To Laura

To My Parents and Grandparents

To All of My Teachers

Acknowledgements

I would like to first of all acknowledge fellowship support from the Boehner Family and the Rackham School of Graduate Studies, and financial support from the Michigan Program in Survey Methodology (MPSM). I would also like to thank Steven Pennell from the Survey Research Center for his assistance with the IRB review of my dissertation proposal.

I am indebted to my colleagues at the Institute for Employment Research (IAB) in Nuremberg, Germany, and especially Dr. Mark Trappmann, who has provided me with tireless assistance working with the PASS survey and the INFAS organization. I also sincerely thank Joerg Heining from the IAB for his constant support working with the MAP from Ann Arbor and for allowing me to access the PASS survey data.

Many of my colleagues in the Survey Methodology Program (SMP), the Joint Program in Survey Methodology (JPSM), and the Survey Research Operations (SRO) group within the Survey Research Center (SRC) have provided me with invaluable support. John Van Hoewyk was always available to help me work with the NSFG data. The NSFG team, led by Nicole Kirgis, Shonda Kruger-Ndiaye, Haley Gu, Frost Hubbard, and James Wagner, has provided amazing support for all topics NSFG-related. Steve Heeringa chaired both of my exam committees, and has been my co-author, co-instructor, colleague, and friend.

Mick Couper served on my comprehensive exam (CE) committee, and could always can be counted on to respond to my crazy ideas and provide comments on my papers. Mike Elliott also served on my CE committee, helped in leading the PhD Seminar where the ideas for this dissertation were formed, and provided me with excellent instruction and advice. Roger Tourangeau, Bob Groves, and Partha Lahiri also led the PhD Seminar, and created an outstanding intellectual environment that gave birth to these ideas. Jill Montaquila served on my CE committee, and was also an excellent instructor. Raghu has always been an amazing instructor and colleague, and provided me with the most entertaining drive from Chicago to Ann Arbor that I've had to date.

I am also grateful for my fellow students in MPSM and JPSM, and their support, their countless conversations and discussions, and their willingness to listen to me go on and on in PhD Seminar. The staff of the Center for Statistical Consultation and Research (CSCAR) has provided me with an amazing support network for more than 10 years, and I consider them to be part of my family. I extend a heartfelt thank-you to my co-chairs Jim Lepkowski and Frauke Kreuter for their support, discussions, encouragement, and friendship. Ed Rothman has been an incredible mentor to me, and is a motivating force for the statistician that I am today. Rod Little has provided non-stop technical expertise and writing assistance, and most importantly has always been willing to put up with my analytical weaknesses. Finally, I thank my family and friends for supporting my decision to go back to school, and especially my wife Laura for lovingly putting up with the back of my head and my late nights for so very long.

Table of Contents

Dedication....ii
Acknowledgements....iii
List of Figures....viii
List of Tables....x
List of Appendices....xii
List of Abbreviations....xiii
Abstract....xv
Chapter
I. Introduction....1
II. Review of the Literature....5
The Effectiveness of Nonresponse Adjustments Based on Interviewer Observations....5
The Error Properties of Interviewer Observations....9
Predictors of Judgment Accuracy from Social Psychology....13
The Implications of Errors in Interviewer Observations for Nonresponse Adjustments....24
A Broader View of the Problem....28
III. The Quality and Utility of Interviewer Observations of Household Characteristics in the National Survey of Family Growth (NSFG)....32
Summary....32
Introduction....32
Background....35
NSFG Data....38
Statistical Analyses....41
Results....46
What are the error properties of the interviewer observations in the NSFG?....46
Are the observations associated with response indicators and key NSFG variables?....48
Do key estimates shift when using the observations to adjust for nonresponse?....52
How do errors in the observations impact nonresponse adjustments?....53
Discussion....56
IV. Factors Impacting the Accuracy of Interviewer Observations in the National Survey of Family Growth (NSFG)....62
Summary....62
Introduction....62
Background....65

Predictors of Accuracy in Behavioral or Personality Judgments……...65
Predictors of Accuracy in Judgments of Household Characteristics…….69
NSFG Data…….70
Capturing the effect of providing interviewers with predictive information……75
Identifying different observational strategies used by NSFG interviewers……...75
Statistical Analyses……80
Remark: Interpenetration……...85
Results……85
Interviewer-specific Measures of Quality……..86
Factors Influencing the Accuracy of Judgments about Presence of Children……..88
Factors Influencing the Accuracy of Judgments about Current Sexual Activity…….106
Do different observational strategies explain the variance in the accuracy of the observations on current sexual activity?.....118
Discussion……120
V. Nonresponse Adjustment Based on Auxiliary Variables Subject to Error….125
Summary……..125
Introduction……..126
Pattern-Mixture Model: Estimation and Inference……..129
Pattern-Mixture Model (PMM) Estimates……...129
Bayesian Inference……...136
Multiple Imputation…….137
Simulation Studies……...139
Simulations Based on the Normal Selection Model……139
Simulations Based on the Pattern-Mixture Model……...142
Results of Simulation Studies……..144
Applications to Real Survey Data……148
The Labor Market and Social Security (PASS) Survey……...148
The Continuous National Survey of Family Growth (NSFG)….151
Generalization to Include Other Fully Observed Auxiliary Variables…154
Discussion……157
VI. Conclusions and Discussion……..161
Synthesis of Study Results……...161
Future Directions for Survey Methodology…….175

Appendices……...178
Bibliography……202

List of Figures

Figure

4.1 Scatter plot examining the association of interviewer gross difference rates (GDRs) on judgments of sexual activity and presence of children under age 15 in the household. Two interviewers with relatively low GDRs on both observations are highlighted with arrows. Sizes of points (weights) are based on the number of initial observations on children under 15 (representing attempted screening interviews)…….87
4.2 Stacked bar chart indicating variance across 94 interviewers in terms of the percentages of judgments of presence of young children that were correct, false positives, or false negatives, with interviewers ordered in terms of percentages of observations that were false positives (from lowest to highest)……....90
4.3 Predicted probabilities of false positives (relative to correct judgments) on presence of young children for combinations of access problems (1 = yes, 0 = no) and prior marital status, based on the final multilevel model for accuracy on this judgment (other covariates fixed to their means)……...101
4.4 Predicted probabilities of false positives for judgments on presence of young children for combinations of many unit buildings (1 = yes, 0 = no) and Black / Non-Black interviewers, based on the final multilevel model for accuracy on this judgment (other covariates fixed to their means)……...102
4.5 Predicted probabilities of false positives for judgments on presence of young children for combinations of many unit buildings (1 = yes, 0 = no) and never having been married for interviewers, based on the final multilevel model for accuracy on this judgment (other covariates fixed to their means)……...……103
4.6 Predicted probabilities of false negatives for judgments on presence of young children for combinations of many unit buildings (1 = yes, 0 = no) and whether interviewers work all three PSU types, based on the final multilevel model for accuracy on this judgment…….…….104
4.7 Predicted probabilities of false negatives for judgments on presence of young children as a function of percentage of ZCTA that is children under 18 (centered within interviewers) and prior marital status for interviewers, based on the final multilevel model for accuracy on this judgment (other covariates fixed to their means)……..……105
4.8 Stacked bar chart indicating variance across 87 interviewers in terms of the percentages of judgments of current sexual activity that were correct, false positives, or false negatives, with interviewers ordered in terms of percentages of observations that were false positives (from lowest to highest)…….…….108

5.1 Missing data pattern under study……..…….127

A.1 Predicted probabilities of false positives for judgments on presence of young children for combinations of physical impediments to housing unit access and variety of PSUs worked, based on the exploratory modeling……182
A.2 Predicted probabilities of false positives for judgments on presence of young children for combinations of days of experience since onset of Cycle 7 data collection (Early / Late), interviewer education (College / No College) and variety of PSUs worked, based on the exploratory modeling……..……..184
A.3 Predicted probabilities of false positives for judgments on presence of young children for combinations of prior NSFG experience and variety of PSUs worked, based on the exploratory modeling……...…188

B.1 Dendrogram showing results of initial hierarchical agglomerative cluster analysis, with evidence of two outliers (Interviewers 36 and 40)……..……192
B.2 Dendrogram showing results of second cluster analysis (excluding the two outliers), with evidence of four distinct groups of interviewers (based on rescaled cluster distances greater than 10)…….…...193
B.3 Screenshot of CAPI application screen where NSFG interviewers could enter open-ended justifications for their sexual activity judgments. The justification was typed into the box labeled “Rsex rel with opposite sex partner,” which was converted to a full text box when interviewers started typing……..……194

List of Tables

Table

3.1 Case counts and overall percentages indicating the error properties of interviewer judgments regarding the presence of children under the age of 15 in selected households (NSFG, Cycle 7)……..……...47
3.2 Case counts and overall percentages indicating the error properties of interviewer judgments of current sexual activity among respondents (NSFG, Cycle 7)……..……47
3.3 Selected main interview response propensity modeling results, showing adjusted relationships of NSFG interviewer observations with response indicators……..….48
3.4 Adjusted estimates of regression parameters for the two interviewer judgments as predictors of five key NSFG variables, contrasted with estimates using the "true" auxiliary variables as predictors instead (NSFG respondents only)……..…50
3.5 Impacts of alternative nonresponse adjustments on NSFG estimates (linearized standard errors reported in parentheses)……..……..52
3.6 Results of simulation study, showing empirical performance of estimators with nonresponse adjustments based on either respondent reports of current sexual activity or interviewer judgments of current sexual activity……...…53

4.1 Predictors at the respondent level and the interviewer level to be considered in the multilevel multinomial logistic regression models of accuracy for the two NSFG interviewer judgments……..…..73
4.2 Descriptive statistics for interviewer justification tendencies (in descending order by mean percentages of justifications) and mean word counts……...…77
4.3 Descriptive statistics for interviewer-level justification tendencies and mean word counts within four distinct clusters of interviewers……..…….79
4.4 Descriptive statistics for available characteristics of the 94 interviewers recording at least 20 judgments on the presence of children under the age of 15 in sampled NSFG households……..……88
4.5 Multilevel modeling results for accuracy of judgments on young children……..…..91
4.6 Descriptive statistics for available characteristics of the 87 interviewers recording at least 20 judgments on the current sexual activity of eligible respondents selected from screening interviews within sampled NSFG households……...……..106
4.7 Multilevel modeling results for accuracy of judgments on current sexual activity……...…109

5.1 Selected simulation results under the normal selection model, with $\alpha = 0$ in the response propensity model……..…….144
5.2 Selected simulation results under the pattern-mixture model……..……..145

5.3 Estimates of mean reported household income and mean housing unit area (in square meters), based on four different nonresponse adjustment methods……..……...149
5.4 Estimates of selected means in the NSFG teenage population, based on four different nonresponse adjustment methods……153

D.1 Selected simulation results under the normal selection model, with α = -1 in the response propensity model……..….201

List of Appendices

Appendix

A Additional Exploratory Analyses from Chapter 4……..…….178
B Cluster Analysis Results from Chapter 4……..……...192
C R Functions Implementing PMM Estimators for Trivariate Normal Incomplete Data……..……195
D Additional Simulation Results from Chapter 5……..……..201

List of Abbreviations

ACASI	Audio Computer Assisted Self Interviewing
ANES	American National Election Study
BGPS	British Gambling Prevalence Survey
CAPI	Computer Assisted Personal Interviewing
CHI	Contact History Instrument
CI	Confidence Interval
CIA	Central Intelligence Agency
EBLUP	Empirical Best Linear Unbiased Predictor
EHCS	English House Condition Survey
EM	Expectation-Maximization
ESS	European Social Survey
FR	Field Researcher
FRQ	Field Researcher Questionnaire
GDR	Gross Difference Rate
HH	Household
HLM	Hierarchical Linear Modeling
HSE	Health Survey for England
IPT	Interpersonal Perception Task
IRB	Institutional Review Board
IWER	Interviewer
LRT	Likelihood Ratio Test
MAR	Missing at Random
MCAR	Missing Completely at Random
MC-SIMEX	Misclassification Simulation-Extrapolation
MI	Multiple Imputation
NHANES	National Health and Nutrition Examination Survey
NHEFS	NHANES Epidemiologic Follow-Up Study
NHIS	National Health Interview Survey
NSDUH	National Survey of Drug Use and Health
NSFG	National Survey of Family Growth
NSR	Non-Self-Representing
PASS	Labor Market and Social Security (German)
PMM	Pattern Mixture Model
PPMA	Proxy Pattern Mixture Analysis
PQL	Penalized Quasi Likelihood
PSU	Primary Sampling Unit
RDD	Random Digit Dialing
RMSE	Root Mean Squared Error
SCF	Survey of Consumer Finances

SR	Self-Representing
TSE	Total Survey Error
UK	United Kingdom
USDA	United States Department of Agriculture
ZCTA	Zip Code Tabulation Area

ABSTRACT

Interviewer observations are an important source of auxiliary information in survey research. Interviewers can record observations for all units in a sample, and selected observations may be associated with both key survey variables and response propensity. Survey statisticians use auxiliary variables with these properties to compute post-survey nonresponse adjustments to survey estimates that reduce both bias and variance in the estimates engendered by nonresponse. Unfortunately, interviewer observations are typically judgments and estimates, making them prone to error. To date, no studies have considered the implications of these errors for the effectiveness of nonresponse adjustments, effective observational strategies leading to reduced error rates, predictors of observation accuracy in face-to-face surveys, or alternative estimation methods for mitigating the effects of the errors on estimates. This dissertation presents results from three research studies designed to fill these important gaps in the existing literature.

The first study 1) analyzes the error properties of two interviewer observations collected in the National Survey of Family Growth (NSFG), finding accuracy rates ranging from 72-78% and evidence of systematic errors; 2) examines the effectiveness of nonresponse adjustments based in part on the observations, finding evidence of associations with key NSFG variables and response propensity but only slight shifts in estimates; and 3)

simulates the implications of errors in the observations for the effectiveness of weighting class adjustments for nonresponse, finding that adjustments based on the error-prone observations attenuate possible reductions in bias. The second study uses multilevel modeling techniques to identify several respondent- and interviewer-level predictors of accuracy in the two NSFG observations, including those supported by social psychological theories of what leads to improved judgment accuracy. The third study develops pattern-mixture model (PMM) estimators of means for the case when an auxiliary variable is error-prone, true values for the variable are collected from survey respondents, and the true values are predictive of unit nonresponse under a non-ignorable missing data mechanism. Simulation studies show that the PMM estimators have several favorable properties in these situations relative to other popular estimators, and R code is provided implementing the PMM approaches.