Attitudes differs from the current survey in many ways that could account for this discrepancy, for example, in sponsor, year, topic, and sample. There is also a difference in survey context:

this item, taken out of its original context in a Census and government-centric survey, may elicit less privacy concern among respondents.[63]

Table 3.11

Privacy, Confidentiality, and Trust Attitudes and Benchmark Comparisons

	All Respondents (n=900)	External Benchmarks
	%	%
Privacy		
Census invasion of privacy *(Agree)*	20.6	31
Gov bothers with requests *(Yes)*	24.3	--[64]
Worry about ID theft		
Frequently	32.3	31
Occasionally	34.6	35
Rarely	22.0	18
Never	11.1	15
Confidentiality		
Medical records not confidential		
Bothered a lot	72.8	53.0
Bothered some	10.8	19.2
Bothered a little	7.5	9.1
Not bothered at all	8.9	18.7
Tax records not confidential		
Bothered a lot	73.2	53.0
Bothered some	9.3	19.2
Bothered a little	6.8	9.1
Not bothered at all	10.7	18.7
Control over personal information		
Strongly agree	40.0	41.4
Somewhat agree	41.6	36.2
Somewhat disagree	10.7	15.2
Strongly disagree	7.7	7.0
Trust		
Trusted/Careful		
Can be trusted	27.0	47.1
Can't be too careful	73.0	52.9[65]
Helpful/Look out for themselves		
Try to be helpful	53.1	51.2

---

[63] The external comparison for the second privacy item, "Do you think the government bothers you too much with requests for information?" is not available.

[64] The external benchmark for this item is not available.

[65] Data from the 2008 GSS is the most recent data available. Compared to the GSS, some Practicum response options were combined ("always trusted" and "usually trusted"; and "usually not trusted" and "always not trusted".)

	Look out for themselves	46.9	48.8[66]
Take Advantage/Fair			
	Take advantage	41.9	43.1[67]
	Try to be fair	58.1	56.9

Note: Weighted estimates.

The third privacy item asks respondents how often they worry about identity theft. Compared to its original administration in a 2009 Gallup survey, respondents are similarly concerned about their identity theft. Although not directly comparable, results to the privacy items suggest that the consent request administered in the Practicum survey only minutes earlier did not exacerbate respondents' privacy concerns.

Two confidentiality items asked how bothered respondents would be if their medical records and income tax records were not kept confidential. These items were asked of all respondents in the sample, regardless of consent request condition. The corresponding "benchmark" item asks respondents about Census responses rather than income tax or medical records. It is reasonable to assume that respondents would consider Census responses less sensitive than medical or income tax records, and responses to the benchmark item are accordingly less extreme. In the 2010 Census Survey of Privacy Attitudes, slightly more than half of the respondents indicated that they would be "bothered a lot" if their Census responses were not kept confidential, compared to almost three-quarters of JPSM respondents who chose that category when asked about both their medical records and their income tax records.

A third item is unmodified from its use in the 2010 Census Survey of Privacy Attitudes and asks respondents how strongly they agree with the statement, "People have lost all control over how personal information about them is used." The distribution of

---

[66] Data from the 2010 GSS is the most recent data available. In 2010, 10% said "it depends". This response distribution was estimated without this 10% for comparability.

[67] Data from the 2010 GSS is the most recent data available. In 2010, 8.6% said "it depends". This response distribution was estimated without this 8.6% for comparability.

Practicum responses to this item is fairly similar to the 2010 Census Survey of Privacy Attitudes.

The three items in the Practicum survey measuring trust asked whether or not respondents think others are generally trustworthy, helpful, and fair. These items have been administered in multiple iterations of the GSS. Compared to the 2010 GSS, similar percentages said that others try to be helpful (53%) and fair (58%).

However, the item assessing trustworthiness deviates somewhat from the most recent GSS administration (2008). In the Practicum survey, 27% of respondents said that most people can be trusted, compared with 47% of GSS respondents. This skew may be explained by a context effect. The trust item in the Practicum survey came after a series of quite sensitive questions concerning privacy and confidentiality. This may have caused respondents to feel particularly vulnerable upon responding.

The unweighted bivariate analyses are overall very similar to the estimates presented in Table 3.11, with the only differences related to the trust items. In the unweighted data, a greater proportion of respondents said that others try to be helpful (62%) and a smaller proportion said that they try to take advantage (34%).

3.7.2.2 Bivariate Analyses: Privacy Attitudes

As shown in Table 3.12, none of the items measuring privacy are related to giving or refusing consent in the total sample. Although the item assessing identity theft concern demonstrates slightly higher consent rates among respondents who say they “rarely” or “never” worry about identity theft, the distribution of responses to this question by consent status is not significant.

Table 3.12

Privacy, Confidentiality, and Trust Attitudes by Consent Status

	Consent (n=269)	Non-Consent (n=631)
	%	%
Privacy		
Census invasion of privacy		
Agree	31.4	68.6
Disagree	33.5	66.5
Gov bothers with requests		
Yes	34.5	65.5
No	33.1	66.9
Worry about ID theft		
Frequently	23.4	76.6
Occasionally	34.1	65.9
Rarely	41.2	58.8
Never	45.5	54.5
Confidentiality		
Medical records not confidential		
Bothered a lot	*23.9*	*76.1*
Bothered some	*39.4*	*60.6*
Bothered a little	*65.8*	*34.2*
Not bothered at all	*77.0*	*23.0*
Tax records not confidential		
Bothered a lot	*26.1*	*73.9*
Bothered some	*32.1*	*67.9*
Bothered a little	*54.1*	*45.9*
Not bothered at all	*71.6*	*28.4*
Control over personal information		
Strongly agree	28.3	71.7
Somewhat agree	36.5	63.5
Somewhat disagree	35.9	64.1
Strongly disagree	41.3	58.7
Trust		
Trusted/Careful		
Can be trusted	31.5	68.5
Can't be too careful	34.3	65.7
Helpful/Look out for themselves		
Try to be helpful	33.8	66.2
Look out for themselves	34.1	65.9
Take Advantage/Fair		
Take advantage	36.9	63.1
Try to be fair	30.2	69.8

Note: Weighted estimates. *Italics* indicate differences by response option $\chi 2$ p<0.01

Analyzing the two additional privacy items from the Census Surveys of Privacy Attitudes by consent request type does not uncover any significant differences (see Table 3.13). Respondents in the health request condition vary minimally in their responses to the identity theft item by consent status. However, worrying “frequently” about identity theft is more prohibitive to consenting to the income and employment consent request: 85% of these respondents withheld consent compared to 53% of those who “never” worry. The difference in responses to this item by consent status is statistically significant ($p<0.10$).

We observe two differences in the privacy items in the unweighted data among all respondents. First, a smaller proportion of respondents who said that the government bothers them with information requests consented to either request compared to the weighted data (25% vs. 35%; significant by consent status in the unweighted data ($p<0.10$) but not the weighted data). In addition, worrying more frequently about identity theft is associated with lower consent rates ($p<0.01$). Also in the unweighted data, fewer respondents who say that the government bothers them with information requests consented to the income and employment request compared to the weighted data (20% vs. 33%).

3.7.2.3 Bivariate Analyses: Confidentiality Attitudes

Table 3.12 presents results for the confidentiality items by consent status for all respondents. About three-quarters of respondents who selected the “bothered a lot” response option to the medical and income tax record items refused consent. Conversely, similar percentages of respondents who selected the “not bothered at all option” to these items consented; the distribution of responses for both items are significant ($p<0.01$).

Respondents did not vary significantly in consent status by their responses to the final confidentiality item, pertaining to whether people have lost control over the use of their personal information.

Table 3.13

Privacy, Confidentiality, and Trust Attitudes by Consent Status and Consent Request Condition

	Health		Income/Employment	
	Consent (n=149)	Non-Consent (n=296)	Consent (n=120)	Non-Consent (n=335)
	%	%	%	%
Privacy				
Census invasion of privacy				
Agree	39.4	60.6	23.7	76.3
Disagree	28.8	71.2	38.4	61.6
Gov bothers with requests				
Yes	36.1	63.9	32.7	67.3
No	29.2	70.8	36.9	63.1
Worry about ID theft				
Frequently	32.1	67.9	15.3	84.7
Occasionally	31.0	69.0	38.2	61.8
Rarely	20.9	79.1	55.9	44.1
Never	44.3	55.7	46.9	53.1
Confidentiality				
Medical records not confidential				
Bothered A lot	*18.3*	*81.7*	**29.5**	**70.5**
Bothered some	*39.7*	*60.3*	**39.1**	**60.9**
Bothered a little	*74.6*	*25.4*	**56.7**	**43.3**
Not bothered at all	*82.6*	*17.4*	**70.2**	**29.8**
Tax records not confidential				
Bothered a lot	**27.7**	**72.3**	*24.2*	*75.6*
Bothered some	**17.9**	**82.1**	*52.0*	*48.0*
Bothered a little	**33.5**	**66.5**	*65.7*	*34.3*
Not bothered at all	**63.4**	**36.6**	*80.5*	*19.5*
Control over personal information				
Strongly agree	33.5	66.5	22.8	77.2
Somewhat agree	28.3	71.7	45.0	55.0
Somewhat disagree	32.9	67.1	40.6	59.4
Strongly disagree	25.4	74.6	48.4	51.6
Trust				
Trusted/Careful				
Can be trusted	31.5	68.5	31.6	68.4
Can't be too careful	30.6	69.4	37.8	62.2
Helpful/Look out for themselves				
Try to be helpful	29.1	70.9	38.1	61.9
Look out for themselves	35.3	64.7	32.6	67.4
Take advantage/Fair				

Take advantage	36.4	63.6	37.6	62.4
Try to be fair	25.1	74.9	35.0	65.0

Notes: Weighted estimates. *Italics* indicates differences by response option $\chi 2$ p<0.01; **bold** indicate differences by response option $\chi 2$ p<0.05; underline indicates differences by response option $\chi 2$ p<0.10

We next examined confidentiality attitudes by consent request type (see Table 3.13). Responses to the medical and income tax record items appear to be significantly related to consent status in both consent request conditions, but responses to the last confidentiality item do not show such a relationship for either type of consent.

We find that, in the unweighted data, smaller proportions of respondents in the overall sample who selected the “not bothered at all” response options to the medical and tax record items consented compared to the weighted data (61% and 55%, respectively).

3.7.2.4 Bivariate Analyses: Trust Attitudes

As shown in Table 3.12, overall, respondents do not vary in their consent status by trust concerns. The differences among the three items by consent status are quite modest and in varying directions.

Examining the relationship between trust attitudes and consent status by type of consent does not further illuminate the relationship between trust attitudes and consent (see Table 3.13). When examined by consent request type, differences in consent by responses to the three items remain small and the direction of these differences is inconsistent.

In the unweighted data, a smaller proportion of respondents who say that others would try to take advantage of you consented to the income and employment request as compared to the weighted data (32% vs. 38%). The difference by consent status is significant in the unweighted data (p<0.10).

3.7.3 Indices

Three three-item scales were created from the privacy, confidentiality, and trust items using data from respondents who provided substantive responses to all three items in each scale.

In developing these scales, an ideal approach would be to field a large number of items assessing each construct, examine the strength of each item relative to others and the dependent variable, and discard weaker items (Churchill, 1979; Saris, 2007). However, the restrictions of the Practicum budget did not allow this approach. Instead nine items were administered to all respondents, each intended to measure one of the three constructs. As mentioned, each of these items or a variant had been successfully fielded in one of more prior national surveys.

3.7.3.1 Reliability

Cronbach's Alpha (Cronbach, 1951) was calculated to assess the reliability and internal consistency for each of the three scales (Churchill, 1979). This measure evaluates the homogeneity of the items within a scale and determines if all scale items measure a single construct. Cronbach's Alpha identifies the proportion of variance in a scale attributable to a common source, presumably the latent variable (DeVellis, 1991). Cronbach's Alpha is calculated as one minus the ratio of noncommunal variance to total variance and is adjusted for the number of items in the scale (DeVellis, 1991).

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_{y_i}^2}\right)$$

$\sum \sigma_i^2$ is the total variance for all individual items (noncommunal variance)

$\sigma_{y_i}^2$ is the total variance in the scale

Cronbach's Alpha can range from 0-1 with higher values indicating greater reliability (DeVellis, 1991). Standards of acceptable Alpha values vary throughout the literature. Nunnally (1967) suggests values of 0.50 or 0.60 as sufficiently reliable, though a later work by the same author cites values as low as 0.70 as unacceptable (Nunnally, 1978). DeVellis (1991) cites values of 0.60 and lower as unacceptable.

The overall Alpha score is influenced by all problematic features of a scale and the items contributing to it. This includes noncentral or extreme item means, low item variability, negative item correlations, low item-scale correlations and weak inter-item correlations (Churchill, 1979; DeVellis, 1991). Scale reliability is also influenced by item covariation within the scale, as well as scale length. Longer scales tend to be more reliable on average (DeVellis, 1991).

3.7.3.2 Privacy Index

The three privacy items were coded so that higher scores represented more privacy concern and all items contributed equal weight to the scale. Possible scores ranged from 0 to 3. The mean index score was 1.1 suggesting a low amount of overall privacy concern among respondents as measured by the three items (see Table 3.14). According to Churchill (1979) and others, the overall alpha for this scale is below minimal acceptable standards (α =0.429). Alpha is at least partly influenced by the small number of items included in the scale and weak correlations among the privacy items (see Table 3.15). The Census invasion of privacy and government information request items correlate strongly with one another though minimally with the identity theft item.

Table 3.14

Distribution of Privacy Index Scores

Index Score	%
0 (*Low Concern*)	6.7
0.3	15.5
0.7	25.1
1.0	24.8
1.3	5.0
1.7	5.4
2.0	5.1
2.3	1.6
2.7	4.3
3 (*High Concern*)	6.5
Mean	1.1
Cronbach's Alpha	0.429

Notes: Cronbach's Alpha was calculated using weighted correlations for each item pair. Scores were calculated on 867 respondents that provided substantive responses to all three privacy items.

Table 3.15

Weighted Correlations among Privacy Items

	Census correlation p-value	Gov Requests	ID Theft
Census	1.000		
Gov Requests	0.479 0.000	1.000	
ID Theft	0.047 0.540	0.074 0.318	1.000

Notes: The significance of each correlation was calculated using simple linear regression accounting for survey weights. The most conservative *p*-value was recorded. Correlations were calculated on 867 respondents that provided substantive responses to all three privacy items.

3.7.3.3 Confidentiality Index

Confidentiality items were also coded so that higher scores represented more confidentiality concerns and all items contributed equal weight to the index. Index scores ranged from 0-3 with a scale mean of 2.4 (see Table 3.16). Cronbach's Alpha ($\alpha = 0.435$) fails to meet accepted reliability standards for the confidentiality scale as well. The small number of items included in the scale, low variance on the medical record and income tax items, and the low inter-item correlations for one of the items contribute to this.

Table 3.16

Distribution of Confidentiality Index Scores

Index Score	%
0 (*Low Concern*)	0.1
0.3	1.0
0.7	1.7
1.0	4.2
1.3	6.4
1.7	6.9
2.0	11.5
2.3	11.4
2.7	29.4
3 (*High Concern*)	27.4
Mean	2.4
Cronbach's Alpha	0.435

Notes: Cronbach's Alpha was calculated using weighted correlations for each item pair. Scores were calculated on 862 respondents that provided substantive responses to all confidentiality items.

Table 3.17 shows that the correlation between the medical record and income tax items is quite strong (r=0.469; p<0.001). This is not unexpected given the similar structure and content of the two items. However, correlations with the third item regarding loss of control of personal information are nominal, suggesting that the three items may not cohesively assess a singular construct.

Table 3.17

Weighted Correlations among Confidentiality Items

	Medical correlation p-value	Tax	Info Control
Medical	1.0000		
Tax	0.469 0.000	1.0000	
Info Control	0.092 0.191	0.051 0.510	1.0000

Notes: The significance of each correlation was calculated using simple linear regression accounting for survey weights. The most conservative *p*-value was recorded. Correlations were calculated on 862 respondents that provided substantive responses to all three confidentiality items.

3.7.3.4 Trust Index

As with the other scales, all trust items were coded so that greater trust concern was coded as higher numbers, all items were given equal weight in the index, and index scores ranged from 0-3 (see Table 3.18). The overall mean of the trust index is near the middle of possible scores (1.6). As shown in Table 3.19, average correlations among the trust items are higher than those in privacy and confidentiality scales. These are reflected in Cronbach's Alpha for the trust index (α =0.610); the highest among the three indices.

Table 3.18

Distribution of Trust Index Scores

Index Score	%
0 *(Low Concern)*	17.4
1	32.1
2	20.2
3 *(High Concern)*	30.3
Mean	1.6
Cronbach's Alpha	0.610

Notes: Cronbach's Alpha was calculated using weighted correlations for each item pair. Scores were calculated on 833 respondents that provided substantive responses to all three trust items.

Table 3.19

Weighted Correlations among Trust Items

	Trust correlation p-value	Help	Fair
Trust	1.000		
Help	0.221 0.011	1.000	
Fair	0.300 0.000	0.508 0.000	1.000

Notes: The significance of each correlation was calculated using simple linear regression accounting for survey weights. The most conservative *p*-value was recorded. Correlations were calculated on 833 respondents that provided substantive responses to all three trust items.

We also estimate the three indices using unweighted data. Overall, we find that mean scores for each of the unweighted indices are very similar to those constructed with the weighted data, Cronbach's alphas are slightly higher with the unweighted data, and correlations are similar, though slightly more significant among the privacy and confidentiality items with the unweighted data.

3.7.4 Logistic Regression Analyses

Three separate regression models incorporated the original privacy, confidentiality, and trust scales as linear predictors into the logistic regression model developed in the Consent Request Variation section of this chapter (see Table 3.20). After controlling for the consent request, demographic and socioeconomic predictors, and the incentive, the confidentiality index is a highly significant predictor of consent in the overall sample of respondents ($p<0.001$), with greater confidentiality concerns associated with consent refusal ($\beta = -1.29$). The trust and privacy indices are not statistically significant predictors of consent.

Table 3.20

Logistic Regression Models Predicting Consent to Record Linkage: Original Privacy, Confidentiality, and Trust Scales

		Privacy (n=867)	Confidentiality (n=862)	Trust (n=833)
Constant		0.30	2.85***	-0.23
Gender	Female (ref.)			
	Male	-0.44	-0.51	-0.56
Age	18-44 (ref.)			
	45-54	-1.07**	-0.99	-1.02**
	55-64	-0.17	-0.03	-0.26
	65+	-0.50	-0.60	-0.30
Education	HS or Less (ref.)			
	Associates/Some College	-0.79*	-0.92**	-0.64
	Bachelors Degree	-0.47	-0.58	-0.33
	Graduate Degree	-1.79***	-1.46**	-1.69**
Race	Non-White (ref.)			
	White	0.45	0.60	0.40

Income	<25K (ref.)			
	25- <75K	0.03	0.25	0.16
	>75K	-0.87	-0.92	-0.64
	Income DK	-0.08	-0.02	0.25
	Income REF	-3.40***	-3.29***	-2.84***
Incentive	No Incentive (ref.)			
	Incentive	0.39	0.33	
Privacy		-0.30		
Confidentiality			-1.29***	
Trust				0.06

Notes: Weighted estimates. *** indicates p<0.01; ** indicates p<0.05; * indicates p<0.10. Archer and Lemeshow Goodness of Fit test for survey data suggests a lack of fit for the Trust model: Privacy [F(9,41) = 1.05; Prob > F = 0.416]; Confidentiality [F(9,41) = 0.880; Prob > F = 0.547]; Trust [F(9,41) = 1.94; Prob > F = 0.073].

The confidentiality index also continues to predict consent in the unweighted data (p<0.001). Additionally in the unweighted data, we find that the privacy index significantly predicts consent in the hypothesized direction (p<0.01). The incentive is also positively associated with consent in the unweighted confidentiality and trust models.

The confidentiality index is a significant predictor of consent to both the health request (p<0.05; see Table 3.21) and the income and employment request (p<0.01; see Table 3.22). The privacy and trust indices are not associated with either request. Predictors of consent are similar among the models presented in Table 3.21 (Health) and 3.22 (Income and Employment), and are similar to the baseline models presented in the Consent Request Variation section of this chapter. One difference between the two models is that race is associated with consent to the income and employment request, with white respondents more likely to consent to this request. We observe similar findings in the baseline logistic regression model.

Table 3.21

Logistic Regression Models Predicting Consent to the Health Request: Original Privacy, Confidentiality, and Trust Scales

		Privacy (n=427)	Confidentiality (n=426)	Trust (n=408)
Constant		0.21	2.52*	-0.37
Gender	Female (ref.)			
	Male	-0.65	-0.60	-0.70
Age	18-44 (ref.)			
	45-54	-0.76	-0.80	-0.87
	55-64	0.71	0.81	0.62
	65+	0.34	0.21	0.41
Education	HS or Less (ref.)			
	Associates/Some College	-1.22**	-1.14*	-1.23**
	Bachelors Degree	-0.39	-0.43	-0.30
	Graduate Degree	-2.43**	-2.32*	-2.35**
Race	Non-White (ref.)			
	White	-0.67	-0.45	-0.73
Income	<25K (ref.)			
	25- <75K	-0.27	-0.26	-0.16
	>75K	0.00	-0.29	0.11
	Income DK	-0.19	-0.59	0.18
	Income REF	-3.96*	-3.88*	-3.60*
Incentive	No Incentive (ref.)			
	Incentive	0.39	0.36	0.53
Privacy		0.12		
Confidentiality			-0.97**	
Trust				0.32

Notes: Weighted estimates. *** indicates p<0.01; ** indicates p<0.05; * indicates p<0.10. Archer and Lemeshow Goodness of Fit test for survey data suggests a lack of fit for the Privacy model: Privacy [F(9,41) = 3.16; Prob > F = 0.006]; Confidentiality [F(9,41) = 1.83; Prob > F = 0.912]; Trust [F(9,41) = 0.82; Prob > F = 0.605].

Table 3.22

Logistic Regression Models Predicting Consent to the Income and Employment Request: Original Privacy, Confidentiality, and Trust Scales

		Privacy (n=440)	Confidentiality (n=436)	Trust (n=425)
Constant		-0.93	2.81**	-0.58
Gender	Female (ref.)			
	Male	0.12	-0.12	-0.25
Age	18-44 (ref.)			
	45-54	-1.30*	-1.30*	-1.27
	55-64	-1.15	-1.23	-1.24*
	65+	-1.28*	-1.61**	-1.06*
Education	HS or Less (ref.)			
	Associates/Some College	-0.51	-0.71	-0.28
	Bachelors Degree	-0.73	-0.92	-0.46
	Graduate Degree	-1.07	-0.47	-1.11
Race	Non-White (ref.)			
	White	1.57*	1.62*	1.51*
Income	<25K (ref.)			
	25- <75K	0.49	0.83	0.58
	>75K	-1.14	-1.08	-0.84
	Income DK	0.16	1.47	1.12
	Income REF	-1.99*	-1.96*	-1.71**
Incentive	No Incentive (ref.)			
	Incentive	0.12	-0.08	-0.05
Privacy		-0.71		
Confidentiality			-1.53***	
Trust				-0.11

Notes: Weighted estimates. *** indicates p<0.01; ** indicates p<0.05; * indicates p<0.10. Archer and Lemeshow Goodness of Fit test for survey data suggests a lack of fit for the Confidentiality and Trust models: Privacy [F(9,41) = 0.52; Prob > F = 0.851]; Confidentiality [F(9,41) = 3.82; Prob > F = 0.001]; Trust [F(9,41) = 3.18; Prob > F = 0.005].

Only a small number of differences emerge when we contrast the unweighted and weighted findings in this subsection. Most notably, we find that the privacy index is a significant predictor of consent to the income and employment request in the hypothesized direction (p<0.05). Also, the incentive positively influences consent to the health request in the models that include the privacy and confidentiality indices.

3.7.5 Factor Analyses

As noted in an earlier section, the privacy, confidentiality, and trust indices demonstrated modest scale reliability and inter-item correlations at best, particularly for the confidentiality and privacy measures. Further, weighted correlations among the scales suggest that relationships exist between the trust and privacy scales (r=0.380; p<0.001) as well as the confidentiality and privacy scales (r=0.254; p=0.003). A lesser relationship exists between the trust and confidentiality scales (r=0.123; p= 0.083). Considering the associations among the indices, factor analysis was used to further explore the loading of these items to the constructs that they were intended to measure, and determine if other relationships among the items existed. Here, we briefly review the findings from these analyses.

A factor analysis with oblique rotation (Promax) to account for covariation among factors results in four factors (see Table 3.23).[68] Factor one (*Trus*t) consists of the three GSS trust items originally intended to measure trust. The second factor comprises two of the three original privacy items (*Privacy 2*), including the Census and government information request items. The third privacy item (frequency of worrying about identity theft) loaded on a third factor along with the confidentiality item regarding control over personal information (*Data Security*). The remaining two confidentiality items (medical and income tax records) loaded on the remaining fourth factor (*Confidentiality 2*).

Logistic regression analyses explored the alternate grouping of items in predicting willingness to consent to record linkage (results not shown). The alternate grouping of items produced results similar to the original scales in predicting consent. Trust (*Trust*)

---

[68] The factor analysis does not account for weights for the following reasons: unweighted item correlations do not vary considerably from those calculated with survey weights (see Appendix), and most commercial software cannot account for survey weights in this analysis.

and privacy concerns, as measured through the alternate privacy scale (*Privacy 2*), do not predict consent, nor does the *Data Security* index. Confidentiality concerns (*Confidentiality 2*) continues to predict consent.

Table 3.23

Alternate Privacy, Confidentiality, and Trust Item Groupings: Rotated Factor Loadings

Question	Factor 1 *Trust*	Factor 2 *Privacy 2*	Factor 3 *Data Security*	Factor 4 *Confidentiality 2*
Trust	0.35			
Help	0.69			
Fair	0.68			
Census		0.60		
Gov Requests		0.57		
ID Theft			0.39	
Info Control			0.42	
Medical				0.63
Tax				0.63
Eigenvalue	1.58	0.81	0.44	0.19

Note: Unweighted estimates. Factor analysis based on 785 cases.

3.7.6 Open-Ended Questions

After each consent request, interviewers asked respondents why they consented or refused record linkage. Responses to these open-ended questions shed additional light on the impact of privacy, confidentiality, and trust concerns on respondents' consent decisions, as well as other impediments and motivations to consenting.

After a set of codes was developed for each of the two questions, the open-ended responses were coded by Practicum students. Each response was coded by two students. Cohen's Kappa was calculated as a measure of inter-rater agreement for all student pairs and all codes (Landis & Koch, 1977).[69] Across the codes for consenting respondents, Cohen's Kappa ranged from .59 to .71 for the student pairs, with 69% of cases in

---

[69] The other client involved in the 2011 JPSM Practicum, Rebecca Medway, contributed substantially to the reliability calculations for the open-ended responses.

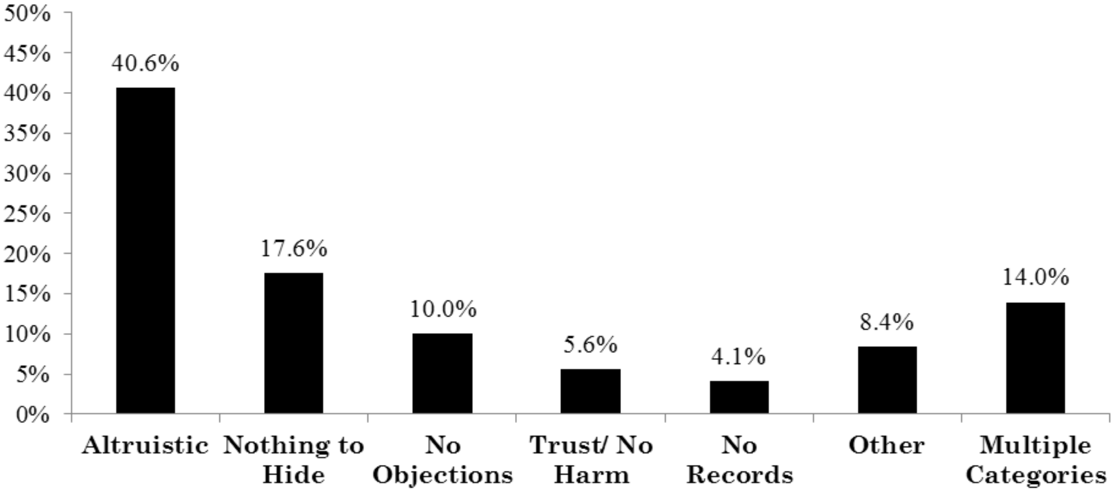
complete agreement. Cohen's Kappa ranged from .48 to .67 for student pairs who coded the responses for the respondents who refused consent, with 60% of cases in complete agreement. According to Landis et al. (1977), Kappa values between 0.61 and 0.81 are considered to be of "good" reliability. Values between 0.41 and 0.60 are considered moderately reliable.

All responses for which the coders did not agree were reviewed. For approximately 85% of cases, we ultimately chose one of the two sets of codes originally assigned by the students. The final distribution of codes is presented in Figures 3.02 and 3.03.

We first review results from respondents who granted consent. By a wide margin, respondents who consented said that they did so for altruistic reasons. Similar findings have been reported by Dunn et al. (2004); Jenkins et al. (2006); and Sala et al. (2010). Over forty percent of consenting respondents generally said that they hoped granting record access would help researchers or benefit the survey overall. For many of these respondents, the verbatim responses suggested that they understood the importance of granting record access to the integrity of the study. For example, one respondent who consented said, "Because I'm just hoping you can use the information to help improve the systems for everyone." Another said: "Maybe it will do some good; anything that can help will be appreciated." Respondents consenting to the income and employment request were significantly more likely to justify their decision with an altruistic explanation ($p<0.01$), perhaps for reasons relating to the current economic crisis.

Figure 3.02

Distribution of Verbatim Responses Provided by Consenting Respondents



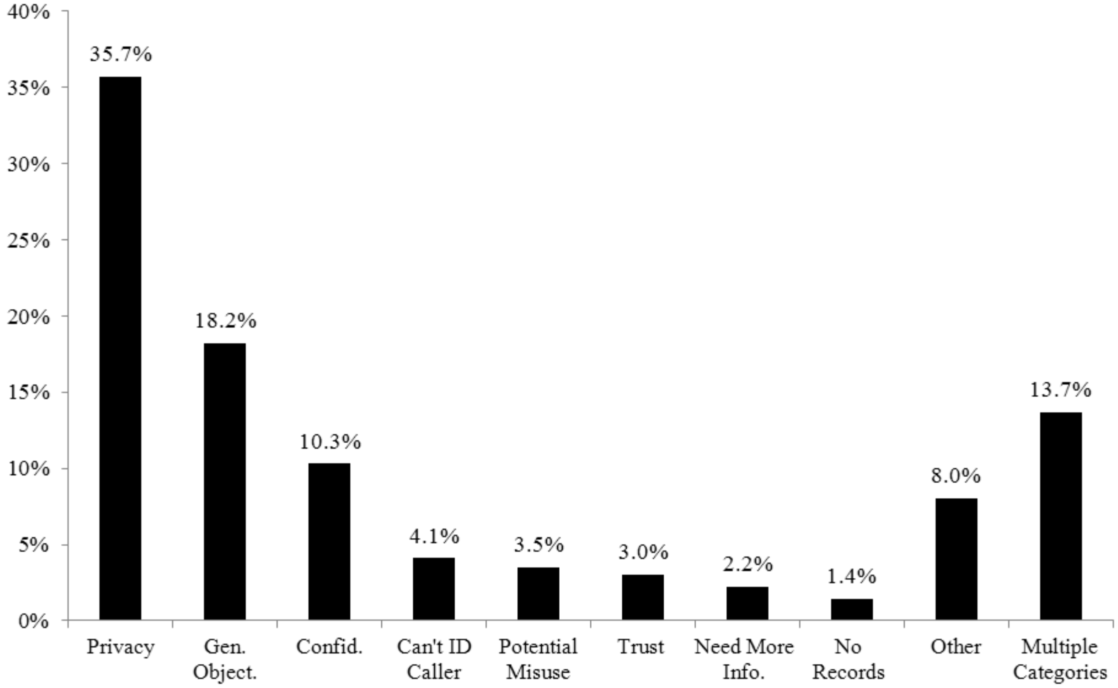
Notes: Weighted estimates. Analysis based on 252 cases, other respondents did not provide substantive response.

A sizable portion of respondents stated that they consented because they had nothing to hide, many using that exact phrase. Also contained within this code, others suggested that their records did not contain information that they deemed too personal or secretive and were therefore comfortable providing access.

The other explanations that respondents provided for consenting included that they generally had no objections to this request (10% of responses); trusted the survey or believed that they could not be harmed by granting record access (6% of responses); or they believed that they did not have any of the requested records and therefore consenting would not provide access to any information (4% of responses). Eight percent of responses did not fall into any of these categories; 14% of responses were coded into multiple categories.

Figure 3.03

Distribution of Verbatim Responses Provided by Nonconsenting Respondents



Notes: Weighted estimates. Analysis based on 618 cases, other respondents did not provide substantive response.

For respondents who did not consent to the request, the most frequently cited reason (36%) was privacy concerns. For example, respondents stated that information contained in their records was "personal information" and the request was an "invasion of privacy." Respondents assigned to the health request type were significantly more likely to provide this explanation for refusing the request than those assigned to the income and employment consent request ($p>0.05$), suggesting that health-related records are considered to be more private than income and employment-related records.

Less than twenty percent of responses given were objections to the request for general reasons. Of these, about 35% indicated that they would not consent to such a request on personal principle ("I just generally don't give that information out") and the remaining 65% were for more unspecific reasons (“I don’t feel it’s necessary”). These codes were combined due to their similarity and the small number of endorsements.

Confidentiality concerns were also mentioned in about 10% of responses. In these responses, respondents expressed concerned about who would be accessing their records. (For example, “Because I don’t know who is privy to it” and “Well, I just don’t need anybody messing around with my records.”)

Respondents also expressed concern about not being able to verify the identity of the caller (4% of responses); the potential for misuse of their records or identity theft (4%); generally lacking trust (3% of responses); needing more information before consenting (2%); and indicating that they did not have the relevant administrative records requested (1%). Respondents in the income and employment condition were more likely to say that they didn’t have the relevant records ($p<0.01$) and that they were refusing consent due to the potential for misuse of their personal records or identity theft ($p<0.01$). Eight percent of responses could not be coded, and about 18% of responses were coded into multiple categories.

The unweighted distribution of open-ended responses is very similar to the weighted distribution, with the only difference in the percentage of responses coded as a “general objection” when refusing consent. In the unweighted data, this is 14% compared to 20% in the weighted data.

3.7.7 Privacy, Confidentiality, and Trust Attitudes: Conclusions

The research in this chapter suggests that confidentiality concerns can negatively affect respondents’ willingness to consent to record linkage; privacy and trust concerns do not show such relationships. External benchmarks for the nine privacy, confidentiality, and trust items suggest that the majority of these items fared well in this survey; however, measures of reliability indicate that the privacy and confidentiality scales did not reliably assess these constructs.

Respondents did not significantly vary in their responses to the privacy items by consent status in the bivariate analyses; however, those who worry “frequently” about identity theft were more likely to refuse the income and employment consent request. As an index, the three privacy items demonstrated less than acceptable reliability, and did not significantly predict consent as hypothesized. The alternate groupings of items as suggested by the factor analysis did not produce stronger relationships with consent.

Although the privacy index was not significantly related to consent, responses to the open-ended items suggest that privacy concerns are an impediment to consenting, and perhaps a greater one for health-related requests. Further, as presented in the *Consent Request Variation* section of this chapter, a relationship exists between refusing consent to record linkage and refusing to provide a substantive response to the income question, considered to be an indicator of privacy concerns. Considering these results, it is possible that the lack of evidence resulting from the privacy index is due to problems with the index itself.

The medical and tax record items included in the confidentiality scale demonstrated strong bivariate associations with consent status by consent request type. Although the confidentiality scale also demonstrated less than ideal reliability, it significantly predicted consent in the hypothesized direction. The alternate grouping of confidentiality items, which removes the question “People have lost all control with how personal information is used about them”, also significantly predicts consent. This suggests that third item may not have contributed any additional information to the confidentiality scale beyond the other two items.

Confidentiality also arises as an obstacle to consenting in the open-ended item following the consent request in nearly one-fifth of responses to this item, providing further evidence that such concerns can prevent respondents from consenting to record linkage.

Trust, as measured through the survey items, does not appear to be associated with consenting to record linkage as hypothesized, though it is worth mentioning that responses to the GSS item specifically measuring trust are not comparable to those in the Practicum survey. Considering that some past research demonstrates a negative relationship between trust concerns and willingness to consent, it is possible that a relationship exists, and the GSS items included in the Practicum survey do not successfully measure the appropriate construct of trust relevant to record linkage. The GSS items measure trust in other people, though items assessing respondents' trust in the agency requesting consent may have provided more useful information in our analyses in demonstrating a relationship with consent. It is also possible that trust is simply less of a factor in a respondents' decision to consent. Further, in their explanation for why they did not consent, far fewer respondents mentioned trust-related rationales than explanations relating to privacy or confidentiality.

Some of the unweighted analyses in this section produced results different from those using weighted data. Most notably, we find that the privacy index is associated with consent in the hypothesized direction among all respondents and among those assigned to the income and employment request. The incentive is positively related to consent in several of these models as well, including among respondents overall and among those assigned to the health request.

3.8 Research Question 3: Consent Request Salience

The third and final research question in this chapter uses non-experimental research methods to explore the relationship between the salience of the consent request and the likelihood that a respondent will permit record linkage. A request for record linkage may be particularly salient to a respondent if features of the request are both prominent and related to respondent characteristics. Some existing research suggests that how salient a consent request is to a respondent can affect their likelihood of consenting, with respondents who find the request to be more salient being potentially more likely to consent (Sala et al., 2010).

For example, several studies requesting consent to health record linkage found that less healthy respondents were more likely to consent, perhaps because this request is more salient to less healthy respondents. This includes a meta-analysis examining predictors of consent that includes over 25,000 combined respondents across seven population surveys (Dunn et al., 2008). The researchers in this study found that respondents with the symptom under investigation in each of the individual surveys (e.g., had joint pain or headaches) were more likely to consent to medical record linkage. A greater proportion of respondents in the Panel Study of Income Dynamics with one or more chronic health conditions allowed their Medicare records to be linked with survey responses (Fulton et al., 2011). Harris et al. (2005) found that significantly more respondents with chronic physical or psychological illnesses allowed their survey data to be linked with their primary care records. Patients with poorer physical functioning were more likely to consent to medical record linkage in research by Woolf et al (2000).[70]

Although the researchers did not propose specific mechanisms contributing to these effects, less healthy respondents may consent at higher rates than their healthier counterparts for a variety of reasons. These respondents may be motivated to help others like themselves, the requested linking information (such as a Medicare number or physician's contact information) might be readily accessible, or they may be more accustomed and comfortable with providing it to others than healthier respondents. Less healthy respondents may have also found the request to be personally relevant, especially when the survey directly pertained to a topic or illness with which they had personal involvement (e.g., Partin et al., 2008). Healthier respondents may refuse consent because they assume that they don't have any information in their medical or health records that would be beneficial to researchers.

The positive effect of increased salience on consent rates extends to income and employment-related records, though existing empirical research in this area is scant. Respondents with no employment history were more likely to refuse consent to link their SSA records in research by Haider et al. (2000:3), who explain that, "It is not so surprising that those who never worked are more likely to decline permission to seek their Social Security records, as they have no covered earnings and may not even have a Social Security Number." In separate research, receiving government benefits positively predicted consent to access benefit records (Sala et al., 2010).

---

[70] Other studies investigating this topic found no relationship between respondents' health status and their likelihood of consenting to health-related records. This includes research by Sala et al. (2010) in which salience to this request was operationalized as having more than one hospital visits in the past year.

Like the research on health-related record linkage requests, these limited findings could also suggest that respondents are more likely to consent if they feel they have records that will benefit others and be of use to researchers. Respondents without an employment history would have limited or no SSA records, and depending on their age, may not even have been issued an SSN. Further, it is possible that individuals who receive government benefits are asked to provide identifying information, such as an SSN, more often, and may be more familiar with and comfortable in doing so.

To evaluate the effect of consent request salience on consent rates, the Practicum survey includes several items that correlate with information that would reasonably be contained in respondents' health and income and employment-related administrative records. For respondents' health-related records, this includes whether or not respondents have a number of chronic health conditions, frequency of physician visits and hospital stays, health-related expenditures, insurance coverage and self-rated health (see Table 3.24). For income and employment-related records, this includes items assessing employment status, hours worked weekly, and sources of income (see Table 3.25). The majority of these salience items were developed for use in nationally-representative surveys such as the ACS, GSS, NHANES and NHIS.

Considering the existing findings on salience of health-related requests, it is hypothesized that respondents in poorer health will be more likely to consent to the request for health-related records. In the Practicum survey, poorer health is operationalized as having a greater number of chronic conditions inquired about in the survey, more doctor and hospital visits, more medical expenditures, and poorer self-rated health. Respondents in poorer health may find this request to be more relevant, for example, if sicker respondents visit the doctor more and therefore have more information in their records. It is also possible that less healthy individuals may be motivated to help others like themselves.

We also include insurance coverage as a potential indicator of salience to the health consent request, also hypothesizing that respondents in poorer health will be more likely to consent. However, some past research finds that the relationship between health status and insurance coverage is not straightforward. An analysis of data from the 2001 SIPP found that respondents who rated their health as “good” reported the lowest rates of insurance coverage (82%), compared to respondents who rated their health as “poor” (85%) or “excellent” (88%). The highest rate of coverage for respondents in “excellent” health appears to be partly moderated by income, with more of these respondents covered by private healthcare (Bhandari, 2006). A similar trend is observed in research by Ross & Mirowsky (2000). Considering what may be a complex relationship between health status and insurance coverage, the inclusion of this variable as an indicator of salience is mainly for exploratory purposes.

Research on the effects of consent request salience on the likelihood to consent to income and employment-related consent requests is more limited and the findings less conclusive.[71] However, like the findings on health-related consent requests, and as suggested by Haider et al. (2000) respondents may be more willing to consent if they believe they have relevant information in their income and employment-related

---

[71] Although the income and employment consent request mentions income, the relationship between income and record linkage consent is likely moderated by the sensitivity and privacy concerns of the income question, especially as the income questions preceded the consent request and potentially made the privacy concerns associated with the record linkage request more salient to the respondent, thus reducing the likelihood that the respondent would consent. Therefore, because of the associated privacy concerns with the income item, we did not include this as a measure of salience.

administrative records. Thus, it is hypothesized that increased salience will be positively associated with increased consent propensity for the income and employment-related consent request as well. For this consent request, increased salience is operationalized as working full-time, working a greater number of hours, and receiving public assistance.

Compared to respondents who are not working, employed respondents may find the income and employment request more relevant, for example, because they would have information in their income and employment-related administrative records, as would respondents receiving public assistance. Further, as the income and employment consent request asks about "employment-related records", respondents who spend more hours working may find this request to be more salient.

Table 3.24

Correlates of Health-Related Records Included in the Practicum Survey

Question	**Response**	**Source**	**Original Wording**	**Q#**
Would you say your health in general is excellent, very good, good, fair, or poor?	Excellent/ Very good/ Good/ Fair/ Poor *(reversed for ½ respondents)*	NHANES		1
Has a doctor or other health professional EVER told you that you have any of the following…				
Diabetes or sugar diabetes*[other than during pregnancy]?*	Yes/No	NHANES		6A1
Hypertension or high blood pressure?	Yes/No	NHIS	Have you EVER been told by a doctor or other health professional that you had ... Hypertension, also called high blood pressure?	6B1
Asthma?	Yes/No	NHANES		6C1
Arthritis?	Yes/No	NHANES *(original sample is ages 20+)*		6D1
Heart disease?	Yes/No	NHANES *(original sample is ages 20+)*	Have you EVER been told by a doctor or other health professional that you had ... Coronary heart disease?	6E1
Anemia?	Yes/No	-----------		6F1
In 2010, were you a patient in a hospital overnight? Do not include an overnight stay in the emergency room.	Yes/No	NHANES		7
How many times were you a patient in a hospital overnight or longer during 2010? Do not count the total number of nights, just the total number of hospital admissions for stays which lasted 1 or more nights.	# of times if Q7=Yes	NHANES	How many different times did {you/SP} stay in any hospital overnight or longer {during the **past 12 months**}? (Do not count total number of nights, just total number of hospital admissions for stays which lasted 1 or more nights.)	8

Table 3.24 continued

Question	Response	Source	Original Wording	Q #
During 2010, how many times did you see a doctor or other health care professional about your health at a doctor's office, a clinic, hospital emergency room, at home or some other place? **[IF HOSPITAL INPATIENT (Q7=1), READ:** Do not include times you were hospitalized overnight.**] ---- if DK/REF ----** Would it be closer to 0 visits, 1 to 3 visits, 4 to 9 visits, or 10 or more visits?	# of times	NHANES	{During the past 12 months, how/How} many times {have you/has SP} seen a doctor or other health care professional about {your/his/her} health at a doctor's office, a clinic, hospital emergency room, at home or some other place? Do not include times {you were/s/he was} hospitalized overnight.	9
The next questions are about health insurance. Include health insurance obtained through employment or purchased directly, as well as government insurance programs like Medicare and Medicaid. Are you covered by any kind of health insurance or health care plan?	Yes/No	NHIS	The next questions are about health insurance. Include health insurance obtained through employment or purchased directly as well as government programs like Medicare and Medicaid that provide Medical care or help pay medical bills. [fill: Are you/Is anyone in the family] covered by any kind of health insurance or some other kind of health care plan?	11
The next question is about money that you have spent on medical and dental care for yourself only. Please do NOT count health insurance premiums, over-the-counter drugs, or costs that you were reimbursed for. In 2010, about how much did you spend for medical and dental care? Would you say it was zero dollars... some money but less than $500... $500 to less than $2,000... $2,000 to less than $3,000... $3,000 to less than $5,000... or $5,000 or more?	Zero dollars/ Some money but less than $500/$500 to less than $2,000/$2,000 to less than $3,000/$3,000 to less than $5,000/$5,000 or more	NHIS	The next question is about money that [fill1: you have/your family has] spent out of pocket on medical care. We do NOT want you to count health insurance premiums, over the counter drugs, or costs that you will be reimbursed for. In the PAST 12 MONTHS, about how much did [fill2: you/your family] spend for medical care and dental care?	14

Table 3.25

Correlates of Income and Employment-Related Records Included in the Practicum Survey

Question	Response	Source	Original Wording	Q #
Last week, were you working full-time, part-time, going to school, keeping house, or what?	Working full-time Working part-time/ With a job, but not at work because of temporary illness, vacation, strike/ Unemployed, laid off, looking for work/ Retired /In school/ Keeping house/ Other	GSS		25
How many hours a week do you usually work, at all jobs?	[# *of hours*]	GSS		27
During 2010, did you receive any income from the following sources: Social Security?	Yes/No	ACS (item modified for phone admin.)	INCOME IN THE PAST 12 MONTHS: *Mark (X) the "Yes" box for each type of income this person received, and give your best estimate of the TOTAL AMOUNT during the PAST 12 MONTHS. (NOTE: The "past 12 months" is the period from today's date one year ago up through today.)Mark (X) the "No" box to show types of income NOT received. (*d. Social Security or Railroad Retirement) *(*g. Retirement, survivor, or disability pensions. *Do NOT include Social Security) (*e. Supplemental Security Income (SSI).*)*	29A
How about other retirement or pensions?	Yes/No	ACS (item modified for phone admin.)		29B
How about public assistance or welfare, including Supplemental Security Income, or SSI?	Yes/No	ACS (item modified for phone admin.)		29C

3.8.1 Consent Request Salience: Analyses

Weighted estimates for questions in Tables 3.24 and 3.25 were calculated for the whole sample and estimates were compared by consent status for all respondents and within consent request conditions. Estimates from external surveys are provided as well. External estimates are not directly comparable; these serve as an external point of reference to the data produced through the Practicum survey.

Logistic regression examines the association between consent request salience and likelihood of consenting for each consent request while controlling for demographic and socioeconomic variables.

3.8.2 Consent Request Salience: Results

3.8.2.1 Benchmark Comparisons: Health

As demonstrated in Table 3.26, compared to estimates produced through the 2009/2010 NHANES, respondents in the Practicum survey rate their overall health as better. As NHANES focuses almost entirely on health, it is possible that after discussing and considering their health in greater detail, respondents felt less confident about their overall wellbeing. In contrast, health was only one of many topics in the Practicum survey and this question came at the very beginning of the interview.

Prevalence of chronic health conditions vary slightly in the Practicum survey as compared to benchmark estimates from NHANES. The Practicum survey has more diabetics than the most recent NHANES estimates, a similar portion of hypertensive and arthritic respondents, and fewer asthmatics. The Practicum survey has substantially more respondents with heart disease. This may be due to slight wording changes between the two surveys: NHANES asks about "coronary heart disease" which respondents may consider more serious and specific than "heart disease", the wording used in the Practicum survey.

Table 3.26

Correlates of Health-Related Records for All Respondents and External Benchmarks

	All Respondents (n=900)	External Benchmarks
	%	%
Mean Self-Rated Health (*1=Poor; 5=Excellent*)[72]	3.2	2.6
Chronic Conditions		
Diabetes	13.6	9.9[73]
Hypertension	33.8	32.6
Asthma	8.0	13.6
Arthritis	25.9	23.9
Heart disease	10.4	3.1
Anemia	11.4	--
1+ Chronic Conditions	54.4	--
Overnight Hospital Patient (*Yes*)	9.2	11.1
If Yes, # of Times	1.7	1.5
# of MD Visits 2010		
0	14.9	15.9
1-3	46.2	46.5
4-9	25.2	23.8
10+	13.8	13.2
Health Insurance (*Yes*)	85.7	--[74]
Healthcare Expenditures		
$0	16.4	12.4
<$500	47.0	34.9
$500 to <$2,000	25.7	31.4
>$2,000	11.0	21.3

Note: Weighted estimates.

---

[72] Mean self-rated health is reported here because only the mean self-rated health benchmark estimate was available. In the Practicum survey, 3.8% of respondents reported their health as poor, 16.3% as fair, 45.3% as good, 25.2% as very good, and 9.4% as excellent.

[73] In the NHANES, 8.8% of respondents report having Diabetes; an additional 1.8% report having borderline Diabetes.

[74] NCHS does not make estimates from this NHIS item publicly-available.

Compared to NHANES estimates, slightly fewer Practicum respondents report staying overnight in the hospital last year. Practicum and NHANES respondents reported similar numbers of doctor visits than NHANES respondents. Finally, healthcare expenditures were similar for Practicum and 2010 NHIS respondents.

As with the other analyses in this chapter, we also examine these items using unweighted data. With the exception of anemia, there is a greater percentage of respondents in the unweighted data with each of the chronic conditions, likely because of the age of the respondents in the sample. Also, a greater proportion of respondents have health insurance in the unweighted data (95%) compared to weighted data (86%).

3.8.2.2 Binary Analyses: Health

As shown in Table 3.27, there are no statistically significant variations in consent rates for respondents in the total Practicum sample by health salience indicators. Self-rated health and the presence of various chronic health conditions do not differ by consent status. Consent rates do not vary by medical expenditures, health insurance coverage, frequency of doctor visits, or whether the respondent stayed in a hospital overnight during 2010.

In the corresponding unweighted data, we find that consent rates are significantly higher among respondents with diabetes ($p<0.05$), hypertension ($p<0.01$), and heart disease ($p<0.01$). Respondents with one or more chronic conditions were more likely to consent, ($p<0.01$), as were respondents with health insurance ($p<0.10$).

Table 3.27

Consent by Health Status for All Respondents

	Consent (n=269)	Non-Consent (n=631)
	%	%
Self-Rated Health		
Poor	43.0	57.0
Fair	20.0	80.0
Good	43.1	56.9
Very Good	23.0	77.0
Excellent	24.6	75.4
Chronic Conditions		
Diabetes		
Yes	30.8	69.2
No	33.8	66.2
Hypertension		
Yes	41.0	59.0
No	29.5	70.5
Asthma		
Yes	25.5	74.5
No	34.1	66.0
Arthritis		
Yes	30.5	69.5
No	34.4	65.6
Heart disease		
Yes	39.6	60.4
No	32.7	67.3
Anemia		
Yes	41.2	58.8
No	32.4	67.6
1+ Chronic Conditions		
Yes	34.0	66.0
No	32.8	67.2
Overnight Hospital Patient		
No	32.3	67.7
Yes	43.0	57.0
If Yes, # of Times	2.2	1.4
# of MD Visits 2010		
0	36.3	63.7
1-3	37.2	62.8
4-9	22.3	77.7
10+	37.4	62.6
Health Insurance		
Yes	32.3	67.7
No	40.1	59.9

Healthcare Expenditures		
$0	40.8	59.1
<$500	32.7	67.3
$500 to <$2,000	25.1	74.9
>$2,000	48.6	51.4

Notes: Weighted estimates. None of the differences are significant ($\chi 2$ p<0.10) by health status.

Table 3.28 focuses on respondents assigned to the health consent request condition. In this sample, we do not see any evidence of a health salience request. None of the items included in the Practicum survey to assess health salience are significant by health consent status at the 0.10 level.

However, in the unweighted bivariate analyses among respondents assigned to the health request, we find that consent rates are significantly higher among respondents who have diabetes (p<0.05), hypertension (p<0.01), heart disease (p<0.05), and anemia (p<0.05), as well as among respondents with one or more chronic conditions (p<0.01).

Table 3.28

Consent by Health Status for Respondents Assigned to the Health Consent Request

	Consent (n=269)	Non-Consent (n=631)
	%	%
Self-Rated Health		
Poor	70.7	29.3
Fair	23.2	76.8
Good	46.2	53.8
Very Good	14.3	85.7
Excellent	12.1	87.9
Chronic Conditions		
Diabetes		
Yes	38.7	61.3
No	29.3	70.7
Hypertension		
Yes	40.2	59.8
No	25.9	74.2
Asthma		
Yes	34.3	65.7
No	30.6	69.4

Arthritis		
Yes	37.5	62.5
No	28.5	71.5
Heart disease		
Yes	37.2	62.8
No	29.8	70.2
Anemia		
Yes	41.5	58.5
No	29.3	70.7
1+ Chronic Conditions		
Yes	34.1	65.9
No	26.9	73.1
Overnight Hospital Patient		
No	42.5	57.5
Yes	29.6	70.4
If Yes, # of Times	2.18	1.45
# of MD Visits 2010		
0	26.3	73.7
1-3	35.9	73.7
4-9	17.4	82.6
10+	40.1	59.9
Health Insurance		
Yes	29.7	70.3
No	41.1	58.9
Healthcare Expenditures		
$0	34.8	65.2
<$500	31.2	68.8
$500 to <$2,000	22.4	77.6
>$2,000	52.4	47.6

Note: Weighted estimates. None of the differences are significant ($\chi 2$ p<0.10) by health status.

3.8.2.3 Benchmark Comparisons: Income and Employment

Table 3.29 shows results for the income and employment salience indicators for the overall sample as well as external benchmark estimates. The employment question aligns relatively closely to the 2010 GSS administration, although the Practicum survey includes slightly more retired respondents and slightly less unemployed respondents. Practicum respondents report working somewhat more hours compared to the 2010 GSS.

Estimates of government-provided benefits in the Practicum survey are quite comparable to the 2010 ACS.

The unweighted employment distribution has less working respondents (41%) and more retired respondents (44%) compared to the weighted estimates. A greater proportion of the unweighted sample receives Social Security (54%) and retirement benefits (41%).

Table 3.29

Correlates of Income and Employment-Related Records for All Respondents and External Benchmarks

		All Respondents (n=900)	External Benchmarks
		%	%
Employment Status			
	Working	60.2	59.6
	Unemployed	3.6	7.3
	Retired	18.0	13.5
	Other	18.2	19.5
Mean Hours Worked/Week		42.3	36.6
Receives Benefits from:			
	Social Security	29.4	28.4
	Other Retirement/Pensions	17.9	17.5
	Public assistance/Welfare/SSI	5.8	2.9
Receives 1+ Benefits		35.7	--

Note: Weighted estimates.

3.8.2.4 Binary Analyses: Income and Employment

Table 3.30 presents consent rates for each of the income and employment salience indicators for the full sample. Employment status overall is not significantly related to consenting, nor is the mean number of hours a respondent works per week. Receiving SSI or other public assistance, or one or more of the benefits inquired about in the survey is positively associated with consenting.

In the unweighted data, respondents receiving Social Security, public assistance, or one or more benefits are more likely to consent. The employment distribution varies by consent ($p<0.01$), but not in the hypothesized direction: consent rates are greater among retired (33%) and unemployed respondents (41%), as compared to working respondents (23%).

Table 3.30

Consent by Income and Employment Status for All Respondents

	Consent (n=269)	Non-Consent (n=631)
	%	%
Employment Status		
Working	30.5	69.5
Unemployed	44.4	55.6
Retired	30.5	69.5
Other	43.8	56.2
Hours Worked/Week		
<20	13.8	86.2
21-40	25.9	74.1
40+	61.4	38.6
Receives Benefits from:		
Social Security		
Yes	41.2	58.8
No	31.1	68.9
Other Retirement/Pensions		
Yes	31.2	68.8
No	34.7	65.3
Public assistance/Welfare/SSI		
Yes	*67.2*	*32.8*
No	*31.9*	*68.1*
Receives 1+ Benefits		
Yes	**43.4**	**56.6**
No	**28.9**	**71.1**

Notes: Weighted estimates. **Bold** indicate differences by income and employment status $\chi 2$ $p<0.05$

We next examine the income and employment salience indicators by consent status, focusing only on respondents assigned to this request. These results are presented in Table 3.31. We find no evidence for a consent request salience effect for items measuring employment, hours worked, or government benefit receipt.

Table 3.31

Consent by Income and Employment Status for Respondents Assigned to the Income and Employment Consent Request

	Consent (n=120)	Non-Consent (n=335)
	%	%
Employment Status		
Working	38.2	61.8
Unemployed	40.9	59.1
Retired	22.0	78.0
Other	38.0	62.0
Mean Hours Worked/Week		
<20	16.5	83.5
21-40	34.3	65.6
40+	46.1	53.9
Receives Benefits from:		
Social Security		
Yes	38.9	61.1
No	35.0	65.0
Other Retirement/Pensions		
Yes	36.0	64.0
No	36.0	64.0
Public assistance/Welfare/SSI		
Yes	53.1	46.9
No	35.4	64.5
Receives 1+ Benefits		
Yes	43.6	56.4
No	32.5	67.5

Notes: Weighted estimates. Italics indicates differences by income and employment status $\chi 2$ p<0.01

In the unweighted data, consent status by employment is significant but not in the hypothesized direction, with consent rates to the income and employment request lower among working respondents ($p<0.01$). Respondents receiving public assistance or SSI are more likely to consent ($p<0.05$), as are respondents receiving more than one government benefit ($p<0.05$).

3.8.2.5 Logistic Regression Analyses

We next used logistic regression to examine the association between the salience of the consent request and respondents' willingness to consent while controlling for demographic and socioeconomic variables. We include the items assessing salience in separate models predicting consent to their respective requests.

The model predicting consent to the health request that includes the indicators of health salience shows no evidence of a salience effect after controlling for respondent demographic and socioeconomic characteristics. This model is presented in Table 3.32.

Table 3.32

Logistic Regression Model Incorporating Indicators of Health Request Salience, and Predicting Consent to the Health Consent Request

		Coefficient	SE	p-value
Constant		0.71	3.10	0.821
Gender	Female (ref.)			
	Male	0.60	0.62	0.335
Age	18-44 (ref.)			
	45-54	-0.82	0.91	0.374
	55-64	1.12	0.77	0.155
	65+	1.12	0.99	0.265
Education	<HS Grad/HS Grad (ref.)			
	Some Col/Assoc Deg	-1.10	0.70	0.123
	Bachelors Deg	-0.45	0.98	0.65
	Graduate Deg	-2.79	1.59	0.086
Race	Non-White (ref.)			
	White	-0.69	0.80	0.392
Income	<25K (ref.)			
	25- <75K	-0.37	0.73	0.611
	>75K	0.55	1.22	0.655
	Income DK	-1.03	1.16	0.375
	Income REF	-4.67	1.55	0.004
Mean Self-Rated Health (*1=Poor; 5=Excellent*)		-0.55	0.37	0.136
1+ Chronic Conditions†	No (ref.)			
	Yes	-0.02	0.92	0.985

Overnight Hospital Patient	No (ref.)			
	Yes	0.21	0.96	0.831
# of MD Visits 2010	0 (ref.)			
	1-3	0.10	1.43	0.945
	4-9	-1.84	1.80	0.309
	10+	-0.22	1.70	0.896
Health Insurance	No (ref.)			
	Yes	-0.67	1.26	0.601
Healthcare Expenditures	$0 (ref.)			
	<$500	0.71	1.72	0.682
	$500 to <$2,000	-0.30	1.77	0.864
	>$2,000	1.68	1.52	0.274
Incentive	No Incentive (ref.)			
	Incentive	0.54	0.68	0.430

Notes: Weighted estimates. Model based on 430 cases; Archer and Lemeshow Goodness of Fit test for survey data suggests a lack of fit [F(9,40)=2.37; Prob>F=0.0297]. [†]The chronic conditions are not individually predictive of consent.

In the analogous unweighted model, one of the health salience indicators is associated with health consent: respondents who have one or more of the chronic conditions inquired about are more likely to consent ($p<0.10$). The incentive is positively associated with consenting in this model ($p<0.05$).

The income and employment salience indicators fare only slightly better (see Table 3.33). Controlling for demographic and socioeconomic predictors, the parameter for having one or more government benefits is significant ($p<0.05$) and positively associated with consenting, similar to the findings reported by Sala et al. (2010). (However, in the unweighted model, this predictor is not significantly associated with consenting.) The other items included in the survey to assess income and employment salience (employment status and hours worked) are not associated with consenting.

Table 3.33

Logistic Regression Model Incorporating Indicators of Income and Employment Salience, and Predicting Consent to the Income and Employment Consent Request

		Coefficient	SE	p-value
Constant		-1.60	1.39	0.257
Gender	*Female (ref.)*			
	Male	0.43	0.45	0.353
Age	18-44 (ref.)			
	45-54	-1.77	0.60	0.005
	55-64	-1.72	0.73	0.024
	65+	-1.99	0.88	0.027
Education	<HS Grad/HS Grad (ref.)			
	Some Col/Assoc Deg	-0.53	0.48	0.278
	Bachelors Deg	-0.48	0.73	0.516
	Graduate Deg	-1.46	1.00	0.151
Race	Non-White (ref.)			
	White	1.24	0.76	0.109
Income	<25K (ref.)			
	25- <75K	0.77	0.58	0.189
	>75K	-0.45	0.86	0.606
	Income DK	1.46	1.06	0.175
	Income REF	-0.89	1.29	0.494
Employment Status	Not Working (ref.)			
	Working	-0.51	1.57	0.746
Mean Hours Worked/Week				
	<20 (ref.)			
	21-40	0.59	1.47	0.689
	40+	1.52	1.45	0.299
Benefits	Receives 0 Benefits (ref.)			
	Receives 1+ Benefits[†]	1.43	0.67	0.038
Incentive	No Incentive (ref.)			
	Incentive	-0.39	0.57	0.497

Notes: Weighted estimates. Model based on 445 cases; Archer and Lemeshow Goodness of Fit test for survey data suggests a lack of fit [F(9,41)=1.93; Prob>F=0.074]. [†]Including these benefits individually was not associated with consent.

3.8.3 Consent Request Salience: Conclusions

It is possible that the design of this study contributed to the primarily null findings in this section. Similar to the explanation considered for the lack of significant differences in consent rates between the two consent request conditions (the first research question discussed in this chapter), it is conceivable that the consent requests and topic of administrative records requested in the Practicum survey were not made prominent and salient enough to respondents. The content of the records requested was stated only once in each of the requests in order to minimize the length and burden of the consent statement for phone administration. Mentioning the topic of the records additional times – either in the consent request or in other survey materials such as the advance letter could have increased the salience of this request.

The unweighted data shows some evidence of a health salience effect, but the evidence is limited to the item indicating whether a respondent has one or more chronic conditions. However, the findings observed in the weighted income and employment model (the positive association between having one or more government benefits and consenting) disappear in the unweighted model.

We proposed that respondents who had greater information in the relevant records would be more willing to consent. This may not be the correct mechanism contributing to a consent salience effect. As evidence for this, not having relevant records was cited as a reason for not consenting by a very small proportion of respondents who refused consent to either request in the privacy section of this chapter. If such a salience effect does indeed exist, it may be less straightforward and more nuanced than predicted.

3.9 Overall Conclusions and Limitations

We began this chapter by hypothesizing that various respondent and consent request characteristics would impact consent rates. This included the content and topic of the administrative records requested, respondents' privacy, confidentiality, and trust concerns, and the salience of the consent request in relation to respondents' characteristics. In this final section, we review the findings of this research and reflect upon its limitations.

Regarding our first research question, we did not find a statistically significant difference in consent rates to the requests for health and income and employment-related administrative records in the weighted bivariate or multivariate analyses. However, the unweighted analyses demonstrate a significant effect of consent request type on consent rates, with greater unweighted consent rates to the health request.

Similarly, we found no effect of the incentive on consent rates in the weighted sample, but we identified a significant effect of the incentive on consent rates in several of the unweighted models. Specifically, the incentive benefitted consent rates in the unweighted models among respondents overall, and among those assigned to the health request.

The second research question, which investigated the effects of respondents' privacy, confidentiality, and trust concerns on their willingness to grant record access, provides some evidence for the proposed hypotheses: confidentiality concerns negatively impact consent to record linkage. Although indicators of scale reliability were less than ideal, the confidentiality index still proved to be a strong predictor of consent in both the weighted and unweighted models. Other indicators of these constructs, including income refusal and open-ended responses suggested a negative relationship between privacy and confidentiality concerns and consenting to record linkage. In addition, the privacy index significantly predicted consent to the income and employment request in the unweighted model. Trust as it was measured in the Practicum survey appears to play less of a role in respondents' decision to consent.

We note that responses to the medical and income tax confidentiality items could have been affected by their similarity to the consent requests. Respondents may have answered these items in such a way to appear consistent with their recent consent decision. Further, given the similar structure of the medical and income tax records items, it is possible that respondents gave each item less than optimal effort and provided the same response to both. Indeed, of the 878 respondents who provided a substantive response to both items, 624, or 71%, provided the same response.

The third and final part of this chapter sought to determine whether the salience of the consent request influenced respondents' consent decision. The health and income and employment items included in the survey to assess salience largely did not demonstrate any relationship with consent in the weighted bivariate and multivariate models. Future qualitative research could be helpful to better understand and define the mechanisms contributing to a salience effect – if one truly exists.

The design of this research may limit the generalizability of findings and could have contributed to the lack of hypothesized results. First, the sampling methodology (a list sample of residential phone numbers) resulted in a skewed sample distribution that required substantial weighting. These weights increased the overall variance of estimates and potentially masked or distorted effects.

Further, the survey suffered from a low response rate (15.7%). Among the survey non-respondents are possibly those sample members with the greatest privacy, confidentiality, and trust concerns. Thus, the results presented in this chapter may not reflect the privacy, confidentiality, or trust attitudes of the target population, or accurately identify predictors of consent. In addition, a substantial portion of consenting respondents provided an altruistic rationale in the open-ended responses, but respondents with some altruistic quality may have been overrepresented among responding sample members.

Another issue to consider is the structure and content of the Practicum instrument. The questionnaire spanned a broad range of topics, including a host of questions preceding the consent request that could have reduced the perceived legitimacy of the survey and decreased consent rates. The survey did not benefit from a well-known sponsor and requesting consent over the telephone appears to have prevented at least some respondents from consenting, as observed through the open-ended responses.

Although we provided interviewers with a one-page document of responses to what we anticipated would be commonly-asked questions regarding the consent request, it is unclear which questions respondents raised, how frequently they were asked, or how well interviewers addressed them. Unaddressed concerns could have prevented respondents from consenting. Further, it is possible that such concerns disproportionately affected one consent treatment. However, only 3.5% of open-ended responses indicate that respondents did not consent because they needed more information about the consent request, and this did not vary by consent request type.

In the Practicum survey, the consent request was limited to one question in the survey, and wasn’t mentioned in any additional survey materials. These design decisions were made in order to accommodate the other research and experiments included in the survey, and because of time and burden constraints given the mode of administration.

This somewhat limits the generalizability of the results from these experiments, as other surveys requesting consent to record linkage tend to emphasize the consent request through either multiple consent questions, a more extensive introduction to the request, or mention of the record linkage in an advance letter.

Further, as mentioned previously, the approach taken in assessing privacy, confidentiality, and trust concerns – including existing items rather than pretesting a larger number of questions that measure these constructs – was a less than ideal method. Overall, the privacy and confidentiality items included in the survey did not reliably assess these constructs.

4 CHAPTER 4: ANALYSIS OF CONSENT RATES IN THE NATIONAL IMMUNIZATION SURVEY

4.1 Introduction

Research shows that interviewers can influence survey responses and response rates. A logical extension of the existing work on interviewer effects is the application of this research to the request for respondents' consent to record linkage. This chapter uses data from the 2009 National Immunization Survey (NIS) to evaluate the effects of interviewers and interviewer characteristics on respondents' willingness to consent to vaccination provider contact. The impact of respondent characteristics on their decision to permit or withhold consent is also examined.

4.2 Background

Interviewers play a central role in survey data collection and act as "the key agent of the researcher" (Durrant, Groves, Staetsky, & Steele, 2010: 2). They can influence the sampled individual's decision to respond to or refuse a survey request (Campanelli et al., 1999; Durrant et al., 2010; Lyberg & Dean, 1992; Lyberg & Lyberg, 1991; O'Muircheartaigh et al., 1998; Snijkers, Hox, & de Leeuw, 1999; West & Olson, 2011). During the interview, interviewers can vary in their ability to conduct the question and answer process and in the quality of responses that they subsequently obtain (Fowler & Mangione, 1990; Hox, de Leeuw, & Kreft, 1991).

Interviewer variability has been observed in both telephone and in-person surveys, in which the responses of different respondents interviewed by the same interviewer are more similar than in the sample overall (Hox, de Leeuw, & Kreft, 1991). This can produce effects on the resulting data similar to those of clustering (Groves et al., 2004; Hox et al., 1991). These interviewer effects can only be accurately calculated if the assignment of interviewers to respondents is comprised of random subsets of respondents

(Durrant et al., 2010; Fowler et al., 1990). Unfortunately, for surveys conducted in-person, interviewer assignment is almost always confounded with geographic areas, preventing the separation of interviewer and area effects (O'Muircheartaigh et al., 1999; Schnell & Kreuter, 2005). Interpenetrated designs commonly used in RDD telephone surveys, for which interviewers are randomly assigned to respondents, facilitate the estimation of interviewer effects on responses and response rates (Groves et al., 2004).

Survey methodology literature suggests that interviewer characteristics can contribute to variation in rates of survey participation and response quality, as well as affect the context or intended meaning of a survey question (Fowler et al., 1990). Past research generally finds that interviewers' demographic characteristics can affect survey responses when they are related to the constructs measured. This includes interviewer age (Durrant et al., 2010; Fowler et al., 1990; Freeman et al., 1976; Hanson & Marks, 1958; Singer et al., 1983; Sudman & Bradburn, 1974), gender (Bradburn 1983; Groves & Fultz, 1985; Kane & Macaulay, 1993; Nealon 1983), and race (Anderson et al., 1988; Hatchett & Schuman, 1975; Hox et al., 1991; Huddy, Billig, Bracciodieta, Hoeffler, Moynihan & Pugliani, 1997; Schaeffer, 1980; Schuman, & Converse, 1971). Prior interviewing experience also affects unit and item nonresponse: existing research demonstrates both positive and negative effects of greater experience (Bailar, Bailey & Stevens, 1977; Durrant et al., 2010; Groves & Couper, 1998; Lipps & Pollien, 2010; Singer et al., 1983).

Although interviewers can affect survey cooperation and response, research has yet to as thoroughly examine whether or not they influence an area of growing importance in many survey interviews: the request for respondents' consent to access and link their responses with personal records. Just as some interviewers are more successful at obtaining unit and item response, it is conceivable that they vary in their ability to obtain respondents' consent to record linkage, due to specific demographic characteristics or differing levels of experience.

A growing number of surveys, such as HRS, MEPS, NHANES, and NHIS, request respondents' consent to record linkage as a way to maximize the amount of information they collect about a single respondent. For these surveys and most others making this request, interviewer-administered data collection is the norm, and therefore concerns regarding interviewer effects on consent to record access and linkage are similarly applicable.

As discussed in Chapter 1, a small but growing body of existing research shows that interviewers can vary in the consent rates they achieve (Cleary 1981; Sakshaug et al., 2011; Sakshaug et al., 2012). While specific interviewer characteristics may be related to this variation, limited research investigates this topic and findings are mixed. Given the small number of studies investigating the effect of interviewer characteristics on consent, it is unclear under what conditions, if any, these traits affect consent.

Most surveys that request consent to record linkage are conducted in-person, and a smaller portion is conducted over the telephone. Of those conducted by telephone, several employ a panel design (ex., PSID and HRS), for which respondents are generally not randomly assigned to interviewers at each wave, complicating the evaluation of interviewer effects on the consent request. Further, in some cases, respondents who initially refuse the request are asked for their consent (and related identifying information) again during later waves.

The research described in this chapter extends existing work by examining data from the 2009 NIS, an RDD and cross-sectional survey, for which available interviewers are randomly assigned to sample cases.[75] The proposed analyses will examine if NIS interviewers vary in their ability to gain respondents' consent to contact vaccination providers and if specific interviewer characteristics are related to consent likelihood, such as interviewer age, gender, race/ethnicity, and NIS interviewing experience. In addition, some information is collected about the child's mother and about the survey respondent if the mother is not the respondent, facilitating a comparison of those who do and do not consent.

4.3 Research Questions

4.3.1 Interviewer Characteristics Associated with Consenting

The primary focus of the research in this chapter investigates the effects of the interviewer and interviewer characteristics on respondents' willingness to grant vaccination provider contact, and therefore provide access to their child's vaccination records. In line with findings from past research, we hypothesize that interviewers will vary in their ability to gain respondents' consent in the 2009 NIS. We will then explore whether interviewer characteristics including age, gender, race/ethnicity, and NIS experience are related to respondents' consent decision.[76]

---

[75] There are some exceptions to the random assignment of interviewers to respondents in NIS: if a respondent refused survey cooperation or consent during a prior contact, a specific interviewer may be assigned to the case in future contacts. Typically, more experienced interviewers are assigned to these cases.

[76] Although characteristics such as age, gender, and race/ethnicity are more discernible in person than over the telephone, it is conceivable that respondents will be able to identify these characteristics for at least a portion of interviewers through their name, accent, and tone of voice (Fowler et al., 1990), particularly an interviewer's gender and race. However, these characteristics can affect responses regardless of respondents' ability to accurately perceive them.

Existing research that investigates the effect of interviewer characteristics on consent appears to be mixed; interviewer characteristics influence consent rates in some studies, but do not affect consent in others. Further, a sizeable portion of relevant research was conducted in the U.K. where national policy regarding the use of administrative data and consent to access records differ from the U.S. (Baker et al., 2000; Dunn et al., 2004), limiting the full generalizability of these findings to the current research. Considering these limitations of the existing research, we do not propose specific hypotheses for the effects of interviewer characteristics on consent in the 2009 NIS. Here, we briefly review findings from the existing literature.

Interviewers' age was unrelated to record linkage consent in research by Sala et al. (2010), conducted in the U.K. The effects of interviewer race on consent likelihood are inconsistent across studies: black interviewers were significantly less likely to obtain SSNs from respondents in the 2004 HRS (Sakshaug et al., 2010), yet race of interviewer had no effect on respondents' willingness to provide SSN in the 2008 administration of the same survey (Sakshaug et al., 2012).

Interviewer gender was unrelated to consent in the 2008 HRS (Sakshaug et al., 2010) and the BHPS (Sala et al., 2010). However, Tate et al. (2005) found that mothers were less likely to grant access to their child's birth records when a male interpreter translated the survey interview. The researchers attributed this finding to increased suspicion of male interpreter.

In research conducted by Cleary (1981), experienced interviewers obtained higher consent rates than less experienced interviewers. Sakshaug et al. (2012) did not observe any relationship between interviewing experience and consent rates, but did find that interviewers who were successful obtaining consent in the early stages of data collection were also successful later on.

4.3.2 Respondent and Mother Characteristics Associated with Consenting

While the primary objective of the research presented in this chapter is to evaluate the effect of interviewers on respondents' consent to vaccination provider contact, we will also examine how characteristics of the survey respondent and sample child's mother affect consent decisions.

As discussed in Chapter 1, the relationship between respondents' demographic characteristics and consent likelihood appears to vary across studies (Sala et al., 2010), and this is no different for the small number of studies in which respondents are asked to permit access to their child's records rather than their own records. In the research examining parental predictors of consent to children's records, there is little evidence of any trends in demographic and socioeconomic predictors of consent.[77] One study, Tate et al. (2005), found that mothers who were single, younger, and had the greatest and least amounts of education were more likely to withhold consent. A second relevant study, Klassen et al. (2005), did not identify any association between age and education and consent likelihood.[78] Further, as consent rates varied by location in research by Tate et al. (2005), it is possible that this area effect was confounded with interviewer effects if a particular interviewer was responsible for completing interviews in a certain area.

---

[77] The two reports were conducted in countries with different policy regarding data linkage and administrative record use, thus limiting their comparability: Klassen et al. (2005) was conducted in the U.S. and Tate et al. (2005) was conducted in the U.K.

[78] Several characteristics examined as predictors of consent in Tate et al., 2005 and Klassen et al., 2005 were not investigated by both studies, preventing a comparison of the effects of these predictors across studies.

In related research using data from the 2002 NIS, Smith, Hoaglin, Battaglia, Khare, and Barker (2005) examined factors associated with respondents who consented to vaccination provider contact and supplied provider contact information for the sample child, but the vaccination provider failed to deliver records. Children with missing vaccination provider data were more likely to have mothers who were younger than thirty years old, Hispanic, non-white, never married, have less than a high school education, prefer to speak Spanish, and live in low income households.[79]

Consented children in the NIS can be missing provider data for several reasons. While provider nonresponse or lack of provider records accounts for a portion of cases with missing vaccination data, missing provider data can also stem from poor quality or incorrect provider contact information supplied by the respondent. Although some mothers unintentionally provide incorrect contact information, others may intentionally provide incorrect information or falsify it to prevent NIS from contacting their healthcare provider – a passive means of refusing the consent request. Thus, we hypothesize that the predictors of missing provider data in the 2002 NIS identified by Smith et al. (2005) similarly predict consent refusal in the 2009 NIS.

In addition to these demographic and socioeconomic characteristics of the child's mother, we also investigate whether similarity of interviewer and respondent demographic characteristics affects consent. Such similarity can positively influence survey participation (Webster, 1996) and may influence consent in the same way, although an analysis of this issue in the 2008 HRS showed no significant effect on consent likelihood (Sakshaug et al., 2012).

---

[79] NIS collects demographic and socioeconomic characteristics about the child's mother, even in cases where the respondent is not the child's mother.

We also predict that respondents who initially refused to participate in NIS will be more likely to refuse the consent request. Related research by Sakshaug et al. (2012) using data from the 2008 HRS, a panel survey, found that indicators of respondent resistance such as number of call attempts and confidentiality concerns in the prior wave were negatively associated with consent refusal in the 2008 wave of the survey. Prior survey refusals can also indicate privacy concerns (Steeh, 1981). The relationship between consent refusal and privacy concerns is explored in Chapter 3 of this dissertation.

NIS sample members with directory listed telephone numbers for whom an address match can be made are mailed an advance letter (Zell, Ezzati-Rice, Battaglia, & Wright, 2000).[80] While the NIS advance letter does not specifically mention the consent request, advance notification of the survey can reduce overall suspicion of the survey request and stress the legitimacy of the research endeavor. Existing research generally concludes a positive effect of such letters on telephone survey response rates (de Leeuw, Callegaro, Hox, Korendijk & Lensvelt-Mulders, 2007; Dillman, Gallegos, & Frey, 1976; Traugott, Groves, & Lepkowski, 1987)[81], including the NIS (Camburn, Lavrakas, Battaglia, Massey, & Wright, 1995). It is hypothesized that advance letters will have a similarly positive influence on respondents' consent behavior.[82]

---

[80] However, past research finds that survey response rates are higher among sample members with listed contact information as reasons for unlisted status may be correlated with response refusal (Traugott, Groves, & Lepkowski, 1987).Thus, any observed effects of the advance letter may be due to inherent differences in sample composition.

[81] At least one study, Singer, Van Hoewyk, and Maher (2000), found no effect of advance letters on telephone survey response rates.

[82] Some NIS sample members are also mailed a monetary incentive. Sample members can only qualify for a monetary incentive of $15 if they refuse to complete the screener or survey after indicating they have an eligible child, or if they refuse the consent request. Because incentive receipt and refusal are confounded, the effect of incentives on consent is not investigated in the present research.

4.4 Methods

4.4.1 NIS

4.4.1.1 Overview

Since 1994, NIS has produced annual estimates of vaccination coverage in U.S. children between the ages of 19 and 35 months.[83] NIS is conducted by the National Center for Immunization and Respiratory Diseases of the Centers for Disease Control (CDC) and the National Center for Health Statistics (NCHS); the 2009 survey was administered by the National Opinion Research Center (NORC). NIS was implemented in order to address the 1992 Childhood Immunization Initiative to monitor vaccine coverage of young children and improve the delivery, awareness, and cost of vaccinations (CDC, 2010).

NIS employs an RDD sample to identify households with children ages 19-35 months old. Although in the first six months of 2009, approximately 21.3% of children lived in cellphone-only households (Blumberg and Luke, 2009), the 2009 NIS was limited to households with landlines. Forty-six percent of sample cases were mailed an advance letter.

Of households completing the screener, 2.8% had age-eligible children. The overall CASRO response rate for this year was 63.7%, resulting in 23,474 completed household interviews regarding 24,809 children. Of children with completed interviews, 79.8% had permission to contact vaccination providers.[84] Providers returned 93.9% of immunization history questionnaires in 2009, and overall, 68.7% of children had

---

[83] Starting in 2009, NIS conducted interviews in the U.S. Virgin Islands. All 432 U.S. Virgin Islands cases are excluded from estimates and figures reported in this chapter.

[84] Weighted consent rate; unweighted consent rate is 79.3%.

adequate vaccination provider data (CDC, 2010). Children could lack adequate provider data because the respondent did not give consent, provided incorrect or insufficient provider contact information, the provider did not have any records for the child, or the provider did not return the survey.

Each NIS interview includes two data collection components: a household CATI survey with respondents who have one or more children in the target population, and a self-administered questionnaire completed by the child's vaccination provider which providers typically return by mail or fax; some provider interviews are completed by telephone. The provider interview is conditional upon respondents' consent, requested during the NIS household survey.

4.4.1.2 Household Survey

The household survey contains a screener -- to identify households with one or more members of the target population -- and a survey interview. Households containing children aged 19-35 months are asked to complete a survey which includes questions on the child's vaccination history (Section A and Section B), demographic and socioeconomic questions regarding the child and mother (Section C), vaccination provider contact information and the consent request (Section D), and if the respondent consents to vaccination provider contact, they are asked about the child's health insurance coverage in a final supplemental module of questions (Section E). If a household has multiple eligible children, respondents are asked the survey questions for each child and are asked to provide consent and vaccination provider contact information for each child individually. An interview is considered complete if respondents complete Sections A, B, and C. Complete interviews may lack the vaccination provider consent and health insurance modules (Sections D and E), although respondents are encouraged to complete them and respondents who initially refused consent may be recontacted.

Because a NIS interview is considered complete before the consent module, it is possible that some cases in the 2009 dataset were not asked for consent. Consent, the dependent variable in the following analyses, lacks any information that would distinguish between respondents who completed the survey but not Section D, and respondents who explicitly refused the direct consent request. As the explicit consent request is the last question in the consent module and item completion information was unavailable, we cannot determine if respondents were asked and refused the explicit consent request, or if the respondent refused an earlier question in the consent module before the explicit consent request was asked. The inability to distinguish between respondents who refused the request from those not asked is because the data owners could not make item and section completion information available for use in this project.[85] Interviewers attempt to gain consent from all NIS respondents, but may not achieve it if there is a hostile refusal prior to the request, or if the case reaches its maximum number of refusals prior to the request.

It is towards the end of the questionnaire that the interviewer requests respondents' oral consent to contact the child's vaccination provider (or providers) (Battaglia et al., 2000; Smith et al., 2001). The consent request is comprised of a series of several related questions. If at any point a respondent refuses a question in the consent

---

[85] Section completion variables and item completion variables with the consent module were requested in the RDC proposal and approved for use in this project. However, after the project's approval, it was determined that they had never been collected in the 2009 NIS.

series, respondents are considered to have withheld consent and the interview is terminated, though these respondents may be recontacted for refusal conversion.[86]

As an introduction to the consent section, the interviewer explains the purpose of this module: NIS is interested in contacting the child's healthcare provider to obtain a copy of their vaccination records. First, interviewers ask respondents for the number of providers who vaccinated their child. Respondents are then asked for the contact information for each of these providers and interviewers look up this information interactively in a database. If interviewers cannot find the provider, they enter their contact information into the database manually.

Respondents are subsequently asked for the child's full name and their own full name to facilitate record identification with each vaccination provider. Once the interviewer collects these pieces of information, they ask if the respondent can authorize the release of the child's immunization records. All respondents are asked the authorization question to ensure that before specifically requesting consent, the interviewer is speaking with someone who can authorize the release of the vaccination records. If the respondent says they have the authorization to release this information, the interviewer assures them that the vaccination records collected from the healthcare provider will be kept in strict confidence.

The section culminates with the request for respondent's consent to contact their healthcare provider(s) for the child's vaccination records. A series of on-screen FAQs help interviewers to address any respondent concerns during the entire consent module.

---

[86] An examination of the disposition codes in the 2009 NIS call records suggest that only 180 completed cases refused consent on a prior contact. Of these, 3 consented on later contacts. It is possible that a greater number of cases refused and then consented, but that this was not indicated in the disposition codes, for example, if interviewers inconsistently recorded this information.

4.4.1.3 Provider Record Check Study

The Provider Record Check Study comprises the second portion of the NIS. If oral consent is obtained, each provider named by the respondent is mailed an Immunization History Questionnaire (IHQ), a short form designed to minimize provider burden and maximize response. The IHQ requests a list of all vaccinations administered to the sample child, and if not returned, providers are recontacted several times. IHQ responses are then entered, edited, and cleaned, and merged with the household survey data. Data from the household telephone survey and provider record check study are aggregated to provide annual national vaccine coverage estimates (Battaglia, 1997).

If available, provider records are preferred over household data as the CDC considers these records to be more accurate than respondent-contributed information, "Thus, the most important subsets of the data consist of children with adequate provider data" (CDC, 2010: 29). However, if respondents refuse consent to contact vaccination providers, estimates of vaccination coverage are produced through respondents' reports of this information in the household interview. Unfortunately, due to parental misreporting of vaccination histories, surveys estimates of childhood vaccinations can be highly inaccurate (Lee et al., 1999).

4.4.2 Data Source

The data used to investigate the research questions in this chapter is the product of two restricted-access data files from the 2009 NIS. The first file contains a record of calls made to all sampled phone numbers, including prior survey refusals, the characteristics of the interviewer making each call, and an indicator of whether or not an advance letter was mailed to the sample member. NORC created this file for this project. CDC produced the second file, which includes only cases with completed interviews, contains information about the respondent and the child's mother, and indicates whether or not consent was obtained. All analyses were conducted in the NCHS Research Data Center (RDC) and RDC staff reviewed all results for disclosure risk.

Merging the call record file with the completed case data required some simplification of the record of calls. Specific information on the particular interviewer that requested consent for a given case was unavailable. As a proxy for this exact information, we assume that the interviewer who made the final call to each completed case in the dataset is also the interviewer who requested consent, as the consent request falls toward the end of the interview. Further, the interviewer may not have reached the individual most knowledgeable about the child's vaccination history on the initial call, and scheduled a call back with the more knowledgeable respondent for a later time (Zell et al., 2000). We merged the two files using child identification number, final call date, and final dial count. Merging the files using this information rather than information identifying the interviewer requesting consent potentially limits the accuracy of the findings in the following analyses.[87]

4.4.2.1 Interviewer Data

In total, 762 interviewers completed interviews on the 2009 NIS. Over three-quarters of interviewers made calls out of a NORC calling center in Chicago, and the remainder made calls from a calling center contracted by NORC in Las Vegas. NORC

---

[87] This assumption may not hold if interviewers recontact respondents with the sole purpose of completing the final section of the survey (Section E) which falls after the consent request. If respondents complete Sections C and D on the same call, Section E will also be attempted on that call. If respondents refuse the consent request and are recontacted, Section E will be attempted during the recontact. However, it is also possible that interviewers may recontact respondents with completed interviews to finish Section E, though it is unknown how frequently this occurs.

Chicago interviewers self-reported their age, gender, and race/ethnicity information when they were hired; however, approximately one-quarter of these interviewers refused to release this information for research purposes. Age, gender and race/ethnicity information could not be released for any of the contracted interviews working in the Las Vegas calling center. Information that identified whether an interviewer was located in Chicago or Las Vegas could not be included on the dataset.

Overall, the Chicago interviewers who refused to release their demographic characteristics account for approximately one-half of missing interviewer data, and the contractor interviewers in Las Vegas account for the remainder (see Table 4.01). In total, age, gender, and race/ethnicity information is available for 60.9% of interviewers. NIS experience information is available for all interviewers.

Table 4.01

Available Interviewer Data

Interviewers	Missing Data Status	% of Interviewers (n)	% of Interviews (n)	Available Data
NORC - Chicago	Released demographics	60.9% (464)	64.5% (15,997)	Experience, age, gender, race/ethnicity
NORC - Chicago	Refused to release demographics	~17.7%† (~135)	35.5% (8,812)	Experience
Contractor - Las Vegas	Contractor could not release demographics	~21.4%† (~163)		Experience

Notes: Unweighted data. †Percent and number of interviewers missing demographic characteristics from each location is estimated based on information provided from NORC.

4.4.3 Analyses

The following analyses examine the previously described characteristics of the mother, respondent, and interviewer on consent to contact vaccination providers in the 2009 NIS. Because NIS collects the demographic and socioeconomic characteristics specifically about the child's mother, even in cases where the respondent is related to the child in another way, all of the following analyses are conducted on the total sample of respondents as well as a sample of respondents limited only to the mothers of sample children.

First, descriptive analyses evaluate the association between mother, respondent, and interviewer covariates and consent to vaccination provider contact, the dependent variable. Next, these relationships are analyzed through a series of logistic regression analyses. First, we evaluate the relationship between characteristics of the child's mother and respondent with consent likelihood. A second set of models incorporates interviewer characteristics, as well as interactions between interviewer and respondent characteristics. We also estimate both sets of regression models excluding interviewers who are missing demographic characteristics to determine whether predictors of consent differ for these interviewers. Wald Tests evaluate the contribution of categorical predictors and Archer and Lemeshow's Goodness of Fit test for survey data assess model fit (Archer et al., 2006). Interviewer variability in obtaining respondent consent is estimated through intraclass correlations, calculated using one-way ANOVAs with random effects.

Unless otherwise noted, all analyses account for interviewer clustering of cases and incorporate weights that reflect the stratified sample design of the NIS, adjusts for unit nonresponse, and are poststratified to population control totals excluding children in non-telephone households (CDC, 2010). We conduct all analyses using Stata 11 software.

4.5 Results

4.5.1 Descriptive Statistics

4.5.1.1 Respondent Data

Seventy-eight percent of 2009 NIS respondents were the child's mother, 14%, were the child's father, 6% was a grandparent, and the remainder consisted of other family and friends of the child. About 95% of respondents completed an interview on one child only.

Overall, 79.8% of respondents and 81.9% of mothers consented to vaccination provider contact. Table 4.02 presents consent rates for all respondents and mothers in the 2009 NIS by characteristics of the child's mother and respondent. Although many significant relationships exist between these characteristics and consent in the bivariate analyses, as discussed below, many of these associations are not in the directions predicted, particularly for the characteristics of the child's mother.

Hispanic ethnicity of the child's mother is positively associated with consenting ($p<0.01$), counter to the hypothesized relationship. This is true for cases when the child's mother is the respondent, and as mothers comprise nearly 80% of completed interviews, we observe a similar trend in the total sample as well. Mother's educational attainment is associated with consenting to vaccination provider contact, but also not in the hypothesized direction. Here, for all respondents and in cases where the respondent is the child's mother, less maternal education is associated with higher rates of consent ($p<0.001$).

Table 4.02

Consent Rates among All Respondents and Mothers, by Sample Characteristics

	All Respondents (n=24,809)	Mothers (n=19,681)
	%	%
Mother Characteristics		
Ethnicity		
Hispanic *(25.4%)*†	**82.9**	**85.3**
Non-Hispanic *(74.6%)*	**78.8**	**80.7**
(n=24,809)		
Race		
White *(76.9%)*	79.6	81.7
Black *(15.5%)*	80.6	82.2
Other *(7.6%)*	80.6	83.3
(n=24,809)		
Education		
<12 years *(19.6%)*	*84.6*	*87.0*
12 years *(30.9%)*	*80.1*	*83.2*
Some College *(19.1%)*	*78.5*	*79.7*
College Grad *(30.4%)*	*77.8*	*79.2*
(n=24,615)		
Marital Status		
Married (67.3%)	**78.9**	*80.5*
Never Married (24.7%)	**83.0**	*86.3*
Widowed/Divorced/Separated (8.0%)	**79.2**	*81.3*
(n=24,715)		
Age[88]		
<30 *(41.0%)*	*82.3*	*85.2*
30+ (59.0%)	*79.1*	*80.4*
(n=24,559)		
2008 Mean Family Income *($ 68,787)* *(n=24,809)*	$65,704 *(consenters)* $80,992 *(non-consenters)*	$60,651 *(consenters)* $63,482 *(non-consenters)*
Respondent Characteristics		
Language of Interview		
English *(83.6%)*	*78.7*	*80.4*
Non-English *(16.4%)*[89]	*85.8*	*88.8*
(n=24,809)		
Prior Survey Refusal		
One or More *(8.5%)*	*20.6*	*21.5*
None *(91.5%)*	*85.3*	*86.9*
(n=24,809)		
Advance Letter		

---

[88] Narrower age categories could not be created due to the small number of respondents in some cells, preventing release of findings from the NCHS RDC.

[89] Less than 1% of cases in the total sample were conducted in a language other than English or Spanish. As these cases could not be analyzed separately given confidentiality concerns, to avoid dropping them from analyses, they are combined with interviews conducted in Spanish.

Mailed *(65.3%)*	80.7	82.5
Not Mailed *(34.7%)*	78.1	80.9
(n=24,809)		

Notes: *Italics* indicates differences by consent status $\chi2$ p<0.001; **bold** indicates differences by consent status $\chi2$ p<0.01; underline indicates differences by consent status $\chi2$ p<0.05. †Distribution and sample size among all respondents.

The relationship between mother's marital status and consent is significant among all respondents (p<0.01) and among mothers (p<0.001), though not in the direction predicted: rates of consent to vaccination provider access are higher for children whose mothers never married. Mother's age is also significantly associated with consenting, yet the direction is inconsistent with our hypothesis: among all respondents and mothers, consent rates are lower among mothers over thirty years old as compared to those under thirty years old (p<0.001). Lastly, no relationship exists in the 2009 NIS between mother's race or income and consent.[90]

Associations between mothers' characteristics evaluated here, including Hispanic origin, education, race, income status, marital status and age, and consent are either nonexistent or not in the predicted direction. As the hypotheses for these analyses were based upon predictors of missing provider data in the 2002 NIS, this suggests that the factors contributing to missing provider records and consent refusal are different. Another explanation is that the factors contributing to missing provider data have changed since the analysis was conducted in 2002. It is possible that an analysis of the 2009 NIS could indicate that factors associated with missing provider data are associated with consent refusal in the more recent data.

---

[90] No relationship between income and consent exists when analyzing income as a categorical or continuous variable, and is thus included in the following models as a continuous predictor. "Don't Know" and "Refused" responses for this variable were imputed to substantive responses by CDC before receiving the data and so we could not examine the relationship between providing a "Don't Know" or "Refused" response to the income item and consent.

We hypothesized that respondents who prefer to speak Spanish (or another language) would be less likely to consent to vaccination provider contact, yet we found the opposite to be true for respondents in the sample overall and for cases where the child's mother is the respondent ($p<0.001$).[91] As predicted, prior survey refusal is positively associated with consent refusal among all respondents and mothers ($p<0.001$). Finally, in line with our hypothesis, the advance letter facilitated consent among all respondents ($p<0.05$), yet it did not have a statistically significant impact on consent rates among mothers.

4.5.1.2 Interviewer Data

Interviewers releasing demographic information were between the ages of 18-75, with a mean age of 43 years old and a median age of 26 years old. Nearly three-quarters of interviewers were female (74%). Seventy-two percent were non-Hispanic black, 13% were Hispanic, 9% were non-Hispanic white, and 7% were of another race.[92]

We quantified NIS experience as the number of days between each completed interview and the interviewers' first dial on the NIS survey (this may be prior to the 2009 administration). NIS experience ranged from 0 to 1,830 days, or 5.0 years, with a mean of 1.8 years and a median of 1.4 years. Figure 4.01 illustrates the distribution of completed interviews by days of interviewing experience. As shown, a disproportionate number of interviews were completed by less experienced interviewers.

---

[91] Excluding interviews conducted in a language other than English or Spanish does not change the direction or significance of this analysis.

[92] All interviewer characteristics reported pertain to NORC Chicago interviewers. As a reference, according to the 2006-2010 ACS 5-year sample, 46% of Chicago adults are White, 32% are Black, and 25% are Hispanic (retrieved from http://usa.ipums.org/usa/sda/).

Figure 4.01

Number of Completed Interviews by Interviewing Experience

Table 4.03 presents consent rates for all respondents and mothers, by characteristics of interviewers who completed interviews on the 2009 NIS. Female interviewers obtain slightly higher consent rates as compared to male interviewers in the bivariate analyses, yet interviewers missing demographic information obtained the highest consent rates. Examining consent rates by interviewer age demonstrates that again, interviewers missing demographic information are most effective in gaining consent. Among those interviewers for whom this information was available, the youngest interviewers are somewhat more likely to obtain consent. Interviewers vary in the consent rates they achieved by their race/ethnicity. White interviewers obtained the lowest consent rates, and the small number of interviewers who identify their racial and

ethnic makeup as "Other" obtained the highest consent rates, even outperforming those with missing demographic information.

Table 4.03

Consent Rates among All Respondents and Mothers, by Interviewer Characteristics

	All Respondents (n=24,809)	Mothers (n=19,431)
	%	%
Gender		
Female *(45.0%)*[†]	79.1	*81.4*
Male *(15.9%)*	76.9	*79.1*
Missing (39.1%)	82.5	*84.2*
(n=762)		
Age		
18-30 *(41.1%)*	79.8	81.9
31-45 *(9.3%)*	76.1	81.1
46+ *(10.5%)*	76.1	77.6
Missing (39.1%)	82.5	84.2
(n=762)		
Race/Ethnicity		
Hispanic *(7.9%)*	80.7	*83.2*
White *(5.3%)*	74.6	*77.3*
Black *(43.8%)*	77.9	*80.1*
Other *(3.9%)*	83.1	*84.6*
Missing (39.1%)	82.5	*84.2*
(n=762)		
Experience		
0-6 mos *(25.1%)*[††]	*90.1*	*91.5*
6 mos - 1 year *(17.1%)*	*82.0*	*84.6*
> 1year *(57.8%)*	*75.1*	*77.3*
(n=24,809)		

Notes: *Italics* indicates differences by consent status $\chi 2$ p<0.001; underline indicates differences by interviewer data missingness $\chi 2$ p<0.05; *italics and underline* indicates differences by interviewer data missingness $\chi 2$ p<0.10. [†] Distribution and sample size of interviews conducted among all respondents. [††] Interviewer experience varies by interview conducted, rather than interviewer.

As illustrated in Table 4.03, when interviewers with missing gender, age, and race/ethnicity information are included in the bivariate analyses, consent rates significantly vary as a function of each of these characteristics. This is true among all respondents as well as for mothers of sample children.

However, when excluding the 298 interviewers with missing data from the bivariate analyses reported in Table 4.03, consent rates do not significantly vary as a function of interviewer characteristics (results not shown). Although excluding these interviewers reduces sample sizes, it appears that the significance of these bivariate analyses in Table 4.03 are largely a function of differences in consent rates contributed by interviewers with missing demographic data. Interviewers with missing demographic characteristics were more effective in obtaining respondents' consent, as compared to interviewers for whom this information is available. As shown in Table 4.04, consent rates obtained by interviewers with missing demographic characteristics are approximately four percentage points higher than those with nonmissing data.[93]

Table 4.04

Consent Rates by Interviewer Missingness, among All Respondents and Mothers

	All Respondents (n=24,809)		Mothers (n=19,431)	
	Int. Data Available (464 interviewers, 15,997 interviews)	Int. Data Missing (298 interviewers, 8,812 interviews)	Int. Data Available (462 interviewers, 12,686 interviews)	Int. Data Missing (293 interviewers, 6,745 interviews)
Consent				
Yes	<u>78.5</u>	<u>82.5</u>	<u>80.8</u>	<u>84.2</u>
No	<u>21.5</u>	<u>17.5</u>	<u>19.2</u>	<u>15.8</u>

Notes: <u>Underline</u> indicates differences by interviewer data missingness $\chi 2$ p<0.05.

We also find that compared to more experienced interviewers, a greater portion of inexperienced interviewers have missing data. Of interviewers with 0-6 months of NIS experience, 45% are missing demographic data, compared to 37% of interviewers with 6 months to 1 year of experience, and 28% of interviewers with more than 1 year of experience (p<0.01). This is likely because characteristics could not be released for any

---

[93] Some characteristics of mothers and respondents vary as a function of missing interviewer demographic characteristics, suggesting that interviewers with and without missing demographic data complete interviews with different respondents (see Table 4.01 in Appendix).

Las Vegas interviewers, and these interviewers tended to have less experience on the NIS than Chicago interviewers (Skalland, B. of NORC, personal communication, July 5, 2012).

Consent rates also vary significantly as a function of NIS interviewing experience (p<0.001). Among all respondents and mothers, interviewers with less prior interviewing experience obtained higher consent rates (p<0.001). Over 90% of respondents consented to vaccination provider contact in interviews conducted with less than six months of prior interviewing experience, and nearly 95% of respondents consented in interviews conducted with fewer than 90 days of interviewing experience.[94] The direction and significance of this bivariate analysis does not change when excluding interviewers with missing demographic information, indicating that the inverse relationship between interviewing experience and consent rates exists for all interviewers in the 2009 NIS, regardless of their location or willingness to provide this information (see Table 4.05).

Table 4.05

Consent Rates by Interviewer Missingness and Prior NIS Experience, among All Respondents and Mothers

	All Respondents (n=24,809)		Mothers (n=19,431)	
	Int. Data Avail. (464 interviewers, 15,997 interviews)	Int. Data Missing (298 interviewers, 8,812 interviews)	Int. Data Avail. (462 interviewers, 12,686 interviews)	Int. Data Missing (293 interviewers, 6,745 interviews)
Experience				
0-6 mos	*88.6*	*91.1*	*90.1*	*91.7*
6 mos - 1 year	*79.8*	*82.6*	*81.9*	*84.2*
> 1year	*73.5*	*76.5*	*75.6*	*78.8*

Notes: *Italics* indicates differences by consent status $\chi 2$ p<0.001.

---

[94] Considering the small number of respondents who refused consent in interviews conducted with less than 90 days of interviewing experience, to adhere to RDC policies this interviewing interval was not used in analyses to protect interviewers' confidentiality.

However, when examining the relationship between interviewing experience and consent only on the 2,117 respondents who previously refused survey participation, there is no association between consent and prior interviewing experience (p=0.155). This is likely due to the targeted assignment of more experienced interviewers to these cases for refusal conversion efforts, and as hypothesized, these respondents may also be less willing to consent to vaccination provider contact.

4.5.2 Logistic Regression Analyses and Intraclass Correlations

4.5.2.1 Logistic Regression Analyses: Respondent Characteristics

We next constructed logistic regression models predicting consent to vaccination provider contact as a function of characteristics of the child's mother (Hispanic ethnicity, race, education, marital status, age, and 2008 family income) and respondent (language of interview, prior survey refusals, and advance letter). The model that incorporates all survey respondents also includes an indicator identifying whether the survey respondent is the child's mother. The dependent variable, consent to vaccination provider contact, is a binary variable that takes the value of one if the respondent granted consent and zero if the respondent withheld consent.

We also assess whether predictors of consent vary in conjunction with whether or not interviewers are missing demographic information. We compare results from models that include all interviewers (Table 4.06) to models that include only interviewers with missing demographic information (Table 4.08), and only interviewers with nonmissing demographic information (Table 4.08). We first discuss results for the set of models that include all interviewers.

Table 4.06

Logistic Regressions Predicting Consent as a Function of Sample Characteristics among All Respondents and Mothers (All Interviewers)

		All Respondents (n=24,354)			Mothers (n=19,233)		
		Coef.	SE	*p*	Coef.	SE	*p*
Constant		1.61	0.20	0.000	2.08	0.22	0.000
Mother Characteristics							
Ethnicity	Non-Hispanic (ref.)						
	Hispanic	-0.04	0.12	0.736	-0.13	0.14	0.357
Race	White (ref.)						
	Black	0.00	0.10	0.963	-0.13	0.10	0.222
	Other	0.16	0.15	0.300	0.07	0.19	0.700
Education	<12 yrs (ref.)						
	12 yrs	-0.31	0.13	0.019	-0.16	0.16	0.296
	>12 yrs, non col. grad	-0.47	0.14	0.001	-0.39	0.16	0.019
	Col. grad	-0.40	0.14	0.005	-0.30	0.17	0.080
Marital Status	Married (ref.)						
	Never Married	0.09	0.15	0.568	0.20	0.17	0.246
	Wid./Div./Sep.	0.03	0.13	0.798	0.01	0.15	0.948
Age	<30 (ref.)						
	30+	-0.12	0.08	0.113	-0.21	0.09	0.021
2008 Family Income		0.00	0.00	0.516	0.00	0.00	0.399
Relationship to Child	Not Mother (ref.)						
	Mother	0.49	0.08	0.000	----	-----	----
Respondent Characteristics							
Language of Int.	English (ref.)						
	Non-English	0.23	0.20	0.243	0.49	0.21	0.019
Prior Survey Refusal	None (ref.)						
	One or More	-3.09	0.12	0.000	-3.15	0.13	0.000
Advance Letter	Not Mailed (ref.)						
	Mailed	0.19	0.07	0.006	0.15	0.07	0.044

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit for either model: All Respondents [F(9,750) = 0.73; Prob > F = 0.684]; Mothers [F(9,744) = 1.15; Prob > F = 0.323].

Table 4.07

Wald Tests for Categorical Parameters in Logistic Regressions Predicting Consent as a Function of Sample Characteristics among All Respondents and Mothers (All Interviewers)

	All Respondents		Mothers	
Categorical Predictor	F-Test Statistic	*p*	F-Test Statistic	*p*
Race	$F_{(2, 757)} = 0.54$	0.585	$F_{(2, 751)} = 0.84$	0.432
Education	$F_{(3, 756)} = 3.56$	0.014	$F_{(3, 750)} = 2.52$	0.057
Marital Status	$F_{(2, 757)} = 0.22$	0.800	$F_{(2, 751)} = 1.40$	0.247

When controlling for other sources of variation, fewer characteristics of the mother and respondent are strong predictors of consent to vaccination provider contact as compared to the bivariate analyses. Of the characteristics associated with consent in these models, once again, several are not in the direction initially hypothesized, including mothers' education in both models in Table 4.06, and mothers' age in cases where the child's mother is the respondent.

Among all respondents and mothers, children whose mothers have more education are more likely to withhold consent to vaccination provider contact, relative to those with less than twelve years of education. As shown in Table 4.07, the Wald X^2 for the multi-category predictor for education is statistically significant among all respondents ($p<0.05$) and mothers only ($p<0.10$), suggesting that overall, mother's education is an important predictor of consent after controlling for relationships with other variables in the model. For cases where the mother is the respondent, older mothers (those 30 years old and older) are less likely to grant consent as compared to younger mothers (those under 30 years old; $p<0.05$).

We also find that the log odds of consenting are significantly higher when the mother is the respondent as compared to when the child has some other relationship with the respondent ($p<0.001$). This is consistent with the differences in consent rates for mothers as compared to non-mothers (81.9% vs. 72.4%, respectively; $p<0.01$).

After controlling for other sources of variation, additional characteristics of the mother, including Hispanic ethnicity, race, marital status, and income, are not statistically significant predictors of consent among all respondents or mothers of sample children.

One or more prior survey refusals is a strong negative predictor of consent for both models presented in Table 4.08 ($p<0.001$). The advance letter positively influences consent decisions among all respondents ($p<0.01$) and mothers of sample children ($p<0.05$). The significance of the coefficient among all respondents as compared to mothers only could suggest that the advance letter has more of a positive impact on consent among non-mothers, although this may also be related to the smaller sample size for the mothers-only analysis. Finally, among mothers only, the likelihood of consenting was greater when the interview was conducted in a language other than English ($p<0.05$).

The next set of regression models include the same predictors, but divide the sample by interviewers missing and not missing demographic information. Here, we investigate whether the characteristics associated with consenting vary in conjunction with whether or not interviewers are missing demographic information. Unfortunately, as this dataset lacks an indicator for interviewers' location, the following analyses cannot assess whether any variation is related to interviewer location (Chicago or Las Vegas) or reason for missing data (refusal to release demographic characteristics or contractor policy to withhold this information). These models are presented in Table 4.08 and demonstrate that predictors of consent do vary by interviewer sample.

Perhaps because interviewers with nonmissing demographic information conducted nearly two-thirds of interviews, the models including interviewers with nonmissing demographic characteristics largely resemble those that include all interviewers. One clear difference is the coefficient for children with mothers who never married. Although not a significant predictor of consent in the models that include all interviewers, among cases with nonmissing interviewer data, consent is significantly higher among these children (compared to married mothers), among all respondents ($p<0.10$) and mothers only ($p<0.05$). However, marital status is not a significant predictor of consent overall in these models (see Table 4.08).

Interviewed by interviewers with nonmissing data, Hispanic mothers of sample children are less likely to consent ($p<0.10$), as are those with a high school education ($p<0.01$). These factors do not predict consent in the model that includes all interviewers.

Compared to the model with all interviewers, predictors of consent vary more substantially in the model that includes only interviewers with missing interviewer data (see Table 4.08). Among all respondents and mother's, education no longer predicts consent, nor does the advance letter. Further, compared to children with married mothers, those with widowed, divorced, and separated mothers are significantly less likely to have consent to provider contact ($p<0.10$). Marital status overall is a significant predictor of consent in the model including mothers only ($p<0.10$; see Table 4.09 for the Wald test for the overall categorical parameter). As with the models that include all interviewers and interviewers with nonmissing demographic characteristics, being the sample child's mother continues to positively predict consent among interviewers with missing demographic characteristics ($p<0.05$), and having one or more prior survey refusals still negatively predicts consent ($p<0.001$).

Table 4.08

Logistic Regressions Predicting Consent as a Function of Sample Characteristics among All Respondents and Mothers (Interviewers with Nonmissing and Missing Demographic Characteristics)

		Nonmissing Interviewer Characteristics						Missing Interviewer Characteristics					
		All Respondents (n=15,690)			Mothers (n=12,548)			All Respondents (n=8,663)			Mothers (n= 6,684)		
		Coef	SE	*p*	Coef	SE	*p*	Coef.	SE	*p*	Coef	SE	*p*
Constant		1.33	0.23	0.000	1.92	0.26	0.000	2.17	0.42	0.000	2.44	0.43	0.000
Mother Chars.													
Ethnicity	Non-Hispanic (ref.)												
	Hispanic	-0.14	0.14	0.308	-0.27	0.16	0.082	0.19	0.24	0.432	0.23	0.30	0.434
Race	White (ref.)												
	Black	-0.07	0.10	0.521	-0.13	0.12	0.293	0.18	0.20	0.368	-0.06	0.18	0.758
	Other	0.19	0.19	0.331	0.13	0.23	0.565	0.03	0.25	0.916	-0.12	0.31	0.691
Education	<12 yrs (ref.)												
	12 yrs	-0.36	0.14	0.008	-0.30	0.18	0.094	-0.12	0.28	0.666	0.21	0.32	0.509
	>12 yrs, non col. grad	-0.54	0.15	0.000	-0.52	0.19	0.007	-0.27	0.31	0.397	-0.05	0.33	0.878
	Col. grad	-0.57	0.14	0.000	-0.56	0.19	0.003	0.01	0.31	0.963	0.30	0.36	0.393
Mar. Status	Married (ref.)												
	Never Married	0.31	0.18	0.077	0.41	0.20	0.042	-0.40	0.31	0.203	-0.28	0.34	0.405
	Wid./Div./Sep.	0.24	0.16	0.125	0.24	0.18	0.188	-0.43	0.24	0.071	-0.52	0.25	0.040
Age	<30 (ref.)												
	30+	-0.10	0.10	0.314	-0.14	0.11	0.205	-0.19	0.12	0.140	-0.41	0.16	0.011
2008 Fam. Inc.		0.00	0.00	0.649	0.00	0.00	0.746	0.00	0.00	0.546	0.00	0.00	0.186
Rel. to Child	Not Mother (ref.)												
	Mother	0.60	0.09	0.000	----	-----	----	0.31	0.14	0.024	----	----	----
Resp. Chars.													
Language of Int.	English (ref.)												
	Non-English	0.25	0.21	0.251	0.50	0.22	0.024	0.41	0.37	0.276	0.65	0.49	0.184

Prior Surv. Refusal	None (ref.)												
	One or More	-3.27	0.15	0.000	-3.21	0.16	0.000	-2.89	0.21	0.000	-3.11	0.21	0.000
Advance Letter	Not Mailed (ref.)												
	Mailed	0.20	0.08	0.011	0.19	0.08	0.025	0.11	0.12	0.356	0.00	0.14	0.982

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests a slight lack of fit only for the model among Missing Interviewer Data - All Respondents. Nonmissing Interviewer Data – All Respondents [F(9,455) = 0.47; Prob > F = 0.893]; Mothers [F(9,453) = 0.49; Prob > F = 0.881]; Missing Interviewer Data – All Respondents [F(9,286) = 1.68; Prob > F = 0.094]; Mothers [F(9,282) = 0.50; Prob > F = 0.875];

Table 4.09

Wald Tests for Categorical Parameters in Logistic Regressions Predicting Consent as a Function of Sample Characteristics among All Respondents and Mothers (Interviewers with Nonmissing and Missing Demographic Characteristics)

	Nonmissing Interviewer Characteristics				Missing Interviewer Characteristics			
	All Respondents		Mothers		All Respondents		Mothers	
Categorical Predictor	F-Test Stat.	*p*	F-Test Stat.	*p*	F-Test Stat.	*p*	F-Test Stat.	*p*
Race	$F_{(2,462)}=0.65$	0.524	$F_{(2,460)}=0.74$	0.479	$F_{(2,293)}=0.41$	0.667	$F_{(2,289)}=0.11$	0.892
Education	$F_{(3,461)}=5.51$	0.001	$F_{(3,459)}=3.57$	0.014	$F_{(2,292)}=1.29$	0.279	$F_{(3,288)}=1.55$	0.202
Marital Status	$F_{(2,462)}=1.59$	0.204	$F_{(2,460)}=2.15$	0.118	$F_{(2,293)}=1.77$	0.172	$F_{(2,289)}=2.67$	0.071

As demonstrated through the logistic regression models presented in Tables 4.06 and 4.08, some sample characteristics associated with consent vary in conjunction with whether interviewers are missing demographic information. More specifically, sample characteristics associated with consent vary more notably in the model with interviewers with missing interviewer characteristics, as compared to the other models. Because of this, the next series of logistic regression models are estimated on all interviewers as well as interviewers with nonmissing interviewer characteristics only. We also calculate the intraclass correlations in the following section using a variety of samples, including interviewers with missing and nonmissing data.

4.5.2.2 Intraclass Correlations

Intraclass correlations (ρ_{int}), or the proportion of between group variance to total variance (Hox et al., 1994), show that interviewers vary in their ability to obtain respondents' consent to vaccination provider contact (see Table 4.10). Further, the size of the intraclass correlation varies slightly by the sample (all respondents or mothers only) and the interviewers included (all interviewers or only those with missing or nonmissing data), with between 4-9% of variance in consent rates due to interviewer variation.

Table 4.10

Intraclass Correlations among All Respondents and Mothers, for All Interviewers and Interviewers with Nonmissing Demographic Characteristics

Sample (# of interviews)	ρ_{int}	SE	95% CI
All Interviewers			
Total (24,809)	0.07	0.01	0.06, 0.08
Experience			
0-6 mos (6,224)	0.08	0.01	0.06, 0.10
6 mos - 1 year (4,254)	0.07	0.01	0.05, 0.09
> 1 year (14,331)	0.05	0.01	0.04, 0.07
Mothers (19,431)	0.06	0.01	0.05, 0.07
Nonmissing Interviewer Characteristics			
Total (15,997)	0.06	0.01	0.05, 0.07
Gender			
Female (12,125)	0.06	0.01	0.05, 0.08
Male (3,872)	0.05	0.01	0.03, 0.08
Race/Ethnicity			
Hispanic (3,092)	0.04	0.01	0.02, 0.07
White (1,533)	0.05	0.02	0.01, 0.08
Black (10,720)	0.07	0.01	0.05, 0.09
Other (652)	0.05	0.03	0.00, 0.10
Age			
18-30 (10,141)	0.07	0.01	0.05, 0.09
31-45 (2,043)	0.06	0.02	0.02, 0.09
46+ (3,813)	0.04	0.01	0.02, 0.07
Mothers (12,686)	0.05	0.01	0.03, 0.06
Missing Interviewer Characteristics	0.09	0.01	0.06, 0.11

Notes: Unweighted data.

As noted in Table 4.10, there is the greatest variability among interviewers with missing interviewer characteristics (ρ_{int} =0.09) and less than six months of NIS interviewing experience (ρ_{int} =0.08); there is likely to be a great deal of overlap between these two groups of interviewers. Interviewers demonstrate less variability in obtaining consent rates as they gain more experience in administering the request. Although a consent request is not a typical survey question and thus not directly comparable, Groves (1989) reports common values of ρ_{int} between 0.01 and 0.02 for survey questions, though values reported are as high as 0.10. Positive intraclass correlations – even small ones can badly inflate the variance of estimates if a particular interviewer handles many cases.

4.5.2.3 Logistic Regression Analyses: Respondent and Interviewer Characteristics

To determine if interviewer characteristics are related to the variation in consent rates, a second set of logistic regression models incorporates information about interviewer characteristics. As with the other analyses in this chapter, the following models cluster responses by interviewer to accommodate the nested structure of the data. Clustering the data by interviewer produces approximately the same coefficients and standard errors as a multilevel regression model, permits the use of weights, and substantially reduces the amount of processing time required.[95] [96]

Table 4.11

Baseline Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers (All Interviewers)

		All Respondents (n=24,354)			Mothers (n=19,233)		
		Coef.	SE	*p*	Coef.	SE	*p*
Constant		2.02	0.28	0.000	2.42	0.29	0.000
Mother Characteristics							
Ethnicity	Non-Hispanic (ref.)						
	Hispanic	0.01	0.12	0.936	-0.07	0.14	0.598
Race	White (ref.)						
	Black	0.01	0.10	0.924	-0.12	0.10	0.263
	Other	0.14	0.15	0.355	0.07	0.19	0.704
Education	<12 yrs (ref.)						
	12 yrs	-0.28	0.13	0.037	-0.14	0.16	0.380
	>12 yrs, non col. grad	-0.46	0.15	0.002	-0.38	0.17	0.023

---

[95] The amount of processing time required was a critical factor in selecting a model, as all analyses were conducted in the NCHS RDC where computing time is limited and data users are charged for their time spent in the RDC.

[96] Of the multilevel modeling packages available in Stata, three accommodate binary dependent variables (*xtlogit, xtmelogit, and gllamm*). Although *gllamm* has the capability to incorporate weights, it must be downloaded which was not feasible given the lack of internet connection in the NCHS RDC.

	Col. grad	-0.39	0.14	0.006	-0.31	0.17	0.081
Marital Status	Married (ref.)						
	Never Married	0.13	0.16	0.427	0.22	0.17	0.208
	Wid./Div./Sep.	0.04	0.14	0.753	0.01	0.16	0.942
Age	<30 (ref.)						
	30+	-0.11	0.08	0.149	-0.20	0.09	0.029
2008 Family Income		0.00	0.00	0.567	0.00	0.00	0.351
Relationship to Child	Not Mother (ref.)						
	Mother	0.51	0.07	0.000	----	-----	----
Respondent Characteristics							
Language of Int.	English (ref.)						
	Non-English	0.34	0.19	0.080	0.65	0.22	0.004
Prior Survey Refusal	None (ref.)						
	One or More	-3.11	0.13	0.000	-3.16	0.13	0.000
Advance Letter	Not Mailed (ref.)						
	Mailed	0.20	0.07	0.003	0.14	0.07	0.053
Interviewer Characteristics							
Gender	Male (ref.)						
	Female	0.24	0.14	0.085	0.28	0.15	0.062
Age	18-30 (ref.)						
	31-45	-0.04	0.18	0.843	0.10	0.20	0.596
	46+	0.05	0.13	0.691	0.02	0.13	0.881
Race/Ethnicity	Hispanic (ref.)						
	White	0.12	0.24	0.629	0.21	0.26	0.420
	Black	0.08	0.19	0.690	0.18	0.20	0.365
	Other	0.26	0.30	0.396	0.44	0.34	0.202
Experience	0-6 mos (ref.)						
	6 mos - 1 year	-0.63	0.14	0.000	-0.59	0.15	0.000
	> 1year	-1.10	0.13	0.000	-1.12	0.14	0.000
Missing Data	Not Missing (Ref.)						
	Missing	0.51	0.22	0.021	0.57	0.23	0.012

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit for either model: All Respondents [F(9,750) = 0.41; Prob > F = 0.928]; Mothers [F(9,744) = 0.40; Prob > F = 0.934].

Table 4.12

Wald Tests for Categorical Parameters in Baseline Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers (All Interviewers)

	All Respondents		Mothers	
Categorical Predictor	F-Test Statistic	*p*	F-Test Statistic	*p*
Race	$F_{(2, 757)} = 0.44$	0.646	$F_{(2, 751)} = 0.72$	0.488
Education	$F_{(3, 756)} = 3.46$	0.016	$F_{(3, 750)} = 2.50$	0.058
Marital Status	$F_{(2, 757)} = 0.49$	0.614	$F_{(2, 751)} = 1.76$	0.172
Int. Age	$F_{(2, 757)} = 0.13$	0.874	$F_{(2, 751)} = 0.14$	0.868
Int. Race/Ethnicity	$F_{(3, 756)} = 0.26$	0.854	$F_{(3, 750)} = 0.60$	0.618
Int. Experience	$F_{(2, 757)} = 34.05$	0.000	$F_{(2, 751)} = 33.98$	0.000

Tables 4.11 and 4.13 present a set of baseline logistic regression models that include both respondent and interviewer predictors. The models in Table 4.11 include data from all interviewers, and the models presented in Table 4.13 exclude interviewers with missing demographic data. Including interviewer characteristics in these regression models does not strongly alter relationships between the mother or respondent characteristics and consent.

Of interviewers for whom demographics characteristics are available, female interviewers achieve higher consent rates than male interviewers ($p<0.10$) in all models presented in Tables 4.11 and 4.13. As demonstrated by the Wald X^2 values for the multi-category predictors for other interviewer demographic characteristics, interviewer age and race/ethnicity are not associated with consent for interviewers releasing those characteristics (see Tables 4.12 and 4.14). However, as illustrated by the indicator for interviewer missing data in the models presented in Table 4.11, interviewers who are missing demographic characteristics achieved higher consent rates ($p<0.05$) among both respondents and mothers, even when controlling for experience.

Interviewer experience is negatively associated with consent ($p<0.001$) in all models presented in Tables 4.11 and 4.13. Although this negative effect may be partly due to non-random assignment of more experienced interviewers to more difficult cases, this result holds when removing respondents with prior survey refusals from the model (see Table 4.2 in Appendix).

Table 4.13

Baseline Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers (Interviewers with Nonmissing Demographic Characteristics)

		All Respondents (n=15,690)			Mothers (n=12,548)		
		Coef.	SE	*p*	Coef.	SE	*p*
Constant		1.84	0.33	0.000	2.40	0.34	0.000
Mother Characteristics							
Ethnicity	Non-Hispanic (ref.)						
	Hispanic	-0.09	0.14	0.510	-0.21	0.15	0.167
Race	White (ref.)						
	Black	-0.07	0.11	0.522	-0.13	0.12	0.300
	Other	0.18	0.19	0.342	0.13	0.23	0.577
Education	<12 yrs (ref.)						
	12 yrs	-0.30	0.13	0.024	-0.24	0.18	0.175
	>12 yrs, non col. grad	-0.51	0.15	0.001	-0.49	0.20	0.013
	Col. grad	-0.54	0.14	0.000	-0.54	0.19	0.005
Marital Status	Married (ref.)						
	Never Married	0.33	0.18	0.072	0.41	0.21	0.048
	Wid./Div./Sep.	0.25	0.17	0.144	0.23	0.19	0.237
Age	<30 (ref.)						
	30+	-0.10	0.10	0.316	-0.13	0.11	0.228
2008 Family Income		0.00	0.00	0.854	0.00	0.00	0.686
Relationship to Child	Not Mother (ref.)						
	Mother	0.61	0.09	0.000	----	-----	----
Respondent Characteristics							
Language of Int.	English (ref.)						
	Non-English	0.37	0.22	0.094	0.70	0.24	0.004
Prior Survey Refusal	None (ref.)						
	One or More	-3.25	0.15	0.000	-3.20	0.16	0.000

Table 4.13 Continued

		All Respondents (n=15,690)			Mothers (n=12,548)		
		Coef.	SE	*p*	Coef.	SE	*p*
Advance Letter	Not Mailed (ref.)						
	Mailed	0.22	0.08	0.007	0.19	0.08	0.024
Interviewer Characteristics							
Gender	Male (ref.)						
	Female	0.25	0.14	0.079	0.28	0.15	0.062
Age	18-30 (ref.)						
	31-45	-0.04	0.18	0.828	0.11	0.20	0.581
	46+	0.05	0.13	0.688	0.03	0.14	0.812
Race/Ethnicity	Hispanic (ref.)						
	White	0.12	0.25	0.616	0.21	0.26	0.413
	Black	0.08	0.19	0.687	0.18	0.21	0.374
	Other	0.28	0.31	0.376	0.44	0.35	0.207
Experience	0-6 mos (ref.)						
	6 mos - 1 year	-0.68	0.17	0.000	-0.72	0.18	0.000
	> 1year	-1.09	0.17	0.000	-1.18	0.18	0.000

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit for either model: All Respondents [F(9,455) = 0.94; Prob > F = 0.492]; Mothers [F(9,453) = 0.72; Prob > F = 0.694].

Table 4.14

Wald Tests for Categorical Parameters in Baseline Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers (Interviewers with Nonmissing Demographic Characteristics)

	All Respondents		Mothers	
Categorical Predictor	F-Test Statistic	*p*	F-Test Statistic	*p*
Race	$F_{(2, 462)} = 0.62$	0.534	$F_{(2, 460)} = 0.71$	0.491
Education	$F_{(3, 461)} = 5.04$	0.002	$F_{(3, 459)} = 3.53$	0.015
Marital Status	$F_{(2, 462)} = 1.62$	0.199	$F_{(2, 460)} = 2.14$	0.118
Int. Age	$F_{(2, 462)} = 0.15$	0.864	$F_{(2, 460)} = 0.15$	0.858
Int. Race/Ethnicity	$F_{(3, 461)} = 0.29$	0.835	$F_{(3, 459)} = 0.58$	0.626
Int. Experience	$F_{(2, 462)} = 20.52$	0.000	$F_{(2, 460)} = 21.75$	0.000

A second series of logistic regression models incorporate several interaction terms (see Tables 4.15 and 4.17). These models include interactions between interviewer race/ethnicity and respondent race, to assess whether racial/ethnic similarity influences consent likelihood[97]; interview language and Hispanic ethnicity of mother, to determine if consent is more likely among Hispanics when the interview is conducted in a language other than English; and interviewing experience and prior survey refusal, to evaluate the effect of interviewing experience among respondents who previously refused survey cooperation. The models that include all respondents (as compared to mothers only) also include interaction terms between respondent's relationship to child and mother's age, and advance letter. Analyses conducted earlier in the chapter suggested that the effect of advance letter and mother's age may vary depending if the child's mother is the survey respondent. An interaction between interviewer gender and respondent's relationship to child is also included in models

---

[97] The ability to address this research question is somewhat limited by the coding of these variables in the existing data. Interviewers' race/ethnicity is coded as one variable (Hispanic, White, Black, Other), and respondents' race and ethnicity are coded as two separate dimensions (Hispanic or not Hispanic; and White, Black, or Other race).

with all respondents to evaluate whether female respondents (mothers) are more likely to consent to female interviewers.[98]

Table 4.15

Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers with Interactions (All Interviewers)

		All Respondents (n 24,353)			Mothers (n=19,233)		
		Coef.	SE	*p*	Coef.	SE	*p*
Constant		1.80	0.32	0.000	2.52	0.30	0.000
Mother Characteristics							
Ethnicity	Non-Hispanic (ref.)						
	Hispanic	-0.04	0.12	0.763	-0.09	0.14	0.505
Race	White (ref.)						
	Black	-0.29	0.22	0.187	-0.53	0.22	0.016
	Other	0.38	0.47	0.413	0.33	0.64	0.607
Education	<12 yrs (ref.)						
	12 yrs	-0.26	0.13	0.039	-0.14	0.16	0.379
	>12 yrs, non col. grad	-0.46	0.14	0.001	-0.38	0.17	0.022
	Col. grad	-0.38	0.14	0.006	-0.29	0.17	0.089
Marital Status	Married (ref.)						
	Never Married	0.13	0.16	0.391	0.22	0.17	0.206
	Wid./Div./Sep.	0.05	0.14	0.689	0.02	0.16	0.872
Age	<30 (ref.)						
	30+	0.17	0.14	0.223	-0.20	0.09	0.031
	30+ X Mother	-0.37	0.14	0.011			
2008 Family Income		0.00	0.00	0.575	0.00	0.00	0.393
Relationship to Child	Not Mother (ref.)						
	Mother	0.89	0.20	0.000	----	-----	----
Respondent Characteristics							
Language of Int.	English (ref.)						
	Non-English	0.04	0.26	0.870	0.34	0.37	0.357
	Non-English X Hispanic	0.38	0.33	0.259	0.33	0.45	0.456
Prior Survey Refusal	None (ref.)						
	One or More	-3.62	0.29	0.000	-3.71	0.29	0.000
Advance Letter	Not Mailed (ref.)						
	Mailed	0.38	0.14	0.005	0.14	0.07	0.061

---

[98] As NIS does not collect information on respondents' gender, mothers are used as a proxy for female respondents. However, this indicator is imperfect and only identifies a portion of female respondents in the sample.

	Mailed X Mother	-0.25	0.15	0.100	----	-----	----
Interviewer Characteristics							
Gender	Male (ref.)						
	Female	0.18	0.19	0.342	0.28	0.15	0.059
	Female X Mother	0.07	0.19	0.697	----	-----	----
	Missing X Mother	-0.19	0.22	0.388			
Age	18-30 (ref.)						
	31-45	-0.05	0.18	0.782	0.10	0.20	0.603
	46+	0.03	0.13	0.798	0.02	0.13	0.910
Race/Ethnicity	Hispanic (ref.)						
	White	0.18	0.24	0.445	0.25	0.26	0.338
	Black	0.09	0.20	0.649	0.13	0.21	0.542
	Other	0.19	0.33	0.554	0.38	0.39	0.338
	White X Black	0.18	0.38	0.634	0.28	0.38	0.454
	White X Other	-0.59	0.79	0.453	-0.85	1.01	0.403
	Black X Black	0.33	0.26	0.202	0.55	0.26	0.033
	Black X Other	-0.23	0.51	0.655	-0.13	0.68	0.847
	Other X Black	0.44	0.38	0.252	0.37	0.42	0.378
	Other X Other	1.31	0.78	0.096	1.60	1.00	0.109
	Missing X Black	0.43	0.33	0.190	0.43	0.30	0.148
	Missing X Other	-0.33	0.53	0.528	-0.41	0.71	0.569
Experience	0-6 mos (ref.)						
	6 mos - 1 year	-0.69	0.15	0.000	-0.69	0.16	0.000
	> 1year	-1.19	0.15	0.000	-1.21	0.15	0.000
	6 mos-1 year X Prior Ref.	0.36	0.36	0.321	0.65	0.39	0.102
	> 1year X Prior Ref.	0.69	0.33	0.036	0.68	0.33	0.043
Missing Data	Not Missing (ref.)						
	Missing.	0.65	0.29	0.028	0.55	0.24	0.020

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests a slight lack of fit for the model including Mothers only: All Respondents [F(9,750) = 0.60; Prob > F = 0.801]; Mothers [F(9,744) = 1.72; Prob > F = 0.080].

Table 4.16

Wald Tests for Categorical Parameters in Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers with Interactions (All Interviewers)

	All Respondents		Mothers	
Categorical Predictor	F-Test Statistic	*p*	F-Test Statistic	*p*
Race	$F_{(2, 757)} = 1.15$	0.318	$F_{(2, 751)} = 0.14$	0.871
Education	$F_{(3, 756)} = 3.65$	0.013	$F_{(3, 750)} = 2.53$	0.056
Marital Status	$F_{(2, 757)} = 0.50$	0.610	$F_{(2, 751)} = 1.62$	0.199
Int. Age	$F_{(2, 757)} = 0.10$	0.903	$F_{(2, 751)} = 0.14$	0.871
Int. Race/Ethnicity	$F_{(4, 755)} = 1.41$	0.289	$F_{(3, 750)} = 0.50$	0.680
Int. Experience	$F_{(2, 757)} = 33.75$	0.000	$F_{(2, 751)} = 31.21$	0.000

Table 4.17

Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers with Interactions (Interviewers with Nonmissing Demographic Characteristics)

		All Respondents (n=15,690)			Mothers (n=12,548)		
		Coef.	SE	*p*	Coef.	SE	*p*
Constant		1.77	0.37	0.000	2.51	0.34	0.000
Mother Chars.							
Ethnicity	Non-Hispanic (ref.)						
	Hispanic	-0.13	0.14	0.349	-0.21	0.16	0.183
Race	White (ref.)						
	Black	-0.33	0.22	0.135	-0.56	0.22	0.011
	Other	0.38	0.47	0.422	0.31	0.63	0.618
Education	<12 yrs (ref.)						
	12 yrs	-0.31	0.13	0.023	-0.26	0.18	0.152
	>12 yrs, non col. grad	-0.52	0.15	0.001	-0.50	0.20	0.011
	Col. grad	-0.55	0.14	0.000	-0.54	0.19	0.004
Marital Status	Married (ref.)						
	Never Married	0.33	0.18	0.068	0.41	0.21	0.047
	Wid./Div./Sep.	0.26	0.17	0.126	0.24	0.19	0.215
Age	<30 (ref.)						
	30+	0.09	0.16	0.576	-0.13	0.11	0.243
	30+ X Mother	-0.24	0.17	0.158			
2008 Family Income		0.00	0.00	0.835	0.00	0.00	0.756
Relationship to Child	Not Mother (ref.)						
	Mother	0.77	0.23	0.001	----	-----	----
Resp.Chars.							
Language of Int.	English (ref.)						
	Non-English	0.11	0.28	0.700	0.63	0.32	0.046
	Non-English X Hispanic	0.31	0.37	0.403	0.03	0.42	0.951
Prior Survey Refusal	None (ref.)						
	One or More	-3.67	0.35	0.000	-3.70	0.40	0.000
Advance Letter	Not Mailed (ref.)						
	Mailed	0.33	0.17	0.057	0.19	0.09	0.031
	Mailed X Mother	-0.15	0.18	0.426	----	-----	----
Int. Chars.							
Int. Gender	Male (ref.)						
	Female	0.19	0.19	0.325	0.28	0.15	0.059
	Female X Mother	0.07	0.20	0.724			
	Missing X Mother				----	-----	----
Age	18-30 (ref.)						
	31-45	-0.05	0.18	0.783	0.11	0.20	0.594
	46+	0.04	0.13	0.771	0.03	0.14	0.843

Race/Ethnicity	Hispanic (ref.)						
	White	0.17	0.24	0.493	0.23	0.26	0.381
	Black	0.07	0.21	0.741	0.10	0.22	0.637
	Other	0.20	0.34	0.559	0.35	0.39	0.378
	White X Black	0.17	0.38	0.647	0.29	0.38	0.455
	White X Other	-0.56	0.81	0.492	-0.80	1.00	0.425
	Black X Black	0.34	0.26	0.191	0.57	0.26	0.030
	Black X Other	-0.24	0.52	0.650	-0.18	0.66	0.787
	Other X Black	0.47	0.38	0.218	0.41	0.41	0.320
	Other X Other	1.35	0.81	0.096	1.65	0.99	0.096
Experience	0-6 mos (ref.)						
	6 mos-1 year	-0.70	0.18	0.000	-0.76	0.19	0.000
	> 1year	-1.14	0.18	0.000	-1.25	0.19	0.000
	6 mos-1 year X Prior Ref.	0.22	0.45	0.630	0.29	0.51	0.575
	> 1year X Prior Ref.	0.55	0.40	0.167	0.63	0.45	0.160

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit for either model: All Respondents [F(9,750) = 0.41; Prob > F = 0.928]; Mothers [F(9,453) = 0.83; Prob > F = 0.588].

Table 4.18

Wald Tests for Categorical Parameters in Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers with Interactions (Interviewers with Nonmissing Demographic Characteristics)

	All Respondents		Mothers	
Categorical Predictor	F-Test Statistic	*p*	F-Test Statistic	*p*
Race	$F_{(2, 757)} = 1.15$	0.318	$F_{(2, 460)} = 3.34$	0.036
Education	$F_{(3, 756)} = 3.65$	0.013	$F_{(3, 459)} = 3.53$	0.015
Marital Status	$F_{(2, 757)} = 0.50$	0.610	$F_{(2, 460)} = 2.13$	0.120
Int. Age	$F_{(2, 757)} = 0.10$	0.903	$F_{(2, 460)} = 0.14$	0.867
Int. Race/Ethnicity	$F_{(4, 755)} = 1.41$	0.229	$F_{(3, 459)} = 0.46$	0.709
Int. Experience	$F_{(2, 757)} = 33.75$	0.000	$F_{(2, 460)} = 22.05$	0.000

Although the interaction terms are somewhat related to consent in the models that contain all interviewers (Table 4.15), they do not predict consent in the models that limit the sample to interviewer with nonmissing demographic characteristics (Table 4.17). In the model with all respondents and all interviewers, the interaction between mother's age and respondent's relationship to the sample child is a significant predictor of consent (<0.05): children with older mothers are significantly less likely to have consent.

Also in this model, mothers who were mailed an advance letter are less likely to consent (compared to non-mothers who were not mailed an advance letter), though this effect is only marginally significant (p=0.100). This finding may provide evidence that the results from earlier models in which the advance letter had a more positive impact on consent in the total sample than mothers-only sample was due to sample size variations rather than differential effects of the advance letter on mothers versus non-mothers.

Lastly, as compared to interviews with no prior survey refusals and 0-6 months of interviewing experience, greater than one year of interviewing experience has a positive effect on consent for interviews conducted with one or more prior survey refusals (p<0.05). For the model estimated on all interviewers including only mothers, this finding is the only significant interaction term.

A third and final set of logistic regression models reduces each of the models in Tables 4.15 and 4.17 to their respective significant predictors of consent only (p<0.10). These final models are presented in Tables 4.19, 4.21, 4.23, and 4.25. As demonstrated in these four models, interviewer and sample characteristics associated with consent vary somewhat depending on whether or not the respondent is the child's mother and the interviewer sample.

There are several predictors of consent common to the reduced models. Greater educational attainment of the mother, one or more prior survey refusals and increased interviewing experience are negatively associated with consent, and the advance letter, and interviewer missing demographic data are positively associated with consent. Further, consent to vaccination provider contact is more likely when the child's mother is the survey respondent.

Some additional characteristics are also related to consent. In the majority of the four models, consent was more likely in interviews conducted in a language other than English and by a female interviewer. In models that include all interviewers (Tables 4.19 and 4.21), consent was associated with children who have younger mothers. Lastly, in the model that includes interviewers with nonmissing data and mothers only, consent was less likely among Hispanic mothers and more likely among mothers who never married.

Table 4.19

Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents (All Interviewers)

		All Respondents (n=24,403)		
		Coef.	SE	*p*
Constant		2.15	0.23	0.000
Mother Characteristics				
Education	<12 yrs (ref.)			
	12 yrs	-0.28	0.13	0.032
	>12 yrs, non col. grad	-0.47	0.14	0.001
	Col. grad	-0.40	0.14	0.004
Age	<30 (ref.)			
	30+	-0.12	0.07	0.098
Relationship to Child	Not Mother (ref.)			
	Mother	0.68	0.12	0.000
Respondent Characteristics				
Language of Int.	English (ref.)			
	Non-English	0.30	0.17	0.078
Prior Survey Refusal	None (ref.)			
	One or More	-3.63	0.30	0.000
Advance Letter	Not Mailed (ref.)			
	Mailed	0.40	0.14	0.004
	Mailed X Mother	-0.28	0.15	0.058
Interviewer Characteristics				
Gender	Male (ref.)			
	Female	0.24	0.14	0.080

Experience	0-6 mos (ref.)			
	6 mos - 1 year	-0.70	0.15	0.000
	> 1year	-1.21	0.14	0.000
	6 mos - 1 year X 1+ Refusal	0.41	0.37	0.270
	> 1year X 1+ Refusal	0.69	0.33	0.039
Missing Data	Not Missing (ref.)			
	Missing	0.42	0.16	0.010

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit [F(9,751) = 0.76; Prob > F = 0.653].

Table 4.20

Wald Tests for Categorical Parameters in Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents (All Interviewers)

Categorical Predictor	F-Test Statistic	*p*
Education	$F_{(3,\ 757)} = 3.86$	0.009
Int. Experience	$F_{(2,\ 758)} = 35.09$	0.000
Int. Experience x 1+ Refusals	$F_{(2,\ 758)} = 2.33$	0.098

Table 4.21

Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among Mothers (All Interviewers)

			Mothers (n=19,263)	
		Coef.	SE	*p*
Constant		2.72	0.23	0.000
Mother Characteristics				
Education	<12 yrs (ref.)			
	12 yrs	-0.15	0.15	0.339
	>12 yrs, non col. grad	-0.41	0.16	0.013
	Col. grad	-0.31	0.16	0.060
Age	<30 (ref.)			
	30+	-0.22	0.09	0.011
Respondent Characteristics				
Language of Int.	English (ref.)			
	Non-English	0.47	0.18	0.008
Prior Survey Refusal	None (ref.)			
	One or More	-3.16	0.13	0.000
Advance Letter	Not Mailed (ref.)			
	Mailed	0.13	0.08	0.076
Interviewer Characteristics				
Gender	Male (ref.)			
	Female	0.27	0.15	0.070
Experience	0-6 mos (ref.)			
	6 mos - 1 year	-0.60	0.15	0.000
	> 1year	-1.14	0.14	0.000
Missing Data	Not Missing (ref.)			
	Missing	0.38	0.16	0.022

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit [F(9,745) = 0.73; Prob > F = 0.682].

Table 4.22

Wald Tests for Categorical Parameters in Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among Mothers (All Interviewers)

Categorical Predictor	F-Test Statistic	*p*
Education	$F_{(3, 751)} = 2.85$	0.037
Int. Experience	$F_{(2, 752)} = 36.14$	0.000

Table 4.23

Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents (Interviewers with Nonmissing Demographic Characteristics)

		All Respondents (n=15,865)		
		Coef.	SE	*p*
Constant		2.32	0.22	0.000
Mother Characteristics				
Education	<12 yrs (ref.)			
	12 yrs	-0.36	0.13	0.006
	>12 yrs, non col. grad	-0.60	0.13	0.000
	Col. grad	-0.65	0.12	0.000
Relationship to Child	Not Mother (ref.)			
	Mother	0.64	0.09	0.000
Respondent Characteristics				
Prior Survey Refusal	None (ref.)			
	One or More	-3.25	0.15	0.000
Advance Letter	Not Mailed (ref.)			
	Mailed	0.21	0.08	0.009
Interviewer Characteristics				
Experience	0-6 mos (ref.)			
	6 mos - 1 year	-0.64	0.18	0.000
	> 1year	-1.06	0.17	0.000

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit [F(9,455) = 0.75; Prob > F = 0.662].

Table 4.24

Wald Tests for Categorical Parameters in Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents (Interviewers with Nonmissing Demographic Characteristics)

Categorical Predictor	F-Test Statistic	*p*
Education	$F_{(3, 461)} = 11.34$	0.000
Int. Experience	$F_{(2, 462)} = 19.52$	0.000

Table 4.25

Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among Mothers (Interviewers with Nonmissing Demographic Characteristics)

		Mothers (n=12,626)		
		Coef.	SE	*p*
Constant		2.51	0.35	0.000
Mother Characteristics				
Ethnicity	Non-Hispanic (ref.)			
	Hispanic	-0.25	0.15	0.083
Education	<12 yrs (ref.)			
	12 yrs	-0.23	0.18	0.215
	>12 yrs, non col. grad	-0.47	0.20	0.017
	Col. grad	-0.56	0.19	0.003
Marital Status	Married (ref.)			
	Never Married	0.44	0.20	0.030
	Wid./Div./Sep.	0.24	0.19	0.200
Respondent Characteristics				
Language of Int.	English (ref.)			
	Non-English	0.61	0.22	0.006
Prior Survey Refusal	None (ref.)			
	One or More	-3.19	0.16	0.000
Advance Letter	Not Mailed (ref.)			
	Mailed	0.19	0.08	0.022
Interviewer Characteristics				
Gender	Male (ref.)			
	Female	0.25	0.15	0.085
Experience	0-6 mos (ref.)			
	6 mos - 1 year	-0.77	0.18	0.000
	> 1year	-1.21	0.18	0.000

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit [F(9,453) = 0.56; Prob > F = 0.826].

Table 4.26

Wald Tests for Categorical Parameters in Reduced Logistic Regression Model Predicting Consent as a Function of Sample and Interviewer Characteristics, among Mothers (Interviewers with Nonmissing Demographic Characteristics)

Categorical Predictor	F-Test Statistic	*p*
Marital Status	$F_{(2, 460)} = 2.64$	0.073
Education	$F_{(3, 459)} = 4.25$	0.006
Int. Experience	$F_{(2, 460)} = 23.59$	0.000

4.6 Discussion

Results from this analysis of the 2009 NIS suggest that certain characteristics of the child's mother, survey respondent, and interviewer are associated with consent to vaccination provider contact and that interviewers vary in their ability to obtain respondents' consent. We also find that the relationship of some predictors of consent vary depending on the interviewer sample, that is, whether demographic information of the interviewer is available.

4.6.1 Sample Characteristics

The analyses in this chapter identify several characteristics of the child's mother and survey respondent that are consistently associated with consent likelihood in the 2009 NIS. Mothers of sample children are more likely to consent than respondents who are related to the child in another way. It is possible that non-mothers do not feel they have the authority to consent, or do not have access to the requested information about the child's vaccination provider. In most analyses, greater educational attainment of the child's mother is significantly and negatively related to consenting.[99]

Prior survey refusals are also consistently related to consenting in this survey, with respondents who refused survey participation one or more times before responding significantly less likely to consent. This relationship may be a function of privacy concerns, general resistance to the survey interview (Sakshaug et al., 2012), or may be related to greater opportunity costs for these respondents (Dillman, Eltinge,

---

[99] As mentioned in previous chapters, the relationship between education and consent tends to vary by studies, although see Chapter 3 of this dissertation for a similar finding.

Groves, & Little, 2002). Given the nature of the target population, all survey respondents have one or more children in their household between the ages of 19-35 months, which likely limits their discretionary time. While agreeing to the survey is an obvious time commitment for respondents, the consent request sequence of the interview requires both respondents' time and efforts as they are asked to retrieve and report the contact information for each vaccination provider. Busier respondents may be more hesitant both to respond initially and expend the efforts associated with consenting.

Respondents for whom an address match could be made were mailed an advance letter, and those mailed the letter were consistently more likely to consent. As the letter was not randomly assigned to households, we cannot be certain that effects on consent are due to the letter itself rather than differences in composition for matched versus unmatched samples. As mentioned, past research identifies differences in survey response rates by listed and unlisted phone numbers (Traugott et al., 1987). Assuming that samples of respondents are equivalent and the effect on consent is due to the advance letter, the letter may have alerted sampled households to the upcoming survey and reinforced the authenticity of the NIS.

In addition to the factors that consistently predicted consent across models, a few additional characteristics of the respondent and child's mother are associated with consenting in some of the models. Consent tends to be more likely when the survey is conducted in a language other than English. Although correlates of interview language including interviewer race/ethnicity and mother's ethnicity do not strongly associate with consent when controlling for interviewer language in the multivariate models, Hispanic ethnicity of the mother is positively related to consent in the bivariate analyses.

Furthermore, consent is somewhat less likely for children of older mothers (those thirty years old and older), as compared to younger mothers. The relationship between mothers' age and consent is stronger when the mother is the survey respondent. An interaction was tested between mother's age and education to examine whether increased education moderated this effect, but was not significantly associated with consent and excluded from the models. Older mothers may have consented at lower rates because they were busier during the survey interview, for example caring for more children.

As noted, several characteristics of the mother and respondent do not predict consent as initially hypothesized. The hypotheses were based on predictors of missing provider data in the 2002 NIS, assuming that at least some respondents were missing provider data because they had purposely misreported provider contact information, thus passively refusing the consent request. As several of the hypotheses were not supported, it appears that factors contributing to missing provider data and consent may not be related. It is also possible that the two are related and the predictors identified in the 2002 report from which the hypotheses were created are out of date.

Finally, the distribution of respondent characteristics varies somewhat by interviewer sample (see Table 4.01 in the Appendix). Without an indicator on the dataset to distinguish whether interviewers are from Chicago or Las Vegas, we cannot determine how this differential recruitment varies by interviewer sample. This could be a function of potentially slightly different working hours between the two calling centers, resulting in an availability of different respondents, or some other unknown variation in distribution of sample to interviewers.

In addition to variations in sample characteristics, predictors of consent vary depending on whether or not demographic information is available for interviewers. For example, the relationship between marital status of the child's mother and consent varies depending on the interviewer sample. Children whose mothers never married are more likely to have consent when demographic information is available for the interviewer, but when this information is missing, children whose mothers are widowed, divorced, or separated are less likely to have consent.

Without additional information on interviewers, it is unclear if the varying relationships between sample characteristics and consent are due to differences in interviewer calling location, reasons for missing demographic information, or some other factor that interacted with respondent characteristics, influencing relationships with consent. For example, interviewers missing demographic characteristics may share some characteristic or characteristics for which information is not included on this dataset, such as personality factors, attitudes towards privacy, or greater experience requesting consent.

4.6.2 Interviewers and Interviewer Characteristics

As suggested by intraclass correlations, interviewers vary in their ability to obtain respondents' consent to vaccination provider contact. We observed the greatest variability in consent rates among interviewers with missing interviewer characteristics, as well as interviewers with less than six months of NIS interviewing experience. It is likely that the variability observed among interviewers with missing demographic data is at least somewhat related to the effects of experience as interviewers missing demographic data appear to be less experienced. If the effects were more closely tied to another characteristic, for example the race of these interviewers with missing data, it is likely that we would have observed larger intraclass correlations for cases where demographic information was available. In addition, with increased experience, interviewers demonstrate less variability in obtaining consent rates, as they are likely to refine their methods over time as they better understand the methods most effective in gaining cooperation to consent.

Some interviewer characteristics are related to respondents' willingness to consent, mainly their level of experience on the NIS and gender. Further, interviewers for whom demographic information was unavailable obtained higher consent rates than interviewers for whom demographic information was available, although as noted, this appears to be at least partially related to the effect of NIS interviewing experience.

Of interviewers for whom demographic information was available, female interviewers obtained higher rates of consent. Respondents may have perceived female interviewers as less threatening than male interviewers (Tate et al., 2005), which could have helped to facilitate consent as respondents are likely protective of the young sample children. The interaction between interviewer gender and respondent's relationship to sample child was not significantly associated with consent.

NIS interviewing experience, for which information was available for all interviewers, was negatively associated with consent overall, and so interviewers with greater amounts of experience on this survey achieved lower consent rates. However, the assignment of interviewers to respondents is not fully random: some more experienced NIS interviewers are designated “refusal converters”, meaning that more experienced interviewers are assigned to respondents who refused participation on prior calls. Thus, if assigned to these more difficult cases that are less likely to consent, some more experienced interviewers will likely obtain lower consent rates.

Although it is possible that the assignment of more difficult cases to more experienced interviewers could account for the negative association between interviewing experience and consent, this result is not borne out in analyses presented in the chapter. Removing the cases identified as prior survey refusals does not change the relationship between consent and interviewing experience (see Table 4.02 in the Appendix). NIS interviewing experience is not significantly associated with consent in bivariate and logistic regression analyses when the sample is limited to only respondents with one or more prior survey refusals.

Additionally, the reduced model of all respondents and all interviewers includes a modestly significant interaction between interviewing experience and prior survey refusal. In this model, among respondents with no prior refusals, increased NIS interviewing experience continues to have a negative effect on consent rates. Alternately, for respondents with one or more refusals, more than one year of NIS experience is positively associated with consent when controlling for other factors in the model. The Wald X^2 value for the multi-category predictor for the interaction is significant at $p<0.10$.

However, it is possible that the more experienced interviewers designated as refusal converters may have been assigned to additional “difficult” cases not flagged in the dataset, for example, if interviewers did not accurately or consistently code this information. If there are a substantial number of additional respondents in the dataset who refused to participate on prior calls, and these respondents were reassigned to more experienced interviewers, the findings presented in this chapter pertaining to the relationship between experience and consent could be biased.

The mostly negative relationship between NIS interviewing experience and consent may be related to how interviewers receive credit in the NIS survey. Although NORC emphasizes that interviewers obtain provider consent, interviewers receive credit in three areas of the interview: finding an eligible household, completing the survey (Sections A through C), and gaining provider consent. These three areas are equally weighted in terms of interviewer credit, and NORC reports that interviewers are aware of the how they receive credit in these areas. With time and experience, interviewers may come to find that the consent request is the most difficult part of the interview, for example, if respondents express unwillingness, hostility, and burden when asked to consent. If so, these interviewers may focus their efforts on recruiting eligible households and completing the interview, and allow the consent request module to be reassigned to another possibly more experienced interviewer, while still earning two-thirds possible credit per interview.

Of the interviewers providing demographic information, interviewer age and race/ethnicity are not associated with respondents’ willingness to consent to vaccination provider contact. However, given the nature of the phone interview, it is difficult to determine if the lack of findings is due to an inability of respondents to distinguish between the different racial/ethnic groups and age categories or a true lack of association between consent and these demographic characteristics of interviewers.

Interviewers for whom demographic information was not available (those who refused to release this information in Chicago, and those making calls out of Las Vegas) achieved significantly higher consent rates than Chicago interviewers releasing this information. Without an indicator identifying which interviews were completed by Chicago interviewers and which were completed by Las Vegas interviewers, we are unable to definitively determine if interviewers from both locations with missing data equally achieved higher consent rates, of if this effect is related to interviewers in only one location. On average, interviewers missing demographic information have less interviewing experience.

As noted previously, Las Vegas interviewers tended to be less experienced than Chicago interviewers (Skalland, B. of NORC, personal communication, July 5, 2012), and also accounted for slightly more than half of the 298 interviewers missing demographic information. The effects of both data missingness and experience on consent rates may be related to this specific subset of interviewers who, in addition to having less experience than the Chicago interviewers, are likely to have different demographics characteristics, different supervisors with different expectations for their productivity, different working hours, and potentially access to different sample telephone numbers – all which may affect consent rates.

However, we note that the indicator for missing data is still a significant predictor of consent in the models presented in this chapter even when controlling for experience, suggesting that although interviewers with missing data tend to be less experienced, there appear to be effects above and beyond experience among interviewers with missing data that predict consent. This may be related to interviewer location, reason for missingness, or some other demographic, personality, or experiential characteristic for which information was unavailable on this dataset.

4.7 Limitations

The research in this chapter has several limitations that restrict the full understanding of the relationship between interviewers and respondents' consent decision. This includes missing interviewer and section completion information and a limited measure of interviewing experience. In addition, analyses were somewhat restricted by RDC usage terms and policy to protect confidential information about interviewers included on the 2009 NIS dataset.

The dataset lacked information on interviewer demographic characteristics for nearly 40% of interviewers. Without demographic characteristics for all interviewers in the sample, or information on interviewer location, we could not determine if the same relationships between interviewer's demographic characteristics and consent existed for the interviewers with missing data as they did for interviewers with nonmissing data.

The dataset was also missing section completion status variables, as well as item completions status variables within the consent request module, which prevented full understanding of the dependent variable. As noted earlier, because surveys were considered complete if respondents answered Sections A through C, it is possible that interviews ended before interviewers initiated the consent module, for example, if a hostile refusal preceded this section. However, as noted above, section completion information was unavailable for this project, and so we cannot determine if respondents completed all, some, or none of the consent module, and if those indicated as having not consented were asked the direct request or did not reach that part of the interview.

Further, without section completion information, we were unable to determine the point at which respondents refused consent within the consent sequence. As noted earlier, the consent module comprised a series of questions that included child and respondent's name, name and contact information for each vaccination provider, authorization to consent to vaccination provider contact, and consent to vaccination provider contact. Refusing any question in the consent module is considered consent refusal. Different respondent and interviewer characteristics may be related to refusing to provide the various pieces of identifying information requested within this sequence. However, the dependent variable does not distinguish the point in the consent series where respondents refuse, for example, if they refuse to provide their name or their child's name, provider information, or refuse when directly asked to consent.

Also, section completion information could be used to confirm or refute the explanation that with increased experience, interviewers may be less likely to administer the consent module. Section completion information could be used to determine if more experienced interviewers are administering the consent module and respondents are refusing, or if they are not administering this series of questions as hypothesized.

The dataset also lacked information that identified the particular interviewer who requested consent. The inclusion of this information would have more accurately identified the relationship between interviewers and consent and strengthened internal validity of all analyses in this chapter. In addition, the assignment of interviewers to respondents in the survey was not fully random which complicated the interpretation of the relationship between experience and consent.

The measure of interviewing experience was quantified as days between the first call and current call, rather than the number of calls, or number of completed interviews, or times that consent was obtained. This measure could indicate that an interviewer has five years of experience on the NIS, but could have only placed two calls that were five years apart.

Further, this measure was limited to interviewing experience on the NIS survey, and did not include interviewers ‘experience on other surveys, nor did it directly translate to experience in requesting consent as consent was not required to complete an interview. It is possible that interviewers with little NIS experience per the measure included in this research had greater overall interviewing experience or greater experience requesting consent. Although not available on this dataset, a more comprehensive measure of interviewing experience and experience requesting consent may have uncovered a different relationship than the measure included in these analyses. Additionally, the measure of prior survey refusal included in these analyses is imperfect as it relates to the household telephone number, rather than any particular respondent.

Finally, the inclusion of interviewer information on the dataset raised confidentiality concerns and required that all analyses be conducted at the NCHS RDC and closely reviewed to prevent disclosure of any sensitive information. This approach had implications for time, cost, and analysis strategy. The amount of processing time required to run multilevel models in Stata as well as inability to use weights necessitated a less desirable approach for regression models that could slightly alter coefficients and standard errors. All data was analyzed in a way that minimized disclosure risk. For example, the coding of some variables, such as mother's age and level of interviewing experience, was dictated by RDC disclosure avoidance policies that forbid release of results with less than a minimal cell size. Although more ideal coding existed for these variables that may have exposed different relationships with the dependent variable, this was not permitted by RDC regulations.

5 OVERALL CONCLUSIONS

5.1 Introduction

The research in this dissertation investigates a range of characteristics of the respondent, interviewer, survey, and consent request that may influence the decision to consent to record linkage. This includes respondents' socioeconomic and demographic characteristics, as well as their privacy, confidentiality, and trust concerns; interviewers' demographic characteristics and level of interviewing experience; and various design characteristics of the survey and consent request, for example, the topic of the records requested, the modes of consent and survey administration, and any identifying information requested to facilitate record linkage.

In this final chapter, we discuss the findings from this research, and consider the current state of knowledge in this area as well as the future research necessary for a full understanding of respondents' decision to consent to record linkage

5.2 Respondent Characteristics

This dissertation explores the relationship between consent and a variety of respondent attributes. This includes demographic and socio-economic characteristics of the respondent, their attitudes, and the perceived salience between personal characteristics of the respondent and the characteristics of the consent request.

Past research investigating the influence of respondent characteristics on consent tends to focus on relationships between demographic and socioeconomic characteristics and consent rates. As described in Chapter 1, in which we summarize the literature on linkage consent, existing research shows significant yet inconsistent relationships with consent, with one such characteristic positively related to consent in one study, and the same characteristic negatively related to consent in another study (Sala et al., 2010). Two chapters in this dissertation further explore the relationship between these factors and consent.

In Chapter 3, which details an analysis of data from the 2011 JPSM Practicum survey, we found that among respondents assigned to the health consent request, education was negatively associated with consent, and categorical predictors for income, age, race, and gender were unrelated to consent. For respondents assigned to the income and employment consent request, age was positively associated with consent, income was negatively associated with consent, and education and gender were unrelated to consent. Compared to non-white respondents assigned to the income and employment consent request, white respondents were less likely to consent to the income and employment consent request.

Chapter 4, which presented an analysis of data from the 2009 NIS, also shows relationships between some of the socio-economic and demographic characteristics examined and consent to vaccination provider records: mother's age and education were negatively associated with consent, however, mother's race and income were unrelated to consent.

As suggested in the literature discussed in Chapter 1, which appears to show no strong relationships between socioeconomic and demographic characteristics and consent, we too find no consistent relationships between these characteristics and consent in this dissertation. This is true even when segmenting the existing research by various design characteristics such as survey topic, consent request, or target population. For example, although minority parents were more likely to grant consent to their child's health records in the NIS, the opposite was true in research by Tate et al. (2005), in which minority parents were less likely to have provided researchers access to their child's health records.

Further, in this dissertation, there was an inverse effect of education for both the health consent request in the Practicum survey and the NIS yet no such result for the Practicum income and employment request. This suggests that the effect of education on consent may be a function of consent request topic. However, existing studies do not consistently suggest such a relationship: Jay et al. (1994) also finds a negative relationship between education and likelihood to consent to a health request, yet Huang et al. (2007) and Cleary et al. (1981) find positive relationships, and Korkeila et al. 2001 find no relationship.

The lack of relationships between these characteristics and consent in the literature and this dissertation could indicate that these characteristics may not be useful in modeling consent. Or, the lack of a consistent relationship may suggest that these results are a function of the scope of the existing studies in this area, which to date, is still a small pool. With more studies, stronger, more consistent patterns may emerge.

It is also possible that it is not respondent characteristics themselves that directly relate to or motivate consent, but the relationships between these characteristics and attributes of the request. Existing research finds that respondents for whom the consent request is more relevant or salient are in some cases more likely to consent. For example, research finds that in some surveys, respondents in poorer health are more likely to consent to health-related consent requests (Dunn et al., 2004; Klassen et al., 2005).

We examined the issue of consent request salience in the 2011 Practicum Survey in Chapter 3 using both health and income and employment-related consent requests, but found no relationship between consent likelihood and measures of health, income, and employment. It is possible that the lack of findings in this research may be a function of the limited salience of the topic of the consent request in the survey, as the topic was mentioned only once, and briefly. Production surveys requesting consent should take advantage of existing measures on the survey that could provide an assessment of consent request salience to gauge if respondents for whom the consent request is more salient may be overrepresented in any linked data.

A secondary focus of past research investigates the influence of respondents' attitudes on consent to record linkage. Existing research generally concludes that respondents with greater privacy concerns are less likely to consent, but past research relies mainly on hypothetical consent scenarios (Singer, 2001; Singer et al., 2011) and indirect indicators of privacy such as refusal to provide a substantive response to the income question (Jenkins et al., 2006; Olson, 1999; Sala et al., 2010).

In this dissertation, we sought to contribute to this body of research by directly assessing respondents' privacy concerns in conjunction with their willingness to consent in the 2011 Practicum survey. We also included several additional indirect measures of privacy in both the 2011 Practicum survey and the analysis of the 2009 NIS, discussed in Chapter 4.

We included three questions in the 2011 Practicum survey with the specific intent of assessing respondent's level of privacy concern. As discussed in Chapter 3, these items did not appear to cohesively assess a single construct of privacy, nor did the questions detect any relationship between respondents' privacy concerns and consent likelihood.

However, other less direct indicators of privacy suggested a relationship between privacy and consent in the 2011 Practicum research. In the open-ended question that followed the consent request in which respondents were asked to provide reasons as to why they did or did not consent, of those who refused consent, the most frequently cited reason was privacy concerns (mentioned by 36%). In addition, in the majority of models where it was evaluated, refusal to provide a substantive response to the income question was positively related to consent refusal.

Results from our analysis of the 2009 NIS may also suggest that respondents' privacy concerns are negatively associated with their willingness to consent to record linkage. Here, respondents who had refused survey participation one or more times prior to completing the interview were significantly less likely to consent. Privacy concerns may be negatively related to both respondents' willingness to participate in the survey and consent to record linkage; past research demonstrates similar findings (Sakshaug et al., 2012).

In addition to privacy attitudes, a small amount of past research assesses how other respondent attitudes and traits, including confidentiality, trust, and altruism, can affect consent. The 2011 Practicum research considers all three of these.

Confidentiality concerns negatively impacted consent in past research (Armstrong et al., 2008); research in 2011 Practicum survey demonstrated similar findings. As an index, the three items included in the survey to measure respondents' confidentiality concerns strongly correlated with consent, with respondents indicating more confidentiality concerns significantly less likely to consent. In the open-ended request, confidentiality concerns were noted as the reason for refusing consent by 10% of respondents, the third largest category. In the Practicum research more broadly, respondents noted confidentiality concerns as prohibitive to linkage during focus groups conducted prior to the survey. As an example, one participant noted:

> "I don't think it would remain confidential, that is my concern. It's linking with other data, I don't know where it's going to end up eventually. I don't know who is going to see it. There is nothing specific that says who else is going to look at it. A lot of this information now is accessible through the internet."

Although limited, past research investigating the effects of trust on consent suggests a relationship (Cleary et al., 1984; Sala et al., 2010), with more trusting respondents being more likely to consent. We evaluated this in the Practicum survey, using the same methods as we did to evaluate privacy and confidentiality. However, the three trust items included in the survey were unrelated to consent, and trust concerns were mentioned by only a small portion of respondents who refused the consent request in the open-ended item (3%). The lack of a relationship between trust and consent in the Practicum survey may be related to the nature of the trust items, which assessed trust in other individuals. A more fitting measure may have been of respondents' trust in the institutions involved in record linkage. Related work by

Cleary et al. (1984) found that respondents who refused consent to medical record linkage were less trusting of their physician, as assessed through a survey item.

Although the Practicum survey did not include a formal or direct measure of altruism, we found that a sizeable percent of consenting respondents said that they did so in the open-ended item for altruistic reasons (41%), such as wanting to help the survey or the researchers involved. Other research reports a connection between altruism and consent as well (Dunn et al., 2004; Jenkins et al., 2006; Sala et al., 2010).

If privacy, confidentiality, and altruism continue to prove to be influences on consent, then researchers may want to leverage these findings to improve consent rates. This could be accomplished, for example, by emphasizing in the consent request how permitting access to administrative data promotes an altruistic goal, or how privacy and confidentiality of this information is protected and maintained.

Although past research and the research in this dissertation tend to focus on a specific set of respondent characteristics and attitudes such as those noted above, it is possible that other respondent factors are related to consent decisions. The open-ended question following the consent request in the 2011 Practicum survey attempted to identify a broader set of motives related to consenting or refusing consent, but these findings were limited given the brevity of the consent request, the context of the open-ended item in a phone survey, and the restriction of the scope of this research to a single item without any probes or follow-up questions by the interviewer.

Future qualitative research that is more extensive in scope may be effective in eliciting other respondent characteristics, qualities, or motives related to agreeing or refusing consent. This could include, for example, follow-up interviews with subsamples of survey respondents who consented and refused consent to identify more nuanced reasons for their decision than the 2011 JPSM Practicum survey permitted. Further, a large portion of refusing Practicum respondents provided general objections as their reason for not consenting. More in-depth qualitative research could help identify the more specific reasons behind their refusal. Survey researchers could then tailor consent requests to address these motives in order to improve consent rates.

Another important area largely ignored in the existing research is respondents' understanding of the consent request; that is, whether respondents comprehend the issue of record linkage well enough to provide their informed consent. Future research such as cognitive interviews should also examine whether respondents' comprehension of the consent request influences their willingness to consent, and if there are specific terms within the request that respondents fail to understand, or are prohibitive or beneficial to their willingness to consent.

5.3 Interviewer Characteristics

As discussed in Chapter 2, consent requests are included primarily in interviewer-administered surveys, and so a solid understanding of how interviewers can influence respondents' willingness to consent to record linkage is critical. A small body of past research investigates the effects of interviewers and interviewer characteristics on the consent request but findings are inconsistent, and sometimes in opposing directions. Existing research is also limited in that it fails to estimate such effects using surveys with interpenetrated sample design, and a large portion of relevant work was conducted outside of the U.S., which may not be fully relevant to domestic research given differences in policies and attitudes towards the statistical use of administrative records.

The fourth chapter of this dissertation extends existing work by examining the effects of interviewers on consent in the 2009 administration of one of the few surveys that both includes a consent request and uses an interpenetrated sample design, the NIS. In this research, we identified that overall, 7% of variation in consent rates was associated with the interviewer.

We also investigated the size of this intraclass correlation for particular subgroups of interviewers and found that the intraclass correlation varied depending on interviewer sample, with between 4-9% of variance in consent rates depending on interviewer subgroup. The intraclass correlation was lowest among the most experienced interviewers, Hispanic interviewers, and oldest interviewers, and greatest among the least experienced interviewers, Black interviewers, youngest interviewers and interviewers for whom demographic characteristics were missing in the dataset. Although it is unclear why interviewer race and ethnicity impacted variance in consent rates in this survey, the variation by interviewer age and experience is more easily understood, as interviewers may have improved or honed their techniques over time.

The NIS collects information from parents and guardians of young children (19-35 months) about the child's vaccination history, and requests access to their vaccination records. The sensitive nature of the survey may explain why female interviewers obtained higher consent rates than male in the 2009 administration if respondents perceived female interviewers as more maternal and less threatening than male interviewers. In a separate study, Tate et al. (2005) found that mothers were less likely to permit linkage of child's birth records to survey responses when a male translated the interview; the researchers proposed a similar explanation for these findings.

The other interviewer characteristics in the 2009 NIS, age and race/ethnicity, were unrelated to consent, perhaps because unlike gender, respondents may not have perceived any connection between the survey topic and these characteristics.

We also examined the relationship between prior NIS interviewing experience and consent rates in the 2009 NIS. The measure of NIS interviewing experience was limited in scope in that it was the number of days between the first and last call on the survey, rather than the number of completed calls or consent requests. The relationship between experience and consent was negative and significant in the 2009 NIS, indicating that interviewers with less prior interviewing experience actually achieved lower consent rates.

Although we proposed several explanations for this negative relationship (incorrect information in the disposition codes and interviewer credit schemes), the dataset truly does not provide enough information to fully understand why less experienced interviewers were more successful obtaining consent in this survey. However, the relationship between experience and consent can be further clarified and better understood if a more comprehensive measure of interviewing experience is used, and if additional studies investigate this relationship in the future. Replicating this analysis with additional datasets can help clarify the nature of the relationship between interviewing experience and consent.

A related area of research that currently remains untouched pertains to interviewer training. To date, existing research has yet to pursue the use of training to improve consent rates (e.g., identifying the optimal length and content of training). It is unknown how surveys typically prepare interviewers to administer consent requests, but at least one survey, NIS, does not formally train interviewers to administer this section. Training related specifically to issues pertaining to the consent request could enhance interviewers' ability to respond to issues or questions raised during the consent request, prepare them to address any respondent discomfort or hostility, conduct refusal confusion, and improve consent rates. Such training could also boost interviewers' comfort and confidence in administering this section and requesting any personally-identifying information, which may also be related to the consent rates they obtain.

5.4 Survey and Consent Request Characteristics

Finally, research in all three empirical chapters of this dissertation examined how design characteristics of the survey and consent request can influence rates of consent to record linkage. As compared to respondent and interviewer characteristics, far less research investigates how these design characteristics can influence consent rates. However, a sizeable amount of analogous research investigates how characteristics of the survey can influence response rates, and thus applying this literature to the consent request is a logical extension.

In Chapter 2 of this dissertation, the Descriptive Analysis, we analyzed the effect of design characteristics of the survey and consent request using data from all available existing surveys with consent requests. Although this research did not use experimental methods, it had the advantage of generalizability, as it was based on existing surveys representing a range of characteristics.

There are several noteworthy conclusions from the Descriptive Analysis. First, by documenting the number of surveys with consent requests over time, we could state that the number of survey administrations that include consent requests is, in fact, increasing with time over the years for which data was available (1982-2010), and that consent rates are declining. Further, we estimated an overall consent rate across the 162 consent requests that were included in this research; the mean was 70.2%, and the median was 73.1%.

We also identified that several characteristics of the survey and consent request were related to consent rates, such as survey mode, even after controlling for the other factors of interest in this research (survey response rate, administration year, survey sponsor, record topic, consent mode, identifier requested, and request type). Most surveys with consent requests are conducted in person, and doing so appears to benefit consent rates.

The mode of consent, that is, whether consent needed to be provided in writing or orally, was unrelated to consent rates in this chapter. However, this characteristic appeared to covary with the placement of the consent request, with written consent requests tending to fall at the end of the survey. Future experimental research could help determine if respondents are equally willing to consent to written and oral consent requests, or if the placement of the request is overriding any effect. For example, a study could use a two-by-two design, varying the placement of the consent request within the survey and the mode of consent (oral or written).

Many surveys request personally-identifying information from respondents to facilitate record linkage. This identifying information requested also appears to be influential in a respondent's decision to consent. Controlling for other sources of variation, consent rates are lower when a health-related identifier is requested (Medicare or Medicaid number, and healthcare or pharmacy provider contact information) as compared to SSN. This may speak to the added burden of retrieving a health-related identifier, which, unlike SSN, is less likely to be accessible by memory, rather than privacy concerns. Past literature primarily considers how the sensitivity of identifiers can affect respondents' willingness to provide them, but future research should also account for burden, and respondents' willingness to retrieve them.

The type of the consent request – whether it is implicit or explicit – demonstrated a strong effect on consent rates, with consent rates obtained through implicit procedures far higher than those obtained through implicit procedures. Given the consistent differences in consent rates obtained through these two procedures, future research should investigate respondents' understanding of the impending record linkage when an opt-out procedure is used through cognitive interviewing procedures; thus, is an implicit consent procedure as informative as an explicit procedure?

When controlling for other characteristics of interest in this research, we found no difference in consent rates for surveys sponsored by government and non-government organizations. Future research should investigate whether respondents are sensitive to or notice a difference in the sponsor of the survey, as well as the sponsor of the records, which may be influential on a respondents' decision to consent. Research by Singer et al. (2011) suggests that this decision is influenced by the sponsor of the administrative record.

Research using data from the 2009 NIS indicated that respondents who refused to participate in the survey on one or more calls were less likely to consent than those with no prior refusals, suggesting a relationship between survey response and consent. However, conducting this research on a larger scale, using the 22 surveys in the descriptive analysis, did not demonstrate any relationship between response rate and consent rate. Future research investigating the mechanisms that contribute to a respondents' willingness to consent may help determine if the same motives contribute to both consent and survey response, or if these are distinct behaviors and decisions with little or no overlap.

Both the Descriptive Analysis and the Practicum survey investigated the role of administrative record topic on consent rates, using different methods. The Descriptive Analysis drew upon the 162 consent rates from existing surveys, using non-experimental methods, and concluded that rates of consent were higher when respondents were asked to provide access to health-related records. The Practicum Survey used experimental methods, and randomly assigned telephone survey respondents to a request for either health or income and employment-related administrative records. Here, consent rates did not significantly vary by record topic.

Findings from both of these chapters contrast to research based on hypothetical consent scenarios included on the 2010 JPSM Practicum survey in which respondents reported that they found their health-related records to be more sensitive than those relating to income and employment, and would be less likely to provide access to health-related records if requested by the Census in a survey as compared to income and employment-administrative records.

Both the Practicum research and Descriptive Analysis have advantages and limitations – the Practicum research uses experimental methods, but lacks generalizability, and the Descriptive Analysis has the benefit of generalizability, yet is not based on experimental methods. However, in both studies, we have little understanding as to *why* these results occurred. For example, do consent rates vary by topic because respondents find the contents of one record more sensitive than the other? Here again, qualitative research could help researchers to understand why respondents may be more willing to allow access to one type of record than another, even when controlling for other factors such as any associated identifying information, and why these results varied across the two dissertation studies examining this issue.

In the Practicum survey, we also evaluated whether consent was influenced by a cash incentive. A random 40% of sample members in the Practicum survey were mailed a prepaid $5 cash incentive. Sample members mailed a cash incentive were no more likely to consent than the sample members not mailed the cash incentive.

Although the incentive was not effective in the 2011 Practicum survey, researchers may be interested in evaluating whether incentives improve consent rates in surveys with other design characteristics, such as in-person surveys, or consent requests in which respondents must provide consent in writing.

A final design characteristic investigated in this dissertation is the use of advance letters, which we investigated using data from the 2009 NIS. Here, we found that respondents for whom an address match could be made and were mailed an advance letter were more likely to consent. Yet because the advance letter was not randomly assigned, we cannot be certain as to whether matched respondents were fundamentally different or the advance letter truly affected consent rates (Traugott et al., 1987).

As demonstrated, design characteristics of the survey and consent request appear to be effective in influencing consent rates. Researchers should try to identify other design characteristics that could similarly influence consent, ideally through experimental methods, in order to develop best practices for researchers and practitioners who are interested in linking survey and administrative data and in obtaining the highest consent rates possible.

5.5 Interplay among Influences on Consent

Finally, we note that the influences on consent discussed in this dissertation – characteristics of the respondent, the interviewer, and design characteristics of the survey and consent request – may be interrelated. Many of the influential characteristics of the interviewer, survey, and consent request affect consent rates by influencing the respondents' attitudes rather than directly influencing consent rates'.

As an example, in the analysis of consent rates in the 2009 NIS, we found that female interviewers obtained higher consent rates and hypothesized that this was a function of the survey and consent request topic and target population. Given the survey topic and target population, NIS respondents may have been more trusting of female interviewers, seeing them as more maternal. Thus, interviewer gender did not directly influence consent rates, but may have affected respondents' level of trust which could have contributed to their consent decision.

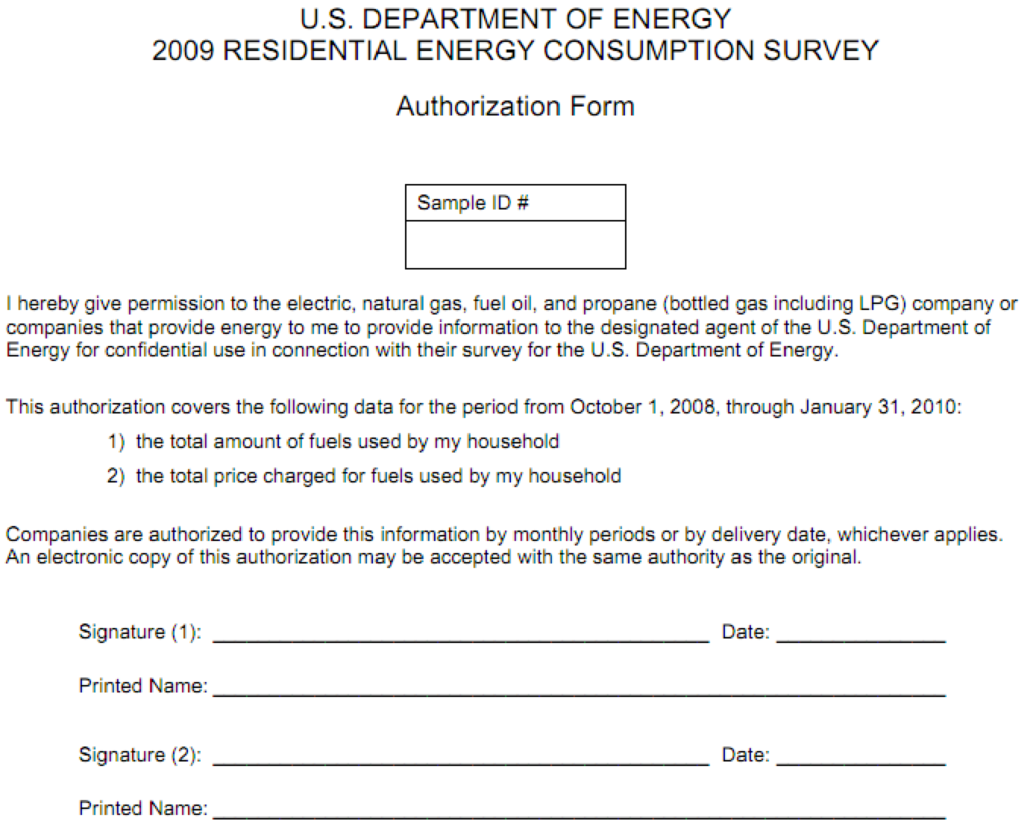
Design characteristics of the survey and consent request may similarly influence respondents' attitudes. Advance letters may reinforce the legitimacy of the research endeavor, increasing respondents' trust, and thus potentially their willingness to consent as well. Although survey sponsorship was unrelated to consent rates in this dissertation, it is feasible to assume that this relationship could act in the same way, by influencing respondents' trust in the research.

Similarly, the mode of survey administration could affect respondents' trust as well, thus leading to variation in consent rates. For example, both trust and consent rates may be greater if the survey was conducted in person, and both trust and consent rates may be lower if the survey was conducted over the telephone where it is more difficult to verify the interviewers' identity. Further, providing consent in a telephone-administered survey may be associated with greater privacy and confidentiality concerns as well if respondents cannot verify who they are sharing their personal information with, which may negatively affect consent decisions.

6 APPENDICES

6.1 Appendix to Chapter 1

U.S. Department of Energy, Energy Information Administration, 2009 Residential Energy Consumption Survey Household Questionnaire Authorization Form



U.S. DEPARTMENT OF ENERGY
2009 RESIDENTIAL ENERGY CONSUMPTION SURVEY

Authorization Form

Sample ID #

I hereby give permission to the electric, natural gas, fuel oil, and propane (bottled gas including LPG) company or companies that provide energy to me to provide information to the designated agent of the U.S. Department of Energy for confidential use in connection with their survey for the U.S. Department of Energy.

This authorization covers the following data for the period from October 1, 2008, through January 31, 2010:

1) the total amount of fuels used by my household
2) the total price charged for fuels used by my household

Companies are authorized to provide this information by monthly periods or by delivery date, whichever applies. An electronic copy of this authorization may be accepted with the same authority as the original.

Signature (1): ______________________________ Date: __________

Printed Name: ____

Signature (2): ______________________________ Date: __________

Printed Name: ____

6.2 Appendix to Chapter 2

6.2.1 Data Sources

1. **AHEAD: Assets and Health Dynamics (Among the Oldest Old)**
    a. Consent Rates:
        i. 1993 Medicaid and Medicare consent rates: Soldo, B.J., Hurd, M.D., Rodgers, W.L., Wallace, R.B (1997). Asset and Health Dynamics Among the Oldest Old: An Overview of the AHEAD Study, *The Journal of Gerontology Series B, 52B, 1-20.*
        ii. 1993 and 1995 SSN consent rates provided by HRS staff.
        iii. 1995 Medicaid and Medicare consent rates retrieved from http://hrsonline.isr.umich.edu/index.php?p=avail (need to create login to access public use data)
    b. Response Rates:
        i. 1993 and 1995 response rates available from http://hrsonline.isr.umich.edu/index.php?p=avail
    c. Survey and Consent Request Characteristics:
        i. Data Description: Assets and Health Dynamics Among the Oldest Old (AHEAD) http://hrsonline.isr.umich.edu/modules/meta/1993/core/desc/ahd93dd.pdf
        ii. HRS Sample Evolution http://hrsonline.isr.umich.edu/sitedocs/surveydesign.pdf
        iii. Soldo et al. (1997)

2. **Cleary & Jette (1984)**
    a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Cleary & Jette (1984).
    b. Additional Notes: We could not determine the exact year of data collection and coded it as 1984, the year the study was published.

3. **Current Population Survey (CPS) Annual Social and Economic Supplement**
    a. Consent Rates: Census Bureau CPS staff provided all CPS consent rates for years 1994 to 2000, and 2007 to 2010 (as percentages). For 2001 to 2006, CPS staff provided the number of consent refusals (numerator of consent rate), and consent rate was calculated using the number of adult respondents as the denominator. Consent rates for 2001 to 2006 were verified with CPS staff.
    b. Response Rates: All response rates were identified in the public use documentation through http://www.nber.org/cps/ (for each individual survey year). Response rates from 1994-1996 are missing from this analysis.
    c. Survey and Consent Request Characteristics: Characteristics were coded from CPS Annual Social and Economic Survey codebooks http://www.nber.org/cps/ (for each survey year).
    c. Additional Notes: As information regarding the transition from the direct SSN request to the opt-out request was largely undocumented online, details of these procedures were mainly provided by CPS staff at the Census Bureau.

4. **Health Interview Evaluation Survey (HIES)**
    a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Edwards et al. (1994).
    b. Additional Notes: No decimal place for consent or response rate was available.

5. **Health and Retirement Study (HRS)**
    a. Consent Rates:
        iv. HRS staff provided all SSN consent rates, noting that the 2010 SSN consent rate provided may not be final.
        v. All Medicaid and Medicare consent rates retrieved from http://hrsonline.isr.umich.edu/index.php?p=avail (need to create login to access public use data)

b. Response Rates: All response rates retrieved from http://hrsonline.isr.umich.edu/sitedocs/sampleresponse.pdf. To reflect changes in HRS sample composition over time, we use response rates from 1992-1996 as reported in Table 1; from 1998 onward, we use response rates from Table 2. Response rates from 2010 have not been released yet and are missing from this analysis.

c. Survey and Consent Request Characteristics:

    i. HRS Sample Sizes: http://hrsonline.isr.umich.edu/index.php?p=shoavail&iyear=LC (specify year)

    ii. IRB Information: http://hrsonline.isr.umich.edu/sitedocs/irb/HRS_IRB_WebPackage-09-09.pdf

    iii. HRS Sample Evolution: http://hrsonline.isr.umich.edu/sitedocs/surveydesign.pdf

    iv. HRS Growing Older in America: http://hrsonline.isr.umich.edu/index.php?p=dbook

6. **Health Field Study (HFS)**

    a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Jay et al. (1994).

    b. Additional Notes: No decimal place for consent or response rate was available. Information on consent mode was obtained from authors.

7. **Longitudinal Study of Aging *II (LSOA II)***

    a. Consent Rates: All consent rates were calculated at NCHS using limited information files created especially for this purpose.

    b. Response Rates: Missing.

    c. Survey and Consent Request Characteristics: Characteristics were coded from the LSOA II survey description found at http://www.cdc.gov/nchs/lsoa/lsoa2.htm

8. **McCarthy et al. (1999)**
    1. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from McCarthy et al. (1999).
    2. Additional Notes: No decimal place for consent or response rate was available.

9. **Medical Expenditure Panel Survey (MEPS)**
    1. Consent Rates: All consent information retrieved from http://meps.ahrq.gov/mepsweb/data_files/publications/annual_contractor_report/hc_ann_cntrct_methrpt.shtml (Table A-3 and A-4).
    2. Response Rates: All response rates retrieved from Table 1. (MEPS-HC overall response rates for public use files* (Point-in-Time)) http://meps.ahrq.gov/mepsweb/survey_comp/hc_response_rate.jsp
    3. Survey and Consent Request Characteristics:
        1. Survey Background: http://meps.ahrq.gov/mepsweb/about_meps/survey_back.jsp
        2. Authorization Forms: http://meps.ahrq.gov/mepsweb/survey_comp/survey.jsp#MPC
    4. Additional Notes:
        1. We calculated consent rate as the number of forms signed/the number of forms requested. As no individual respondent information was available, the data is not clustered by respondent.
        2. We aligned MEPS data to calendar year in order to compare consent rates with other surveys. In waves where panel round 3 overlaps two years, we assign this wave to the following year.
        3. The number of completed cases for 2010 is missing.

10. **Murdoch et al. (2010)**
    a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Murdoch et al. 2010
    b. Additional Notes: No decimal place for consent or response rate was available.

11. **National Health and Nutrition Examination Survey (NHANES)**
    a. Consent Rates: All consent rates were calculated at NCHS using limited information files created especially for this purpose.
    b. Response Rates: All response rates retrieved from http://www.cdc.gov/nchs/nhanes/response_rates_CPS.htm (interviewed sample)
    c. Survey and Consent Request Characteristics:
        i. NHANES Overview: http://www.cdc.gov/nchs/nhanes/about_nhanes.htm
        ii. Instruments by year: http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm
    d. Additional Notes: We code the year of administration for each NHANES survey as the final year the survey was fielded.

12. **National Health and Nutrition Examination Survey III (NHANES III)**
    a. Consent Rates: All consent rates were calculated at NCHS using limited information files created especially for this purpose.
    b. Response Rates: Response rates were retrieved from http://www.cdc.gov/nchs/data/nhanes/response_rates_cps/nh3_rr.pdf
    c. Survey and Consent Request Characteristics: Coded from NHANES III Documentation retrieved from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/nhanes/nhanes3/1A/ADULT-acc.pdf

d. Additional Notes: Although administration of NHEFS covered multiple years we code the year of administration as the final year the survey was fielded (1994).

13. **NHANES I Epidemiologic Follow-up Study (NHEFS)**
   a. Consent Rates: All consent rates were calculated at NCHS using limited information files created especially for this purpose.
      Response Rates
      http://www.cdc.gov/nchs/nhanes/response_rates_CPS.htm
   b. Survey and Consent Request Characteristics:
      i. NHANES main web page, retrieved from http://www.cdc.gov/nchs/nhanes/about_nhanes.htm#content
      ii. NHEFS operations document http://www.cdc.gov/nchs/data/nhanes/nhefs/sr01_022.pdf
   c. Additional Notes: Although administration of NHEFS covered multiple years we code the year of administration as the final year the survey was fielded (1984).

14. **National Health Interview Survey (NHIS)**
   a. Consent Rates: All consent rates were calculated at NCHS using limited information files created especially for this purpose.
   b. Response Rates, and Survey and Consent Request Characteristics: Response rates and survey characteristics were coded from the Data Description documents for each NHIS survey year, retrieved from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/(year)/srvydesc.pdf
   c. Additional Notes: The sample of NHIS respondents asked for consent varies by year. To standardize consent rates across years, we include consent rates only from the sample adult.

15. **National Immunization Survey (NIS)**

   a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: Identified through the Data User's Guide for each survey year accessed through http://www.cdc.gov/nchs/nis/data_files_04_prior.htm and http://www.cdc.gov/nchs/nis/data_files.htm

16. **National Immunization Survey Teen (NIS Teen)**

   a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: Identified through the Data User's Guide for each survey year accessed through http://www.cdc.gov/nchs/nis/data_files_teen.htm

17. **Partin et al. (2008)**

   a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Partin et al. (2008).

18. **Panel Study for Income Dynamics (PSID)**

   a. Consent Rates: Calculated using public-use data http://simba.isr.umich.edu/VS/s.aspx (need to create login)
   b. Response Rates: Response rates identified in http://psidonline.isr.umich.edu/data/Documentation/UserGuide2009.pdf
   c. Survey and Consent Request Characteristics:
      PSID Main Interview Manual:
      http://psidonline.isr.umich.edu/data/Documentation/UserGuide2009.pdf

19. **Residential Energy Consumption Survey (RECS)**

   a. Consent Rates and Response Rates: EIA RECS staff provided all RECS consent rates and response rates.
   b. Survey and Consent Request Characteristics:
      i. RECS 2009 Methodology: http://www.eia.gov/consumption/residential/methodology/2009/brief.cfm

        ii. Archived RECS Publications: http://www.eia.gov/consumption/residential/data/archive.cfm

20. **Survey of Health Insurance and Program Participation (SHIPP)**
    a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Pascale (2011).

21. **Survey of Income and Program Participation (SIPP)**
    a. Consent Rates and Response Rates: Census Bureau SIPP staff provided all SIPP consent rates and response rates.
    b. Survey and Consent Request Characteristics:
        i. SIPP Users Guide (by year): http://www.census.gov/sipp/usrguide.html
        ii. SIPP Technical Documentation (by year): http://www.census.gov/apsd/techdoc/sipp/sipp.html
        iii. SIPP Survey Contents: http://www.census.gov/sipp/content.html
    c. Additional Notes: For some of the SIPP panels included in this analysis, data no longer existed from some individual waves. Missing wave data includes the following:

1990 Panel: Wave 1
1991 Panel: Wave 2-4
1996 Panel: Wave 6-9
2001 Panel: Wave 1-5
(Consent rates from the other waves were included in this research.)

As information regarding the transition from the direct SSN request to the opt-out request was largely undocumented online, details of these procedures were mainly provided by SIPP staff at the Census Bureau.

22. **Woolf et al. (2000)**

a. Consent Rate/Response Rate/Survey and Consent Request Characteristics: We coded consent rate, response rate, and survey and consent request characteristics from Woolf et al. (2000).

b. Additional Notes: No decimal place for consent or response rate was available.

6.2.2 Additional Tables

Table 2.13

Linear Regression Predicting Consent as a Function of Survey and Consent Request Characteristics – Expansion of Health Identifiers

		Coef.	SE	*p*
Constant		121.66	10.90	0.000
Year		-0.70	0.32	0.040
Survey Mode†	In-Person (ref.)			
	Mail	-27.99	10.76	0.018
	Phone: Panel	-5.40	8.95	0.553
	Phone: Cross-Sectional	-10.73	3.59	0.008
Survey Sponsor	Government (ref.)			
	Other	-3.72	5.84	0.532
Record Topic	Health (ref.)			
	Income/Employment	-37.25	9.68	0.001
	Other (*utility*)	-2.34	6.49	0.722
Consent Mode	Oral (ref.)			
	Written	-23.62	4.41	0.000
Identifier Requested†	SSN (ref.)			
	Medicare	-48.42	11.65	0.001
	Medicaid	-47.97	11.36	0.000
	MD Authorization	-15.84	7.59	0.051
	Pharm. Authorization	-8.48	7.64	0.281
	None	-0.58	6.52	0.930
Request Type	Explicit (ref.)			
	Implicit	30.65	9.08	0.003
Survey Response Rate		0.07	0.06	0.273

Notes: N=154; r^2=0.53; adjusted Wald test for all parameters: $F(9,7) = 4196.75$, $p<0.001$. † SAQ (Survey Mode) was dropped from the model; Other (Identifier Requested) was dropped from the model.

6.3 Appendix to Chapter 3

6.3.1 Practicum Questionnaire

PROGRAMMER: FOR ALL EXPERIMENTS, CREATE FLAG VARIABLES TO SHOW WHICH RESPONDENTS WERE ASSIGNED TO WHICH CONDITIONS. FOR ALL RANDOMIZATIONS, INCLUDE RANDOMIZATION ASSIGNMENT VARIABLES. ALL EXPERIMENTS AND RANDOMIZATIONS ARE INDEPENDENT OF ONE ANOTHER. INCLUDE TIME STAMPS NOTING START AND END TIME FOR ENTIRE INTERVIEW, AS WELL AS FOR EACH SECTION.

MAIN INTRODUCTION
Hello, my name is _______, and I am calling on behalf of the University of Maryland. May I please speak with **[RESPONDENT NAME]**?

ONCE TARGET RESPONDENT IS ON THE PHONE:
[IF RESPONDENT DID NOT ANSWER PHONE, REPEAT: Hello, my name is _______, and I am calling on behalf of the University of Maryland.**]** We're conducting a nationwide study on health, economic and other issues and we would like to include your opinions. Your participation is voluntary, and your responses will be used for research purposes only. To begin...
[READ IF NECESSARY: The interview will only take about 20 minutes to complete.**]**

VOICEMAIL MESSAGE (LEAVE ONLY ONCE -- THE FIRST TIME A CALL GOES TO VOICEMAIL):
Hello, I am calling on behalf of the University of Maryland. We're conducting a nationwide study on health, economic and other issues. This is NOT a sales call. I am sorry we missed you today and will try to reach you again. If you would like, please call us back at 1-800-887-3150 Monday through Friday 9 AM- 11:00 PM Eastern Daylight Time or 10:00 AM -10:00PM Eastern Daylight Time on Saturday and Sunday to schedule the interview. Have a good (day/evening). **[SCHEDULE CALLBACK]**

ASK ALL
SEX Respondent's sex **(DO NOT ASK; RECORD BY OBSERVATION)**
{Formerly Q65}
1 Male
2 Female

SECTION 1: GENERAL HEALTH
ASK RANDOM HALF SAMPLE (EXPERIMENT1=1)
Q1A Would you say your health in general is excellent, very good, good, fair, or poor?

1 Excellent
2 Very good
3 Good
4 Fair
5 Poor
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT1=2)

Q1B Would you say your health in general is poor, fair, good, very good, or excellent?

1 Poor
2 Fair
3 Good
4 Very good
5 Excellent
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF DESCRIBED HEALTH STATUS (Q1A=1-5 or Q1B=1-5)

Q2 Why do you feel that your health is **[INSERT RESPONSE FROM Q1a or Q1b]**? **[OPEN-END; RECORD VERBATIM RESPONSE]**

1 **[RECORD VERBATIM RESPONSE]**
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT2=1)

Q3A People do different things in order to stay healthy. Which of the following do you think is the MOST important thing for a person to do in order to stay healthy – eat right, get enough sleep, reduce stress, have a yearly physical, or get regular exercise?

[READ IF NECESSARY: If you had to choose just one, which do you think is most important?**]**

1 Eat right
2 Get enough sleep
3 Reduce stress
4 Have a yearly physical
5 Get regular exercise
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT2=2)

Q3B People do different things in order to stay healthy. Which of the following do you think is the MOST important thing for a person to do in order to stay healthy – get regular exercise, have a yearly physical, reduce stress, get enough sleep, or eat right?

[READ IF NECESSARY: If you had to choose just one, which do you think is most important?**]**

1 Get regular exercise
2 Have a yearly physical
3 Reduce stress
4 Get enough sleep
5 Eat right
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

[READ TO ALL:] The next questions are about foods you may have eaten in the past 7 days.

[RANDOMIZE Q4A1/Q4B1 WITH Q5A1/Q5B1; RANDOM HALF WILL GET Q4A1/Q4B1 FIRST (EXPERIMENT16=1) AND RANDOM HALF WILL GET Q5A1/Q5B1 FIRST (EXPERIMENT16=2)]

ASK RANDOM HALF SAMPLE (EXPERIMENT3=1)

Q4A1 In the past 7 days, how many servings of fruit did you eat? **[READ IF NECESSARY:** A serving of fruit is equal to about one half cup of fruit.**]**
[IF PROVIDES "PER DAY" RESPONSE, ASK: So how many servings of fruit would that be in the past 7 days?**]**

___ **[RECORD EXACT NUMBER 0-97]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT3=2)

Q4B1 In the past 7 days, how many servings of fruit did you eat? Please do not include apples, bananas or oranges. **[READ IF NECESSARY:** A serving of fruit is equal to about one half cup of fruit.**]**
[IF PROVIDES "PER DAY" RESPONSE, ASK: So how many servings of fruit would that be in the past 7 days?**]**

___ **[RECORD EXACT NUMBER 0-97]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

INT_4 INTERVIEWER: DO NOT ASK; CODE RESPONDENT'S APPROACH TO ANSWERING QUESTION: IF R DID NOT SAY HOW, RECORD AS "DON'T KNOW"; RECORD AS MANY AS APPLY

1 Counted each individual serving (e.g., 1 serving Tuesday plus 3 servings Friday)
2 Used average daily servings to arrive at answer (e.g., I had 2 per day)
3 Thought about types of fruits and added them up
4 Other **(SPECIFY)**

98 **(DO NOT READ)** Don't know

ASK RANDOM HALF SAMPLE (EXPERIMENT4=1)

Q5A1 In the past 7 days, how many servings of vegetables did you eat? **[READ IF NECESSARY:** A serving of vegetables is equal to about one half cup of vegetables.**]**

[IF PROVIDES "PER DAY" RESPONSE, ASK: So how many servings of vegetables would that be in the past 7 days?**]**

___ **[RECORD EXACT NUMBER 0-97]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT4=2)

Q5B1 In the past 7 days, how many servings of vegetables did you eat? Please do not include carrots, beans, or lettuce. **[READ IF NECESSARY:** A serving of vegetables is equal to about one half cup of vegetables.**]**

[IF PROVIDES "PER DAY" RESPONSE, ASK: So how many servings of vegetables would that be in the past 7 days?**]**

___ **[RECORD EXACT NUMBER 0-97]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

INT_5 INTERVIEWER: DO NOT ASK; CODE RESPONDENT'S APPROACH TO ANSWERING QUESTION: IF R DID NOT SAY HOW, RECORD AS "DON'T KNOW"; RECORD AS MANY AS APPLY

1 Counted each individual serving (e.g., 2 servings Tuesday plus 3 on Friday)
2 Used average daily servings to arrive at answer (e.g., I had 2 per day)
3 Thought about types of vegetables and added them up
4 Other **(SPECIFY)**
98 **(DO NOT READ)** Don't know

[RANDOMIZE ORDER OF Q6A1 / B1 / C1 / D1 / E1 / F1]

[READ TO ALL:] Has a doctor or other health professional EVER told you that you have any of the following?

ASK ALL

Q6A1 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Diabetes or sugar diabetes **[IF FEMALE (SEX=2), INSERT:** other than during pregnancy**]**? **[READ IF NECESSARY:** Has a doctor or other health professional EVER told you that you have this condition?**]**

[INTERVIEWER NOTE: Do not accept self-diagnosed or diagnosed by a person who is not a doctor or other health professional**]**

1 Yes

2 No
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF DIAGNOSED WITH DIABETES (Q6A1=1)

Q6A2 How old were you when you were first diagnosed with diabetes or sugar diabetes?
[READ IF NECESSARY: Just your best guess is fine.**]**
___ years old **[RECORD EXACT AGE 1-96]**
0 Less than 1 year old
97 97 years old or older
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q6B1 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Hypertension or high blood pressure? **[READ IF NECESSARY:** Has a doctor or other health professional EVER told you that you have this condition?**]**
[INTERVIEWER NOTE: Do not accept self-diagnosed or diagnosed by a person who is not a doctor or other health professional**]**
1 Yes
2 No
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF DIAGNOSED WITH HYPERTENSION OR HIGH BLOOD PRESSURE (Q6B1=1)

Q6B2 How old were you when you were first diagnosed with hypertension or high blood pressure?
[READ IF NECESSARY: Just your best guess is fine.**]**
___ years old **[RECORD EXACT AGE 1-96]**
0 Less than 1 year old
97 97 years old or older
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q6C1 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Asthma? **[READ IF NECESSARY:** Has a doctor or other health professional EVER told you that you have this condition?**]**
[INTERVIEWER NOTE: Do not accept self-diagnosed or diagnosed by a person who is not a doctor or other health professional**]**
1 Yes
2 No
98 **(DO NOT READ)** Don’t know

99 **(DO NOT READ)** Refused

ASK IF DIAGNOSED WITH ASTHMA (Q6C1=1)

Q6C2 How old were you when you were first diagnosed with asthma?
[READ IF NECESSARY: Just your best guess is fine.**]**
___ years old **[RECORD EXACT AGE 1-96]**
0 Less than 1 year old
97 97 years old or older
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q6D1 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Arthritis?
[READ IF NECESSARY: Has a doctor or other health professional EVER told you that you have this condition?**]**
[INTERVIEWER NOTE: Do not accept self-diagnosed or diagnosed by a person who is not a doctor or other health professional**]**
1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF DIAGNOSED WITH ARTHRITIS (Q6D1=1)

Q6D2 How old were you when you were first diagnosed with arthritis?
[READ IF NECESSARY: Just your best guess is fine.**]**
___ years old **[RECORD EXACT AGE 1-96]**
0 Less than 1 year old
97 97 years old or older
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q6E1 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Heart disease? **[READ IF NECESSARY:** Has a doctor or other health professional EVER told you that you have this condition?**]**
[INTERVIEWER NOTE: Do not accept self-diagnosed or diagnosed by a person who is not a doctor or other health professional**]**
1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF DIAGNOSED WITH HEART DISEASE (Q6E1=1)

Q6E2 How old were you when you were first diagnosed with heart disease?
[READ IF NECESSARY: Just your best guess is fine.**]**
___ years old **[RECORD EXACT AGE 1-96]**

0 Less than 1 year old
97 97 years old or older
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL
Q6F1 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Anemia?
[READ IF NECESSARY: Has a doctor or other health professional EVER told you that you have this condition?**]**
[INTERVIEWER NOTE: Do not accept self-diagnosed or diagnosed by a person who is not a doctor or other health professional**]**
1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF DIAGNOSED WITH ANEMIA (Q6F1=1)
Q6F2 How old were you when you were first diagnosed with anemia?
[READ IF NECESSARY: Just your best guess is fine.**]**
___ years old **[RECORD EXACT AGE 1-96]**
0 Less than 1 year old
97 97 years old or older
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL
Q7 In 2010, were you a patient in a hospital overnight? Do not include an overnight stay in the emergency room.
1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF HOSPITAL INPATIENT IN 2010 (Q7=1)
Q8 How many times were you a patient in a hospital overnight or longer during 2010? Do not count the total number of nights, just the total number of hospital admissions for stays which lasted 1 or more nights.
___ **[RECORD EXACT NUMBER 1-97]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL
Q9A During 2010, how many times did you see a doctor or other health care professional about your health at a doctor's office, a clinic, hospital emergency room, at home or some other place? **[IF HOSPITAL INPATIENT (Q7=1), READ:** Do not include times you were hospitalized overnight.**]**

[READ IF NECESSARY: How many times would that be for all of 2010?**]**
[IF DK or REF, READ: Just your best guess is fine.**]**
___ **[RECORD EXACT NUMBER 0-97]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF DK OR REF IN Q9A (Q9A=98,99)
Q9B Would it be closer to 0 visits, 1 to 3 visits, 4 to 9 visits, or 10 or more visits?
1 0 visits/None
2 1 to 3
3 4 to 9
4 10 or more
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE WHO VISITED A DOCTOR IN 2010 (EXPERIMENT5=1 and [Q9A=1-97 or Q9B=2-4])
Q10A Which of the following describes how you came up with your answer? Did you think about EACH INDIVIDUAL VISIT; did you think about HOW OFTEN you usually go to the doctor; did you think about TYPES of visits; or did you estimate based on a GENERAL IMPRESSION? **[ALLOW MULTIPLE RESPONSES]**
1 Think about each visit
2 Think about how often you usually go to the doctor
3 Think about types of visits
4 Estimate based on a general impression
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE WHO VISITED A DOCTOR IN 2010 (EXPERIMENT5=2 and [Q9A=1-97 or Q9B=2-4])
Q10B Which of the following describes how you came up with your answer? Did you estimate based on a GENERAL IMPRESSION; did you think about TYPES of visits; did you think about HOW OFTEN you usually go to the doctor; or did you think about EACH INDIVIDUAL VISIT? **[ALLOW MULTIPLE RESPONSES]**
1 Estimate based on a general impression
2 Think about types of visits
3 Think about how often you usually go to the doctor
4 Think about each visit
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

SECTION 2: HEALTH INSURANCE

ASK ALL

Q11 The next questions are about health insurance. Include health insurance obtained through employment or purchased directly, as well as government insurance programs like Medicare and Medicaid. Are you covered by any kind of health insurance or health care plan?
[INTERVIEWER NOTE: Health insurance and health care plans include private health insurance, Medicare, Medi-gap, Medicaid, SCHIP/CHIP, military healthcare (TRI-CARE/VA/CHAMP-VA), Indian health service, state-sponsored health plan, other government program, a single service plan (e.g. dental, vision, or prescription)**]**

1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE WHO ARE INSURED (EXPERIMENT6=1 and Q11=1)

Q12A Would you rate your health insurance as excellent, very good, good, fair, or poor?

1 Excellent
2 Very good
3 Good
4 Fair
5 Poor
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE WHO ARE INSURED (EXPERIMENT6=2 and Q11=1)

Q12B Would you rate your health insurance as poor, fair, good, very good, or excellent?

1 Poor
2 Fair
3 Good
4 Very good
5 Excellent
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF RATED HEALTH INSURANCE (Q12A=1-5 or Q12B=1-5)

Q13 Why do you feel that your health insurance is **[INSERT RESPONSE FROM Q12A or Q12B]**? **[OPEN-END; RECORD VERBATIM RESPONSE]**

1 **[RECORD VERBATIM RESPONSE]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q14 The next question is about money that you have spent on medical and dental care for yourself only. Please do NOT count health insurance premiums, over-the-counter drugs, or costs that you were reimbursed for. In 2010, about how much did you spend for medical and dental care? Would you say it was zero dollars... some money but less than $500... $500 to less than $2,000... $2,000 to less than $3,000... $3,000 to less than $5,000... or $5,000 or more?

1 Zero dollars
2 Some money but less than $500
3 $500 to less than $2,000
4 $2,000 to less than $3,000
5 $3,000 to less than $5,000
6 $5,000 or more
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

SECTION 3: POLITICS

[READ TO ALL:] The next questions are about government involvement in health care. Please tell me whether you agree or disagree with the following statements.

ASK RANDOM HALF SAMPLE (EXPERIMENT7=1)

Q15A Increasing government involvement in health care will improve the quality of care. Do you agree or disagree?

1 Agree
2 Disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT7=2)

Q15B Increasing government involvement in health care will hurt the quality of care. Do you agree or disagree?

1 Agree
2 Disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT8=1)

Q16A It is a violation of individual rights for the federal government to require that everyone have health insurance. Do you agree or disagree?

1 Agree
2 Disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT8=2)

Q16B It is the responsibility of the federal government to require that everyone have health insurance. Do you agree or disagree?

1 Agree
2 Disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

Q17-21 Now I'd like to ask you about some institutions in American society. As I read each one, please tell me how much confidence you have in that institution using a scale from 1 to 10, where 1 means "no confidence at all" and 10 means "great confidence." First, on a scale of 1 to 10, how much confidence do you have in **[RANDOMIZE ORDER OF Q17-Q21]**? (Next,) how about…**[INSERT NEXT ITEM]**?

[READ IF NECESSARY: On a scale from 1 to 10, how much confidence do you have in this institution?**]**

[READ IF NECESSARY: You can use any number between 1 and 10, where 1 means "no confidence at all" and 10 means "great confidence."**]**

Q17. Congress
Q18. The news media
Q19. The public school system
Q20. The criminal justice system
Q21. The health care system

CATEGORIES

___ **[RECORD EXACT NUMBER 1-10]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT9=1)

Q22A Which of the following do you think is the MOST important thing for Congress to concentrate on right now: the wars in Iraq and Afghanistan, the gap between the rich and the poor, climate change, illegal immigration, or dependence on foreign oil?

[READ IF NECESSARY: If you had to pick from just these 5 choices, which ONE do you think should be the top priority?**]**

1 The wars in Iraq and Afghanistan
2 The gap between the rich and the poor
3 Climate change
4 Illegal immigration
5 Dependence on foreign oil
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT9=2)

Q22B Which of the following do you think is the MOST important thing for Congress to concentrate on right now: dependence on foreign oil, illegal

immigration, climate change, the gap between the rich and the poor, or the wars in Iraq and Afghanistan?
[READ IF NECESSARY: If you had to pick from just these 5 choices, which ONE do you think should be the top priority?**]**

1 Dependence on foreign oil
2 Illegal immigration
3 Climate change
4 The gap between the rich and the poor
5 The wars in Iraq and Afghanistan
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

[READ TO ALL:] Now I'm going to read a few statements that some people agree with but others disagree with.

ASK RANDOM HALF SAMPLE (EXPERIMENT10=1)
Q23A Do you agree or disagree: Economic growth should be given priority, even if the environment suffers to some extent.

1 Agree
2 Disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT10=2)
Q23B Do you agree or disagree: Protection of the environment should be given priority, even at the risk of slowing economic growth.

1 Agree
2 Disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT11=1)
Q24A Do you agree or disagree: Global warming has been proven.

1 Agree
2 Disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT11=2)
Q24B Do you agree or disagree: Global warming has not been proven.

1 Agree
2 Disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

SECTION 4: EMPLOYMENT

[READ TO ALL:] Now I'm going to ask you some questions about your current employment situation.

ASK ALL

Q25 Last week, were you working full-time, part-time, going to school, keeping house, or what? **[CODE ONE RESPONSE ONLY; IF MORE THAN ONE RESPONSE, GIVE PREFERENCE TO FIRST MENTION]**

1 Working full-time
2 Working part-time
3 With a job, but not at work because of temporary illness, vacation, strike
4 Unemployed, laid off, looking for work
5 Retired
6 In school
7 Keeping house
8 Other **(SPECIFY)**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK EMPLOYED FT/PT OR WITH JOB BUT NOT AT WORK (Q25=1,2,3)

Q26 I have a few questions about your current job. Do you work for a private company, a non-profit organization, or for the government or a government agency?

[INTERVIEWER NOTE: If R has more than 1 job, they should answer about the job where they work the most hours.**]**

1 Private company
2 A non-profit organization
3 For the government or a government agency
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK EMPLOYED FT/PT OR WITH JOB BUT NOT AT WORK (Q25=1,2,3)

Q27 How many hours a week do you usually work, at all jobs? **[INTERVIEWER:** If R gives a partial hour (e.g. "15 minutes" or "an hour and a half," please round up to the nearest whole number.**]**

___ hours **[RECORD EXACT NUMBER 0-96]**
97 97 hours or more
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

[READ TO ALL:] We are trying to understand how people all over the country are getting along financially, so now I have some questions about earnings and income.

ASK RANDOM HALF SAMPLE (EXPERIMENT12=1)

Q28A In 2010, how much was your total family income, from all sources, before taxes? Total income includes interest or dividends, rent, Social Security, other

pensions, alimony or child support, unemployment compensation, public aid or welfare, armed forces or veteran's allotment.
[INTERVIEWER NOTE: If R refuses once, READ: "Information about your income is very important. We greatly appreciate your response and will keep it strictly confidential." **IF STILL REFUSED, CODE AS REFUSED. IF R GIVES RANGE, PROBE FOR A DOLLAR AMOUNT.]**

_____ dollars **[RECORD EXACT NUMBER 0-499,999]**
500000 $500,000 or more
777777 **(DO NOT READ)** Don't know
888888 **(DO NOT READ)** Refused

ASK IF REFUSED IN Q28A (Q28A=888888)
Q28D1 In 2010, was your total family income from all sources, before taxes, more than $50,000?
1 Yes, more than $50,000
2 No, under $50,000 (incl. exactly $50,000)
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS $50,000 OR LESS (Q28D1=2)
Q28D2 Was it more than $25,000?
1 Yes, more than $25,000
2 No, under $25,000 (incl. exactly $25,000)
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS $25,000 OR LESS (Q28D2=2)
Q28D3 Was it more than $10,000?
1 Yes, more than $10,000
2 No, under $10,000 (incl. exactly $10,000)
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS MORE THAN $50,000 (Q28D1=1)
Q28D4 Was it more than $75,000?
1 Yes, more than $75,000
2 No, under $75,000 (incl. exactly $75,000)
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM QUARTER SAMPLE (EXPERIMENT12=2)
Q28B1 In 2010, was your total family income from all sources, before taxes, more than $50,000? Total income includes interest or dividends, rent, Social Security, other pensions, alimony or child support, unemployment compensation, public aid or welfare, armed forces or veteran's allotment.

[INTERVIEWER NOTE: If R refuses once, READ: “Information about your income is very important. We greatly appreciate your response and will keep it strictly confidential.” **IF STILL REFUSED, CODE AS REFUSED]**

1 Yes, more than $50,000
2 No, under $50,000 (incl. exactly $50,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS $50,000 OR LESS (Q28B1=2)
Q28B2 Was it more than $25,000?

1 Yes, more than $25,000
2 No, under $25,000 (incl. exactly $25,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS $25,000 OR LESS (Q28B2=2)
Q28B3 Was it more than $10,000?

1 Yes, more than $10,000
2 No, under $10,000 (incl. exactly $10,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS MORE THAN $50,000 (Q28B1=1)
Q28B4 Was it more than $75,000?

1 Yes, more than $75,000
2 No, under $75,000 (incl. exactly $75,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK RANDOM QUARTER SAMPLE (EXPERIMENT12=3)
Q28C1 In 2010, was your total family income from all sources, before taxes, more than $25,000? Total income includes interest or dividends, rent, Social Security, other pensions, alimony or child support, unemployment compensation, public aid or welfare, armed forces or veteran's allotment. **[INTERVIEWER NOTE: If R refuses once, READ:** “Information about your income is very important. We greatly appreciate your response and will keep it strictly confidential.” **IF STILL REFUSED, CODE AS REFUSED]**

1 Yes, more than $25,000
2 No, under $25,000 (incl. exactly $25,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS $25,000 OR LESS (Q28C1=2)
Q28C2 Was it more than $10,000?

1 Yes, more than $10,000
2 No, under $10,000 (incl. exactly $10,000)
98 **(DO NOT READ)** Don’t know

99 **(DO NOT READ)** Refused

ASK IF INCOME IS MORE THAN $25,000 (Q28C1=1)
Q28C3 Was it more than $50,000?
1 Yes, more than $50,000
2 No, under $50,000 (incl. exactly $50,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK IF INCOME IS MORE THAN $50,000 (Q28C3=1)
Q28C4 Was it more than $75,000?
1 Yes, more than $75,000
2 No, under $75,000 (incl. exactly $75,000)
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL
Q29A During 2010, did you receive any income from the following sources: Social Security?
1 Yes
2 No
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL
Q29B How about other retirement or pensions? **[READ IF NECESSARY:** During 2010, did you receive any income from this source?**]**
1 Yes
2 No
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL
Q29C How about public assistance or welfare, including Supplemental Security Income, or SSI? **[READ IF NECESSARY:** During 2010, did you receive any income from this source?**]**
1 Yes
2 No
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

SECTION 5: CONSENT REQUEST

ASK RANDOM HALF SAMPLE (EXPERIMENT13=1)
CR1A We would like to understand how the use of health care may change as people age. To do that, we need to obtain information about vital statistics, health care costs and diagnoses from your health-related records. In order for us to

retrieve these records, we need your consent. This will allow us to conduct more research without asking additional questions. Your consent is voluntary and the information that you provide will be kept completely confidential. May I have your consent to access these records?

1 Consents
2 Declines to consent
98 **(DO NOT READ)** Don't know

ASK RANDOM HALF SAMPLE (EXPERIMENT13=2)

CR1B We would like to understand how people's income changes as they age. To do that, we need to obtain information about income and employment from your income and employment-related records. In order for us to retrieve these records, we need your consent. This will allow us to conduct more research without asking additional questions. Your consent is voluntary and the information that you provide will be kept completely confidential. May I have your consent to access these records?

1 Consents
2 Declines to consent
98 **(DO NOT READ)** Don't know

ASK THOSE WHO CONSENTED (CR1A=1 or CR1B=1)

CR2A Can you tell me why you decided to consent to this request to access your records? **[OPEN-END; RECORD VERBATIM RESPONSE]**

1 **[RECORD VERBATIM RESPONSE]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK THOSE WHO DECLINED TO CONSENT / DK (CR1A=2,98 or CR1B=2,98)

CR2B I appreciate your patience and I indicated that you do NOT consent. Before we move on to the next section, can you tell me why you decided not to consent to this request to access your records? **[OPEN-END; RECORD VERBATIM RESPONSE]**

1 **[RECORD VERBATIM RESPONSE]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

INT_CR1 INTERVIEWER: DO NOT ASK; Please note any reactions the respondent had to the consent request; Record as many as apply.

1 Hostile
2 Confidentiality concerns
3 Needed clarification
4 Respondent had no reaction
5 Other **(SPECIFY)**
98 **(DO NOT READ)** Don't know

SECTION 6: PATRIOTISM
THERE ARE NO Q30 THRU Q35

[READ TO ALL:] Now I am going to read you a few statements. After each one, please tell me how proud you are of America in that area. First...

[RANDOMIZE ORDER OF Q36-Q45]

ASK ALL

Q36 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** The way democracy works.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q37 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its political influence in the world.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q38 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** America's economic achievements.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud

4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q39 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its social security system.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**
1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q40 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its scientific and technological achievements.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**
1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q41 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its achievements in sports.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**
1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q42 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its achievements in the arts and literature.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q43 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** America's armed forces.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q44 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its history.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q45 **[IF RANDOMIZED 2ND-THRU-LAST, INSERT:** How about**]** Its fair and equal treatment of all groups in society.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED ITEMS IN THE SERIES, THEN AS NECESSARY: Would you say you are very

proud, somewhat proud, not very proud, or not proud at all (of America in this area)?**]**

1 Very proud
2 Somewhat proud
3 Not very proud
4 Not proud at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

SECTION 7: PRIVACY

[READ TO ALL:] The next questions are about the collection of information by government and businesses.

ASK ALL

Q46 Every ten years, including 2010, most households are sent a Census questionnaire that includes a few questions about everyone living there. Would you agree or disagree that the Census is an invasion of privacy?

1 Agree
2 Disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q47 Do you think the government bothers you too much with requests for information?

1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

[RANDOMIZE ORDER OF Q48-49]

ASK ALL

Q48 **[IF RANDOMIZED SECOND, READ:** What about your medical records?**]** How much would it bother you if your medical records were not kept confidential? Would it bother you a lot, some, a little, or not at all?

1 A lot
2 Some
3 A little
4 Not at all
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q49 **[IF RANDOMIZED SECOND, READ:** What about your income tax records?**]** How much would it bother you if your income tax records were not kept confidential? Would it bother you a lot, some, a little, or not at all?

1 A lot
2 Some
3 A little
4 Not at all
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q50 Please tell me if you strongly agree, somewhat agree, somewhat disagree, or strongly disagree: People have lost all control over how personal information about them is used.

1 Strongly agree
2 Somewhat agree
3 Somewhat disagree
4 Strongly disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q51 Generally speaking, would you say that most people can be trusted, or that you can’t be too careful in dealing with people?

1 Most people can be trusted
2 You can’t be too careful
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q52 Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?

1 Try to be helpful
2 Just look out for themselves
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q53 Do you think most people would try to take advantage of you if they got a chance, or that they would try to be fair?

1 Would try to take advantage of you
2 Would try to be fair
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q54 How often do you worry about being a victim of identity theft – frequently, occasionally, rarely, or never?

1 Frequently
2 Occasionally
3 Rarely
4 Never
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

SECTION 8: CONSCIENTIOUSNESS

[READ TO ALL:] Now I am going to read a few statements that may or may not describe you. For each statement, please tell me whether you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree. First…

[RANDOMIZE ORDER OF Q55-Q60D]
ASK ALL

Q55 I am always prepared.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q56 I carry out my plans.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q57 I pay attention to details.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q58 I waste my time.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q59 I do just enough work to get by.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q60 I don't see things through.

[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q60A I make plans and stick to them.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q60B I have difficulty getting started doing work.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q60C I avoid my duties.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree

3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q60D I get chores done right away.
[INTERVIEWER: READ FOR FIRST 2 RANDOMIZED STATEMENTS IN THE SERIES, THEN AS NECESSARY: Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?**]**

1 Strongly agree
2 Somewhat agree
3 Neither agree nor disagree
4 Somewhat disagree
5 Strongly disagree
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

SECTION 9: BACKGROUND

[READ TO ALL:] And now just a few background questions.

ASK ALL

QTV LAST WEEK, how many hours did you spend watching television?
[INTERVIEWER: If R gives a partial hour (e.g. "15 minutes" or "an hour and a half"), please round up to the nearest whole number.**]**

____ **[RECORD EXACT NUMBER 0-168]**
998 **(DO NOT READ)** Don't know
999 **(DO NOT READ)** Refused

ASK IF QTV=1-168

INT_TV INTERVIEWER: DO NOT ASK; CODE RESPONDENT'S APPROACH TO ANSWERING QUESTION: IF R DID NOT SAY HOW, RECORD AS "DON'T KNOW"; RECORD AS MANY AS APPLY

1 Thought about specific days of the week and added them up (e.g., 2 hours Monday plus 3 hours Friday)
2 Thought about how many hours usually watch per day and used that as a point of reference (e.g., I usually watch 2 hours a day)
3 Thought about types of shows (e.g., news, movies) and added them up
4 Other **(SPECIFY)**
98 **(DO NOT READ)** Don't know

[RANDOMIZE Q61A1/Q61B1 WITH Q62A1/Q62B1; RANDOM HALF WILL GET Q61A1/Q61B1 FIRST (EXPERIMENT17=1) AND RANDOM HALF WILL GET Q62A1/Q62B1 FIRST (EXPERIMENT17=2)]

ASK RANDOM HALF SAMPLE (EXPERIMENT14=1)

Q61A1 In a TYPICAL week, how many hours do you spend using a computer?

[INTERVIEWER NOTE: Accept responses in hours or in minutes; in Q61A2, note whether response was reported in hours or minutes.**]**

____ **[RECORD EXACT NUMBER 0-9997]**
9998 **(DO NOT READ)** Don't know
9999 **(DO NOT READ)** Refused

ASK IF SPECIFIED COMPUTER TIME (Q61A1=0-9997)

Q61A2 **INTERVIEWER: If R already stated that time spent on computer was in hours or minutes, do not ask and enter 1 or 2. Otherwise, ASK:** Would you say that time is in hours or minutes?

[PROGRAMMER: Auto-punch Q61A2=3 (not applicable) when Q61A1=0]

1 Hours
2 Minutes
3 **(DO NOT READ)** Not applicable **[PROGRAMMER: Punch 3 only for Q61A1=0]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT14=2)

Q61B1 In a TYPICAL week, how many hours do you spend using a computer? Please do not include any time spent writing or reading emails.

[INTERVIEWER NOTE: Accept responses in hours or in minutes; in Q61B2, note whether response was reported in hours or minutes.**]**

____ **[RECORD EXACT NUMBER 0-9997]**
9998 **(DO NOT READ)** Don't know
9999 **(DO NOT READ)** Refused

ASK IF SPECIFIED COMPUTER TIME (Q61B1=0-9997)

Q61B2 **INTERVIEWER: If R already stated that time spent on computer was in hours or minutes, do not ask and enter 1 or 2. Otherwise, ASK:** Would you say that time is in hours or minutes?

[PROGRAMMER: Auto-punch Q61A2=3 (not applicable) when Q61B1=0]

1 Hours
2 Minutes
3 **(DO NOT READ)** Not applicable **[PROGRAMMER: Punch 3 only for Q61B1=0]**

98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT15=1)

Q62A1 In a TYPICAL week, how many hours do you spend talking on the telephone?
[INTERVIEWER NOTE: Accept responses in hours or in minutes; in Q62A2, note whether response was reported in hours or minutes.**]**

____ **[RECORD EXACT NUMBER 0-9997]**
9998 **(DO NOT READ)** Don't know
9999 **(DO NOT READ)** Refused

ASK IF SPECIFIED TELEPHONE TIME (Q62A1=0-9997)

Q62A2 **INTERVIEWER: If R already stated that time spent talking on the telephone was in hours or minutes, do not ask and enter 1 or 2. Otherwise, ASK:** Would you say that time is in hours or minutes?
[PROGRAMMER: Auto-punch Q62A2=3 (not applicable) when Q62A1=0]

1 Hours
2 Minutes
3 **(DO NOT READ)** Not applicable **[PROGRAMMER: Punch 3 only for Q62A1=0]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK RANDOM HALF SAMPLE (EXPERIMENT15=2)

Q62B1 In a TYPICAL week, how many hours do you spend talking on the telephone? Please do not include time spent speaking with family members.
[INTERVIEWER NOTE: Accept responses in hours or in minutes; in Q62B2, note whether response was reported in hours or minutes.**]**

____ **[RECORD EXACT NUMBER 0-9997]**
9998 **(DO NOT READ)** Don't know
9999 **(DO NOT READ)** Refused

ASK IF SPECIFIED TELEPHONE TIME (Q62B1=0-9997)

Q62B2 **INTERVIEWER: If R already stated that time spent talking on the telephone was in hours or minutes, do not ask and enter 1 or 2. Otherwise, ASK:** Would you say that time is in hours or minutes?
[PROGRAMMER: Auto-punch Q62B2=3 (not applicable) when Q62B1=0]

1 Hours
2 Minutes
3 **(DO NOT READ)** Not applicable **[PROGRAMMER: Punch 3 only for Q62B1=0]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

THERE IS NO Q63
ASK ALL

Q64 How many years have you been living in your current home?
[INTERVIEWER: IF R says "All my life"**, PROBE FOR NUMBER OF YEARS]**

___ years **[RECORD EXACT NUMBER 0-97]**
0 Less than 1 year
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

THERE IS NO Q65
ASK ALL

Q66 In what month and year were you born?

1 **[RECORD RESPONSE IN THIS FORMAT: MM/YYYY]**
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q67 Are you Spanish, Hispanic, or Latino?

1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q68 I am going to read you a list of five race categories. Please choose one or more races that you consider yourself to be: White; Black or African-American; American Indian or Alaska Native; Asian; OR Native Hawaiian or Other Pacific Islander.
[INTERVIEWER NOTE: DO NOT PROBE UNLESS RESPONSE IS HISPANIC OR A HISPANIC ORIGIN; ALLOW MULTIPLE RESPONSES
IF R SAYS HISPANIC OR LATINO, PROBE: Do you consider yourself a WHITE (Hispanic/Latino) or a BLACK (Hispanic/Latino)? **IF R DOES NOT SAY WHITE, BLACK OR ONE OF THE RACE CATEGORIES LISTED, RECORD AS "OTHER" (CODE 6)]**

1 White
2 Black or African-American
3 American Indian or Alaska native
4 Asian
5 Native Hawaiian or other pacific islander
6 Other
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK ALL

Q69 What is the highest level of school you have completed or the highest degree you have received? **[DO NOT READ BUT CAN PROBE FOR CLARITY IF NECESSARY]**

1 Less than high school
2 High school graduate, High school diploma or the equivalent (for example: GED)
3 Some college but no degree
4 Associate degree
5 Bachelor's degree (for example: B.A., A.B., B.S.)
6 Graduate degree [master's degree, professional school degree, or doctorate degree]
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL

Q70 What state do you currently live in?

1 Alabama
2 Alaska
3 Arizona
4 Arkansas
5 California
6 Colorado
7 Connecticut
8 Delaware
9 District of Columbia
10 Florida
11 Georgia
12 Hawaii
13 Idaho
14 Illinois
15 Indiana
16 Iowa
17 Kansas
18 Kentucky
19 Louisiana
20 Maine
21 Maryland
22 Massachusetts
23 Michigan
24 Minnesota
25 Mississippi
26 Missouri
27 Montana
28 Nebraska
29 Nevada
30 New Hampshire
31 New Jersey
32 New Mexico
33 New York
34 North Carolina
35 North Dakota
36 Ohio
37 Oklahoma
38 Oregon
39 Pennsylvania
40 Rhode Island
41 South Carolina
42 South Dakota
43 Tennessee
44 Texas
45 Utah
46 Vermont
47 Virginia
48 Washington State
49 West Virginia
50 Wisconsin
51 Wyoming
98 **(DO NOT READ)** Don’t know
99 **(DO NOT READ)** Refused

ASK ALL (DO NOT ASK IF PRETEST)

Q71 A letter describing this study may have been sent to your home recently. Do you remember seeing the letter?

1 Yes
2 No
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

ASK IF SAW LETTER (Q71=1)

Q72 Do you happen to remember if there was anything else in the envelope with the letter? **[IF YES AND SAID ANY AMOUNT OF MONEY WITHOUT PROMPTING, ENTER CODE=1; IF YES AND DID NOT SPECIFY, PROBE:** Could you please tell me what was included with the letter?**]**

1 Yes, money
2 Yes, something other than money **(SPECIFY)**
3 No, nothing was included with the letter
98 **(DO NOT READ)** Don't know
99 **(DO NOT READ)** Refused

THANK AND END INTERVIEW:] These are all of the questions we have for you. Thank you very much for your time. Good-bye.

<u>POST-INTERVIEW OBSERVATIONS TO BE ANSWERED BY INTERVIEWER</u>

Q73 **INTERVIEWER, PLEASE ANSWER:** The respondent answered the survey questions to the best of his or her ability.

1 Not at all
2 Not that often
3 Somewhat often
4 Pretty often
5 Very often

Q74 **INTERVIEWER, PLEASE ANSWER:** The respondent was reluctant to answer the survey questions.

1 Not at all
2 Not that often
3 Somewhat often
4 Pretty often
5 Very often

Q75 **INTERVIEWER, PLEASE ANSWER:** The respondent had trouble understanding the survey questions.

1 Not at all
2 Not that often
3 Somewhat often
4 Pretty often
5 Very often

6.3.2 Interviewer FAQ

Health Care Consent Request (CR1A)

Interviewer Note: Only read the following if needed:

IF RESPONDENT WANTS CLARIFICATION

An important research issue is how the use of health care may change as people age. Granting access to these records will provide researchers with the missing data they need to understand this relationship without taking up more of your time. All of your information will be kept strictly confidential and used for statistical purposes only. Researchers at the University of Maryland understand people's concerns about releasing such information. Everyone involved in the project guarantees their commitment to protecting the data and confidentiality of respondents. Please be assured that we take these issues as seriously as you do.

IF RESPONDENT ASKS IF THEIR BENEFITS WILL BE AFFECTED

Your benefits will not be affected in any way by your decision.

IF RESPONDENT ASKS HOW THEIR RECORDS WILL BE ACCESSED AND LINKED

With your consent, your records will be accessed using information such as your name, address, sex and date of birth. This information will be kept strictly confidential and used for statistical purposes only.

IF RESPONDENT ASKS ABOUT WHAT RECORDS WILL BE ACCESSED

With your consent, the University of Maryland may access information about your vital statistics and health care costs and diagnoses from your health-related records. This information will be kept strictly confidential and used for statistical purposes only.

Income/Employment Consent Request (CR1B)

Interviewer Note: Only read the following if needed:

IF RESPONDENT WANTS CLARIFICATION

An important research issue is how income may change as people age. Granting access to these records will provide researchers with the missing data they need to understand this relationship without taking up more of your time. All of your information will be kept strictly confidential and used for statistical purposes only. Researchers at the University of Maryland understand people's concerns about releasing such information. Everyone involved in the project guarantees their commitment to protecting the data and confidentiality of respondents. Please be assured we that take these issues as seriously as you do.

IF RESPONDENT ASKS IF THEIR BENEFITS WILL BE AFFECTED

Your benefits will not be affected in any way by your decision.

IF RESPONDENT ASKS HOW THEIR RECORDS WILL BE ACCESSED AND LINKED

With your consent, your records will be accessed using information such as your name, address, sex and date of birth. This information will be kept strictly confidential and used for statistical purposes only.

IF RESPONDENT ASKS ABOUT WHAT RECORDS WILL BE ACCESSED

With your consent, the University of Maryland may access information about your income and employment from your income and employment-related records. This information will be kept strictly confidential and used for statistical purposes only.

6.3.3 Advance Letter

Dear [INSERT NAME],

Researchers at The University of Maryland are conducting an important nationwide study about Americans' health and their views on various social issues. We are asking a scientific random sample of individuals to take part in a short interview. A few days from now you will receive a phone call from Princeton Data Source. If the call comes at an inconvenient time, the interviewer will be happy to set an appointment to call back at a better time.

Your help is voluntary but very important. The answers you give will be confidential, and we will take all possible steps to protect your privacy. Your answers will be used for research only.

[INCENTIVE: We have included a token of our appreciation for your participation. / CONTROL: We thank you in advance for your participation.] Your assistance is crucial to the success of this research.

Sincerely,

Stanley Presser

Dr. Stanley Presser
University of Maryland

6.3.4 Unweighted Tables and Figures

(Table and figure numbers correspond with weighted tables and figures in Chapter 3)

Table 3.05

Percent of All Respondents Who Consent to Linkage and Percent to Each Request, by Demographic Category

		Consent Request	
	All Respondents (n=900)	Health (n=445)	Income/Employment
	%	%	%
Male *(605)*	**27.8**	31.8	23.9
Female *(295)*	**34.2**	37.0	31.5
18-44 *(101)*	*28.7*	*27.5*	**30.0**
45-54 *(139)*	*18.0*	*17.2*	**18.7**
55-64 *(216)*	*27.8*	*30.3*	**25.2**
65+ *(444)*	*34.9*	*41.2*	**28.7**
<HS Grad/HS Grad	**35.3**	*41.5*	28.9
Some Col/Assoc Deg	**30.1**	*33.6*	27.1
Bachelors Deg (194)	**28.9**	*33.0*	24.2
Graduate Deg (148)	**20.3**	*17.1*	23.1
White (764)	26.5	39.3	**16.0**
Non-White (136)	30.5	32.6	**28.4**
<25K (165)	*43.0*	*48.2*	*37.5*
25- <75K (341)	*34.0*	*39.0*	*29.4*
>75K (237)	*24.5*	*28.5*	*20.7*
Income DK (49)	*30.6*	*32.1*	*28.6*
Income REF (108)	*8.3*	*3.9*	*12.5*
No Incentive (334)	**25.8**	*25.7*	25.8
Incentive (566)	**32.3**	*38.3*	26.7

Notes: Unweighted estimates. *Italics* indicates differences by demographics $\chi 2$ p<0.01; **bold** indicate differences by demographics $\chi 2$ p<0.05; underline indicates differences by demographics $\chi 2$ p<0.10.

Table 3.06

Consent Request Variation Baseline Logistic Regression Model Predicting Consent to Record Linkage

		Coefficient	SE	p-value
Constant		-0.77	0.36	0.031
Consent Request	Income/Employment (ref.)			
	Health	0.34	0.15	0.028
Gender	Female (ref.)			
	Male	-0.22	0.17	0.198
Age	18-44 (ref.)			
	45-54	-0.71	0.32	0.027
	55-64	-0.01	0.28	0.970
	65+	0.25	0.26	0.326
Education	<HS Grad/HS Grad (ref.)			
	Some Col/Assoc Deg	-0.11	0.19	0.568
	Bachelors Deg	-0.12	0.22	0.578
	Graduate Deg	-0.59	0.26	0.023
Race	Non-White (ref.)			
	White	0.28	0.22	0.213
Income	<25K (ref.)			
	25- <75K	-0.25	0.21	0.235
	>75K	-0.55	0.25	0.030
	Income DK	-0.63	0.36	0.076
	Income REF	-2.08	0.39	0.000
Incentive	No Incentive (ref.)			
	Incentive	0.32	0.16	0.045

Notes: Unweighted estimates.

Table 3.07

Wald Tests for Categorical Parameters in Baseline Logistic Regression Model

Categorical Predictor	F-Test Statistic	p-value
Age	$F_{(3, 47)} = 14.82$	0.002
Education	$F_{(3, 47)} = 5.36$	0.147
Income	$F_{(4, 46)} = 31.35$	0.000

Table 3.08

Consent Request Variation Baseline Logistic Regression Model Predicting Consent to Record Linkage, by Request Type

		Health (n=445)			Income and Employment (n=455)		
		Coefficient	SE	p-value	Coefficient	SE	p-value
Constant		-0.68	0.61	0.271	-1.39	0.64	0.031
Gender	Female (ref.)						
	Male	0.13	0.25	0.590	0.28	0.24	0.238
Age	18-44 (ref.)						
	45-54	-0.75	0.48	0.116	-0.71	0.44	0.11
	55-64	0.15	0.40	0.709	-0.21	0.39	0.591
	65+	0.68	0.38	0.073	-0.17	0.36	0.636
Education	<HS Grad/HS Grad (ref.)						
	Some Col/Assoc Deg	-0.20	0.28	0.477	-0.03	0.28	0.926
	Bachelors Deg	-0.20	0.31	0.519	-0.10	0.33	0.768
	Graduate Deg	-1.24	0.39	0.002	-0.08	0.36	0.816
Race	Non-White (ref.)						
	White	-0.26	0.32	0.419	0.87	0.35	0.013
Income	<25K (ref.)						
	25- <75K	-0.11	0.30	0.720	-0.36	0.30	0.235
	>75K	-0.31	0.36	0.394	-0.74	0.37	0.042
	Income DK	-0.79	0.49	0.105	-0.36	0.55	0.513
	Income REF	-3.01	0.76	0.000	-1.47	0.48	0.002
Incentive	No Incentive (ref.)						
	Incentive	0.57	0.23	0.014	0.02	0.23	0.941

Notes: Unweighted estimates.

Table 3.09

Wald Tests for Categorical Parameters in Baseline Logistic Regression Model Predicting Consent to Record Linkage, by Request Type

	Health		Income and Employment	
Categorical Predictor	F-Test Statistic	p-value	F-Test Statistic	p-value
Age	$F_{(3, 47)} = 15.95$	0.001	$F_{(3, 47)} = 3.15$	0.370
Education	$F_{(3, 47)} = 10.38$	0.016	$F_{(3, 47)} = 0.11$	0.991
Income	$F_{(4, 46)} = 17.90$	0.001	$F_{(4, 46)} = 10.92$	0.038

Table 3.11

Privacy, Confidentiality, and Trust Attitudes and Benchmark Comparisons

	All Respondents (n=900)	External Benchmarks
	%	%
Privacy		
Census invasion of privacy (*Agree*)	18.4	31
Gov bothers with requests (*Yes*)	24.6	--[100]
Worry about ID theft		
Frequently	27.4	31
Occasionally	38.4	35
Rarely	23.6	18
Never	10.6	15
Confidentiality		
Medical records not confidential		
Bothered a lot	69.2	53.0
Bothered some	13.5	19.2
Bothered a little	6.9	9.1
Not bothered at all	10.4	18.7
Tax records not confidential		
Bothered a lot	68.4	53.0
Bothered some	13.4	19.2
Bothered a little	7.8	9.1
Not bothered at all	10.4	18.7
Control over personal information		
Strongly agree	41.4	41.4
Somewhat agree	41.3	36.2
Somewhat disagree	12.3	15.2
Strongly disagree	5.0	7.0
Trust		
Trusted/Careful		
Can be trusted	31.5	47.1
Can't be too careful	68.5	52.9[101]
Helpful/Look out for themselves		
Try to be helpful	62.4	51.2
Look out for themselves	37.6	48.8[102]
Take Advantage/Fair		
Take advantage	34.3	43.1[103]
Try to be fair	65.7	56.9

Notes: Unweighted estimates.

---

[100] The external benchmark for this item is not available.

[101] Data from the 2008 GSS is the most recent data available. Compared to the GSS, some Practicum response options were combined ("always trusted" and "usually trusted"; and "usually not trusted" and "always not trusted".)

[102] Data from the 2010 GSS is the most recent data available. In 2010, 10% said "it depends". This response distribution was estimated without this 10% for comparability.

[103] Data from the 2010 GSS is the most recent data available. In 2010, 8.6% said "it depends". This response distribution was estimated without this 8.6% for comparability.

Table 3.12

Privacy, Confidentiality, and Trust Attitudes by Consent Status

	Consent (n=269)	Non-Consent (n=631)
	%	%
Privacy		
Census invasion of privacy (*Agree*)		
Agree	28.4	71.6
Disagree	30.2	69.8
Gov bothers with requests (*Yes*)		
Yes	25.1	74.9
No	31.3	68.7
Worry about ID theft		
Frequently	*21.2*	*78.8*
Occasionally	*30.6*	*69.4*
Rarely	*32.2*	*67.8*
Never	*44.2*	*55.8*
Confidentiality		
Medical records not confidential		
Bothered a lot	*21.5*	*78.5*
Bothered some	*39.2*	*60.8*
Bothered a little	*49.2*	*50.8*
Not bothered at all	*60.9*	*39.1*
Tax records not confidential		
Bothered a lot	*23.8*	*76.2*
Bothered some	*36.1*	*63.9*
Bothered a little	*37.7*	*62.3*
Not bothered at all	*55.4*	*44.6*
Control over personal information		
Strongly agree	*23.3*	*76.7*
Somewhat agree	*34.9*	*65.1*
Somewhat disagree	*31.5*	*68.5*
Strongly disagree	*40.9*	*59.1*
Trust		
Trusted/Careful		
Can be trusted	30.3	69.7
Can't be too careful	29.9	70.1
Helpful/Look out for themselves		
Try to be helpful	31.6	68.4
Look out for themselves	28.7	71.3
Take Advantage/Fair		
Take advantage	33.5	66.6
Try to be fair	28.0	72.0

Note: Unweighted estimates. *Italics* indicate differences by response option $\chi 2$ p<0.01; underline indicates differences by response option $\chi 2$ p<0.10

Table 3.13

Privacy, Confidentiality, and Trust Attitudes by Consent Status and Consent Request Condition

	Health		Income/Employment	
	Consent (n=149)	Non-Consent (n=296)	Consent (n=120)	Non-Consent (n=335)
	%	%	%	%
Privacy				
Census invasion of privacy				
Agree	31.4	68.6	25.0	75.0
Disagree	34.1	65.9	26.5	73.5
Gov bothers with requests				
Yes	30.8	69.2	19.6	80.4
No	34.4	65.6	71.7	28.3
Worry about ID theft				
Frequently	**26.3**	**73.7**	*16.5*	*83.5*
Occasionally	**35.1**	**64.9**	*26.0*	*74.0*
Rarely	**30.9**	**69.2**	*33.3*	*66.7*
Never	**47.3**	**52.7**	*40.0*	*60.0*
Confidentiality				
Medical records not confidential				
Bothered A lot	*23.0*	*77.0*	*20.0*	*80.0*
Bothered some	*42.4*	*57.6*	*36.1*	*63.9*
Bothered a little	*62.5*	*37.5*	*34.5*	*65.5*
Not bothered at all	*75.0*	*25.0*	*47.9*	*52.1*
Tax records not confidential				
Bothered a lot	*28.4*	*71.6*	*19.4*	*80.6*
Bothered some	*32.2*	*67.8*	*40.0*	*60.0*
Bothered a little	*40.5*	*59.5*	*34.4*	*65.6*
Not bothered at all	*60.0*	*40.0*	*50.0*	*50.0*
Control over personal information				
Strongly agree	29.4	70.7	*17.1*	*82.9*
Somewhat agree	35.7	64.3	*34.1*	*65.9*
Somewhat disagree	33.9	66.1	*28.9*	*71.2*
Strongly disagree	46.7	53.3	*37.9*	*62.1*
Trust				
Trusted/Careful				
Can be trusted	29.9	70.1	30.7	69.3
Can't be too careful	34.8	65.25	24.92	75.1
Helpful/Look out for themselves				
Try to be helpful	36.8	63.2	26.5	73.6
Look out for themselves	30.3	69.8	27.1	72.9
Take advantage/Fair				
Take advantage	35.1	64.9	31.8	68.2
Try to be fair	32.9	67.2	23.5	76.6

Notes: Unweighted estimates. *Italics* indicates differences by response option $\chi 2$ p<0.01; **bold** indicate differences by response option $\chi 2$ p<0.05; underline indicates differences by response option $\chi 2$ p<0.10

Table 3.14

Distribution of Privacy Index Scores

Index Score	%
0 (*Low Concern*)	6.6
0.3	18.8
0.7	27.6
1.0	19.0
1.3	3.5
1.7	6.5
2.0	7.2
2.3	1.7
2.7	4.5
3 (*High Concern*)	4.7
Mean	1.0
Cronbach's Alpha	0.470

Notes: Unweighted estimates. Cronbach's Alpha was calculated using unweighted correlations for each item pair. Scores were calculated on 867 respondents that provided substantive responses to all three privacy items.

Table 3.15

Item Correlations among Privacy Items

	Census correlation p-value	Gov Requests	ID Theft
Census	1.000		
Gov Requests	0.459	1.000	
	0.000		
ID Theft	0.042	0.136	1.000
	0.212	0.000	

Notes: Unweighted estimates. Correlations were calculated on 867 respondents that provided substantive responses to all three privacy items.

Table 3.16

Distribution of Confidentiality Index Scores

Index Score	%
0 (*Low Concern*)	0.2
0.3	1.4
0.7	2.6
1.0	4.1
1.3	4.8
1.7	7.2
2.0	12.2
2.3	14.4
2.7	25.1
3 (*High Concern*)	28.5
Mean	2.3
Cronbach's Alpha	0.509

Notes: Unweighted estimates. Cronbach's Alpha was calculated using unweighted correlations for each item pair. Scores were calculated on 862 respondents that provided substantive responses to all confidentiality items.

Table 3.17

Item Correlations among Confidentiality Items

	Medical correlation p-value	Tax	Info Control
Medical	1.000		
Tax	0.543 0.000	1.000	
Info Control	0.111 0.001	0.072 0.034	1.000

Notes: Unweighted estimates. Correlations were calculated on 862 respondents that provided substantive responses to all three confidentiality items.

Table 3.18

Distribution of Trust Index Scores

Index Score	%
0 *(Low Concern)*	26.3
1	30.4
2	19.7
3 *(High Concern)*	23.7
Mean	1.41
Cronbach's Alpha	0.674

Notes: Unweighted estimates. Cronbach's Alpha was calculated using unweighted correlations for each item pair. Scores were calculated on 833 respondents that provided substantive responses to all three trust items.

Table 3.19

Item Correlations among Trust Items

	Trust correlation p-value	Help	Fair
Trust	1.000		
Help	0.344 0.000	1.000	
Fair	0.339 0.000	0.534 0.000	1.000

Notes: Unweighted estimates. Correlations were calculated on 833 respondents that provided substantive responses to all three trust items.

Table 3.20

Logistic Regression Models Predicting Consent to Record Linkage: Original Privacy, Confidentiality, and Trust Scales

		Privacy (n=427)	Confidentiality (n=426)	Trust (n=408)
Constant		-0.18	1.55***	-0.54
Gender	Female (ref.)			
	Male	-0.14	-0.28	-0.21
Age	18-44 (ref.)			
	45-54	-0.72**	-0.71	-0.72**
	55-64	-0.03	0.04	-0.08
	65+	0.20	0.12	0.31
Education	HS or Less (ref.)			
	Associates/Some College	-0.16	-0.16	-0.09
	Bachelors Degree	-0.16	0.03	-0.14
	Graduate Degree	-0.56**	-0.45	-0.47*
Race	Non-White (ref.)			
	White	0.28	0.39	0.28
Income	<25K (ref.)			
	25- <75K	-0.33	-0.25	-0.30
	>75K	-0.70***	-0.76***	-0.64**
	Income DK	-0.88**	-0.74*	-0.55
	Income REF	-2.12***	-2.27***	-2.02***
Incentive	No Incentive (ref.)			
	Incentive	0.26	0.39**	0.30*
Privacy		-0.30***		
Confidentiality			-0.95***	
Trust				-0.02

Notes: Unweighted estimates.

Table 3.21

Logistic Regression Models Predicting Consent to the Health Request: Original Privacy, Confidentiality, and Trust Scales

		Privacy (n=427)	Confidentiality (n=426)	Trust (n=408)
Constant		-0.06	1.84***	-0.37
Gender	Female (ref.)			
	Male	-0.08	-0.34	-0.70
Age	18-44 (ref.)			
	45-54	-0.72	-0.84*	-0.87
	55-64	0.16	0.31	0.62
	65+	0.64*	0.57	0.41
Education	HS or Less (ref.)			
	Associates/Some College	-0.27	-0.16	-1.23**
	Bachelors Degree	-0.16	0.12	-0.30
	Graduate Degree	-1.21***	-1.17***	-2.35**
Race	Non-White (ref.)			
	White	-0.31	-0.14	-0.73
Income	<25K (ref.)			
	25- <75K	-0.14	-0.13	-0.16
	>75K	-0.42	-0.59	0.11
	Income DK	-0.93*	-0.90	0.18
	Income REF	-2.74***	-2.70***	-3.60*
Incentive	No Incentive (ref.)			
	Incentive	0.47**	0.74***	0.53
Privacy		-0.21		
Confidentiality			-1.02***	
Trust				0.32

Notes: Unweighted estimates.

Table 3.22

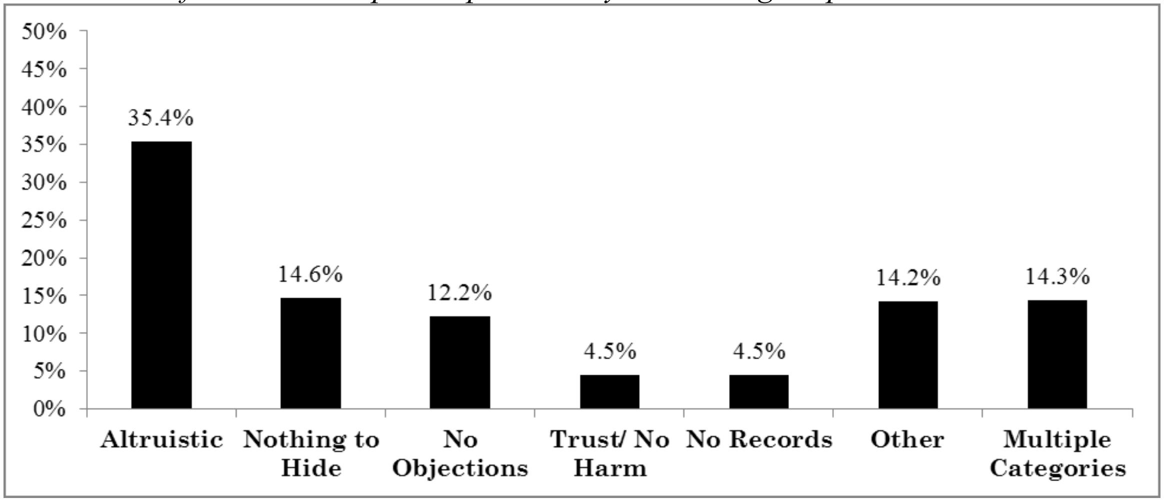
Logistic Regression Model Incorporating Indicators of Health Request Salience, and Predicting Consent to the Health Consent Request

		Privacy (n=440)	Confidentiality (n=436)	Trust (n=425)
Constant		-0.40	1.33	-0.84
Gender	Female (ref.)			
	Male	-0.14	-0.24	-0.32
Age	18-44 (ref.)			
	45-54	-0.72	-0.66	-0.57
	55-64	-0.27	-0.29	-0.21
	65+	-0.22	-0.35	-0.02
Education	HS or Less (ref.)			
	Associates/Some College	0.02	-0.13	0.04
	Bachelors Degree	-0.24	-0.10	-0.02
	Graduate Degree	-0.01	0.12	0.04
Race	Non-White (ref.)			
	White	0.94**	0.93**	0.89**
Income	<25K (ref.)			
	25- <75K	-0.46	-0.31	-0.47
	>75K	-0.98**	-0.94**	-0.84**
	Income DK	-1.02	-0.56	-0.62
	Income REF	-1.71**	-1.87***	-1.64***
Incentive	No Incentive (ref.)			
	Incentive	-0.01	0.06	-0.06
Privacy		-0.42**		
Confidentiality			-0.93***	
Trust				0.03

Notes: Unweighted estimates.

Figure 3.02

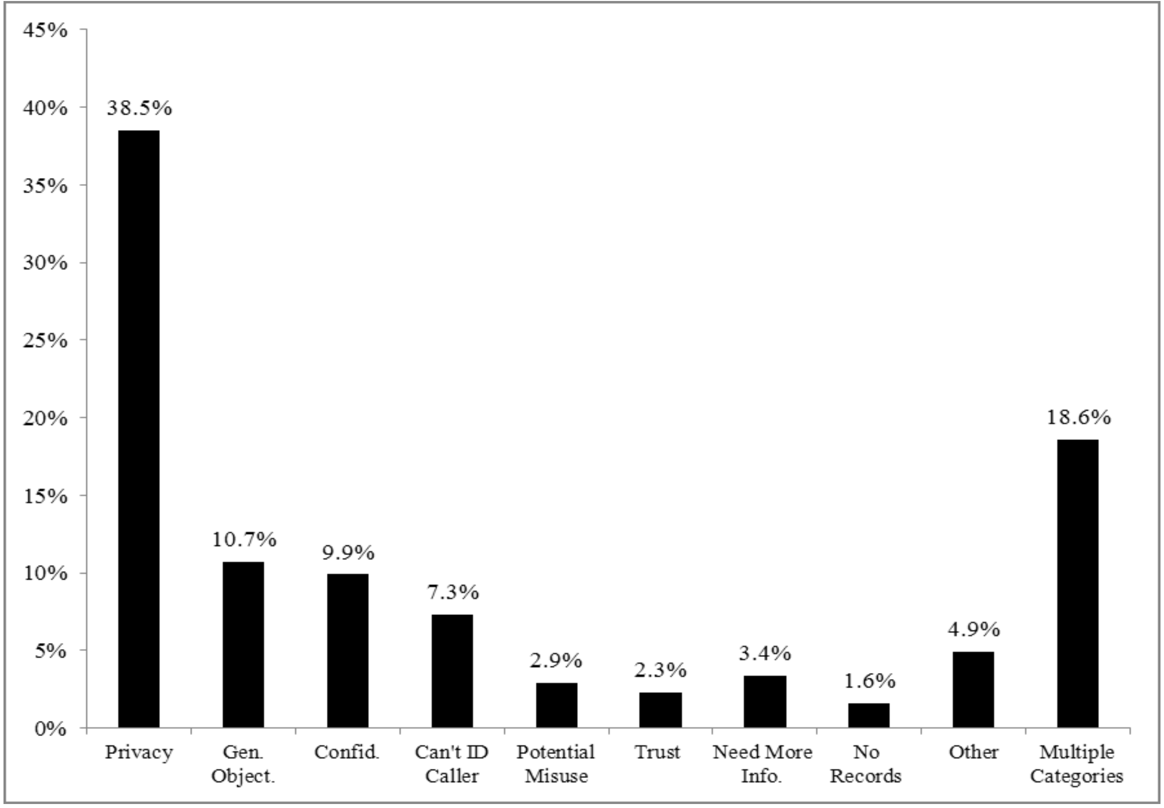
Distribution of Verbatim Responses provided by Consenting Respondents



Notes: Unweighted estimates. Analysis based on 252 cases, other respondents did not provide substantive response.

Figure 3.03

Distribution of Verbatim Responses provided by Nonconsenting Respondents



Notes: Unweighted estimates. Analysis based on 618 cases, other respondents did not provide substantive response.

Table 3.26

Correlates of Health-Related Records for all Respondents and External Benchmarks -

	All Respondents (n=900)	External Benchmarks
	%	%
Mean Self-Rated Health (*1=Poor; 5=Excellent*)[104]	3.2	2.6
Chronic Conditions		
Diabetes	16.0	9.9[105]
Hypertension	47.9	32.6
Asthma	10.5	13.6
Arthritis	34.6	23.9
Heart disease	17.0	3.1
Anemia	10.6	--
1+ Chronic Conditions	69.7	--
Overnight Hospital Patient (*Yes*)	16.0	11.1
If Yes, # of Times	1.6	1.5
# of MD Visits 2010		
0	7.8	15.9
1-3	45.2	46.5
4-9	30.9	23.8
10+	16.2	13.2
Health Insurance (*Yes*)	94.7	--[106]
Healthcare Expenditures		
$0	10.4	12.4
<$500	41.7	34.9
$500 to <$2,000	33.0	31.4
>$2,000	14.9	21.3

Notes: Unweighted estimates.

---

[104] Mean self-rated health is reported here because only the mean self-rated health benchmark estimate was available. In the Practicum survey, 5.4% of respondents reported their health as poor, 19.6% as fair, 41.2% as good, 22.1% as very good, and 11.5% as excellent.

[105] In the NHANES, 8.8% of respondents report having Diabetes; an additional 1.8% report having borderline Diabetes.

[106] NCHS does not make estimates from this NHIS item publicly-available.

Table 3.27

Correlates of Health-Related Records for All Respondents, by Consent Status

	Consent (n=269)	Non-Consent (n=631)
	%	%
Mean Self-Rated Health		
Poor	41.7	58.3
Fair	32.2	67.8
Good	30.6	69.4
Very Good	22.5	77.6
Excellent	29.4	70.6
Chronic Conditions		
Diabetes		
Yes	**38.2**	**61.8**
No	**28.3**	**71.7**
Hypertension		
Yes	*34.7*	*65.4*
No	*25.4*	*74.6*
Asthma		
Yes	30.9	69.2
No	29.6	70.4
Arthritis		
Yes	28.3	71.7
No	30.6	69.3
Heart disease		
Yes	*41.2*	*58.8*
No	27.5	72.5
Anemia		
Yes	35.8	64.2
No	29.2	70.8
1+ Chronic Conditions		
Yes	31.9	68.1
No	25.5	74.5
Overnight Hospital Patient		
No	28.7	71.3
Yes	34.3	65.7
If Yes, # of Times	1.9	1.5
# of MD Visits 2010		
0	28.6	71.4
1-3	30.4	69.6
4-9	27.8	72.2
10+	31.7	68.3
Health Insurance		
Yes	29.2	70.8

No	41.7	58.3
Healthcare Expenditures		
$0	31.5	68.5
<$500	30.4	69.7
$500 to <$2,000	27.1	73.0
>$2,000	32.6	67.4

Notes: Unweighted estimates. *Italics* indicates differences by response option $\chi 2$ p<0.01; **bold** indicate differences by response option $\chi 2$ p<0.05; underline indicates differences by response option $\chi 2$ p<0.10.

Table 3.28

Correlates of Health-Related Records for Respondents Assigned to Health Consent Request, by Consent Status

	Consent (n=269)	Non-Consent (n=631)
	%	%
Mean Self-Rated Health		
Poor	55.6	44.4
Fair	46.5	53.5
Good	38.5	61.5
Very Good	26.7	73.3
Excellent	33.3	66.7
Chronic Conditions		
Diabetes		
Yes	**43.4**	**56.6**
No	**31.3**	**68.7**
Hypertension		
Yes	*41.2*	*58.8*
No	*25.9*	*74.1*
Asthma		
Yes	30.6	69.4
No	33.6	66.4
Arthritis		
Yes	32.9	67.1
No	33.9	66.1
Heart disease		
Yes	**45.0**	**55.0**
No	**30.8**	**69.2**
Anemia		
Yes	**46.8**	**53.2**
No	**31.6**	**68.4**
1+ Chronic Conditions		
Yes	*38.0*	*62.0*
No	*23.5*	*76.5*

Overnight Hospital Patient		
No	31.9	68.1
Yes	40.0	60.0
If Yes, # of Times	1.7	1.4
# of MD Visits 2010		
0	23.5	76.47
1-3	34.3	65.7
4-9	31.6	68.4
10+	37.3	62.7
Health Insurance		
Yes	33.0	67.0
No	42.9	57.1
Healthcare Expenditures		
$0	37.2	62.8
<$500	32.4	67.6
$500 to <$2,000	30.3	69.7
>$2,000	40.3	59.7

Notes: Unweighted estimates. *Italics* indicates differences by response option $\chi 2$ p<0.01; **bold** indicate differences by response option $\chi 2$ p<0.05.

Table 3.29

Correlates of Income and Employment-Related Records for all Respondents and External Benchmarks

	All Respondents (n=900)	External Benchmarks
	%	%
Employment Status		
Working	41.3	59.6
Unemployed	2.45	7.3
Retired	43.9	13.5
Other	12.4	19.5
Mean Hours Worked/Week	42.1	36.6
Receives Benefits from:		
Social Security	54.0	28.4
Other Retirement/Pensions	41.4	17.5
Public assistance/Welfare/SSI	5.2	2.9
Receives 1+ Benefits	37.9	--

Note: Unweighted estimates.

Table 3.30

Correlates of Income and Employment-Related Records for All Respondents, by Consent Status

	Consent (n=269)	Non-Consent (n=631)
	%	%
Employment Status		
Working	*22.9*	*77.1*
Unemployed	*40.9*	*59.1*
Retired	*33.0*	*67.0*
Other	*40.5*	*59.5*
Mean Hours Worked/Week	40.8	42.4
Receives Benefits from:		
Social Security		
Yes	*36.0*	*64.0*
No	*23.4*	*76.6*
Other Retirement/Pensions		
Yes	30.7	69.3
No	29.8	70.2
Public assistance/Welfare/SSI		
Yes	*47.8*	*52.2*
No	*29.0*	*71.0*
Receives 1+ Benefits		
Yes	*34.8*	*65.2*
No	*22.7*	*77.3*

Notes: Unweighted estimates. *Italics* indicates differences by response option $\chi2$ p<0.01.

Table 3.31

Correlates of Income and Employment-Related Records for Respondents Assigned to Income and Employment Consent Request, by Consent Status

	Consent (n=120)	Non-Consent (n=335)
	%	%
Employment Status		
Working	22.3	77.7
Unemployed	33.3	66.7
Retired	26.2	73.9
Other	39.3	60.7
Mean Hours Worked/Week	40.5	42.1
Receives Benefits from:		
Social Security		
Yes	**30.7**	**69.3**
No	**22.1**	**77.9**
Other Retirement/Pensions		
Yes	27.9	72.1
No	25.6	74.4
Public assistance/Welfare/SSI		
Yes	40.0	60.0
No	25.7	74.3
Receives 1+ Benefits		
Yes	**20.8**	**79.2**
No	**30.2**	**70.8**

Note: Unweighted estimates. Italics indicates differences by response option $\chi 2$ $p<0.01$; **bold** indicate differences by response option $\chi 2$ $p<0.05$.

Table 3.32

Logistic Regression Model Incorporating Indicators of Health Request Salience, Among Respondents Assigned to the Health Consent Request

		Coefficient	SE	p-value
Constant		-0.94	1.215	0.437
Gender	Female (ref.)			
	Male	0.08	0.263	0.775
Age	18-44 (ref.)			
	45-54	-0.97	0.504	0.054
	55-64	-0.06	0.433	0.891
	65+	0.47	0.426	0.269
Education	<HS Grad/HS Grad (ref.)			
	Some Col/Assoc Deg	-0.19	0.292	0.525
	Bachelors Deg	-0.16	0.326	0.628
	Graduate Deg	-1.21	0.413	0.003
Race	Non-White (ref.)			
	White	-0.26	0.338	0.442
Income	<25K (ref.)			
	25- <75K	0.02	0.318	0.948
	>75K	-0.08	0.392	0.829
	Income DK	-0.92	0.552	0.097
	Income REF	-3.45	1.047	0.001
Mean Self-Rated Health (*1=Poor; 5=Excellent*)		-0.08	0.136	0.555
1+ Chronic Conditions	No (ref.)			
	Yes	0.55	0.307	0.075
Overnight Hospital Patient	No (ref.)			
	Yes	-0.12	0.336	0.714
# of MD Visits 2010	0 (ref.)			
	1-3	0.28	0.486	0.559
	4-9	0.02	0.525	0.964
	10+	0.12	0.582	0.832
Health Insurance	No (ref.)			
	Yes	0.64	0.540	0.238
Healthcare Expenditures	$0 (ref.)			
	<$500	-0.39	0.415	0.346
	$500 to <$2,000	-0.60	0.434	0.169
	>$2,000	0.05	0.485	0.921
Incentive	No Incentive (ref.)			
	Incentive	0.57	0.245	0.021

Notes: Unweighted estimates. Model based on 430 cases.

Table 3.33

Logistic Regression Model Incorporating Indicators of Income and Employment Request, Among Respondents Assigned to the Income and Employment Consent Request

		Coefficient	SE	p-value
Constant		-1.38	0.70	0.049
Gender	Female (ref.)			
	Male	0.25	0.24	0.313
Age	18-44 (ref.)			
	45-54	-0.81	0.45	0.074
	55-64	-0.40	0.42	0.339
	65+	-0.58	0.46	0.213
Education	<HS Grad/HS Grad (ref.)			
	Some Col/Assoc Deg	-0.05	0.28	0.856
	Bachelors Deg	-0.10	0.33	0.756
	Graduate Deg	-0.12	0.37	0.748
Race	Non-White (ref.)			
	White	0.86	0.35	0.015
Income	<25K (ref.)			
	25- <75K	-0.35	0.31	0.263
	>75K	-0.67	0.38	0.08
	Income DK	-0.30	0.55	0.594
	Income REF	-1.44	0.51	0.005
Employment Status	Not Working (ref.)			
	Working	0.14	0.53	0.786
Mean Hours Worked/Week				
	<20 (ref.)			
	21-40	-0.09	0.60	0.882
	40+	-0.20	0.62	0.748
Benefits	Receives 0 Benefits (ref.)			
	Receives 1+ Benefits	0.48	0.37	0.189
Incentive	No Incentive (ref.)			
	Incentive	0.00	0.23	0.991

Notes: Unweighted estimates. Model based on 445 cases.

6.4 Appendix to Chapter 4

6.4.1 Additional Tables

Table 4.01

Sample Characteristics among All Respondents and Mothers, by Availability of Interviewer Demographic Characteristics

	All Respondents		Mothers	
	Int. Chars. Available 464 interviewers 15,997 interviews	Int. Chars. Missing 298 interviewers 8,812 interviews	Int. Chars. Available 462 interviewers 12,686 interviews	Int. Chars. Missing 293 interviewers 6,745 interviews
	%	%	%	%
Mother Characteristics				
Ethnicity				
Hispanic	*27.6*	*20.9*	29.2	22.1
Non-Hispanic	*72.4*	*79.1*	70.8	77.9
Race				
White	*76.0*	*78.6*	76.9	80.8
Black	*16.2*	*14.2*	15.8	12.9
Other	*7.8*	*7.2*	7.3	6.3
Education				
<12 years	*20.9*	*17.1*	*20.6*	*16.2*
12 years	*31.2*	*30.0*	*30.2*	*28.8*
Some College	*18.8*	*19.7*	*20.0*	*20.8*
College Grad	*29.1*	*33.2*	*29.2*	*34.2*
Marital Status				
Married	66.4	69.1	67.8	71.3
Never Married	26.0	22.2	24.8	20.3
Wid./Div./Sep.	7.6	8.7	7.4	8.4
Age				
<30	41.4	40.1	*40.2*	*37.7*
30+	58.6	59.9	*59.8*	*62.3*
2008 Family Income (mean)	$67,161	$72,041	$60,200	$63,178
Respondent Characteristics				
Language of Interview				
English	81.2	88.3	79.8	86.9
Non-English	18.8	11.7	20.2	13.1
Prior Survey Refusal				

One or More	8.4	8.6	7.8	7.4
None	91.6	91.4	92.2	92.6
Advance Letter				
Mailed	64.4	67.2	*64.4*	*32.8*
Not Mailed	35.6	32.8	*35.6*	*67.2*
Relationship to Child				
Mother	79.3	75.8	---	---
Non-Mother	20.7	24.2	---	---
Consent				
Yes	78.5	82.5	80.8	84.2
No	21.5	17.5	19.2	15.8

Notes: Bold indicates differences by interviewer data missingness $\chi2$ p<0.01; underline indicates differences by interviewer data missingness $\chi2$ p<0.05; *italics and underline* indicates differences by interviewer data missingness $\chi2$ p<0.10.

Table 4.02

Baseline Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents, By Prior Refusal Status (All Interviewers)

		No Prior Refusals (n=23,319)			One or More Prior Refusals (n=2,034)		
		Coef.	SE	*p*	Coef.	SE	*p*
Constant		2.05	0.29	0.000	-1.83	0.64	0.005
Mother Characteristics							
Ethnicity	Non-Hispanic (ref.)						
	Hispanic	-0.07	0.12	0.559	0.66	0.33	0.046
Race	White (ref.)						
	Black	-0.10	0.09	0.288	0.84	0.32	0.008
	Other	0.23	0.18	0.191	-0.41	0.64	0.522
Education	<12 yrs (ref.)						
	12 yrs	-0.20	0.13	0.141	-0.55	0.31	0.075
	>12 yrs, non col. grad	-0.38	0.15	0.013	-0.85	0.31	0.006
	Col. grad	-0.30	0.15	0.045	-1.00	0.33	0.002
Marital Status	Married (ref.)						
	Never Married	0.11	0.16	0.489	0.35	0.44	0.419
	Wid./Div./Sep.	0.05	0.14	0.736	0.30	0.40	0.454
Age	<30 (ref.)						
	30+	-0.15	0.08	0.058	0.18	0.24	0.449
2008 Family Income		0.00	0.00	0.276	0.00	0.00	0.723
Relationship to Child	Not Mother (ref.)						
	Mother	0.54	0.07	0.000	0.48	0.25	0.053
Respondent Characteristics							
Language of Int.	English (ref.)						
	Non-English	0.40	0.20	0.044	0.12	0.43	0.786
Prior Survey Refusal	None (ref.)						
	One or More						
Advance Letter	Not Mailed (ref.)						
	Mailed	0.14	0.07	0.057	0.61	0.23	0.007
Interviewer Characteristics							
Gender	Male (ref.)						

	Female	0.28	0.15	0.053	-0.14	0.32	0.665
Age	18-30 (ref.)						
	31-45	-0.08	0.19	0.663	0.46	0.42	0.270
	46+	0.06	0.14	0.681	0.16	0.35	0.661
Race/Ethnicity							
	Hispanic (ref.)						
	White	0.22	0.26	0.401	-0.95	0.62	0.126
	Black	0.12	0.20	0.558	-0.21	0.30	0.481
	Other	0.21	0.29	0.464	0.72	0.81	0.379
Experience							
	0-6 mos (ref.)	-0.70	0.15	0.000	-0.18	0.39	0.633
	6 mos - 1 year	-1.21	0.15	0.000	-0.45	0.33	0.172
	> 1year						
Missing Data	Not Missing (ref.)						
	Missing	0.51	0.23	0.028	0.41	0.37	0.268

Notes: Archer and Lemeshow Goodness of Fit test for survey data suggests no evidence of lack of fit for either model: No Prior Refusals [F(9,748) = 0.53; Prob > F = 0.851]; One or More Prior Refusals [F(9,341) = 0.73; Prob > F = 0.684].

Wald Tests for Categorical Parameters in Baseline Logistic Regression Models Predicting Consent as a Function of Sample and Interviewer Characteristics, among All Respondents and Mothers, By Prior Refusal Status (All Interviewers)

	All Respondents		Mothers	
Categorical Predictor	F-Test Statistic	*p*	F-Test Statistic	*p*
Race	$F_{(2, 755)} = 1.36$	0.257	$F_{(2, 348)} = 3.75$	0.024
Education	$F_{(3, 754)} = 2.30$	0.076	$F_{(3, 347)} = 3.44$	0.017
Marital Status	$F_{(2, 755)} = 0.32$	0.723	$F_{(2, 348)} = 0.35$	0.706
Int. Age	$F_{(2, 755)} = 0.26$	0.770	$F_{(2, 348)} = 0.66$	0.516
Int. Race/Ethnicity	$F_{(3, 754)} = 0.30$	0.824	$F_{(3, 347)} = 1.30$	0.274
Int. Experience	$F_{(2, 755)} = 34.47$	0.000	$F_{(2, 348)} = 1.45$	0.236

6.4.2 2009 NIS Consent Request Module

SECTION D[107]

D5. (If respondent has "shot cards" or records of the child's vaccinations available):

> **To get a complete picture of the vaccinations received by your (children/child), we would like to contact doctors or health clinics to obtain a copy of the vaccination records. These records contain only the immunizations and dates of the immunizations for your (children/child).**
>
> READ IF NECESSARY:
> **Information we collect from you and your health care provider will be used to monitor and report on childhood immunizations. Last year, over 21,000 providers participated in this study. Participation by you and your child's provider helps the CDC understand the potential for childhood diseases.**

D5. (If respondent does not have "shot cards" or records of the child's vaccinations available):

> **Thank you for the valuable information you've shared with us. We find that it's often difficult to remember specifics about vaccinations. We'd like to collect the dates and types of vaccinations your (children have/child has) received by contacting the doctors or health clinics who provided them.**
>
> READ IF NECESSARY:
> **Information we collect from families like yours is used to develop health care policies and to determine where funding is most needed for vaccination programs such as Vaccines for Children.**
> **- - Since 1994, the Vaccines for Children (VFC) program has helped families of children who may not otherwise have access to vaccines by providing free vaccines to doctors who serve them.**
> **- - Children who are uninsured, (Medicaid recipients, Native Americans, Alaska Natives), can receive the necessary CDC recommended immunizations as part of routine health care in their doctor's office if their doctor is part of this program. Also, some state or local health departments have special programs for other groups of children.**

---

[107] Retrieved from ftp://ftp.cdc.gov/pub/health_statistics/nchs/Dataset_Documentation/NIS/NISPUF09_HHQUEX.pdf on July 8, 2012.

Confidentiality is mandated by law and I can assure you that the data is reported only in summary form and neither you nor the child will be identified as a participant.
When you give us permission to contact your child's provider to collect specific dates and types of shots, we also take the opportunity to ask the provider a few questions about the medical practice or clinic.

D6_X. How many locations have provided vaccinations for your child named [NAME OF (FIRST) ELIGIBLE CHILD] whose birth date is [DATE OF BIRTH OF (FIRST) ELIGIBLE CHILD]?

ENTER NUMBER .... __ GO TO D6A_1_X
ZERO .... 0 GO TO D6AA_X
DON'T KNOW .... 77 GO TO D6AA_X
REFUSED .... 99 GO TO SECT_D_TERM

D6AA_X. How many locations have provided health care for your child? Please include the hospital or birthing center where [he/she] was born, and any other clinics or doctor's offices that have seen [him/her].

ENTER 0 IF CHILD HAS NEVER SEEN A DOCTOR OR OTHER HEALTH CARE PROVIDER.
ENTER 77 FOR DON'T KNOW AND 99 FOR REFUSED
ENTER NUMBER .... GO TO D6A_1_X
ZERO .... 0 INS_INTRO
DON'T KNOW .... 77 GO TO SECT_D_TERM
REFUSED .... 99 GO TO SECT_D_TERM

D6 A_1_X. Starting with the most recent, please tell me the contact information for each location. (Would you take a moment to find shot records, appointment cards, or other records you may have?)

Yes, continue on .... 1 GO TO PLU
No, can't find, continue .... 2 GO TO PLU
Refused .... 99 GO TO SECT_D_TERM

D8_X. [ASK IF D6_X GE 1] In order to help the doctor or clinic locate your child's vaccination records, we need to know the child's full name - first, middle and last name

IF RESPONDENT REFUSES WE CAN ACCEPT A FIRST INITIAL AND FULL LAST NAME.

Continue .... 1 GOT TO D8A_1

Refused .... 99 GO TO SECT_D_TERM

D9. Could I know…what is your full name – first, middle, and last?
IF RESPONDENT REFUSES WE CAN ACCEPT A FIRST INITIAL AND FULL LAST NAME.

Continue .... 1 GO TO D9A
Refused .... 99 SECT_D_TERM

D9D_X. I need to verify that I am speaking with someone who can authorize the release of immunization records for [NAME OF ELIGIBLE CHILD(REN)]. Are you that person?

YES .... 1 GO TO D6_C
NO .... 2 GO TO D9D1
REFUSED .... 99 GO TO SECT_D_TERM

D6_C. The vaccination records collected from the provider(s) will be kept in strict confidence.

D7_ID Capture Interviewer ID upon entering question D7

D7_X Do we have your permission to contact the provider(s) named in this interview, give the provider(s) basic information that identifies (Fill Var: name of first/second/...ninth child, from S3_5), and request that information relevant to (his/her) immunization history be sent to the Centers for Disease Control and Prevention or its contractors for study purposes only?

YES .... 1 GO TO DCG [OR D7G if registry flag=1]
NO (Only choose this when you have made all appropriate aversion attempts) .... 2 GO TO SECT_TERM_D

D7G_X. Sometimes to get a complete record of your child(ren)'s vaccinations it would be helpful to contact your local immunization registry. This registry has information on children's vaccinations. The information we collect will be about your child(ren)'s vaccinations only.

Do we have your permission to contact your local immunization registry, give them basic information that identifies your child(ren), and request that information relevant to your child(ren)'s immunization history be sent to the Centers for Disease Control and Prevention or its contractors for study purposes only?

YES 1
NO 2
DON'T KNOW 77

REFUSED 99
(SUGGESTED TEXT IF THE RESPONDENT HAS A QUESTION)
WHAT IS A REGISTRY?

Immunization registries are confidential, population-based, computerized information systems that attempt to collect vaccination data about all children in a geographic area.

WHY DO YOU NEED TO CONTACT A REGISTRY?

Vaccination information from doctors and clinics sometimes is not complete or available. So, in order to get the most complete information possible about children's vaccinations, we also need to contact local registries to collect vaccination information.

D7_R. We appreciate the information you have already provided, but without your consent, we cannot contact your health care provider. We are only requesting the dates and types of vaccinations your child(ren) has received and I can assure you that no further information will be provided to us. All information collected is kept confidential under federal law and the names of you and your child(ren) will be completely separated from the data released in study results. The doctor or health clinic will receive 2 forms, one that I have signed indicating your consent to collect immunization information, and one that looks similar to a shot record with only the names of the vaccines listed and blank spaces for the dates to be filled in.

Continue .... 1 GO TO D7_1
Respondent still refuses .... 2 GO TO SECT_D_TERM

7 BIBLIOGRAPHY

Abreu, D., Daniel, K., Iwig, B., & Hoge, S. (2009). Four-digits or no-digit Social Security Numbers: Impact on the National Agricultural Statistics Service record linkage maintenance processes. *Proceedings of the Federal Committee on Statistical Methodology Research Conference.*

Anderson, B.A., Silver, B., D., & Abramson, P.R. (1988). The Effects of the race of the interviewer on race-related attitudes of black respondents in SRC/CPS National Election Studies. *Public Opinion Quarterly,* 52(3), 289-324.

Angus, V.C., Entwistle, V.A., Emslie, M.J., Walker, K.A., & Andrew, J.E. (2003). The requirement for prior consent to participate on survey response rates. *Biomedical Health Services Research*, 3 (1), 21.

Archer, K.J., & Lemeshow, S. (2006). Goodness-of-fit test for a logistic regression model fitted using survey sample data. *The Stata Journal*, 6(1), 97–105

Armstrong, V., Barnett, J., Cooper, H., Monkman, M., Moran-Ellis, J. & Shepherd, R. (2008). *Public Perspectives on the Governance of Biomedical Research: A qualitative study in a deliberative context*. London, England: University of Surrey.

Bailar, B., Bailey, L., & Stevens, J. (1977). Measures of interviewer bias and variance. *Journal of Marketing Research,* 14, 337-343.

Baker, R., Shiels, C., Stevenson, K., Fraser, R. & Stone, M. (2000). What proportion of patients refuse consent to data collection from their records for research purposes? *British Journal of General Practice,* 50, 655–656.

Bates, N. (2005). Development and testing of informed consent questions to link survey data with administrative records. *Proceedings of the American Statistical Association, Minneapolis, Minnesota, 3786-3793.*

Battaglia, M.P., Huggins, V. Rodén, A. Wright, R.A., Piccinino, L., & Dennis, J. M. (2000). Evaluating the impact of a new CATI screener in the National Immunization Survey. *Proceedings of the American Statistical Association Section on Survey Research Methods*, *Alexandria, Virginia,* 703-708.

Bethlehem, J.G. (2002). Weighting nonresponse adjustments based on auxiliary information. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J. Little, (Eds.), *Survey Nonresponse* (275-288). New York: Wiley.

Berk, M.L., & Bernstein, A.B. (1988). Interviewer characteristics and performance on a complex health survey. *Social Science Research,* 17(3), 239-251.

Bhandari, S. (2006). Health status, health insurance, and health services utilization: 2001. Current Population Reports (P70-106). Washington, DC: U.S. Census Bureau.

Biemer P. & Lyberg, L. (2003). *Introduction to survey quality*. New York: John Wiley & Sons, Inc.

Blumberg, S.J., & Luke, J.V. (2009). Wireless substitution: Early release of estimates from the National Health Interview Survey, January-June 2009. National Center for Health Statistics. Retrieved from: http://www.cdc.gov/nchs/nhis.htm.

Bohensky, M., Jolley, D., Sundararajan, V., Evans, S., Pilcher, D.V., Scott, I., & Brand, C.A. (2010). Data linkage: A powerful research tool with potential problems. *Biomedical Health Services Research*, 22(10), 346-352.

Boyd, H.W. & Westfall, R. (1955). Interviewers as a source of error in surveys. *Journal of Marketing Research,* 19(4), 311-324.

Bradburn, N. M. (1983). Response effects. In: P. Rossi, J. Wright, and A. Anderson (Eds.) *Handbook of Survey Research.* New York: Academic Press.

Calderwood, L. & Lessof, C. (2009) Enhancing longitudinal surveys by linking to administrative Data. In P. Lynn (Ed.), *Methodology of longitudinal surveys*. Chichester: John Wiley & Sons, Ltd.

Camburn, D., Lavrakas, P.J., Battaglia, M.P., Massey, J.T., & Wright. R.A. (1995). Using advance letters in random-digit-dialing telephone surveys. *Proceedings of the Section on Survey Research Methods, Lauderdale, Florida*, *969-974.*

Campanelli, P., & O'Muircheartaigh, C. (1999). Interviewers, interviewer continuity, and panel survey nonresponse. *Quality and Quantity* 33(1), 59-76.

Cannell, C., Groves, R., Magilavy, L., Mathiowetz, N., & Miller, P. (1987). An experimental comparison of telephone and personal health surveys. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, St Petersburg, Florida.

Centers for Disease Control and Prevention. (2011). National Immunization Survey: A user’s guide for the 2010 public-use data file. National Opinion Research Center: Chicago, IL.

Centers for Disease Control Prevention. (2011) National Immunization Survey: NIS public-use data files, reports and publications: Household Interview Questionnaire 2011. Retrieved from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NIS/NIS_Child_HHQuex_Q3_2011.pdf

Churchill, G.A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*,16(1), 64-73.

Cialdini, R.B., & Goldstein, N.J. (2004). Social influence: Compliance and conformity. *Psychology,* 55, 591-621.

Cleary, P.D. & Jette, A.M. (1984). The validity of self-reported physician utilization measures. *Medical Care*, 22(9), 796-803.

Cleary, P., Mechanic, D. & Weiss, N. (1981). The effect of interviewer characteristics on responses to a mental health interviewer, *Journal of Health and Social Behavior,* 22(2), 183-193.

Collins, M. (1980). Interviewer variability: A review of the problem. *Journal of Market Research Society*, (22), 239-247.

Couper, M.P. (2002). New technologies and survey data collection: Challenges and opportunities. Paper presented at the International Conference on Improving Surveys, Copenhagen, Denmark.

Couper, M.P., Singer, E., & Kulka, D. (1998). Nonresponse to the 1990 Census: Politics, privacy, and/or pressures. *American Politics Quarterly,* 26, 59-80.

Crimmins, E., Guyer, H., Langa, K., Ofstedal, M.B., Wallace, D., & Weir, D. (2009). Documentation of biomarkers in the Health and Retirement Study. Documentation report prepared for the University of Michigan.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16(3), 297-334.

Curtin, R., Presser, S. & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly,* 64, 413-28.

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

Dahlhamer, J.M., & Cox, C.S. (2007). Respondent consent to link survey data with administrative records: Results from a split-ballot field test with the 2007 National Health Interview Survey. *Proceedings of the Federal Committee on Statistical Methodology Research Meeting, Washington, DC.*

Das, M. (2011). To link or not to link: Why and when respondents give consent. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ.

Davern, M., Roemer, & Thomas, M. (2006). Incorporating linked survey and administrative data files into policy research. Paper presented to the Federal Committee on Statistical Methodology, Washington, DC.

De Leeuw, E., Callegaro, M., Hox, J., Korendijk, E., & Lensvelt Mulders, G. (2007). The influence of advance letters on response in telephone surveys: A meta-analysis. *Public Opinion Quarterly,* 71(3), 413–43.

De Leeuw, W. & De Heer, E. (2002) Trends in household survey non-response: a longitudinal and international comparison. In Groves, R. (ed.) *Survey nonresponse*. New York: John Wiley.

DeVellis, R.F. (1991). *Scale development.* Newbury Park, CA: Sage.

Dillman, D.A., Eltinge, J.L., Groves, R.M., & Little, R.J. (2002). Minimizing unit and item nonresponse. In R.M. Groves (Eds.), *Survey Nonresponse* (pp. 3-26). New York, NY, Wiley.

Dillman, D.A., Gallegos, J.G., & Frey, J.H. (1976). Reducing refusal rates for telephone interviews. *Public Opinion Quarterly,* 40(1), 66-78.

Dillman, D.A., Sinclair, M.D., & Clark, J.R. (1993). Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates. *Public Opinion Quarterly,* 57(3), 289-304.

Doll, R. & Hill, A.B. (1956). *Lung cancer and other causes of death in relation to smoking. A second report on the mortality of British doctors. British Medical Journal, 2(5001),* 1071-*1076*.

Drazga-Maxfield, L. (2008). Uses of administrative data at the U.S. Social Security Administration. Paper presented at the International Seminar on the Use of Administrative Data for Economic Statistics and the Register-Based Population and Housing Census, Korea.

Droitcour, J. (2001). Record Linkage and Privacy: Issues in creating new federal statistical information. Washington, DC: Government Printing Office.

Dunn, K.M., Jordan, K., Lacey, R.J., Shapley, M., & Jinks, C. (2004). Patterns of consent in epidemiologic research: Evidence from over 25,000 responders. *American Journal of Epidemiology,* 159(11), 1087-1094.

Durrant, G.B., Groves, R.M., Staetsky, L. & Steele, F. (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly*, 74(1), 1-36.

Feldman, J.J., Hyman, H., & Hart, C.W. (1951). A field study of interviewer effects on the quality of survey data. *Public Opinion Quarterly,* 15(4), 734-762.

Fienberg, S.E. (2006). Privacy and confidentiality in an e-commerce world: Data mining, data warehousing, matching and disclosure limitation, *Statistical Science,* 21(2), 143-154.

Finkelstein, M.M. (2001). Do factors other than need determine utilization of physicians services in Ontario? *Canadian Medical Association Journal,* 165(5), 565-570.

Fowler, F. J., & Mangione, F.J. (1990). *Standardized survey interviewing.* Newbury Park: Sage.

Freeman, F. & Butler, E.W. (1976). Some sources of interviewer variance in surveys. *Public Opinion Quarterly,* 41(1), 79-91.

Fulton, J., Schoeni, R. & Freedman, V. (2011). Respondent and Interviewer Predictors of Consent to Medicare Linkage in the Panel Study of Income Dynamics. Presentation at the Institute for Social Research, University of Michigan.

Fulton, J. & Tourangeau, R. (2011). Respondent Willingness to Grant Record Access or Provide Identifying Information in a Survey. Presentation at the American Association for Public Opinion Research Conference, Phoenix.

Gastwirth, C.J. (1986). Ethical Issues in Access to and Linkage of Data Collected by Government Agencies, *Proceedings of the American Statistical Association, Social Statistics Section,* 6-13.

Goyder, J. (1985). Face-to-face interviews and mailed questionnaires: the net difference in response rate. *Public Opinion Quarterly*, 49(2), 234-252.

Goyder, J. (1987). *The silent minority: Nonrespondents on sample surveys*. Boulder, CO: Polity Press/Westview Press.

Gray, M., Constantine, R., d'Ardenne, J., Blake, M. & Uhrig, S.C.N. (2008). Cognitive testing of Understanding Society. The UK Household Longitudinal Study questionnaire. Understanding Society Research Report No. 2008-04.

Groves, R.M. (1989). *Survey errors and survey costs*. New York: Wiley.

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. & Couper, M. (1998). *Nonresponse in household interview survey*. New York: Wiley.

Groves, R.M., Fowler, F., Couper, M.P., Lepkowski, J., Singer, R., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: John Wiley and Sons.

Groves, R.M. & Fultz, N.H. (1985). Gender effects among telephone interviewers in a survey of economic attitudes. *Sociological Methods and Research,* 14(1), 31-52.

Groves, R.M., & Kahn, R.L. (1979). *Surveys by telephone*. New York: Academic Press.

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

Groves, R.M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly,* 64(3), 299-308.

Guarino, J.A., Hill, J.M., & Woltman, H.F. (2001). The effect of SSN requests and notification of administrative record use on response behavior in Census 2000. *Proceedings of the American Statistical Association, 3336-3341.*

Gustman, A.L. & Steinmeier, T.L. (1999). What people don't know about their pensions and Social Security: An analysis using linked data from the Health and Retirement Study. NBER Working Paper No. 7368.

Haider, S., & Solon, G. (2000). *Nonrandom selection in the HRS Social Security earnings sample* (Labor and Population Program Working Paper Series: No. 00-01, 2000). California: RAND.

Haines, D.E. & Greenberg, B. (2005). *Statistical uses of Social Security administrative data*. Paper presented at American Statistical Association, Minneapolis, MN.

Hanson, R.H. & Marks, E.S. (1958). Influence of the interviewers on the accuracy of survey results. *Journal of the American Statistical Association,* 53, 635-55.

Harris, T., Cook, D., Victor, C., Beighton, C., DeWilde, S., & Carey, I. (2005). Linking questionnaires to primary care records: Factors affecting consent in older people. *Journal of Epidemiology and Community Health*, 59(4), 336-338.

Hatchet, S. & Schuman, H. (1975). White respondents and race of interviewer effects. *Public Opinion Quarterly,* 39, 523-528.

Heberlein, T.A., & Baumgartner, R. (1978). Factors affecting response rates of mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review,* 43, 446-462.

Hess, M., Hoffer, T., Lee, L., Loew, D., Webber, K., & Williams, K. (2005). Qualitative testing for the Survey of Earned Doctorates questionnaire. Report available upon request from the National Science Foundation.

Hoover, D.R., Crystal, S., Kumar, R., Sambamoorthi, U. & Cantor, J.C. (2002). Medical Expenditures during the Last Year of Life: Findings from the 1992-1996 Medicare Current Beneficiary Survey. *Health Services Research,* 37(6), 1625-1642.

Hox, J.J, de Leeuw, E.D. & Kreft, G.G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In: P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (Eds.) *Measurement Errors in Surveys.* New York: Wiley.

Huang, N., Shih, S.F, Chang, H.Y., & Chou, Y.J. (2007). Record linkage research and informed consent: Who consents? *Biomedical Health Services Research,* 7, 7-18.

Huddy, L., Billig, J., Bracciodieta, J., Hoeffler, L., Moynihan, P., & Pugliani, P. (1997). The effects of interviewer gender on the survey response. *Political Behavior,* 19,197-220.

Hurd, M. D., Juster, F. & Smith, J.P. (2003). Enhancing the quality of data on income: Recent developments in survey methodology. *Journal of Human Resources*, 38(3), 758-772.

Jabine, T. & Scheuren, F. (1986). Record linkages for statistical purposes: Methodological issues. *Journal of Official Statistics,* 2(3), 255-277.

Jay, G.M., Belli, R.F., & Lepkowski, J.M. (1994). Quality of doctor visit reports: A comparison of medical record and survey data. *Proceedings of the Americal Statistical Association, 362-367.*

Jenkins, S.P., Cappellari, L., Lynn, P., Jackle, A., & Sala, E. (2006) Patterns of consent: Evidence from a general household survey. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 169(4), 701-711.*

Jenkins, S.P., Lynn, P., Jackle, A., & Sala, E. (2008). The feasibility of linking household survey and administrative record data: New evidence for Britain. *International Journal of Social Research Methodology*, 11(1), 29-44.

Jones, P. & Elias, P. (2006). *Administrative data as a research resource: A selected audit*, London, England: National Data Strategy.

Juster, T. F. & Smith, J.P. (1997). Improving the quality of economic data: Lessons from HRS and AHEAD. *Journal of the American Statistical Association*, 92(440), 1268-1278.

Kahn, R.L. & Cannell, C.F. (1957). *The dynamics of interviewing: Theory, technique, and cases.* New York: Wiley.

Kane, E.W. & Macaulay, L.J. (1993). Interviewer gender and gender attitudes. *Public Opinion Quarterly*, 57(1), 1-28.

Klassen, A.F., Lee, S.K., Barer, M., & Raina, P. (2005). Linking survey data with administrative health information. *Canadian Journal of Public Health,* 96(2), 151-154.

Koenig, M.L. (2003). An assessment of the Current Population Survey and the Survey of Income and Program Participation using Social Security Administrative data. Paper presented at the 2003 Federal Committee on Statistical Methodology Research Conference.

Korkeila, K., Suominen, S., Ahvenainen, J., Ojanlatva, A., Rautava, P., Helenius, H., & Koskenvuo, M. (2001). Non-response and related factors in a nation-side health survey. *European Journal of Epidemiology,* 17(11), 991-999.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Lane, J. (2010). Linking administrative and survey data. In P. Marsden & J. Wright (Eds.), *The Oxford Handbook of Survey Research*. (pp. 659-680). Oxford: Oxford University Press.

Lee, L., Brittingham, A., Tourangeau, R., Rasinski, K., Willis, G., Ching, P., Jobe, J., & Black, S. (1999). Are reporting errors due to encoding limitations or retrieval failure? Surveys of child vaccination as a case study. *Journal of Applied Cognitive Psychology*, 13, 43-63.

Lewin, K. (1942). Field theory of learning. *Yearbook of National Social Studies of Education* 41, 215-242.

Lillard, L.A. & Farmer, M.M. (1997). Linking Medicare and national survey data. *Annals of Internal Medicine*, 127, 691-669

Linsky, A. S. (1975). Stimulating responses to mailed questionnaires: A review. *Public Opinion Quarterly,* 39, 82-101.

Lipps, O. & Pollien, A. (2011). Effects of interviewer experience on components of nonresponse in the European Social Survey. *Field Methods*, 23(2), 156-172.

Lyberg, I. & Dean, P. (1992). Methods for reducing nonresponse rates: A review. Paper presented at the Annual Meeting of the American Association for Public Opinion Research.

Lyberg, I. & Lyberg, P. (1991). Nonresponse Research at Statistics Sweden. Paper presented at the Annual Meeting of the American Association for Public Opinion Research.

Machlin, S. & Taylor, A. (2000). *Design, methods, and field results of the 1996 Medical Expenditure Panel Survey Medical Provider Component* (*MEPS Methodology Report* No. 9. AHRQ Pub. No. 00-0028). Retrieved from http://www.ahrq.gov/about/cfact/cfactbib29.htm

Martinez, R.W. (2010) Integrating administrative records into the federal statistical system 2.0. Paper presented at the National Academies Committee on National Statistics Workshop on the Future of Federal Household Surveys.

McCarthy, D.B., Shatin, D., Drinkard, C., Kleinman, J.H., & Gardner, J. S. (1999). Medical records and privacy: Empirical effects of legislation. *Health Services Research,* 34(1): 1 Part 2, 417-425.

McNabb, J., Timmons, D., Song, J., & Puckett, C. (2009). Uses of administrative data at the Social Security Administration. *Social Security Bulletin,* 69(1), 75-84.

Melton, L. (1997). The threat to medical-records research. *New England Journal of Medicine,* 337(20), 1466-1770.

Meyer, P. S., Dahlhamer, J. M., & Pleis, J. (2006). Developing new methods and questions for improving response and measurement on sensitive questions on the National Health Interview Survey, *Proceedings of the American Statistical Association, 3426-3429.*

Michaud, S., Dolson, D., Adams, D., & Renaud, M. (1995). Combining administrative and survey data to reduce respondent burden in longitudinal survey (*SLID Research Paper Series*). Ottawa: Statistics Canada.

Miller, D.M., Gindi, R.M., & Parker, J.D. (2011). Trends in record linkage refusal rates: Characteristics of National Health Interview Survey respondents who refuse record linkage. Paper presented at American Statistical Association, Miami Beach, FL.

Morton-Williams, Jean. (1993). *Interviewer Approaches.* Brookfield, VT: Dartmouth.

Murdoch, M., Pietila, D.M., & Partin, M.R. (2010). Randomized trial showed that an "embedded" survey strategy optimized authorization rates compared with two "after survey" strategies in veterans with PTSD. *Journal of Clinical Epidemiology,* 63(6), 665-671.

National Academy of Sciences. (1979). Privacy and Confidentiality as Factors in Survey Response. Washington, D.C.: National Research Council.

National Center for Health Statistics (2009). National Health Interview Survey Questionnaire. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved from http://www.cdc.gov/nchs/nhis/nhis_questionnaires.htm

National Center for Health Statistics (2010). National Health and Nutrition Examination Survey Questionnaire. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved from http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm

National Consumers League Survey, April, 1999. Retrieved August 31, 2011 from the iPOLL Databank, The Roper Center for Public Opinion Research, University of Connecticut. http://www.ropercenter.uconn.edu/data_access/ipoll/ipoll.html

Nealon, J. (1983). The effects of male vs. female telephone interviewers. *Proceedings of the American Statistical Association Section on Survey Research Methods.*

Nelson, K., Garcia, R.E., Brown, J., Mangione, C.M., Louis, T., Keeler, E. & Cretin, S. (2002). Do patient consent procedures affect participation rates in health services research? *Medical Care, 40(4),* 283-288.

Nunnally, J.C. (1967). *Psychometric theory.* New York, NY: McGraw-Hill.

Nunnally, J.C. (1978). *Psychometric theory* (2nd ed). New York, NY: McGraw-Hill.

Olsen, R.J. (2005). The problem of respondent attrition: Survey methodology is key. *Bureau of Labor Statistics Monthly Labor Review,* February, 63-70.

Olson, J.A. (1999). Linkages with data from Social Security Administration records in the Health and Retirement Study. *Social Security Bulletin.* 73(2), 73-85.

O'Muircheartaigh, C. & Campanelli, P. (1998). The relative impact of interviewer effects and sample design effects on survey precision. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 161, 63–77.

Partin, M.R., Burgess, D.J., Halek, K., Grill, J., Vernon, S.W., Fisher, D.A., Griffin, J.M., & Murdoch, M. (2008). Randomized trial showed requesting medical records with a survey produced a more representative sample than requesting separately. *Journal of Clinical Epidemiology,* 61(10), 1028-1035.

Pascale, J. (2011). Requesting Consent to Link Survey Data to Administrative Records: Results from a Split-Ballot Experiment in the Survey of Health Insurance and Program Participation (SHIPP). Study Series in Survey Methodology #2011-03. Retrieved from http://www.census.gov/srd/papers/pdf/ssm2011-03.pdf.

Petty, D.R., Zermansky, A.G., Raynor, D.K., Lowe, C.J., Buttress, A.D., Vail, A., & Freemantle, N. (2001). "No thank you": Why elderly patients declined to participate in a research study. *International Journal of Clinical Pharmacy*, 23(1), 22-27.

Pew Internet Survey, Feb, 2007. Retrieved April 26, 2011 from the iPOLL Databank, The Roper Center for Public Opinion Research, University of Connecticut. http://www.ropercenter.uconn.edu/data_access/ipoll/ipoll.html

Potok, N. (2011). Expanded use of administrative records in the 2020: Can the non-technical issues be resolved early in the decade? Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ.

Powell, A.A., Burgess, D.J., Vernon, S.W., Griffin, J.M., Grill, J.P., Noorbalooch, I.S., & Partin, M.R.. (2009). Colorectal cancer screening mode preferences among U.S. veterans. *Preventive Medicine,* 49(5), 442–448.

Rasinski, K.A., Baldwin, A.K., Willis, G.B., & Jobe (1994). Risk and loss perceptions associated with survey reporting of sensitive behavior. Paper presented at the annual meeting of the American Statistical Association, Toronto, Canada.

Redfern, P. (1986). Which countries will follow the Scandinavian lead in taking a register-based Census of population? *Journal of Official Statistics*, 2(4), 415–424.

Ross, C.E. & Mirowsky, J. (2000). Does medical insurance contribute to socioeconomic differentials in health? *The Milbank Quarterly,* 78(2), 291-321.

Rubin, D.R. (1987). *Multiple imputation for nonresponse in surveys.* New York: Wiley.

Sakshaug, J. (2011). To link or not to link. Presentation to the Washington, D.C. Chapter of AAPOR.

Sakshaug, J.W., Couper, M.P., & Ofstedal, M.B. (2010). Characteristics of physical measurement consent in a population-based survey of older adults. *Medical Care*, 48(1), 64-71.

Sakshaug, J.W., Couper, M.P., Ofstedal, M.B., & Weir, D. (2012 - in press). Linking survey and administrative Records: Mechanisms of consent. *Sociological Methods and Research*.

Sakshaug, J. & Kreuter, F. (2011). Assessing the magnitude of administrative non-consent biases in the German PASS Study. Paper presented at the European Survey Research Association Conference, Lausanne, Switzerland.

Sala, E., Burton, J., Knies, G. (2010). Correlates of obtaining informed consent to data linkage: Respondent, interview and interviewer characteristics. *ISER Working Paper,* No. 2010-2028.

Saris, W E. (2007). *Design, evaluation, and analysis of questionnaires for survey research.* Hoboken: John Wiley & Sons, Inc., (2007).

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Sayer, B. & Cox, C.S. (2003). How many digits in a handshake? National Death Index matching with less than nine digits of the Social Security Number. *Proceedings from the American Statistical Association, 3666-3672*

Schaeffer, N.C. (1980). Evaluating race-of-interviewer effects in a national survey. *Sociological Methods and Research,* 8(May), 400-419.

Scheuren, F. (1997) Linking health records: Human rights concerns. In A.R. Chapman (Ed.), *Health care and information ethics: Protecting fundamental human right*s. Kansas City: Sheed and Ward.

Schnell, R. & Kreuter, F. (2005). Separating interviewer and sampling-point effects. *Journal of Official Statistics,* 21(3), 389–410.

Schuman, H. & Converse, J.M. (1971). The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly,* 35(1), 44-68.

Shah, S., Harris, T.J., Rink, E., DeWilde, S., Victor, C.R., & Cook, D.G. (2001). Do income questions and seeking consent to link medical records reduce survey response rates? A randomized controlled trial among older people. *British Journal of General Practice,* 51(464), 223-225.

Shell Poll, Mar, 1999. Retrieved March 21, 2011 from the iPOLL Databank, The Roper Center for Public Opinion Research, University of Connecticut. http://www.ropercenter.uconn.edu/data_access/ipoll/ipoll.html

Silva, M. S., Smith, W.T. & Bammer, G. (2002) The effect of timing when seeking permission to access personal health services utilization records. *Annals of Epidemiology*, 12(5), 326–330.

Singer, E. (1978). Informed consent: Consequences for response rate and response quality in social surveys. *American Sociological Review*, 43(April), 144-162.

Singer, E. (2001). Public perceptions of confidentiality and attitudes toward data sharing by federal agencies. In P. Doyle, J. Lane, J. Theeuwes, & L. Zayatz (Eds.), *Confidentiality, disclosure, and data access*. New York: Elsevier.

Singer, E. (2011). Toward a benefit-cost theory of survey participation: Evidence, further tests, and implications. *Journal of Official Statistics*, 27(2), 379-392

Singer, E., Bates, N. & Miller, E.R. (1992). Memorandum for Susan Miskura, Bureau of the Census, July 15, 1992.

Singer, E., Bates, N. & Van Hoewyk, J. (2011). Concerns about privacy, trust in government, and willingness to use administrative records to improve the decennial census. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ.

Singer, E., Frankel, M.R., & Glassman, M.B. (1983). The effect of interviewer characteristics and expectations on response. *Public Opinion Quarterly,* 47(1), 68-83.

Singer, E., Mathiowetz, N.A., &. Couper, M.P. (1993). The impact of privacy and confidentiality concerns on survey participation: The case of the 1990 U.S. Census. *Public Opinion Quarterly,* 57(4), 465-482.

Singer, E., Van Hoewyk, J. & Neugebauer, J. (2003). Attitudes and behavior: The impact of privacy and confidentiality concerns on participation in the 2000 Census. *Public Opinion Quarterly* 67(3), 368-384.

Singer, E. & Ye, C. (2010). The use and effects of incentives in surveys. In D. Massey and R. Tourangeau (Eds.), *The Future of Social Science Surveys: Opportunities and Challenges*. New York: Russell Sage.

Smith, P.J., Hoaglin, D.C., Battaglia, M.P., Khare, M., & Barker, L.E. (2005). Statistical methodology of the National Immunization Survey, 1994-2002. *Vital Health Statistics,* 2 March (138), 1-155.

Snijkers, G., Hox, J.J. & de Leeuw, E.D. (1999). Interviewers' tactics for fighting survey nonresponse. *Journal of Official Statistics*, 15(2), 185-198.

Steeh, C.G. (1981). Trends in nonresponse rates, 1952-1979. *Public Opinion Quarterly,* 45(1), 40-57.

Steeh, C.G., Kirgis, N., Cannon, B., & DeWitt, J. (2001). Are they really as bad as they seem? Nonresponse rates at the end of the twentieth century. *Journal of Official Statistics*, 21(2), 227-247.

Sudman, S. & Bradburn, N.M. (1974). *Response effects in surveys: A review and synthesis*. Chicago, IL: Aldine Publishing Co.

Sykes, W. & Collins, M. (1988). Effects of mode of interview: Experiments in the UK. In: R.M. Groves, *Telephone survey methodology*. New York: Wiley

Tate, A.R., Calderwood, L., Dezateux, C., & Joshi, H. (2005). Mother's consent to linkage of survey data with her child's birth records in a multi-ethnic national cohort study. *International Journal of Epidemiology,* 35(2), 294-308.

Tourangeau, R., Groves, R.M., & Redline, C.D. (2010). Sensitive topics and reluctant respondents: Demonstrating a link between nonresponse bias and measurement error. *Public Opinion Quarterly,* 74(3), 413-432

Tourangeau, R., Rips, L.J., & Rasinski, K.A. (2000). *The psychology of survey response*. England: Cambridge University Press.

Tourangeau, R., Singer, E. & Presser, S. (2003). Context effects in attitude surveys: Effects on remote items and impact on predictive validity. *Sociological Methods & Research,* 31(4), 486-513.

Traugott, M.W., Groves, R.M, & Lepkowski, J.M. (1987). Using dual frame designs to reduce nonresponse in telephone surveys. *Public Opinion Quarterly*, 51(4), 522-539.

Tu, V., Willison, D., Silver, F., Fang, F., Richards, J., Laupacis, A., & Kapral, M. (2004). Impracticability of informed consent in the Registry of the Canadian Stroke Network. *New England Journal of Medicine,* 350, 1414–1421.

U.S. Census Bureau. (2011). *History of the 2007 Economic Census*. Washington, DC: Government Printing Office.

U.S. Department of Energy, Energy Information Administration. (2009). Residential Energy Consumption Survey Household Questionnaire. Retrieved from http://www.eia.gov/EMEU/recs/recs09/2009RECSHousehold_Form457A.pdf

U.S. Department of Energy, Energy Information Administration. (2011). How does the EIA estimate energy consumption and end use in U.S. homes? Retrieved from http://www.eia.gov/consumption/residential/methodology/2009/brief.cfm

U.S. Government Accountability Office. (2001). Record linkage and privacy: Issues in creating new federal research and statistical information. Washington, D.C.: United States General Accounting Office.

Wallman, K.K., & Coffey, J.L. (1997). Sharing Statistical Information for Statistical Purposes. In W. Alvey & B. Jamerson (Eds.), *Record Linkage Techniques* (pp. 268-278). Washington, DC: National Academy Press.

Webster, C. (1996). Hispanic and Anglo interviewer and respondent ethnicity and gender: The impact on survey response quality. *Journal of Marketing Research*, Feb, 62-72.

Welch, V., Groenhout, B., Webber, K., & Williams, K. (2008). Report of the impact of the introduction of a four digit SSN item on the Survey of Earned Doctorates. Report available upon request from the National Science Foundation.

Weir, D. (2007). Health and Retirement Study Institutional Review Board information. Retrieved from http://hrsonline.isr.umich.edu/sitedocs/irb/HRS_IRB_WebPackage-09-09.pdf.

West, B.T., & Olson, K. (2011). How much of interviewer variance is really nonresponse error variance? *Public Opinion Quarterly,* 74(5), 1004-1026.

Willis G. (2006). Cognitive interviewing as a tool for improving the informed consent process. *Journal of Empirical Research on Human Research Ethics*, 1(1), 9-24.

Willis, G.B., Sirkin, N.G., & Nathan., G. (1994). The cognitive aspects of responses to sensitive survey questions. *Cognitive methods staff working paper series,* 9.

Woolf, S.H., Rothemich, R.E, & Marsland, D.W. (2000). Selection bias from requiring patients to give consent to examine data for health services research. *Archives of Family Medicine,* 9(10), 1111-1118.

Yaffe, R., Shapiro, S., Fuchsberg, R., Rohde, C. & Corpeno, H. (1978). Medical economics survey-methods study: Cost-effectiveness of alternative survey strategies, *Medical Care,* 16(8), 641-659.

Yang, Q., Liu, T., Valdez, R., Moonesinghe, R., & Khoury, M.J. (2010). Improvements in ability to detect undiagnosed diabetes by using information on family history among adults in the United States. *American Journal of Epidemiology*, 15, 1079-1089.

Zell, E.R., Ezzati-Rice, T.M., Battaglia, M.P., & Wright, R.A. (2000). National Immunization Survey: The methodology of a vaccination surveillance system. *Public Health Reports*, 115(1), 65-77.