MISSING CASES: AN EVALUATION OF ABORTION UNDERREPORTING IN ADD HEALTH

Katherine I. Tierney

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Sociology in the College of Arts and Sciences.

Chapel Hill
2017

Approved by:

Kathleen Mullen Harris

Samuel Philip Morgan

Yong Cai

ABSTRACT

Katherine I. Tierney: Missing Cases: An Evaluation of Abortion Underreporting in Add Health
(Under the direction of Kathleen Mullen Harris and S. Philip Morgan)

The underreporting of abortions on surveys is a problem that limits the ability of researchers to understand who has abortions, under what circumstances, and with what consequences. This paper is the first to evaluate the quality of the abortion data in the National Longitudinal Study of Adolescent to Adult Health (Add Health). The overall quality of these data and patterns of underreporting are evaluated. The paper also uses multi-level analyses to uncover the characteristics associated with reporting over time. The accuracy of Add Health abortion data varies across years and method of evaluation used. We find that Add Health captures 35% of expected abortions when using an abortion rate comparison, 43% when using a ratio comparison, and 20% when using a modeling technique. Additionally, we find that the consistency of reporting over time is influenced by the circumstances under which women have an abortion. Theoretical and practical implications are discussed.

ACKNOWLEDGEMENTS

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

Note: Use of this acknowledgment requires no further permission from the persons named.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1650116. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

This paper uses data from NCHS. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of NCHS, which is responsible only for the initial data.

This paper uses data from the 2008 Abortion Patient Survey, which was collected by the Guttmacher Institute. Funding for this project was provided by an anonymous donor.

This paper uses data from the *2000-01 National Survey of Contraceptive Use Among women Having Abortions* and *1994-1995 National Survey of Contraceptive Use Among Women Having Abortions.* These data were made available by the Data Archive on Adolescent Pregnancy and Pregnancy Prevention, Sociometrics Corporation, 170 State Street, Suite 260, Los Altos, CA 94022-2812. The study entitled *2000-01 National Survey of Contraceptive Use Among women Having Abortions* was conducted by The Alan Guttmacher Institute. Funding for the data collection was provided by the Alan Guttmacher Institute. Funding for the work done by the Data Archive on Adolescent Pregnancy and Pregnancy Prevention to prepare the data for public use was provided by the National Institute for Child Health and Human Development (NICHD) under Contract No. N01-HD-9-3306 to Sociometrics Corporation. The study entitled the *1994-1995 National Survey of Contraceptive Use Among Women Having Abortions* was conducted by Stanley K. Henshaw of the Alan Guttmacher Institute. Funding for data collection was provided by the National Institute of Child Health and Human Development. Funding support for preparing the revised documentation for public distribution was provided by a contract (282-94-0018) between the U. S. office of Population Affairs and Sociometrics Corporation. The original investigators, funding agency and DAAPPP are not responsible for the analyses or interpretations presented here.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: AN EVALUATION OF MISSING CASES IN ADD HEALTH

Introduction

Is it possible for survey researchers to gather accurate abortion histories from women in the United States? Evidence of substantial underreporting has been found in all nationally representative surveys assessed (Fu et al. 1998; Jones and Forrest 1992; Jones and Kost 2007; Tennekoon 2017). This underreporting is patterned by race, socioeconomic status, religious affiliation, education, and marital status (Fu et al. 1998; Jones and Forrest 1992; Jones and Kost 2007; Tennekoon 2017). Further, small studies comparing abortion reports to medical records suggest abortion reporting is related to age, time since the abortion, fertility intentions, and attitudes about abortion (Jagannathan 2001; Udry et al. 1996). Unfortunately, the only large-scale reliable data on abortions has limited demographic information (abortion surveillance data from the Centers for Disease Control and Prevention) or asks limited questions (Guttmacher Institute data); and while these data are representative, neither source is longitudinal. Without accurate longitudinal or extensive cross-sectional data on women who have abortions, researchers cannot reasonably make inferences about who has abortions, under what conditions, and with what consequences. This information is vital in the U.S., where abortion is controversial and policy decisions and laws are implemented with scant data.

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a unique data source that may have a more complete reporting of abortions. Given Add Health's collection methods, longitudinal design, and socio-historical collection period, there are reasons to speculate that abortion reporting may be better in this survey than in those previously analyzed. However, to date, researchers have not evaluated the quality of Add Health's abortion data.

The purpose of this paper is twofold. First, this paper seeks to comprehensively analyze the quality of the abortion data in Add Health. The paper uses several approaches to evaluate both the overall quality as well as any patterns of underreporting by race and age at the time of the abortion (gold-standard analyses). The gold-standard analyses first compared the Add Health estimates of the abortion rates to an external "gold standard" rate. While prior studies have used a comparison of counts of abortions, Add Health's cohort design makes such comparisons unwise because of differential experiences of risk. Next, this paper compares Add Health's abortion rate ratios to an external gold-standard. The quality analyses conclude with the use of a binomial-thinned zero-inflated Poisson model developed by Tennekoon (2017) to measure underreporting in the absence of a gold-standard (Poisson model analyses). While these analyses can identify missing data in the aggregate, they offer limited insights into the individuals who are underreporting.

As a result, the second goal of this paper is to gain insights about reporting over time. Because of the longitudinal nature of Add Health, we can observe women's reporting over time. Women who report an abortion at one wave, but not another wave are individuals who have previously underreported. This paper uses multilevel analyses to study these individuals (consistency analyses). The goal of these analyses is to uncover the stable individual (e.g., race) and abortion-related (e.g., marital status at the time of the abortion) characteristics associated with reporting over time. These analyses hold the promise of assisting researchers in developing improved survey methods that could increase reporting by improving our understanding of women who report consistently and inconsistently.

Section 1.1: Background and Literature Review

<u>The Issue and Consequences of Underreporting</u>

Abortion underreporting is a widely-acknowledged limitation of fertility surveys. In fact, all large national surveys in the United States that have been analyzed are inaccurate in estimating abortions. That is, when researchers compare nationally representative survey estimates to data collected by the Centers for Disease Control and Prevention (CDC) or by the Guttmacher Institute, a research and policy institute

focused on reproductive health, surveys consistently fall short. for example, Fu et al. (1998) found that the 1995 National Survey of Family Growth (NSFG) captured only six of ten abortions, despite methodological changes aimed at improving reporting. Similarly, Jones and Kost (2007) found that about half of abortions were unreported in the face-to-face interviews of the 2002 NSFG. Studies on abortion underreporting also identify wide variation in the amount of underreporting across specific years in the survey. for instance, Jones and Kost (2007) found that across 1997-2001 reporting ranged from 42% to 55%. Similarly, Fu et al. (1998) found yearly estimates of abortions capture by the survey ranging from 25% to 60%.

Additionally, two medical record linkage studies have found that women are likely to underreport abortions in survey collection. for instance, in their study of 104 women, (Udry et al. 1996) found that 19% of women surveyed failed to report one or more abortions on a survey compared to their medical records. More recently, (Jagannathan 2001) found women surveyed reported only 29% of abortions that occurred in their medical records. Finally, a recent paper using multivariate modeling techniques, rather than the previous gold-standard approach, found that that the average reporting rate in the NSFG between 2006-2013 was 35.4%.

Underreporting has serious implications for the field. for example, conclusions about the impacts of abortion on short- or long-term outcomes are suspect with underreporting issues, especially when such data are systematically missing, as I discuss below. As a result of this issue, limited insights about who has abortions, under what conditions, and with what consequences can be made from survey research.

The lack of robust and valid research results in, at best, under-informed, and, at worst, misinformed conclusions, which are often used to advocate for specific policies or laws that have far-reaching implications for women in the United States. for example, numerous pro-life organizations draw upon the findings Coleman et al. (2009) to argue that abortions cause long-term psychological distress and, thus, women seeking abortions ought to be subject to mandatory counseling or waiting times. However, such findings are not only contested in the field (e.g., Biggs et al. 2016; Charles et al. 2008; Coleman 2011), but Coleman et al.'s (2009) study, which uses the National Comorbidity Survey, does not

address the impact of reporting on the results. In other words, the study does not discuss how underreporting patterns may indicate that women reporting abortions are qualitatively different from women who do not report. Quantifying abortion underreporting, then, is important in ensuring research findings are appropriately contextualized and understood.

Who Underreports: Patterns in Survey Data & Evidence from Abortion Stigma Research

In addition to the poor overall quality of abortion data in national surveys, patterns of underreporting are also present. Prior research on national surveys and medical record linkage studies provides information about reporting in relation to race, age, marital status, income, religion, education and time since the abortion at the time of the survey. In addition, the medical record linkage studies provide unique information about the relationship between reporting behaviors and women's fertility intentions and attitudes about abortion.

Overall, however, much of the findings in this area are conflicted due to differences in the operationalization of groups, lack of statistical tests of significance, or reports of difficulties concluding statistical significance due to small sample sizes. Moreover, the methodologies of Jones and Kost (2007), Fu et al. (1998), and Jones and Forrest (1992) differ markedly from those of Tennekoon (2017), which may make their findings incomparable. Table 1 summarizes the findings across the six major studies in this area.

The strongest evidence of differential reporting exists for differences by race and age. First, a number of studies suggest that White women have the more accurate reporting rates relative to Black women (Jagannathan 2001; Jones and Forrest 1992; Tennekoon 2017). Similarly, Udry et al. (1996) found Whites had the higher reporting rates than non-White women. The relationship between reporting and Hispanic identification, however, is not clear largely due to different operationalization; some studies compare Hispanic and non-Hispanics, while others compare non-Hispanic Whites/Blacks/Other race and Hispanics (Jones and Kost 2007; Tennekoon 2017). Moreover, Jones and Forrest (1992) found mixed evidence of reporting by Hispanic origin across multiple waves of two national surveys.

Second, nearly all the studies that analyzed the relationship between age and reporting found evidence for or trends suggesting a U-shaped relationship; that is, women in younger and older age groups had better reporting rates than women in the middle of their childbearing years (Fu et al. 1998; Jones and Forrest 1992; Jones and Kost 2007; Tennekoon 2017). Interestingly, however, Jagannathan (2001) found a negative relationship between reporting and age.

In contrast to these clearer trends, the findings regarding marital status, income, religion, education and time since the abortion are mixed. for instance, both Tennekoon (2017) and Jones and Forrest (1992) found unmarried women had more accurate reporting than unmarried women. by contrast, Fu et al. (1998) and Jones and Kost (2007) found trends in the opposite direction, though these findings were not statistically significant or were reported without statistical tests. To make matters more complex, Jagannathan (2001) found no differences in reporting by marital status. Similarly, any potential relationship between income and reporting is muddled. Fu et al. (1998) suggested a positive relationship between reporting and income, while Jones and Kost (2007) found trends supporting a U-shaped relationship. Meanwhile, Tennekoon (2017) found non-linear effects, and Jagannathan (2001) found that receipt of welfare had no impact on reporting. These findings suggest that there may be either no relationship or some kind of non-linear relationship between income and reporting.

Further, Jones and Kost (2007) and Fu et al. (1998) found trends suggesting Protestant women have the highest reporting rates. Meanwhile, Tennekoon (2017) found Catholics had the highest reporting rates. Interestingly, Fu et al. (1998) and Tennekoon (2017) found women in other religions had the lowest reporting rates, while Jones and Kost (2007) found trends suggesting Catholic women had the lowest reporting rates.

Finally, with regard to education, two studies found evidence of a linear relationship (Fu et al. 1998; Udry et al. 1996), while Jones and Kost (2007) suggest a U-shaped relationship. Tennekoon's (2017) study suggests that measurement of this variable may be central to the different findings. for example, Tennekoon (2017) used two educational attainment variables. The first variable was a measure of educational attainment in years, and the second was binary variable comparing more than 12 years of

education to less than 12 years. The continuous measure yielded a positive relationship with reporting, while the categorical variable revealed no relationship. Finally, limited information is available about the role of the time since an abortion and reporting. Udry et al. (1996) found that reporting rates decreased as time since the abortion increased, while Jones and Kost (2007) did not find this effect.

Additional differences between other subgroups have been found in analyses that have not yet been replicated. for example, Jagannathan (2001) found that women who currently want children and women with more restrictive attitudes about abortion at the time of interview were less likely to report prior abortions. However, no previous studies have analyzed these subgroups, so these findings have not been broadly validated.

Abortion stigma research provides greater insights into the mixed or singular findings described above. Specifically, Astbury-Ward, Parry, and Carnwell's (2012) in-depth interviews with women who have had an abortion suggest that abortion stigma functions as a barrier to disclosure. Given this finding, understanding which groups are impacted by abortion stigma may lend support for some of the findings reviewed. for instance, Cockrill et al. (2013) used content analysis of public abortion stories and cognitive testing to generate a scale of abortion stigma. They developed four subscales including worries about judgment, isolation, self-judgement, and community condemnation. They analyzed experiences of stigma across these subscales. Notably, the authors found Catholic and Protestant women experience the most perceived stigma on the full-scale; Black women scored lower than Whites on worries about being judged, but higher on isolation subscales, and women aged 19-24 scored highest on the full scale relative to older women. These findings provide support for the prior findings that Black women are more likely to underreport than White women (Fu et al. 1998; Jones and Kost 2007; Tennekoon 2017). In addition, they provide support for Jones and Kost's (2007) preliminary insights into the relationship between religious affiliation and abortion reporting. Further, they support the age-related findings of Jagannathan (2001).

Overall, prior research suggests that Black women and women in the middle of their childbearing years at the time of an abortion underreport abortions and provide mixed or preliminary evidence to

suggest reporting may be patterned by characteristics such as socioeconomic status, Hispanic identification, religious affiliation, current fertility intentions, and attitudes about abortion.

Section 1.2 Theoretical Framework

While underreporting is clearly stratified in significant ways, there is limited theorizing on these findings. Authors of this research have suggested the following explanations as underlying causes of underreporting: 1) experiences of stigma (Jones and Kost 2007; Astbury-Ward, Parry and Carnwell 2012); 2) feelings of guilt (Jones and Kost 2007); 3) social desirability (Jones and Kost 2007; Astbury-Ward, Parry, and Carnwell 2012); 4) faulty memory (Jones and Kost 2007 & Udry et al. 1996); 5) "cultural, historical and political differences" in cross-country research (Udry et al. 1996: 229); 6) institutional distrust, especially among Black women (Jagannathan 2001); and 7) attitudes about abortion (Jagannathan 2001; Udry et al. 1996).

Research on abortion stigma suggests different experiences of stigma may contribute to the patterns of reporting described (Cockrill et al 2013). However, abortion stigma research does not provide a mechanistic explanation for differing levels of stigma among these groups nor how such experiences might impact reporting. Building upon the prior theoretical mechanisms, I propose a broader theoretical understanding of how differential stigma may be triggered in survey research (Figure 1). Drawing on identity and identity threat theories, I suggest that questions about socio-demographics prime identities of personal importance for respondents of surveys, as evidenced in the stereotype threat literature (e.g., Fine 2010; Logel et al. 2009; Steele, Spencer, and Aronson 2002). Upon reaching questions about prior abortions, the primed identity makes salient particular sociocultural meanings of abortion that lead to either a perceived identity threat or lack of perceived identity threat.

Meanings of abortion that may negatively impact reporting might include abortion as denying or declining motherhood, which may threaten a woman's identity as female[1]. This identity threat may be particularly salient for low-income women because they may perceive fewer alternatives to motherhood (Edin and Kefalas 2011). It may also include sociohistorical meanings such as abortion as genocide against Black bodies, which may particularly threaten Black women's racial identities (Bird and Bogart 2003; Ross 1993). Additionally, sociocultural meanings related to abortion as a sin or intrinsic evil as is common in conservative religious rhetoric, which could threaten religious identities of women in conservative religions (O'Brien 2010). Moreover, the specific circumstances of an abortion, and whether the woman feels it was "justifiable" or not given their sociocultural positioning, may contribute to reporting.

Overall, these internalized meanings of abortion and the conditions under which an abortion occurred may cause particular groups of women, especially those who underreport, to perceive an identity threat and experience self-stigma, fear of judgment, or concerns about privacy breaches that would identify them as having an abortion. As a result, and to preserve their social identity in the context of the social interaction of a survey, women may fail to report a prior abortion. This idea is consistent with research on social identity that suggests that individuals seek to preserve social identities during socialization and that threats to identity can result in lowered self-esteem as well as deviant thoughts and actions, such as lying or cheating (Belmi et al. 2015; Scheepers and Ellemers 2005; Snow and Anderson 1987).

[1] Scholars have described the pressures and expectations that women will be both willing and enthusiastic mothers as compulsory motherhood or the motherhood mandate (e.g., Hertz 2006; Russo 1976). This mandate is so ingrained that some women lack for any explanation for why they want children or to become mothers (McMahon 1995). McMahon (1995) explains: "where childbearing is culturally normative, women do not need to explain why they want children – only why they do not." (p.51). Such a desire for children is so strong that a rise in single-motherhood among middle-class women has been documented, despite strong norms for two-parent households (Hertz 2006). Thus, as Kumar, Hessini, and Mithcelle (2009) explain: "to choose to avert a specific birth, counters prevailing views of women as perpetual life givers and asserts women's moral autonomy in a way that can be deeply threatening. Women, who seek induced abortions, either clandestinely or through established health systems, may be perceived as challenging the inescapability of maternity and defying reproductive physiology (Rylko-Bauer 1996, 480, Erviti, Castro, and Collado 2004, Bradshaw et al. 2008)" (Kumar, Hessini, and Mitchell 2009: 628).

The present study cannot differentiate between simple and protective underreporting distinguished in Figure 1. However, it has the potential to help researchers to understand how individual and abortion-specific characteristics may impact the likelihood of reporting for individuals, which can inform researchers attempting to collect better quality abortion data in the future.

Abortion Reporting and The National Longitudinal Survey of Adolescent to Adult Health: Theoretical Justification

Although the CDC and Guttmacher Institute's aggregate data about abortions provides accurate information about women who obtain abortions in a given year, these estimates do not allow for inferences about causal mechanisms nor the impacts of different outcomes over the life-course. As a result, evaluating the quality of abortion data from surveys, especially longitudinal surveys, continues to be worthwhile. The National Longitudinal Survey of Adolescent to Adult Health (Add Health) is one such survey (Harris 2009). Add Health is a school-based longitudinal study of adolescents in grades 7-12 in the United States in 1994-1995. Four waves of the survey have been completed to date, covering the ages of 12-20 in Wave 1 to 24-33 in Wave 4.

All waves of the survey include questions about pregnancy outcomes, including live births, abortions, and miscarriages. The survey also includes questions on a range of topics such as mental health status, high school grades, drug and sexual risk behaviors, educational attainment, income, work history, social networks, and neighborhood and school characteristics, along with a myriad of additional demographic and health-related data. Thus, depending on the accuracy of abortion reporting, the Add Health study could provide powerful information to improve our understanding of the impact of abortions over the life-course.

Based on the theoretical framework developed in Figure 1, at least three rationales can be put forward for why abortion reporting in Add Health may be more accurate than previous surveys and studies. First, from a data collection standpoint, Add Health interviews were completed using Computer-Assisted Personal Interview or Audio Computer-Assisted Self-Interview, with more sensitive questions, including pregnancy histories, asked in a self-administered portion of the interview (Harris et al. 2009).

Prior research notes the importance of data collection methodology for reporting (e.g., (Fu et al. 1998; Huntington, Mensch, and Toubia 1993; Jagannathan 2001; Lara et al. 2004; Rasch et al. 2000). Specifically, Lara et al. (2004) found that self-administered questionnaires yield higher reporting rates of abortion compared to face-to-face interviews. Fu et al. (1998) also found that self-reported versions of the NSFG yielded more accurate abortion reporting than face-to-face interviews.

In addition to having a self-administered portion of the questionnaire for these data, the collection methods in Add Health are unique relative to other social science surveys that ask about abortion. Add Health's data collection has become increasingly medicalized over time. In all waves, both height and weight of participants were collected, however, in Waves 3 and 4 more intensive biomarker collection was completed. Specifically, beginning in Wave 3 urine and saliva samples were collected. In Wave 3 the collection of this information was directly related to sexually transmitted infections and human immunodeficiency virus (HIV) testing. In Wave 4, saliva and blood specimens were collected as were cardiovascular measures such as blood pressure and pulse. The medicalized nature of the data collection process and the medicalized nature of abortion may reduce the perceived threats of disclosure, such as fear of judgment, thereby improving reporting rates (Cockrill et al. 2013; Halfmann 2011). Importantly, all the biometric data are collected after the interview component of the survey; as a result, we might expect improved abortion reporting starting in Wave 4, given that after Wave 3, the participants might perceive the study to be more medicalized. These two features of Add Health's data collection process may impact reporting by reducing the fear of stigma and judgment.

Second, the longitudinal nature of Add Health may improve reporting (and provides the opportunity to conduct reliability analyses). for instance, pregnancy histories have been collected over the course of four waves, which may allow women who were reluctant to report abortions in early waves to report them at later waves. Two possible mechanisms may underlie improved reporting due to longitudinal design. First, reluctance to disclose an abortion due to mistrust or fear of confidentiality breaches may be assuaged over the course of multiple waves, thereby increasing trust with the survey teams and leading to an increased willingness to disclose a prior abortion. Relatedly, sociological theory

on stigma suggests the pressure for disclosure increases as a relationship lengthens because the threat of being discovered, or "outed," as "discreditable" rises (Astbury-Ward et al. 2012; Cockrill et al. 2013; Goffman 1963). Second, the meaning of an abortion to a woman changes over time (Avalos 1999). That is, a woman who initially feels ashamed or stigmatized following an abortion, which may reduce her willingness to report it on a survey, may come to view the abortion with satisfaction as time passes (Avalos 1999). This change in personal meaning may decrease the identity threat, thereby increasing her likelihood of reporting an abortion in later waves. The longitudinal nature may work through either reducing fear of stigma, judgment, or privacy breaches or through a reduction in the activation of internalized sociocultural meanings of abortion and the subsequent identity threats to improve reporting rates.

Third, the data collection methods and longitudinal nature may work in concert with Add Health's socio-historical period of data collection to improve reporting, which differentiates it from previously analyzed longitudinal surveys. The research on the National Longitudinal Surveys of Work Experience of Youth that identified abortion underreporting was from data collected between 1979 and 1984 when prior abortions may have been obtained illegally (e.g., prior to the Supreme Court decision on Roe v. Wade in 1973). This collection period may have increased underreporting. Furthermore, the women in the Add Health study never lived as adults in the pre-Roe V. Wade era, suggesting they may have qualitatively different ideas about abortion disclosure. In fact, in their study of 14-21-year-olds who had an abortion (Smith, Adler, and Tschann 1999) found these young women were more willing to disclose their abortion than family income, engagement in oral or anal sex, but were less likely to be willing to report their abortion compared to cigarette smoking. This study suggests that adolescents of a similar age and in a similar period as those in Add Health may perceive abortion as less sensitive than other topics. Thus, the sociohistorical data collection period may impact the types of available internalized meanings of abortion.

Section 1.2: Hypotheses

This study aimed to answer the following questions: Is there evidence of abortion underreporting in the Add Health study? If so, what is the extent and pattern of this underreporting? What factors impact individual reporting over time?

Drawing upon prior research and the conceptual framework developed above, the project's specific hypotheses and expected findings are:

Gold-standard analyses

Hypothesis 1: The percentage of the abortion rate estimated by Add Health will be less than the external estimates from the CDC and the Guttmacher Institute, but will capture more abortions than prior surveys.

Hypothesis 2: Underreporting of abortions on Add Health will be patterned by race and age at the time of an abortion[2] in the direction implied by prior research or the theoretical framework:

2A: Black women will underreport more than White women.

2B: Hispanic women will not differ in underreporting relative to White women. Given the mixed findings in this area, this hypothesis is a test of the null.

2C: Underreporting by age will follow a reverse J-shape, with younger women being more likely to report than older women in the sample[3].

Poisson Model Analyses

Hypothesis 3: The Poisson model analyses will follow the patterns described in the gold-standard analyses, with the addition of the following hypotheses[4]:

3A: Underreporting will not differ by marital status. Given the mixed results in this area, this hypothesis will be a test of the null.

[2] It is clear that underreporting may also be patterned by marital status, income, and religious affiliation, however, due to limited external data, these hypotheses cannot be tested using the gold-standard quality approaches (rate and ratio comparisons).

[3] While the literature suggests a U-relationship, the age-structure of the Add Health data prevents us from being able to test for this shape. Instead, we would expect, based upon the Add Health age structure, a reverse-J relationship.

[4] Because the Poisson model analyses do not rely on external data, we are able to test for more of the differences in underreporting.

3B: Women with lower incomes will more likely to underreport than women with higher incomes.

3C: Women who identify with evangelical or conservative religions, such as Protestant and Catholic denominations, will be more likely to underreport than women with other or no religious affiliations.

3D: Women with higher educational attainment will be more likely to report accurately than women with lower educational attainment.

Consistency Analyses

Hypothesis 4: Women who continually report an abortion will be qualitatively different than women who do not report an abortion consistently in individual and abortion-specific characteristics. Specifically:

Individual characteristics

4A: White women will be more likely to consistently report an abortion than Black women.

4B: Women with more positive attitudes about abortion will be more likely to report consistently.

4C: Women with higher early life socioeconomic status will be more likely to report consistently.

Abortion-specific characteristics

4D: Women who are older at the time of the abortion will be more likely to report consistently.

4E: Women who have higher incomes at the time of the abortion will be more likely to report consistently.

4F: There will be no effect of marital status at the time of abortion on the consistency of reporting.

4G: Women who are not religiously affiliated at the time of the abortion will be more likely to report consistently relative to Catholic and Protestant women.

4H: Women with more recent abortions at the time of the survey will be more likely to report consistently.

Hypothesis 5: The association between individual characteristics and reporting consistency will be impacted by abortion-specific characteristics.

Section 1.3: Data and Measures

Data

National Longitudinal Study of Adolescent to Adult Health (Add Health)

Add Health is a school-based longitudinal survey that currently includes four waves. Add Health includes complete pregnancy histories, and a myriad of other demographic, health, and social functioning data. Wave 1 was collected 1994-1995, Wave 2 was collected in 1996, Wave 3 was collected in 2001-2002, and Wave 4 was collected in 2008-2009. The response rates for Waves 1, 2, 3, and 4 were 79%, 88.6%, 77.4%, and 80.3%, respectively (Harris et al. 2009). Add Health used a systematic cluster design with unequal probability of selection. The sample included 80 high schools and 52 middle schools from the United States. The study design has a number of sampling weights that can be used to generate inferences about the population of United States' schools in terms of region of the country, urbanicity, school size, school type, and ethnicity (Harris et al. 2009). Thus, data analyses must be weighted in order to be nationally representative. The present study uses the Wave 4 cross-sectional weights (GSWGT4_2) for the quality and Poisson analyses. The consistency analyses are unweighted due to the structure of the analyses (this decision is discussed in more detail in the sections that follow). The Add Health restricted-use dataset includes 10,486 women at Wave 1, 7,556 women at Wave 2, 7,167 women at Wave 3, and 8,352 women in Wave 4 out of total samples of 20,745, 14,738, 15,197, and 15,701 individuals, respectively.

The gold-standard analyses and Poisson used only the pregnancy histories collected in Wave 4 of Add Health. Known instrumentation issues and changes in data collection over time suggested that analyses of Wave 4 data would provide the most accurate and high-quality pregnancy histories possible. In addition, Wave 4 includes the most number of potential pregnancies as wave 4 is the most recently collected wave of data. Moreover, in-depth analyses of the Wave 4 data would be of most use to researchers aiming to leverage the longitudinal nature of these data. by contrast, the consistency analyses used abortions reported in wave 4 and wave 2. Known instrumentation issues in wave 1 and wave 3

prevent more comprehensive analyses. These pregnancy histories were combined with other data collected across waves to identify individual's characteristics and characteristics of a specific abortion (e.g., religion at the time of an abortion).

External Estimates

Although women underreport abortions on surveys, two reliable estimates of the number of abortions in the United States exist. The first set of counts come from the Centers for Disease Control and Prevention (CDC). The CDC has compiled data on abortion procedures received from state health departments since 1979 (Guttmacher Institute 2011). The CDC requests aggregate data from state health agencies about abortions; data collection is "facilitated by the legal requirement for hospitals, facilities, and physicians to report all abortions to a central health agency" (Pazol et al. 2011:2). This source is subject to some error and is somewhat less reliable than the Guttmacher Institute's data. These estimates are especially unreliable in early years, as states and areas voluntarily report abortions, and early estimates are based on limited data. However, a recent study found the abortion surveillance data published by the CDC has improved over time; the CDC estimates are nearly identical to the sales of mifepristone, a medication used to induce abortions before 7 weeks of gestation[5], indicating this measure likely captures the majority of medical abortions (Pazol, Creanga, and Zane 2012).

The second set of estimates come from two data sources collected by the Guttmacher Institute. First, the Guttmacher Institute surveys all abortion providers in the United States to gather data on the number of abortions provided. Importantly, these data may also be subject to some unquantified error due to unknown abortion providers, "non-response, and responses based on estimates rather than records" (Fu et al. 1998:130). Second, the Guttmacher Institute has conducted four nationally representative surveys of women who have had abortions. The surveys were fielded during the following time frames: 1988; 1994-1995; 2000-2001; and 2008. These data have been used by researchers in conjunction with the estimates from the abortion provider census and the CDC counts to estimate the number of abortions obtained by

[5] A large majority of abortions occur within this time frame.

subgroups of women (e.g., women with different religious affiliations, incomes, educational attainment, and ethnicities). This study used the Guttmacher Institute's surveys from 1994-1995, 2000-2001 and 2008. These surveys are known as 2008 Survey of Abortion Patients, the 2000-2001 National Survey of Contraceptive Use Among Women Having Abortions, and the 1994-1995 National Survey of Contraceptive Use Among Women Having Abortions, respectively (Jones 2008; Jones, Darroch, and Henshaw 2004; Henshaw 1997). While Add Health data from 2008 was not used in this paper, the Guttmacher Institute's survey from 2008 was used for interpolation purposes. Following prior research, both the CDC and Guttmacher Institute's data were used to generate comparison estimates in this study.

In order to generate external estimates of abortion rates and ratios, two additional external sources were used. First, for the abortion rates, age and sex-specific population counts were gathered from the United States Census website for each of the study years (United States Census Bureau 2016). Second, the number of births by age and race was obtained using the Natality Data from the National Vital Statistics of the National Center for Health Statistics for each of the study years (National Center for Health Statistics and The National Bureau of Economic Research 2016).

Measures from Add Health

Below is a detailed description of the measures that were used from Add Health (all measures are summarized in Table 2).

Outcome Variables

Abortions, Abortion Rates, and Abortion Ratios

The gold-standard analyses used data on abortions and births from Wave 4 of Add Health. In Wave 4, participants filled out a table of pregnancies, and were asked: "How did this pregnancy end?" The following response options were given: 1) Pregnancy has not yet ended; 2) live birth; 3) single, stillbirth; 4) miscarriage; 5) multiple, involving both a live birth and another outcome; 6) multiple, no live birth; and 7) abortion. Participants were also asked to provide the year and month for each pregnancy outcome. Using responses to this question, a total number of abortions and births were compiled for each

of the years under study (1994-2007). These counts were used to generate abortion rates and rate ratios, as described in the analytic methods section below.

The Poisson analyses use the pregnancy histories provided in Wave 4 as well. However, the analyses include women who reported no abortions in the modeling process, and the outcome variable is a count of abortions, rather than related to a ratio. The number of births between 1994 and 2007 was used as a covariate in the Poisson analyses. More details on this process are provided in the analytic methods section.

for the consistency of reporting analyses, waves 2 and 4 of the survey were used to assess reporting of abortions over time. Unlike other longitudinal surveys, Add Health asked for complete pregnancy histories in multiple waves of the survey (in Waves 3 and 4). Additionally, data from Wave 1 and 2 asked about prior pregnancies in 1994 and 1994-1996, respectively. In Wave 1 and Wave 2, participants were asked if they had even been pregnant. They were then asked for each pregnancy reported, "How did this pregnancy end?" The participants were given the following response choices in Wave 1: 1) It has not ended; you are still pregnant; 2) a live birth; 3) stillbirth or miscarriage; 4) an abortion. In Wave 2, the following response options were provided: 1) It has not ended; you are still pregnant; 2) a live birth; 3) stillbirth; 4) miscarriage; and 5) an abortion. Women were also asked to provide the month and year of each pregnancy outcome. These pregnancy histories were to trace reporting of abortions reported in Wave 4 and Wave 2.

Consistency of Abortion Reporting

The present paper analyzed the consistency of abortions reported in Wave 2 and Wave 4 using a binary outcome variable, measured at the abortion-level (not a person-level). This consistency variable serves as the outcome variable for the consistency analyses. An abortion was coded as consistently reported if it occurred between 1994 and 1996 and was reported in both wave 2 and wave 4 or if an

abortion occurred in or after 2003 (after wave 3 data collection) and was reported in wave 4[6]. An abortion was coded as 1 if it was reported consistently and a 0 if reported inconsistently. Abortions that occurred to the same individual during the same year and month are considered the same abortion. In the event that the month of an abortion was missing over two reports, abortions reported in the same year were considered the same abortion. The consistency of reporting is the dependent variable for the consistency analyses.

Independent Variables

Race/Ethnicity

Add Health collected information about race and ethnicity from participants at Wave 1 and Wave 3. Participants were asked, "What is your race?" And were able to check as many race categories as applied to them. Possible answers included: white; Black or African American; Asian or Pacific Islander; American Indian or Native American; and other. Participants were also asked, "Are you of Hispanic or Spanish/Latino origin?" Participants that indicated multiple races were also asked, "Which one category best describes your racial background?" Responses included: white; Black or African American; Asian or Pacific Islander; American Indian or Native American; and other.

The present study used the race and ethnicity reported in Wave 1 unless these data were missing. When these were missing in Wave 1, later waves were used to determine the correct classification for participants. Individuals who identified as only one race were coded that race. Individuals who identified as Hispanic alone or Hispanic and any other race(s) were coded, "Hispanic." This coding was consistent with the coding used by the Guttmacher Institute's datasets. Individuals who selected more than one race were coded with the single-best race question above. If an individual selected "other" in the single-best race question or had non-response to this question and identified as two or more races, they were coded as "Other/Multiracial," and were excluded from comparisons with external counts as no such category exists

[6]Individuals who report an abortion as occurring after 2003 are necessarily only able to report once. As a result, they may not consistently report in the future. However, these individuals were included in order to conserve statistical power. Without these individuals included, there would be insufficient power to run the multi-level analyses required for these analyses.

in the external data. The gold-standard analyses used the following categories for analysis: 1) non-Hispanic White, 2) non-Hispanic Black; 3) non-Hispanic Native American; 4) non-Hispanic Asian or pacific Islander; 5) Hispanic; and 6) Other or multiracial The Poisson analyses used a simpler race variable following prior research, which included: 1) non-Hispanic White, 2) non-Hispanic Black; 3) non-Hispanic Other; and 4) Hispanic (Tennekoon 2017).

The consistency analyses also used this simplified race category in order to preserve statistical power, while still contributing substantively to the key research issues in the field.

Age at the time of abortion

Month and year of the participant's birth were collected at all waves of Add Health. Although both month and year of an abortion were collected, data on the month of abortion was frequently missing (approximately 10% of months were missing in Wave 4). To maximize the sample size, an approximate age at abortion was calculated using the year of respondent birth and year of the abortion for the gold-standard and consistency analyses.

Age at Wave 4

Participant's age at the time of the survey was also calculated using the participant's date of birth and survey date for the Poisson model analyses. Following prior research, the age at the time of Wave 4 survey was collapsed into the following categorical variable ages 24-29 and ages 30-34.

Marital status at time of abortion & marital history at Wave 4

Marital status was not used in the gold-standard analyses due to difficulty accurately identifying married person-years with the Add Health data. for the Poisson analyses, marital history at the time of Wave 4 was used to approximate the marital history variables used by Tennekoon (2017). Marital history was measured using the following question from Wave 4: "How many persons have you ever married? Be sure to include your current spouse if you are married now." Women who listed 0 marital partners were coded as never being married at Wave 4, all others were coded as having been married at least once.

for the consistency analyses, the marital status at the time of the abortion was used to understand factors associated with consistency of reporting. In Wave 4, participants were asked about each

pregnancy, “At the time you got pregnant, were you and he married to each other?” This question was used to create a binary variable indicating whether a woman was married at the time of an abortion. Women who had an abortion and were coded as “legitimate skip” were coded as unmarried. This coding of marital status at the time of the abortion was also used for the quality and consistency analyses.

Income at the time of abortion & income at Wave 4

The gold-standard analyses do not include an income evaluation because external data are not available to generate an abortion rate or ratio. by contrast, the Poisson analyses analyze the impact of income at the time of survey on underreporting. Following Tennekoon (2017), income reported at wave 4 (reported in whole dollars) was translated into a percentage of the poverty limit and entered into the model as a categorical variable that included the categories: Below 100% of the poverty limit, 100%-199%, and 200% and more of the poverty limit. This percentage was calculated using the 2007 federal poverty limit guidelines and the respondent’s report of the number of people other than themselves in their household (U.S. Department of Health and Human Services 2007). Tennekoon (2017) used five categories, but limited sample sizes led to the use of the three categories described.

The consistency analyses aim to understand the impact of income at the time of an abortion on reporting. Because Add Health does not include information about income for each individual year, prior waves were used to estimate income at the time of an abortion; abortions that occurred between 2005 and 2007 used income reported at wave 4, abortions that occurred between 2000 and 2004 used income reported in wave 3, abortions that occurred between 1996 and 1999 used personal income reported at wave 2, and abortions that occurred earlier than 1996 used wave 1 data. Two alternative coding methods were also used. Further details about these alternative coding schemes can be found in Appendix 1.

Differences were observed across the coding processes. As a result, the analyses of income and underreporting will be analyzed using all three methods, though only the coding scheme described above is reported in the results section as analyses of the three approaches did not substantively impact the

results (see Appendix 1). for all of the coding processes, missing data issues were attenuated by going back one wave prior if the preferred data were missing[7].

Due to instrument changes across waves, several different questions were to measure income at the time proceeding an abortion. In Wave 1 and Wave 2, participants were asked two questions that will be used to calculate an estimate of yearly income. First, participants were asked, "How much money do you earn in a typical non-summer week from all your jobs combined?" Second, participants were asked, "How much money do you earn in a typical summer week from all your jobs combined?" Both questions allowed for a numerical response (e.g., not a range). Assuming a traditional academic calendar, summer weekly income will be multiplied by 12 and the remainder of weeks in the year (40) will be considered non-summer weeks In Wave 3 participants were asked, "Including all the income sources you reported above, what was your total personal income before taxes in {2000/2001}? Please include all of the income sources you identified in the previous question," Participants gave their income in whole dollars. Finally, in Wave 4 participants were asked "Now think about your personal earnings. In {2006/2007/2008}, how much income did you receive from personal earnings before taxes, that is, wages or salaries, including tips, bonuses, and overtime pay, and income from self-employment?" Participants gave their income in whole dollars.

for the consistency of reporting a transformed whole-dollar income variable was used. Given the skewed distribution of this raw variable (likely due to the age of the sample), the data was inverse hyperbolic sine transformed for analyses to better approximate normality. This transformation was used in favor of a log or natural log transformation to ensure individuals that reported zero income were not dropped.

Reported religious denomination at time of abortion & religion at Wave 4

The Poisson model and consistency analyses included measures of religious affiliation. In the Poisson analyses, participants reported religious denomination in Wave 4 was used. by contrast, religion

[7] for example, under this coding scheme if an abortion occurred in 2002-2006 and no income data was reported in Wave 3, wave 2 data was used. If no wave 2 data were available, the income was considered unknown.

at the time of the abortion was approximated for the consistency analyses. In order to best estimate a woman's religious affiliation prior to or at the time of an abortion, the same "best guess" coding scheme described above was used. That is, abortions that occurred between 2005 and 2008 used religion reported at wave 4, abortions that occurred between 2000 and 2004 used religion reported in wave 3, abortions that occurred between 1996 and 1999 used personal religion reported at wave 2, and abortions that occurred earlier than 1996 used wave 1 data.

In some cases, interviewees in Wave 4 had not been interviewed at Wave 2 or Wave 3. In these cases, the last reported religious affiliation prior to the abortion was used. This variable was used for the consistency analyses. An alternative coding scheme using the temporal sequencing coding scheme was also tested, but no significant differences were found in the distributions.

Add Health participants were asked their religious denomination in each wave of the Add Health study. Specifically, participants were asked, "What is your religion?" Responses to this question were numerous and were categorized into groups based on the available categories from the external datasets (Protestant, Catholic, other, or none).

Time since the abortion at first survey

The month and year of survey interview were recorded for each wave. This information along with the year of the abortion was used to generate a categorical variable approximating the time between the abortion and the wave at which women had the first opportunity for reporting (wave 2 or wave 4, depending on the year at which the abortion occurred) for the consistency analyses. In order to address the potential issue of one-way causation (that is, the fact that women who were further away from a past abortion could only potentially report it at Wave 4 due to the age structure of the data), only two categories were used to describe time since abortion, including 1) one-year or less, and 2) more than one-year past.

Early Life Attitudes about Abortion

The consistency analyses included the impact of early life attitudes about abortion on reporting over time. In Wave 2 of the survey, all participants were asked about the morality of having an abortion

and the likelihood of having one (or encouraging a partner to have one) if faced with a pregnancy[8]. Participants were asked two questions pertaining to their attitudes about abortion. First, they were asked: "Do you agree or disagree with the following statement? You are morally opposed to abortions." Second, they were asked: "Do you agree or disagree with the following statement? [If R is male:] If you got someone pregnant, you would encourage the girl to get an abortion. [If R is female:] If you got pregnant, you would consider getting an abortion." Possible answers to both questions included strongly agree, agree, neither agree nor disagree, disagree and strongly disagree. The two scales were added to represent an index of abortion attitudes, where a higher number is indicative of more negative attitudes about abortion.

Early Life Socioeconomic Status

Early life socioeconomic status was measured with a proxy variable about early life welfare receipt, and was used for the consistency analyses only. In Wave 4, participants were asked:" Before you were 18 years old, did anyone in your household ever receive public assistance, welfare payments, or food stamps?" Responses included Yes, No, Refused, legitimate skip, and don't know. All legitimate skips were recoded to "No."

Educational Attainment at Wave 4

Educational attainment at Wave 4 was used in the Poisson analyses following prior research (Tennekoon 2017). In Wave 4 participants were asked, "What is the highest level of education that you have achieved to date?" Following prior research, the twelve categories used in Add Health were collapsed to a binary variable for the Poisson analyses. The categories include less than a college degree and a college degree or more.

[8] It is possible that some women in the sample may have already had an abortion prior to the administration of this question. That is, women who reported an abortion in Wave 1 or women who had an abortion prior to 1996 may be influenced by their experience of an abortion in their responses to this question. However, given the small number of abortions expected to occur to women in this age group, the effect of this bias is likely minimal.

Exposure to Risk for Abortion – Years Sexually Active Between 1994-2007

A critical component of a Poisson regression is the exposure to risk. Tennekoon (2017) used years of sexual activity as their exposure to risk. Similarly, the models used in this paper used the number of years a woman was sexually active between 1994 and 2007 (matching the outcome variable for these analyses).

The exposure to risk was calculated using a woman's birth date, age at wave 4, and the following questions regarding sexual activity and menarche: "How old were you the first time you ever had vaginal intercourse?" And "How old were you when you had your very first menstrual period?" Only women who answered in the affirmative to the following question in Wave 4 were included in these analyses: "Have you ever had vaginal intercourse? (Vaginal intercourse is when a man inserts his penis into a woman's vagina.)." Women who had intercourse prior to menarche did not contribute to the sexually active risk years until after menarche if reported.

Measures from External Estimates

External Estimates of the Number of Abortions Pooled and by Individual Year

The counts of the total number of abortions for the study period years come from published information from the Guttmacher Institute's census of abortion providers (Finer and Henshaw 2003; Jones and Kavanaugh 2011). Following prior research, I used the CDC's age distribution for each year to calculate the expected number of abortions that occurred in the age distribution of Add Health for each year, except in the years when an age distribution was collected by the Guttmacher Institute's survey of abortion patients (Kost and Jones 2007; Fu et al. 1998). Because the CDC data are provided in age ranges (under 15; 15-19; 20-24; 25-29; 30-34; 35-39; 40+), the Add Health data was restricted to best match these ranges (see Appendix 2 for further details).

for each individual year (e.g., 1994, 2001, etc.) External estimates were age-matched to Add Health. for example, in 2005 women in the Add Health sample would be between 21 and 30. Women 30 and over were excluded from the analyses and the comparison group was aged-matched based upon the

CDC estimates for women aged 20-24 and 25-29. Spline and linear interpolated age estimates from the Guttmacher Institute's survey of abortion patients were compared with the estimates used to test the robustness of these analyses (see Appendix 3A and Appendix 3B). Only one significant difference was observed for the interpolation methods, in the year 2006, though the degree of difference is still minimal considering the 95% confidence intervals. Overall, these external estimates were treated as known population counts, despite the fact that some unquantified error does exist.

Importantly, the Add Health estimates and person-year estimates retain this age matching over time. for instance, in 1994 the external estimates (and Add Health counts) include only women between 15 and 21. When compiling a total number of abortions, this subset of abortions obtained is added to the next subset (e.g., in 1995 women 15 to 22). As a result, the total number of abortions obtained between 1994 and 2007 retains the age structure of Add Health as closely as possible throughout.

External Estimates of the Number of Abortions by Subgroup

The subgroup analyses were estimated following the methods of Jones and Kost (2007) and Fu et al. (1998). These authors estimated the subgroup abortion counts by using 1) the proportions of abortions obtained by women in different groups derived from the Guttmacher Institute's nationally representative surveys of abortion patients and 2) the estimated total number of abortions from the Guttmacher Institute's census of abortion providers. for example, if the Guttmacher Institute's census were to indicate that there were a million abortions in a given year, and the survey of abortion patients indicated that 25% of women obtaining abortions identified as Catholic, the authors would estimate that a quarter of a million abortions were obtained by Catholic women, and compare the data being analyzed for underreporting to this figure.

The present paper followed this procedure. for the years between the Guttmacher study years, linear or spline interpolation was used to estimate the proportions of abortions obtained by different sub-groups. Multiple modes of interpolation were tested to ensure best-fit. The interpolated group proportions were further adjusted to reflect the age of Add Health participants in the given year.

for example, to generate the external estimate of the number of abortions obtained in 1994-2007 by Black women ($Estimate_{1994\text{-}2007Black}$) several steps occurred:

1) Identify the total number of abortions obtained between 1994-2007 ($T_{94\text{-}07}$) using the Guttmacher Institute's published counts.
2) Calculate the proportion of abortions obtained by Black women in 1994-2007 ($P_{black94\text{-}07}$) using the interpolated estimates from the Guttmacher Institute's survey of abortion patients.
3) Calculate the proportion of abortions obtained by women in the Add Health age range in 1994-2007 using age-range counts provided by the CDC ($p_{addheathage}$). Importantly, the external count does not include all women 15-29 for all years between 1994-2007, instead, it includes only the age ranges captured by Add Health bounded between 15-29 in each of the individual years.
4) Adjust the total number of abortions reported in 1998 by age and race:

$$Estimate_{1994\text{-}2007Black} = T_{94\text{-}07} * P_{black94\text{-}07} * p_{addheathage}$$

External Estimates of Person-Years Lived and Births

for the external abortion rates, person-years were calculated following the methods described below, drawing upon data from the United States Census (United States Census Bureau, 2016). for the external abortion ratios, the number of live births to women by individual ages were obtained from the Natality Data from the National Vital Statistics of the National Center for Health Statistics for each year under study (National Center for Health Statistics & The National Bureau of Economic Research, 2016).

Section 1.4 Sample Restrictions and Missing Data

The data was restricted in several ways for the gold-standard, Poisson and consistency analyses. for the gold-standard analyses, the sample was restricted in three ways. First, only responses from female participants were included. Second, only abortions reported as occurring between 1994 and 2007 were included; the final year of Wave 4 (2008) was omitted from these analyses as collection reflects only a portion of a year. Third, age-restrictions were imposed each year to match available data (see Appendix 2 for further details). The use of an age restriction in each year was made to facilitate the creation of

appropriate external estimates. As a result, the final sample size for the gold-standard analyses was 1,019 women with 1,460 abortions. The Poisson analyses were limited to women and abortions between 1994 and 2007 and respondents who reported being sexually active with non-missing data on predictor variables (N=7,375). The Poisson analyses include women of any age at the time of the abortion.

for the consistency analyses, the sample was restricted in five ways. First, the present study only compares reporting of abortions in wave 4 to abortions reported in wave 2; abortions reported in waves 1 and 3 are excluded from these analyses due to known instrumentation issues. Second, only women interviewed in both wave 2 and wave 4 are included in the analyses. Third, only abortions reported by women have been included for analysis. Fourth, only abortions that were reported with a year are included in the analyses. Finally, only women with non-missing data on all variables are included in these analyses. The final sample represents 829 women with 1,248 abortions.

An analysis of missing data for the consistency analyses revealed 96% of the cases were complete for the sample. Missing data were primarily limited to refusals or non-response to questions about income or religion at the time of an abortion. List-wise deletion of the 49 cases with missing data was used due to the centrality of these variables, difficulty identifying appropriate predictors for imputation, and the small number of missing cases.

As a result of these different data restrictions, three distinct abortion-related variables were used. The gold-standard analyses use a yearly and an overall rate and rate-ratio for abortions that occurred between 1994 and 2007 to an age-restricted group (see description in Appendix 2). The Poisson modeling analysis uses the number of abortions obtained by women of any age between the years of 1994 and 2007. Finally, the consistency analyses measured the consistency of abortions with a reported year to women who were in both Wave 2 and wave 4.

Section 1.5: Analytic Methods

<u>Gold-standard analyses</u>

The gold-standard analyses were conducted in three parts. First, abortion rates for each year under study (1994-2007) and across years were computed for Add Health and the external estimates. Abortion rates were computed using the following standard formula:

1.1 $$\frac{\textbf{\textit{Number of Abortions Obtained by Women ages } x-x+n \textbf{\textit{ in Time T}}}}{\textbf{\textit{Person Years Lived by Women } x-x+n \textbf{\textit{ in Time T}}} \; \boldsymbol{X\ 1000}$$

Person-years were calculated for Add Health for each individual in each year. for example, in 1994, Add Health included abortions obtained by 15 to 21-year-olds. A woman who turned 15 in June of 1994, would contribute .5 person-years, while a woman who was at least 15 and younger than 21, would contribute a whole person-year. for the external estimates, mid-year population, from the U.S. Census, of women in the age-group matched to Add Health were used (e.g., the mid-year population of women 15-21 in 1994). To calculate the total person-years in the restricted Add Health sample, the number of person-years contributed to each individual year (1994-2007) were summed. As a result, the total person-years lived 1994-2007 is age-matched relative to the ages included in each year under study. The number of abortions and person-years lived were weighted using the sample weights, and then used to compute the rates.

Second, abortion rate ratios for the years 1994 through 2007 were estimated for Add Health (Equation 1.2) and a standard abortion ratio was calculated for the external data (Equation 1.3). The Add Health abortions, person-years, and births were weighted to create weighted abortion rate ratios. The following formula defines an abortion rate ratio and an abortion ratio:

1.2 $$\frac{\frac{\textbf{\textit{Number of Abortions Obtained by Women ages } x-x+n \textbf{\textit{ in Time T}}}}{\textbf{\textit{Person Years Lived by Women } x-x+n \textbf{\textit{ in Time T}}}}{\frac{\textbf{\textit{Number of Live Births to women ages } x-x+n \textbf{\textit{ in Time T}}}}{\textbf{\textit{PErson}}-\textbf{\textit{Years Lived by Women inTime T}}}} \; \boldsymbol{X\ 1000}$$

1.3 $$\frac{\textbf{\textit{Number of Abortions Obtained by Women ages } x-x+n \textbf{\textit{ in Time T}}}}{\textbf{\textit{Number of Live Births to Women ages } x-x+n \textbf{\textit{ in Time T}}}} \; \boldsymbol{X\ 1000}$$

The final step of the gold-standard analyses were the subgroup analyses. The subgroup analyses evaluated underreporting by age and race/ethnicity using the rate-ratio approach described above. The rate-ratio approach was used because external data on the number of live births by age and race are available to compute external estimates. Given that the subdivisions of data resulted in small sample sizes, these analyses used the pooled years of abortion data rather than a year-by-year approach.

Poisson Model Analyses

Given the limited sub-group analyses that can be conducted using a rate or rate-ratio approach due to lack of external data, I used a recently published procedure aimed at conducting abortion underreporting analyses without the use of a gold-standard. The new methodology, described in detail in Tennekoon (2017), uses a binomial-thinned zero-inflated Poisson model to estimate 1) the likelihood that a report of 0 abortions was incorrect/excessive, 2) the predicted number of abortions to an individual, and 3) the reporting rates. The approach can maximize the available data since women who reported zero abortions are included in the modeling. It is important to note that these models are highly sensitive to the parameters included in the three stages. In fact, the various robustness models provided by Tennekoon (2017) yield reporting rates between 25.82% and 67.85%. As a result, the final model used in this paper was parameterized to match Tennekoon's (2017) preferred model as closely as possible. Although the impact of the variables on the number of reported abortions and excess zeros are included in Appendix 4 (Table B and C), these findings will not be reported in this paper as they are beyond its scope. Additionally, Appendix 4 (Table A) provides details about the additional measures used for the count and excess zero models. All the Poisson models were weighted using the existing Add Health sample weights and Stata 14's svy estimation suite.

Consistency Analyses

All of the consistency analyses were conducted with multi-level logistic regression models. Importantly, the sampling weights available for Add Health were not used for these analyses. The weights were not used because the analyses were conducted at the abortion, rather than individual-level. Currently, no sampling weights for this level of analysis are available. In order to address this issue, two

levels were used to account for the clustering of errors in the dataset. The first level was an individual level to account for abortions obtained by the same woman. The second level was the school level to account for correlated errors due to sampling design. The paper presents bivariate and multivariate, multi-level logistic regression models. All of the tables for the consistency analyses provide exponentiated betas and standard errors for the fixed effects. Random effects for school and individual are reported in the tables provided, but not interpreted in the results as these factors are not of central interest in the present study.

Section 1.6: Results

Gold-standard analyses

Descriptive Statistics

Before implementing any sample restrictions, wave 4 of Add Health included a total of 1,833 abortions reported by 1,219 unique women. When abortions were limited to women of any age between 1994 and 2007, 1,584 abortions reported by 1,084 women remained. Once abortions were limited to women between the ages of 15 and 29 (irrespective of the age-matching procedures) between 1994 and 2007, a total of 1,506 abortions reported by 1,041 abortions were reported. Finally, with all the sample restrictions in place, a total of 1,460 abortions were reported by 1,019 women. Specifically, the final sample includes abortions obtained between 1994 and 2007 to women age-matched for each year of Add Health. The unweighted summary statistics for each of these iterations are shown in Table 3. Table 3 includes the features of 1) all abortions reported in wave 4; 2) abortions reported between 1994-2007 to women of any age; 3) abortions reported between 1994-2007 to women 15-29 at the time of abortion; and 4) abortions reported between 1994-2007 using yearly age-restrictions described in Appendix 2. The fourth category is used for the following analyses.

In the fully restricted sample, the number of abortions ranged from 0 to 17. The majority of women in the subsample had not had an abortion (87%). of those who had at least one abortion, 71% reported a single abortion, 21% reported two abortions, 6% reported three abortions, and 2% reported four

or more abortions. In the restricted sample women were, by definition, between 15 and 29 at the time of the abortion, while in the fully unrestricted sample age abortion ranged from 8 to 31. The unweighted descriptive statistics showed that the majority of abortions were reported by non-Hispanic Black women (40%), followed by non-Hispanic whites, Hispanics, Asian and Pacific islanders, Native American, and women of other races (34%, 19%, 5%, .8% and .7%, respectively). Finally, the majority of abortions were reported by women who were between 20 and 24 at the time of their abortion (46%).

Overall Reporting and Reporting by Year, Abortion Rates

The overall Add Health abortion rate for 1994 through 2007 captures 35.3% of the external estimate [95% CI: 29.7%-41.0%]. Table 4 shows the individual yearly capture rate, and Figure 2 compares this capture rate to the capture rate of the rate-ratio method. The highest and lowest capture rates were observed in the last and first year of the study period. In 2007 Add Health captured a high of 50.9% of the external rate [95% CI: 35.5%-66.2%], while the Add Health rate in 1994 captured 20.2% of the external rate [95% CI: 8.2%-32.2%]. Importantly, these estimates are age-specific. for instance, in the year 2000, Add Health captures 36.2% of expected abortions to women 17-27. Moreover, the overall capture rate does not reflect all abortions from 1994-2007 but reflects the proportion of the external rate age-matched to the Add Health data for each year understudy.

The percentage captured varies somewhat from year to year, and there may be a slight upward trend over time. In four of the fourteen years examined, Add Health captured between 20 and 29 percent of the external estimate (1994, 1995, 1996, 1997). In six of the fourteen years, Add Health captured between 30 and 39 percent of the external estimate (1998, 2000, 2001, 2003, 2005, and 2006). In four of the years, Add Health captured more than 40 percent of the external estimate (1999, 2002, 2004, 2007). The results of these analyses suggest that Add Health did not out-perform prior studies, inconsistent with hypothesis 1.

Overall Reporting and Reporting by Year, Abortion Rate Ratios (Method 1)

The abortion rate ratios provide more information about the quality of these data (Table 5, Figure 2). Again, these data are age-specific. The weighted Add Health estimates suggest a population ratio of

164.7 abortions per 1,000 live births between 1994 to 2007 to women age-matched to the individual years in Add Health. This rate ratio captures 42.9% [95% CI: 35.5% - 50.3%] of the external ratio (376 abortions per 1,000 live births). The year-by-year the findings mirror the ebb and flow in accuracy observed in the rate analyses. Using this approach, the percentage of external estimates captured had a high of 74.9% [95% CI: 25.3% - 124.4%] in 1994, though the 95% confidence interval reveals the volatility of this figure. The lowest percentage captured using this method was observed in 2004 at 32.5% [95% CI: 21.6% - 43.4%]. In nine of the years, the Add Health ratio estimates captured less than 50% of the external estimates (1996, 1997, 1998, 1999, 2000, 2001, 2003, 2005, 2006). There does not appear to be a pattern by year in these data. In all the years analyzed, the capture rate is not statistically different from the capture rate observed using the rate method. These findings do not support hypothesis 1.

Subgroup Analyses by Race and Age

The prior section addressed the overall accuracy of the Add Health data overall and by year. The present section uses the pooled years to evaluate underreporting of abortions across race and age. The central hypotheses are provided at the beginning of each subsection. The analyses are summarized in Table 6, Figure 3, and Figure 4.

Race

2A: Black women will underreport more than White women.

2B: Hispanic women will not differ in underreporting relative to White women. Given the mixed findings in this area, this hypothesis is a test of the null.

The race-stratified findings indicate that there was no difference in reporting rates between White and Black women in the Add Health sample. The analyses of the pooled data reveal that Add Health captures approximately 39.9% of external abortion ratio for White women [95% CI: 33.2% - 46.6%] and 36.6% of the external ratio for Black women [95% CI: 23.5%-49.7%]. A conservative approach, comparing the 95% confidence intervals, shows an overlap between the two intervals, suggesting the estimates are not significantly different from one another. Thus, we reject hypothesis 2A, in favor of the null.

Further, there were no differences in reporting between Hispanic women and White women. The Add Health dataset captures 63.3% [95% CI: 42.2% - 69.1%] of the external ratio for Hispanic women. The 95% confidence intervals for Hispanic and White women overlap, indicating reporting is not significantly different between these two groups. As a result, we fail to reject the null hypothesis proposed in 2B. Overall, we find no significant differences in race reporting across the groups included in our analyses with appropriate external reference groups. However, the variability of the estimates varies widely, with large intervals for Native American and Asian women, due to small counts in the sample.

Age

2C: Underreporting by age will follow a reverse J-shape, with younger women being more likely to report than older women in the sample

The rate-ratio analyses revealed no significant differences in the percentage of the external ratio captured by age-group. While the trend seems to mirror the expected reverse J-shape, the 95% confidence intervals overlap across all three age groups analyzed. Specifically, Add Health captured 58.4% [95% CI: 47.6%-69.1%] of the external abortion ratio for women 15-19, 58.3% [95% CI: 45.7%-68.9%] of the abortion ratio for 20-24-year-olds, and 40.0% [95% CI: 29.4%-50.7%] of the external abortion ratio for women 25-29. These findings do not support hypothesis 2C, though they are trending in the direction expected.

Poisson Analyses

The Poisson analyses for underreporting are reported in Tables 7-9. The findings from the excess zero and count analyses can be found in Appendix 4 (Tables B and C). The hypotheses for the Poisson analyses were as follows: *Hypothesis 3:* The Poisson model analyses will follow the patterns described in the gold-standard analyses, with the addition of the following hypotheses:

2A: Black women will underreport more than White women.

2B: Hispanic women will not differ in underreporting relative to White women. Given the mixed findings in this area, this hypothesis is a test of the null.

2C: Underreporting by age will follow a reverse J-shape, with younger women being more likely to report than older women in the sample.

3A: Underreporting will not differ by marital status. Given the mixed results in this area, this hypothesis will be a test of the null.

3B: Women with lower incomes will more likely to underreport than women with higher incomes.

3C: Women who identify with evangelical or conservative religions, such as Protestant and Catholic denominations, will be more likely to underreport than women with other or no religious affiliations.

3D: Women with higher educational attainment will be more likely to report accurately than women with lower educational attainment.

Descriptive Statistics

Table 7 provides the summary and descriptive statistics for the variables used in the Poisson analyses. In the weighted sample, the mean number of abortions reported was 0.18 (*SE*=0.01), while the mean number of births was 1.05 (*SE*=0.04*)*. On average, women in the sample had been exposed to the risk of pregnancy for 10.91 years (*SE*=0.10). Whites comprised the majority of the weighted and unweighted sample (55% and 68%, respectively) as did women aged 24-29 at the time of the survey (72%). About half of the women in the sample had been married at least once (57%) at wave 4. Approximately a third of the women fell into each of the three income levels, while the majority had less than a college education (66%). The sample was primarily comprised of heterosexual (or straight) individuals and those who had worked full-time.

Reporting Rates Overall and by Group

Overall, the Poisson analyses found that Add Health captures 20.3% of abortions with a standard error of .59 for the years of 1994 to 2007 (Table 8). This finding is similar, though statistically less than other methods employed. This finding provides evidence against hypothesis 1 and hypothesis 3; Add Health reporting rates are not better than prior surveys.

The unconditional reporting rates reported in Table 8, revealed differences from the gold-standard analyses. The mean reporting rate for Hispanic women was 54.4% (*SE*=2.76), followed by 25.5% (*SE*=4.38) for women of other races, 24.0% (*SE*=0.97) for Black women, and 17.3%(*SE*=0.56) for White women. These findings suggest that Whites had significantly lower reporting rates than Black women. Moreover, the results show that Hispanic women have significantly higher reporting rates than women in all other groups. These findings do not support hypotheses 2A (which predicts White women have higher rates of reporting) or 2B (which predicts that Hispanic and White women will have the same reporting rates).

The reporting rate estimates further show a significant difference in unconditional reporting rates by marital history. Women who have never been married had higher average reporting rates than women who had been married at least once (*M*=26.9%, *SE*=0.64, compared with *M*=15.2, *SE*=0.58 for women who had been married at least once). This finding does not support the null hypothesis proposed in hypothesis 3A.

The relationship with income and reporting shows that reporting improves as income increases. The mean reporting rate for women in the lowest income category was 17.8% (*SE=0.63*) compared 19.9% (*SE*=0.88) for women in the middle-income category, and 23.2% (*SE*=0.93) for women in the highest income category. Although the two confidence intervals for the low and middle-income groups overlap, an F-test revealed that the mean reporting rates were, in fact, statistically different (*F*=63.32, *p*=.01). These findings provide support for hypothesis 3B; women with lower incomes are more likely to underreport than women with higher incomes.

Further, the analyses show that women in the other religion category had the highest reporting rates (*M*=75.3%, *SE*=.73*)*, followed by Catholic women (*M*=31.0%, *SE*=.78*)*, women without a religion (*M*=23.8%, *SE*=.56*)*, and Protestant women (*M*=6.3%, *SE*=.30*)*, and all four groups were statistically different from one another. These findings provide mixed support for hypothesis 3C. On the one hand, these analyses do suggest underreporting is substantial among Protestants, however, it does not support this finding among Catholics.

Finally, no significant differences were observed by age at the survey nor education at wave 4. These null findings do not support hypothesis 2C nor 3D.

Although these unconditional reporting rates show statistical differences, the conditional likelihood of reporting revealed no significant differences in reporting across groups (Table 9). That is, when the features of underreporting are modeled and controlled for, no differences in the likelihood of reporting is observed.

Consistency of Abortion Reporting

The following section reports the findings from the consistency analyses. The central hypothesis for these analyses was: women who continually report an abortion will be qualitatively different than women who do not report an abortion consistently in individual and abortion-specific characteristics. Table 10 and Table 11 report the unweighted descriptive statistics of the sample used for the consistency analyses. Importantly, this data is presented at the abortion, not the person, level. Table 12 provides the bivariate models, and Table 13 provides the multivariate model findings.

Descriptive Statistics

of the total 1,248 unique abortions in the sample, 36% were reported consistently. That is 36% of abortions reported by participants were either reported in both Wave 2 and Wave 4 or occurred after Wave 3 data collection and reported in Wave 4. of the 1,248 abortions, 35% were obtained by Whites, 39% by Blacks, 19% by Hispanics, and 6% by other race or multiracial women. A large majority of the participants had never received welfare prior to age 18 (94%). by contrast, the mean income at the time of the abortion was $10,117.79 (*SD*=$13,647.44). Additionally, the majority of abortions were obtained by women who identified as Protestant (53%) and unmarried (93%) at the time of the abortion. Further, 23% of the abortions were obtained by women who had had an abortion within a year or less between their abortion and the first opportunity for reporting. The average age at which women had abortions was 21.7 (SD=3.87), with a range of 13-31. Finally, the mean score on the abortions attitude scale was 6.16 (SD=1.37) with a range of 2-10.

Bivariate Analyses

Table 12 presents the findings from bivariate multilevel models analyzing the relationship between the consistency variable and individual and abortion-specific characteristics.

Individual characteristics

4A: White women will be more likely to consistently report an abortion than Black women.

4B: Women with more positive attitudes about abortion will be more likely to report consistently.

4C: Women with higher early life socioeconomic status will be more likely to report consistently.

A significant association between consistency of reporting and race was found (Table 12, Model 1), providing evidence against hypothesis 4A. Specifically, holding random effects of person and school constant, Black women were 47% more likely to report consistently compared with White women. No other significant associations among the race categories were observed. Similarly, neither the relationship consistency of reporting and the abortion attitude scale (Table 12, Model 2) nor early life socioeconomic status (Table 12, Model 3) was significant, providing no evidence in support of hypothesis 4B or 4C.

Abortion-specific characteristics

4D: Women who are older at the time of the abortion will be more likely to report consistently.

4E: Women who have higher incomes at the time of the abortion will be more likely to report consistently.

4F: There will be no effect of marital status at the time of abortion on the consistency of reporting.

4G: Women who are not religiously affiliated at the time of the abortion will be more likely to report consistently relative to Catholic and Protestant women.

4H: Women with more recent abortions at the time of the survey will be more likely to report consistently.

With regard to the abortion-specific characteristics, several significant associations were found. First, holding random effects of person and school constant, age at abortion showed a strong association

with the consistency of reporting status (Table 12, Model 4), consistent with hypothesis 4D. A one unit increase in age at the time of the abortion increased the likelihood of consistent reporting by 2.78 times, and this effect is exponential in nature. Second, holding random effects constant, a one unit increase in the inverse hyperbolic sine transformed dollar increases the likelihood of consistent reporting by 27% (Table 12, Model 5). Though not directly translatable to a dollar figure, it clearly indicates that as income at the time of the abortion rises, the likelihood of consistent reporting also rises. This finding provides evidence in support of hypothesis 4E.

Third, after controlling for random effects, there was no significant difference in the likelihood of reporting between women who were Catholic and Protestant nor women with no religion and Protestants at the time of their abortion (Table 10, Model 6). However, women who identified as an "Other" religion were 73% more likely than Protestant women to report consistently, however, this finding was only significant at the .10-level. This finding does not support hypothesis 4G. Fourth, women who were married at the time of the abortion were 2.30 times as likely to report the abortion consistently than unmarried women (Table 12, Model 7), which is inconsistent with hypothesis 4F. Finally, holding random effects of school and person constant, women who had an abortion more than a year prior to being surveyed at wave 2 or wave 4 were 39% less likely to report consistently than those who had the abortion within a year or less (Table 12, Model 8) – a finding consistent with hypothesis 4H.

Multivariate Analyses

Table 13 presents the individual characteristics, abortion-specific, and joint multivariate models.

Model A: Individual characteristics

Model A, which incorporates all three individual characteristics, showed almost no changes in the magnitude of the relationships observed in the bivariate models. Specifically, I found that holding all other variables constant and holding the random effects of school and individual constant, Black women were 48% more likely to report consistently than White women. No additional race categories, nor the abortion scale nor the early socioeconomic status indicator were significantly associated with the consistency of reporting. These findings suggest that these factors do not mediate one another.

Model B: Abortion-specific characteristics

Model B, the abortion-specific characteristics model, showed several differences when compared to the bivariate models. After controlling for random effects of school and individual as well as the fixed effects of income, religion, marital status and time since the abortion to reporting, the effect of age at abortion increases substantially when compared with the bivariate model (Table 12, Model 4). Specifically, each year is associated with an increase in the likelihood of consistent reporting by 6.34 times (an odds ratio of 7.34 versus 3.78 in the bivariate model). The relationship between consistency and time since the abortion report also increased dramatically from model 5 to model B. Specifically, holding all fixed effect variables and random effects constant, women who had an abortion more than a year prior to the survey were 99.9% less likely to report consistently than those who were surveyed within a year or less of the abortion. The previous significant relationships observed between religion, income, and marital status and consistency of reporting were no longer statistically significant in the multivariate model.

Model C: Joint Model

Model C combined the individual and abortion-specific characteristic models. Several interesting changes occurred as a result. First, the significant association between consistency of reporting and race observed in model 5 and model A was eliminated, suggesting hypothesis 4A be rejected in the multivariate framework. Second, a significant association between consistency of reporting and the abortion attitudes scale was uncovered, in support of hypothesis 4C. Specifically, each point toward more negative views about abortion decreased the likelihood of consistent reporting by 26%, holding all else constant. Third, the effect age on consistency was slightly attenuated, though remained strong moving from an odds ratio of 7.34 in model B to 7.12 in Model C, consistent with hypothesis 4D. Finally, no change in the magnitude of the significant association between time since abortion and consistency was found when moving from model B to model C, which provides continued support for hypothesis 4H. That is, holding the random effects and fixed effects of individual characteristics and other abortion characteristics, women who had an abortion more than a year prior to the survey were still 99.9% less likely to report consistently than those who were surveyed within a year or less of the abortion. No further

evidence was found to support hypothesis 4C or 4E, but support for 4F was found as no relationship between consistency of reporting and marital status was found.

Diagnostics

for the bivariate analyses, LR-tests comparing the bivariate model to the null model showed significant improvements in fit for all of the bivariate models. for the multivariate analyses, two sets of LR-tests were performed. First, LR-tests between the multivariate model and the null model showed significant improvements in fit across all models. Second, a series of LR tests were used to compare models A and B to Model C. The LR-tests revealed 1) the joint model (C) was a significant improvement in fit than the individual characteristics model (A) ($\chi2$=1084.08, p<.001), but 2) the joint model (C) did not significantly improve the fit of the abortion-specific characteristics (B) ($\chi2$=8.38, p=.14).

Section 1.7: Limitations

This paper has several important limitations. The first set of limitations is related to the methods used in this paper. To begin with, this paper was unable to use the count comparison methodology used in other papers (e.g., Jones and Forrest 1992) because of the longitudinal design and age-structure of the Add Health dataset. As a result, this study used rates and rate-ratios to account for the differential exposure experienced by women in Add Health. Although these approaches better match the Add Health design, they are not without fault and are still period measures. The differences, challenges, and approaches to addressing age, period, and cohort effects are well documented elsewhere (e.g., Hobcraft, Menken, and Preston 1982; Preston et al. 2001; Yang and Land 2013).

While the Poisson regression had the potential to address such issues, the overall reporting rate is slightly lower than anticipated when compared with the other findings. However, the use of an abortion measure with unrestricted ages may have contributed to these differences. Moreover, the Poisson analyses are sensitive to the variables included in each stage of the analysis. As a result, it might be that a model more consistent with these analyses could be found, but that was not the primary task of this paper. Further, not all the iterations of the model described in Tennekoon (2017) could be tested here due to

different collinearity of variables and smaller sample sizes. In addition to these limitations, the present study relied heavily on interpolation to estimate the external estimates due to the age distribution of the Add Health sample. Several different interpolation methods were used to ensure robustness, but due to limited availability of data, these estimates may be inaccurate or mismatched in important ways.

The second set of limitations is related to the availability and quality of the available external data. First, due to limited availability of external data, this paper is unable to make comparisons between Add Health and all known areas of underreporting. However, the consistency analyses of this paper aim to better uncover factors, such as attitudes about abortion, for which no external data exists, but the prior evidence is available (Jagannathan 2001). Second, the present study relied on sample data about abortion patients to identify the number of abortions obtained nationally by different subgroups. Although this approach has been used in prior literature, it is still potentially subject to error. However, to address this concern, I used the weights designed by the study authors in order to address possible biases in response rates. Third, the external estimates are likely subject to error, which is unaccounted for in the present study. However, no alternative to this particular issue is forthcoming.

The final set of limitations is related to the consistency analyses conducted in this paper. While this approach allows us to learn about reporting over time, the known instrumentation issues limited the analyses to wave 2 and wave 4. This restriction resulted in small sample sizes, which results in less sensitive analyses due to limited statistical power. Further, the analyses did not use sampling weights to address the impact of survey design on the findings. Thus, the results should be treated as somewhat preliminary. Despite these known limitations, this paper still provides the first comprehensive analysis of abortion reporting in the Add Health study as well as providing insights into abortion reporting over time.

Section 1.8: Discussion

This paper began with the question: is it possible for survey research to gather accurate abortion histories from women in the United States? Reliable and accurate survey data on abortions could provide information about who has abortions, under what circumstances, and with what short and long-term consequences. In the United States, abortion regulations and restrictions have been climbing in recent

years and changes political leadership will likely increase these efforts (Nash et al. 2015). However, these policies are largely implemented with limited empirical research. As a result, this study sought to 1) evaluate the quality of the abortion data in the National Longitudinal Study of Adolescent to Adult Health because it has unique features that could have theoretically improved reporting rates; and 2) to use the longitudinal nature of this survey to better understand which individuals are likely to consistently or inconsistently report. The goals of this work were to evaluate a survey with potentially more accurate data on abortions that could be used to better understand the impact of abortions across the life course and to advance our understanding of the factors that impact reporting over time, which could improve future research design and data collection.

Overall, this paper found that Add Health demonstrated underreporting consistent with other nationally representative surveys. Using an abortion rate approach, Add Health captured approximately 35% of the expected external abortion rate from an age-matched population. The abortion rate-ratio captured approximately 42% of the expected abortion ratio, though this figure is not statistically different from the rate method. The overall Add Health capture rate was on the lower end of the spectrum of prior findings but was not unprecedented. for instance, Jagannathan (2001) found women only reported 29% of abortions, while some prior waves and years of the NSFG captured 25% of estimates (Jones and Forrest 1992). It is unclear, however, if the results of this study are comparable to prior research because prior research has not employed the rate or rate-ratio methods. Moreover, the Add Health accuracy varied across years. This finding is consistent with analyses of other large national surveys that have found that the degree of underreporting varied widely from year-to-year (Fu et al. 1998; Jones and Forrest 1992; Jones and Kost 2007).

There are several potential reasons why Add Health did not perform as expected. First, the age-structure of the data and the need to create imperfectly matched external estimates may have caused distortions in the capture rate. Alternatively, returning to the theoretical framework, it may be that the use of computer-assisted interviews does not reduce the experience of an identity threat sufficiently to increase reporting. While Fu et al. (1998) found self-reporting on the NSFG improved reporting (by about

10%). That is, the self-reporting method still only captured 50% of abortions. In fact, Lara et al. (2004) found that using a random response technique (an indirect survey method) outperformed self-reporting in the accuracy of recall about prior abortions. Moreover, Tennekoon (2017) analyzed recent waves of the NSFG that uses this collection methodology and concluded only 35% of abortions were reported. Second, the theoretical mechanism underlying the longitudinal design may not function as theorized. I theorized that trust built with the survey team would increase trust and potentially increase reporting (Goffman 1963). However, it may be the changing of interviewers/data collection personnel prevents the usual functioning of this mechanism. Alternatively, collecting lifetime pregnancy histories may simply result in a faulty recall, as hypothesized by Udry et al. (1996).

Third, it is likely that the sociohistorical setting did not have a substantial impact on the internalized meanings of abortion for respondents in the way theorized. Although Smith, Alder, and Tschann's (1999) research suggested potential differences among a population similar to the Add Health sample, both the Pew Research Center and Gallop Poll have shown attitudes about the legality of abortion has remained relatively stagnant, if divided, since the late 1970s (Newport, Moore, and Saad 1999.; Pew Research Center 2016). Moreover, from the numerous evaluations of the NSFG, it seems that reporting of abortions may be declining over time. for instance, Jones and Kost (2007) found that reporting in the 2002 NSFG was 47% and Tennekoon (2017) found the reporting rate for the NSFG surveys between 2006 and 2013 was only 35%. This trend may suggest that we are currently in a sociohistorical time period when accurate reporting is less likely to occur. It's possible one, all, or some combination of these theoretical failures resulted in the underperformance of Add Health relative to expectations.

In addition to generalized underreporting, there was also mixed evidence of underreporting among different groups in Add Health. This component of the discussion first reviews the findings from the gold-standard analyses, followed by the Poisson analyses, and the consistency analyses in light of one another and the literature. Theoretical implications of these findings are taken up in the next section.

To begin with, the gold-standard analyses of Add Health showed no evidence of underreporting by Black women, in contrast with some findings from prior research (Fu et al. 1998; Jagannathan 2001;

Jones and Kost 2007; Tennekoon 2017). An intriguing possible reason for this finding is the fact that Wave 4 of Add Health did not ask about race. That is, from a theoretical perspective, race may not have been primed as explicitly, thereby reducing the effect of the identity threat on reporting, per the theoretical framework presented.

Second, the gold-standard analyses did not find any statistically significant differences in reporting by age at abortion. However, the downward trend observed is consistent with the expected reverse-J shape expected based upon prior research (Fu et al. 1998; Jones and Forrest 1992; Jones and Kost 2007; Tennekoon 2017). Thus, the findings of the study provide some weak support for this observed pattern, though cannot provide strong support due to lack of statistical significance. Generally, the gold-standard analyses suggest that while underreporting in Add Health exists, the patterns observed vary from prior research. Prior research has primarily focused on the National Survey of Family Growth, which may have particular emphases that Add Health does not. That is, the prior findings may reflect the similarity in salience or priming in a particular survey rather than a static nature of underreporting.

The goal of the Poisson analyses was to use a non-period methodology to better gauge the quality of the Add Health data. However, as mentioned in the limitations section, the model's sensitivity to additional and varied parameters makes the results volatile. In addition, the results may not be directly comparable to the measures of religion, income, marital history, age, and educational attainment are all Wave 4 measures, rather than measures at the time of the abortion.

The patterns observed in the Poisson analyses did vary somewhat from the patterns in the gold-standard analyses. First, the Poisson analyses found that Black women had improved reporting relative to White women. This finding is different both from the gold-standard analyses and prior research. However, it may be possible that the lack of priming resulted in improved reporting among this group, which has higher rates of abortion in general (Jones and Kavanaugh 2011). Second, the Poisson model found no effect of age at wave 4 on reporting rates. This finding is not directly comparable to prior research nor the gold-standard findings as it is a contemporary measure of age. Moreover, the findings are

difficult to compare with Tennekoon's (2017) study as the age ranges in the Add Health sample are narrower than those used in Tennekoon's (2017) analyses.

In addition to these findings, the Poisson model provided unique information about reporting rates across different subgroups. Consistent with the theoretical framework, the unconditional mean reporting rate was highest among women with the highest incomes at the time of the survey. This finding is consistent with the trends reported by Fu et al. (1998) but differs from the other findings in this area (Jones and Kost 2007; Tennekoon 2017). The theoretical framework, however, suggested that women in lower income categories may be particularly sensitive to identity threats to their identity as female. Thus, these findings provide some support for the theoretical framework proposed.

Similarly, the Poisson model provided mixed support regarding the role of religious affiliation on reporting rates. The model found that Protestant women had the lowest rates of reporting, which is consistent with the idea that these women experience high levels of stigma (Cockrill et al. 2013). However, these very low reporting rates among Protestants is inconsistent with other research that finds their rates are highest (Fu et al. 1998; Jones and Kost 2007). Moreover, Catholic women did not show this same trend, despite indicators from prior research and the theoretical framework proposed (Cockrill et al. 2013; Jones and Kost 2007). The findings on this topic remain ambiguous. It may be that the observed patterns are the result of the simplified categories used for religious affiliation. for instance, the Poisson model found that women with an "other" religion had the highest rates of reporting. However, this group is likely quite heterogeneous. As a result, this finding doesn't provide much information about the underlying mechanisms that may link religious affiliation with reporting.

Additionally, the Poisson model revealed that women who had never been married had higher reporting rates than women who had been married at least once. This finding was consistent with the work of Jones and Forrest (1992) and Tennekoon (2017) but inconsistent with other studies that suggested the opposite pattern (Fu et al. 1998; Jones and Kost 2007). Importantly, the two studies that suggested the opposite pattern by marital status showed no statistical tests or found only trends (rather than significant differences). In other words, the Poisson model findings regarding marital status are consistent with

studies that report and make statistical inferences about group differences. Moreover, this finding is potentially consistent with the experience of stigma related to unwed motherhood, which will be discussed in the theoretical implications section below (Bock 2000; Ellison 2003).

Finally, the Poisson model provided information about the role of educational attainment at Wave 4 on reporting. The model found no significant effect of educational attainment. This finding essentially splits the differences among prior research which finds positive, negative, and U-shaped effects of education (Fu et al. 1998; Jones and Kost 2007; Tennekoon 2017). Further, based upon the findings of Tennekoon (2017) it seems that the coding of the education variable makes a substantive difference as they found a difference in reporting not based upon years of education but between a binary education variable. This difference would be challenging to test using the Add Health data, which does not include a continuous years of education variable. In contrast to the significant findings reported, the conditional model estimating the likelihood of reporting revealed no significant differences in reporting across any subgroup, suggesting that controlling for various factors eliminates group differences in reporting. In sum, the Poisson model did differ somewhat from the gold-standard analyses and provided new insights into reporting by subgroups both unconditionally and conditionally.

The objective of the consistency analyses was to better understand the factors associated with individual reporting over time. These women are necessarily women who have reported at least once. However, they vary in the consistency of their reporting. Consistent reporters offer researchers insights into ways to improve reporting, while inconsistent reporters could be viewed as women who might underreport, in that they underreported at least once.

The consistency analyses contributed unique information about consistent reporters. for instance, more positive attitudes about abortion were associated with consistent reporting in the multivariate analyses, but not the bivariate analyses. The multivariate finding is consistent with Jagannathan's (2001) findings related to attitudes. However, because only multivariate analyses on this topic were included in Jagannathan's (2001) paper, it is unknown if the results are truly analogous. Further, these analyses showed age at the time of the abortion and time since the abortion to first survey had strong effects in the

predicted directions, consistent with prior findings (Jagannathan 2001; Udry et al. 1996). The interaction and changing effects observed when the individual and abortion-characteristic models were combined also strongly suggest that the two interact with one another in substantive ways. Interestingly, these results are consistent with the elimination of underreporting differences by subgroup observed in the conditional Poisson Model. The elimination of the effects of race, income, religion, and marital status at the time of the abortion suggests that these factors are being moderated through women's age at the time of the abortion, the timing of the interview, and early life attitudes about abortion. That is, these two abortion-specific and one individual characteristic change the degree to which women are willing to report consistently. It seems likely that the attitudinal component may be central as a significant effect is revealed when incorporated into the full model. However, the precise nature of these relationships requires further study

Taken all together the consistency findings suggest that for some subset of women who underreport, their early life abortion attitudes, age at abortion, and timing of survey moderate the relationship between both individual and abortion-specific characteristics and consistency of reporting. In other words, the decision to report is related to the totality of the circumstances of the abortion, especially the woman's age, her attitudes, and timing of the survey, contribute to reporting. This idea is consistent with the theoretical framework developed and will be discussed further in the following section. Moreover, the results are largely consistent with those from the medical record linkage surveys (Udry et al. 1996; Jagannathan 2001).

In summary, the findings from the gold-standard and Poisson analyses suggest Add Health's unique features did not prevent extensive underreporting of abortions, though the precise extent of this underreporting varies by subgroup, year, and across methods used to analyze these data. In addition, the consistency analyses revealed that consistency of reporting is influenced strongly by women's age at abortion, the timing of data collection relative to the abortion, and early life attitudes about abortion, which is potentially useful information for future researchers as will be discussed in the conclusion.

Section 1.9: Theoretical Implications

The findings of this paper have added complexity and new evidence on the nature of underreporting of abortions. Several of the findings were consistent with the proposed theoretical framework. for example, no differences in underreporting by race were found in the gold-standard analyses, and race was not asked in wave 4 of the survey. Moreover, the Poisson model showed that Black women, who have higher rates of abortion to begin with, had higher reporting rates than White women. These two pieces of information may imply that priming may function as predicted.

In addition, the higher reporting rates among women in the highest income group is consistent with the sociocultural meanings mechanisms proposed. I theorized in the beginning of this paper that women who have lower incomes may be more sensitive to threats to their identity as female, given that they may perceive fewer alternatives to motherhood (Edin and Kefalas 2011). The reporting rate patterns observed are consistent with the theorized mechanism. Further, the interaction between individual and abortion-specific characteristics in the consistency analyses suggest that the personal identities activated may be impacted by both sociocultural meanings of abortion (e.g., attitudes about abortion) and the circumstances of the abortion (e.g., age at the time of abortion).

Despite these findings, several of the theorized relationships did not function as expected. for instance, Add Health did not perform as well as expected. In addition, unanticipated relationships, such as the relationship between underreporting and age (null finding), religion, and never-married status have important implications for theory. Thus, at least three such implications can be suggested: 1) the types of socio-cultural meanings that may be activated and by whom needs to be further refined, 2) the type of stigma experienced may be more important in understanding underreporting than generalized experiences of stigma, and 3) the role of individual meaning of an abortion needs to be better incorporated into the framework. Each of these implications will be briefly discussed in light of the findings.

To begin with, the unanticipated findings related to differences in reporting by marital status and religion suggest the sociocultural meanings of abortion proposed requires further work. Prior research has been mixed on its findings regarding the relationship between marital status and underreporting. This

study tested the null hypothesis but found that never-married women had higher reporting rates than women who had been married at least once. The mixed findings in the literature may be the result of multiple types of meanings and identities being activated. for example, some women in Add Health were fairly young at Wave 4. It may be that women who had been married at least once at Wave 4 are also more likely to be a member of a conservative religious group, as individuals who grew up as conservative Protestants, Mormons, and other religions are more likely to marry earlier than mainline Protestants, though evidence in this area is somewhat conflicted (Lehrer 2004; Uecker and Stokes 2008). Consequently, it may be that overlapping identities lead to the unexpected patterns observed.

In addition, it may be that unmarried women experience less stigma for having an abortion than married women. This could function in several ways. First, women who are unmarried may experience more stigma in the face of a pregnancy (Bock 2000; Ellison 2003). Thus, unmarried women may face less stigmatization for having an abortion than married women. Alternatively, it may be that married women experience a unique stigma for having an abortion. Marriage is associated with increased economic and social resources (e.g., Lupton and Smith P. 2003; Ross, Mirowsky, and Goldsteen 1990; Umberson 1992; Waite and Lehrer 2003). Married women, then, may feel that their reasons for pursuing an abortion are not acceptable or will not be considered acceptable by others, resulting in reduced reporting rates. Overall, these conclusions suggest that unmarried women may be considered "unfit" mothers. In other words, the sociocultural meanings internalized by these women may be more positively associated with choosing abortion rather than parenting. However, women considered "fit" mothers may experience increased identity threats and underreport.

Second, the null findings related to reporting by age and the mixed findings related to religion suggest that the amount of overall stigma perceived by a specific group may not be the primary mechanism responsible for underreporting. for example, the Poisson analyses found Protestant women had high rates of underreporting, while Catholics had moderate rates. by contrast, Cockrill et al. (2013) found increased perceptions of stigma among both Catholic and Protestant woman. However, Cockrill et al.'s (2013) study also found that only Protestant women had elevated risk for increased scores on the

worries about judgment stigma subscale. Similarly, while younger women experience more perceived stigma overall, there are no differences between 19-24-year-old women and any other group of women except those over 40 on the worries about judgment subscale (Cockrill et al. 2013). Moreover, although no differences were observed between races on the overall stigma scale, White women did experience significantly more worries about judgment than Black women. Thus, it might be that concern about judgment is central to underreporting, rather than overall experiences of stigma. Given these insights, it might be that the nature of the perceived stigma should be incorporated into the theoretical model more fully. Specifically, the findings imply that it is not simply high levels of stigma that cause protective underreporting, but it is the experience of concerns about judgment that triggers protective underreporting.

Finally, the consistency analyses identified a unique subset of women who underreport abortions. for these women, being younger at their time of the abortion, being interviewed more than a year after the abortion, and having negative attitudes about abortion makes them more likely to have underreported at least once. From a theoretical perspective, this finding has two potential explanations. These inconsistent reporters are women who either 1) reported accurately in wave 2 and inaccurately in wave 4 or 2) reported inaccurately in wave 2 and accurately in wave 4. The first scenario could be evidence of simple underreporting; that is, women who had an early life abortion and recalled it at a time period closer to the abortion despite negative attitudes about abortion early in life, but did not recall it at a later time. However, it could also be evidence of the changing meaning of abortion over time. It may be that the experience of having the abortion deepened negative attitudes or feelings about the abortion that led to protective underreporting when surveyed later.

The second scenario also demonstrates the possibility of the meaning of abortion changing over time. for example, if a woman does not report the abortion in wave 2, but does in wave 4, her first decision not to report the abortion may be protective underreporting, considering the early life negative attitudes about abortion. Meanwhile, over time her attitudes about abortion or feelings about her abortion might have changed, resulting in an accurate report in the second survey. While prior research has shown

the meaning of abortion changes over time, this personal meaning is not well incorporated into the theoretical model (Avalos 1999). As a result, the consistency analyses suggest that the theoretical model should be modified to include both sociocultural meanings about abortion as well as the specific personal meanings a woman attributes to abortion. The most logical placement for this feature is in tandem with the activation of internalized sociocultural meanings of abortion as it seems likely that the two meaning structures inform and shape one another. A revised framework is presented in Figure 5.

Although the theoretical framework proposed did not function exactly as anticipated, the model does provide a useful way of considering, theorizing about, and identifying areas where survey research could intervene in order to improve reporting.

Section 2.0: Conclusions, Implications for Practice, and Future Directions

This paper contributes to our understanding of abortion reporting in two ways. First, in characterizing the abortion reporting in Add Health, we have observed that obtaining accurate information about abortions is still a real challenge to researchers. It is worth debating whether or not such approaches are truly useful, given their expense, at this current sociological period of time in the United States. However, airing toward optimism, the consistency analyses of this paper do suggest potential paths forward. Specifically, this paper has identified features associated with consistent and inconsistent reporting. Women who report inconsistently are willing to report an abortion under some circumstances. Inconsistent reporters may be a subset of underreporters, as they have underreported at least once. Thus, inconsistent reporters represent a potential group where the design and collection of a cross-sectional or longitudinal survey could be improved to capture more accurate reports of prior abortions. Moreover, this paper has used and compared three types of methodologies to analyze underreporting in a large nationally representative survey.

From the consistency findings, I suggest three specific recommendations for researchers hoping to obtain more accurate information about abortion histories. First, it may be fruitful to obtain accurate information about abortions from women who have just experienced this event. Researchers may want to work with women's health clinics that provide abortions to make survey requests at more opportune

times. Second, collection efforts should acknowledge and incorporate the role of different types of stigma in data collection. It may be that fear of judgment or engagement in self-judgement contributes to underreporting. As a result, the use of an empathetic setting, evidenced to improve self-reports of abortion in research by Rasch et al. (2000), may help improve abortion reporting. Moreover, it may be useful to consider adding an open response box where women can (if they desire) provide a rationale or explanation for their decision to have an abortion. Finally, careful attention to the role of identity priming and survey design could be helpful in improving reporting. for instance, moving demographic and socioeconomic questions after abortion histories may improve reporting by minimizing the potential of increasing identity salience on factors such as race, class, or religion.

In addition to efforts to improve reporting, this paper offers insights into methodologies for evaluating abortion underreporting. Evaluations of underreporting remain an important indicator of the quality of the data researchers collect. Indeed, this paper shows that despite careful design and theoretical advantages, our best surveys can be fraught with underreporting of abortions. Therefore, evaluation of data before use is necessary.

Researchers seeking to complete abortion underreporting evaluations should consider 1) the goal of their analyses and 2) the structure of data being evaluated to pick the most appropriate methodology. for instance, researchers seeking to compare information across surveys and time should use the traditional gold-standard count approach because it has been widely used by prior research. However, this approach is not always appropriate for cohort studies as seen in this study. Under these circumstances, researchers hoping to evaluate one survey internally would find a rate or rate-ratio approach most useful. by contrast, researchers most interested in understanding the individual likelihood of underreporting would be best served by Tennekoon's (2017) modeling approach.

The choice of methodology should also be related to the structure of the sample. The use of the gold-standard methods (count, rate, and ratio) is complicated for longitudinal surveys or surveys that do not include women throughout their reproductive years. As a result, it may be advantageous for such studies to use Tennekoon's (2017) method. Similarly, researchers should also consider the importance of

the survey weights when selecting a method. Some studies require the use of weighting due to under-sampling or over-sampling of groups. However, the use of weights which are not designed for these purposes may give too much weight to a specific individual's report (or lack thereof). The gold-standard methods, and especially the count method, relies heavily on the accuracy of the weighting processes. Thus, researchers need to consider how the underlying representativeness of their sample without weights may impact the accuracy of their estimates.

Even with these initial findings and recommendations, it is clear further research is needed in this area. Future research should aim to better identify how individual and abortion-specific characteristics interact with one another. The use of interactive models and stepwise entry may be helpful in these endeavors. Additionally, research extending the findings of Cockrill et al. (2013) to abortion-specific characteristics and experiences of stigma may be useful in improving our understanding the types of stigma that impact reporting. This information, in turn, could enable researchers to better proactively address reporting issues as well as better contextualizing findings. Another area of potential research could be investigating the how a survey's structural components (e.g., ordering of questions, medicalized nature or not, the period of collection), may independently impact reporting of abortions. Pursuing these lines of research could help fill in the missing cases that characterize fertility surveys. Ultimately, improving reporting and data collection about abortion has the potential to allow thoughtful empirical work to contribute to the political, policy, and personal debates about reproductive health in the United States.

TABLE 1: SUMMARY OF LITERATURE

	Jones and Forrest (1992)	Fu et al. (1998)	Jones & Kost (2007)	Tennekoon (2017)	Udry et al. (1996)	Jagannathan (2001)
Notes	*Includes multiple surveys and waves of surveys*	*No significance tests*	*Findings below are not statistically significant, trends only*	*Use of modeling*	*Medical Record Linkage Study*	*Medical Record Linkage Study*
Race	Whits>Non-White ---------- Hispanic - Mixed	Other>White> Black ---------- *No Hispanic*	Other>White> Black> Hispanic	White=Other> Black ---------- Hispanic>Non-Hispanic	Whites>Non-Whites ---------- *No Hispanic*	White> Hispanic> Black ---------- *No Other*
Age	U-Shape-Mixed	U-Shape	U-Shape	U-shape		Increased Age→ Increased Reporting
Marital Status	Unmarried> Married	Married> Unmarried	Married> Unmarried	Never Married> Married= Cohabitating> Previously Married		Unmarried= Married
Income		Increased Income → Increased Reporting	U-Shape	Non-linear effects		No effect of welfare receipt
Religion		Protestant> None> Catholic> Other	Protestant> None> Other> Catholic	Catholic> None =Protestant> Other		
Education		Increased Education → Reduced Reporting	U-Shape	Mixed	Increased Education → Reduced Reporting	
Attitudes About Abortion						Increased Positive Attitudes → Increased Reporting
Time Since Abortion			No Effect		Increased Time Since Abortion → Reduced Reporting	

NOTES: Empty cells indicate untested relationships.

TABLE 2: DESCRIPTION AND SUMMARY OF ADD HEALTH MEASURES

Variable	Gold-standard analyses	Poisson Model Analyses	Consistency Analyses
Abortions and Births	Wave 4 Count and Population Estimations (Converted to Rates and Rate-Ratios)	Wave 4 Counts and Estimations at the Individual Level (Dependent Variable)	Wave 2 and Wave 4 Counts
Consistency of Reporting	N/A	N/A	Dichotomous: *1= Consistent reports between wave 2 and wave 4 and/or consistent report for abortions reported after 2002 in Wave 4*
Race/Ethnicity	Categorical *White, Black, Asian/Pacific Islander, American Indian/Native American, Other, Hispanic*	Categorical *White, Black, Hispanic, Other/Multiracial*	Categorical (Individual Characteristic) *White, Black, Hispanic, Other/Multiracial*
Age	Categorical *15-19, 20-24, 25-29* *At time of Abortion*	Categorical *0=24-29, 1=30-34* *At time of survey*	Continuous (Abortion-specific characteristic) *At time of Abortion*
Marital Status/Marital History	*N/A*	Ever married at wave 4 *1=Yes*	Dichotomous (Abortion-specific Characteristic) *1=Married*
Income	*N/A – No External Estimate Available*	Continuous *Inverse hyperbolic sine transformed dollars*	Categorical *Under 100% Federal, Poverty Line, 100-199% FPL, and 200%+ FPL*
Religious Affiliation	*N/A – No External Estimate Available*	Categorical *Protestant, Catholic, Other, None*	Categorical *Protestant, Catholic, Other, None*
Time Since Abortion at Opportunity for First Report	N/A	N/A	Dichotomous (Abortion-specific Characteristic) *0= One year or less* *1= More than one-year past*
Attitudes about Abortion	N/A	N/A	Continuous (Individual Characteristic)
Early Life Socioeconomic Status	N/A	N/A	Dichotomous, *Ever on welfare?* *1=Yes*
Education at Wave 4	N/A	Categorical *0= Less than College* *1=College or More*	N/A
Exposure to Risk for Abortion (Years)	N/A	Continuous *Number of years sexually active 1994-2007*	N/A

TABLE 3: UNRESTRICTED AND RESTRICTED SAMPLE CHARACTERISTICS FOR THE GOLD-STANDARD ANALYSES

Characteristics	All Women & Abortions in Wave 4	Total Abortions between 1994-2007 to women of any age	Total Abortions between 1994-2007 to women 15-29 at time of abortion	Total Abortions Between 1994-2007 using yearly age-restrictions
Total Number	1,833	1,584	1,506	1,460
Unique Women	1,219	1,084	1,041	1,019
Number of Abortions (Min-Max)	0-17	0-17	0-17	0-17
Age Range at Abortion	8-31	12-31	15-29	15-29
% With No Abortions	85%	87%	87%	87%
% 1 Abortions (of women with an abortion)	66%	71%	70%	71%
% 2 Abortions (of women with an abortion)	23%	20%	21%	21%
% 3 Abortions	8%	7%	7%	6%
% 4 or More Abortions	3%	2%	2%	2%
Race				
White	34%	35%	34%	34%
Black	39%	40%	41%	40%
Native American	0.8%	0.8%	0.8%	0.8%
Asian/Pacific Islander	5%	5%	5%	5%
Hispanic	19%	19%	19%	19%
Other	0.7%	0.7%	0.7%	.07%
Age at Time of Abortion				
15-19	25%	29%	30%	31%
20-24	39%	45%	48%	46%
24-29	21%	24%	23%	23%
Outside 15-29 Range	15%	1.2%	--	--

TABLE 4: ADD HEALTH AND EXTERNAL ABORTION RATES, AND PERCENTAGE OF EXTERNAL RATE CAPTURED, OVERALL AND BY YEAR

Year *Age Group*	Unweighted Rate	Weighted Rate, per 1000	External Rate[a] , per 1000	% Weighted Rate Captured[b]
1994-2007	15.1	12.9	35.9	35.3%
19-29[c]				*[29.7% -41.0%]*
1994	6.1	6.4	31.7	20.2%
15-21				*[8.2% -32.2%]*
1995	9.6	8.6	32.1	26.7%
15-22				*[15.8% -37.6%]*
1996	10.5	8.5	38.9	21.8%
15-24				*[14.4% -29.2%]*
1997	12.3	8.6	37.4	23.0%
15-24				*[15.2% -30.7%]*
1998	16.5	12.6	36.1	35.0%
15-24				*[24.6% -45.4%]*
1999	16.9	15.5	37.5	41.3%
16-24				*[31.2% -51.4%]*
2000	18.3	14.4	39.9	36.2%
17-27				*[26.3% -46.2%]*
2001	19.2	15.7	39.4	39.7%
18-28				*[29.1% -50.4%]*
2002	19.3	16.9	36.9	46.0%
19-29				*[33.3% -58.6%]*
2003	13.2	11.5	36.2	31.8%
20-29				*[21.7% -41.9%]*
2004	16.4	14.2	35.0	40.5%
21-29				*[28.2% -52.8%]*
2005	12.9	10.4	34.1	30.6%
22-29				*[21.0% -40.3%]*
2006	13.1	9.8	29.9	32.7%
25-29				*[21.2% -44.3%]*
2007	14.0	14.5	28.6	50.9%
25-29				*[35.5% -66.2%]*

NOTES: 95% Confidence intervals in brackets.

a. The external counts are age-matched using the CDC age distribution data, except for the years 1994, 1995, 2000 and 2001 which use available data from the Guttmacher Institute's survey of abortion patients.

b. The percentage of the external rate captured uses the weighted Add Health rates

c. The total counts do not include all women 15-29 during 1994-2007, instead, they include only women included in each year of analysis. See Appendix 2 for further details.

Source for birth data: National Center for Health Statistics (1994-2007)

TABLE 5: ADD HEALTH ABORTION RATE-RATIOS, EXTERNAL ABORTION RATIOS, AND PERCENTAGE OF EXTERNAL WEIGHT CAPTURED, OVERALL AND BY YEAR

Year *Age Group*	Unweighted Rate Ratio	Weighted Rate Ratio, Per 1,000 Live Births	External Ratio[a] , per 1000 Live Births	% Weighted Rate-Ratio Captured[b]
1994-2007	194.9	164.7	375.9	42.9%
19-29[c]				*[35.5% -50.3%]*
1994	357.1	327.2	436.9	74.9%
15-21				*[25.3% -124.4%]*
1995	297.8	231.0	429.5	53.8%
15-22				*[29.1% -78.5%]*
1996	278.9	234.4	485.8	48.2%
15-24				*[28.6% -67.9%]*
1997	220.9	172.8	478.0	36.2%
15-24				*[24.0% -48.4%]*
1998	257.4	198.1	461.0	43.0%
15-24				*[29.3% -56.6%]*
1999	216.6	195.1	442.5	44.1%
16-24				*[31.7% -56.5%]*
2000	213.6	177.4	404.6	43.8%
17-27				*[30.8% -56.9%]*
2001	209.7	162.4	382.5	42.5%
18-28				*[29.0% -55.9%]*
2002	227.2	192.7	346.1	55.7%
19-29				*[39.9% -71.4%]*
2003	141.1	120.8	339.7	35.6%
20-29				*[23.2% -47.9%]*
2004	181.4	164.0	321.1	51.1%
21-29				*[34.1% -68.0%]*
2005	126.6	101.5	312.4	32.5%
22-29				*[21.6% -43.4%]*
2006	132.0	98.5	252.7	39.0%
25-29				*[24.4% -53.5%]*
2007	147.1	153.4	241.6	63.5%
25-29				*[43.3% -20.2%]*

NOTES: 95% Confidence intervals in brackets.

a. The external counts are age-matched using the CDC age distribution data, except for the years 1994, 1995, 2000 and 2001 which use available data from the Guttmacher Institute's survey of abortion patients.

b. The percentage of the external ratio captured uses the weighted Add Health rate-ratios.

c. The total counts do not include all women 15-29 during 1994-2007, instead, they include only women included in each year of analysis. See Appendix 2 for further details.

Source for birth data: National Center for Health Statistics (1994-2007)

TABLE 6: ADD HEALTH SUB-GROUP ABORTION RATE-RATIOS, EXTERNAL ABORTION RATIOS, AND PERCENTAGE CAPTURED

	Add Health Raw Ratio per 1000 Live Births	Add Health Ratio Estimate, per 1000 Live Births	External Ratio per 1000 Live Births	Ratio Captured [95% CI]
Race				
White	128.0	191.7	300.4	39.9%
				[33.2%-46.6%]
Black	300.1	237.1	648.0	36.6%
				[23.5%-49.7%]
Native American	200.0	209.3	324.0	64.6%
				[44.3%-68.3%]
Asian	258.9	294.9	701.9	42.0
				[18.7%-65.3%]
Hispanic	205.2	213.1	336.6	63.3%
				[42.2%-84.4%]
Other Race[a]	263.2	202.4		--
Age at Time of Abortion				
15-19	293.8	244.6	418.5	58.4%
				[47.7%-69.2%]
20-24	197.4	156.8	273.5	57.3%
				[45.7%-68.9%]
25-29	132.8	104.9	261.9	40.0%
				[29.4%-50.7%]

NOTES: This table looks at all years 1994-2007 combined.
a. No external counts are available for the other or multiple race category
Source for birth data: National Center for Health Statistics (1994-2007)

TABLE 7: DESCRIPTIVE STATISTICS FOR CONTINUOUS AND CATEGORICAL VARIABLES FOR THE POISSON ANALYSES, WEIGHTED AND UNWEIGHTED MEANS AND STANDARD ERRORS (N=7375)

Variable	Mean	Standard Error	Weighted Mean	Weighted Standard Error
Number of Abortions Reported (1994-2007)[1]	0.21	0.01	0.18	0.01
Number of Births Reported (1994-2007)	1.04	0.01	1.05	0.04
Number of Marriages at Wave 4	0.61	0.01	0.63	0.01
Exposure (Years)	10.92	0.03	10.91	0.10
Race				
White	0.55		0.68	
Black	0.22		0.16	
Other	0.07		0.04	
Hispanic	0.05		0.03	
Age at Wave 4				
24-29	0.69		0.72	
30-34	0.31		0.28	
Marital History at Wave 4				
Never Married	0.44		0.43	
Married at least 1	0.56		0.57	
Income at Wave 4				
>100% Poverty Line	0.37		0.38	
100-199% Poverty Line	0.25		0.25	
<=200% Poverty Line	0.33		0.32	
Religion at Wave 4				
Protestant	0.54		0.53	
Catholic	0.22		0.20	
Other	0.08		0.08	
None	0.16		0.18	
Education at Wave 4				
Less than College	0.65		0.66	
College/More than College	0.35		0.34	
Work History at Wave 4				
Ever Worked Full Time	0.95		0.94	
Sexual Orientation				
Heterosexual	0.96		0.96	

1. Includes abortions obtained by women of any age.

TABLE 8: POISSON MODEL ESTIMATED UNCONDITIONAL REPORTING RATES (%), OVERALL AND BY GROUP, WEIGHTED

	Estimate	Linearized Standard Error[1]	95% Confidence Intervals[2]
Overall Reporting Rage	20.25	0.59	19.1-21.4
Race			
White	17.3	0.56	16.2-18.4
Black	24.0	0.97	22.0-25.9
Other	25.5	4.38	16.9-34.2
Hispanic	54.4	2.76	48.8-59.9
Age at Wave 4			
24-29	20.5	0.65	19.4-21.7
30-34	19.4	0.93	17.6-21.3
Marital History at Wave 4			
Never Married	26.9	0.65	25.7-28.2
Married at least 1	15.2	0.58	14.0-16.4
Income at Wave 4			
Less than 100% Poverty Line	17.8	0.63	16.5-19.0
100%-199% Poverty Line	19.9	0.88	18.2-21.7
Greater than or Equal 200% Poverty Line	23.2	0.83	21.5-24.8
Religion at Wave 4			
Protestant	6.3	0.30	5.7-6.9
Catholic	31.0	0.78	30.2-33.3
Other	75.3	0.73	73.9-76.8
None	23.8	0.56	22.7-24.9
Education at Wave 4			
Less than College	20.0	0.62	18.7-21.2
College/More than College	20.8	0.85	19.2-22.5

Notes:

1. The linearized standard error used by Stata 14's svy estimation suite is the Huber-White sandwich estimator.
2. All non-overlapping confidence intervals all indicate significant differences at least at the .05-level using the Wald test. Overlapping confidence intervals indicate non-significant differences unless otherwise noted in the text.

TABLE 9: POISSON ANALYSES OF THE CONDITIONAL REPORTING RATES BY GROUP, WEIGHTED

	Estimate[1]	Linearized Standard Error[2]
Race (*Ref: White)*		
Black	1.02	1.02
Other	0.07	1.60
Hispanic	1.66	1.42
Age at Wave 4 (*Ref 24-29)*		
30-34	-0.60	0.57
Marital History at Wave 4*(Ref: Never Married)*		
Married at least 1	-0.65	0.50
Income at Wave 4 (*Ref: 100-199% Poverty Line)*		
Less than 100% Poverty Line	-0.09	0.24
Greater than or Equal 200% Poverty Line	0.14	0.41
Religion at Wave 4(*Ref: Protestant*)		
Catholic	2.22	7.90
Other	4.04	3.19
None	1.73	6.18
Education (*Ref: Less than College)*		
College/More than College	0.52	2.19
Constant	-2.75	10.79
Number of observations	7375	
Log pseudolikelihood	-4618330.3	

*p<.05, **p<.01, ***p<.001

Notes:

1. The estimate coefficients indicate the impact of the group on the likelihood of underreporting.
2. The linearized standard error used by Stata 14's svy estimation suite is the Huber-White sandwich estimator.

TABLE 10: DESCRIPTIVE STATISTICS FOR CATEGORICAL VARIABLES IN THE CONSISTENCY ANALYSES, UNWEIGHTED **(N=1248)**

	Count	Percentage
Consistency of Reporting		
Consistent Report	452	36
Inconsistent Report	796	64
INDIVIDUAL CHARACTERISTICS		
Race		
White	442	35
Black	448	39
Hispanic	242	19
Other/Multiracial	76	6
Ever on Welfare		
No	1176	94
Yes	72	6
ABORTION-SPECIFIC CHARACTERISTICS		
Religion at Abortion		
Protestant	659	53
Catholic	281	23
Other	88	7
None	220	18
Marital Status at Abortion		
Unmarried	1162	93
Married	86	7
Time Since Abortion at First Reporting Opportunity (Wave 2 or 4)		
1 Year or Less	286	23
More than 1 year	962	77

NOTES: These data are presented at the abortion level, not the individual level.

TABLE 11: DESCRIPTIVE STATISTICS FOR THE CONTINUOUS VARIABLES IN THE CONSISTENCY ANALYSES, UNWEIGHTED (N=1248)

	Mean (*SD*)	Range
INDIVIDUAL CHARACTERISTICS		
Abortion Attitude Scale	6.16 (*1.37*)	2-10
ABORTION SPECIFIC CHARACTERISTICS		
Income at Abortion (Best-Guess Coding)	$10,117.79 (*13647.44*)	$0-$86,000
Income at Abortion (Inverse Hyperbolic Sine Transformation)	7.38 (*4.09*)	0-12.06
Age at Abortion	21.7 (*3.87*)	13-31

TABLE 12: MULTILEVEL BIVARIATE LOGISTIC REGRESSION MODELS OF CONSISTENCY OF REPORTING, UNWEIGHTED
(Number of Abortions=1248, Number of unique women=826, Number of Schools=122)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
INDIVIDUAL CHARACTERISTICS								
Race (Reference: White)								
Black	1.47* (0.28)							
Hispanic	1.24 (0.27)							
Other/Multiracial	1.07 (0.39)							
Abortion Attitude Scale		0.99 (0.06)						
Ever on Welfare Before 18 (Reference: No)								
Yes			0.63 (0.24)					
ABORTION-SPECIFIC CHARACTERISTICS								
Age at Abortion				3.78*** (0.51)				
Income at Abortion (IHS Transformed)					1.27*** (0.04)			
Religion at Abortion (Reference: Protestant)								
Catholic						0.78 (0.15)		
Other						1.73^ (0.53)		
None						1.24 (0.26)		
Marital Status at Abortion (Reference: Unmarried)								
Married							2.30** (0.70)	
Time Since Abortion (Reference: 1 year or less)								
More than 1 year past								0.61** (0.11)
Constant	0.40*** (0.06)	0.50^ (0.19)	0.49*** (0.05)	0.00*** (0.000)	0.07*** (0.02)	0.47*** (0.05)	0.45*** (0.04)	0.70** (0.11)
Individual Effect (Estimate)	1.24 (0.20)	1.26 (.20)	1.26 (0.20)	0.21 (1.24)	0.27 (0.31)	1.23 (0.20)	1.27 (0.20)	1.25 (0.20)
School Effect (Estimate)	0.00 (0.43)	0.00 (0.29)	0.00 (0.34)	1.68 (0.28)	1.47 (0.24)	0.00 (0.25)	0.00 (0.36)	0.00 (0.47)
Log Likelihood	-804.15	-806.22	-805.47	-287.58	-749.90	-800.20	-802.44	-802.38

^p<.10, *p<.05, **p<.01, ***p<.001.

NOTES:

Exponentiated coefficients presented, standard errors in parentheses

LR-tests comparing the multilevel model to the marginal logistic model were all significant at the .001-level.

TABLE 13 MULTILEVEL: MULTIVARIATE LOGISTIC REGRESSION MODELS OF CONSISTENCY OF ABORTION REPORTING, UNWEIGHTED
(Number of Abortions=1248, Number of unique women=829, Number of Schools=122)

	Model A	Model B	Model C
INDIVIDUAL CHARACTERISTICS			
Race (Reference: White)			
Black	1.48*		2.05
	(0.28)		(1.05)
Hispanic	1.25		1.57
	(0.29)		(0.92)
Other/Multiracial	1.05		1.61
	(0.38)		(1.29)
Abortion Attitude Scale	0.99		0.74*
	(0.06)		(0.10)
Ever on Welfare Before 18 (Reference: No)			
Yes	0.60		0.52
	(0.22)		(0.39)
ABORTION-SPECIFIC CHARACTERISTICS			
Age at Abortion		7.34***	7.12***
		(2.22)	(2.11)
Income at Abortion (IHS Transformed)		0.98	.98
		(0.05)	(0.04)
Religion at Abortion (Reference: Protestant)			
Catholic		0.44	0.49
		(0.23)	(0.26)
Other		01.42	1.39
		(0.92)	(0.57)
None		1.58	1.39
		(0.82)	(0.81)
Marital Status at Abortion (Reference: Unmarried)			
Married		0.78	0.72
		(0.49)	(0.44)
Time Since Abortion (Reference: 1 year or less)			
More than 1-year past		0.001***	0.001***
		(.001)	(.002)
Constant	0.43*	0.00***	0.00***
	(0.17)	(0.00)	(0.00)
Individual Effect (Estimate)	1.24	1.36	1.19
	(0.20)	(0.54)	(0.56)
School Effect (Estimate)	0.00	2.49	2.46
	(0.63)	(0.47)	(0.47)
Log Likelihood	-800.84	-263.00	-258.80

*p<.05, **p<.01, ***p<.001
NOTE: Exponentiated coefficients presented, standard errors in parentheses
LR-tests comparing the multilevel model to the marginal logistic model were all significant at the .001-level.

FIGURE 1: THEORETICAL FRAMEWORK FOR UNDERREPORTING OF ABORTIONS

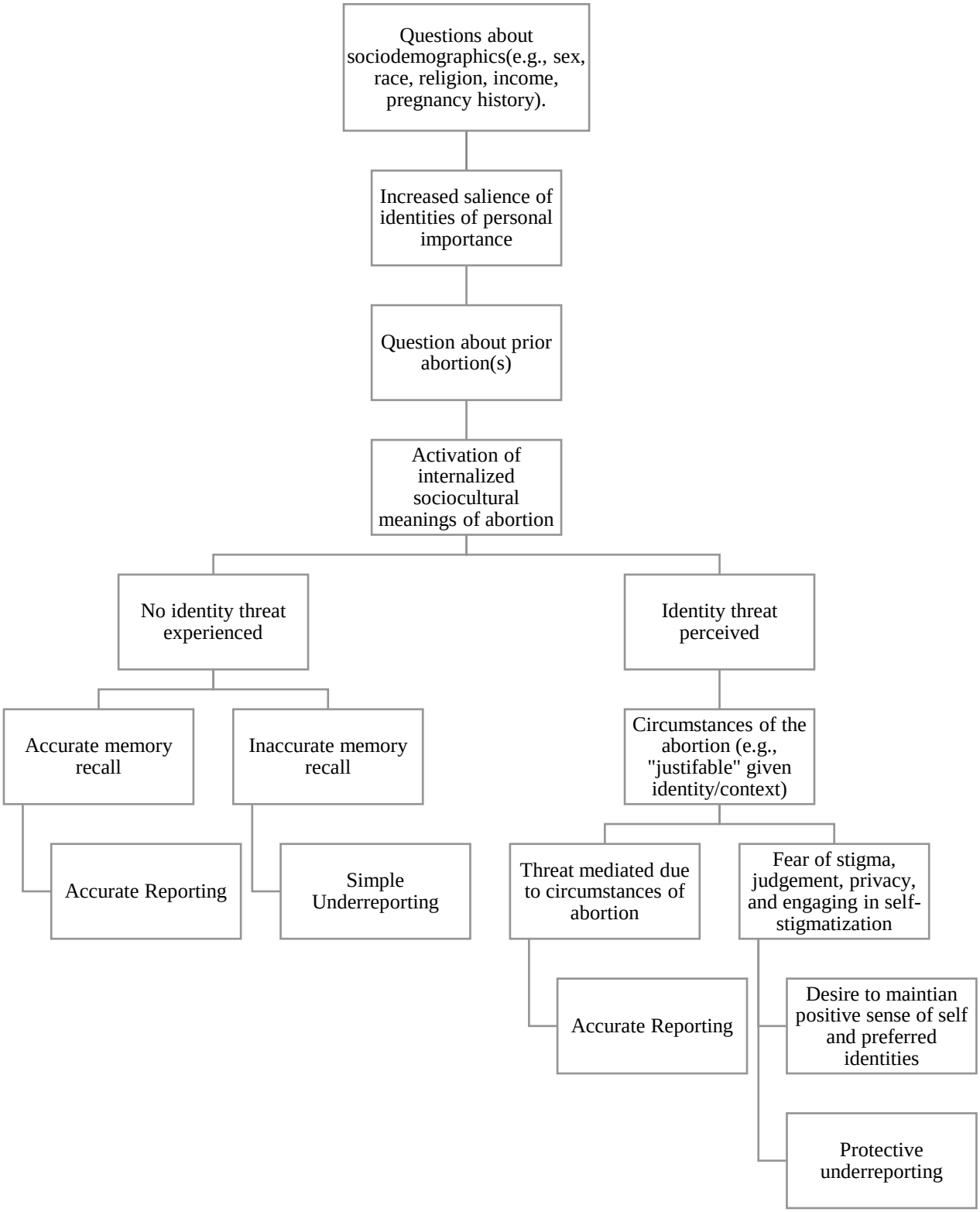
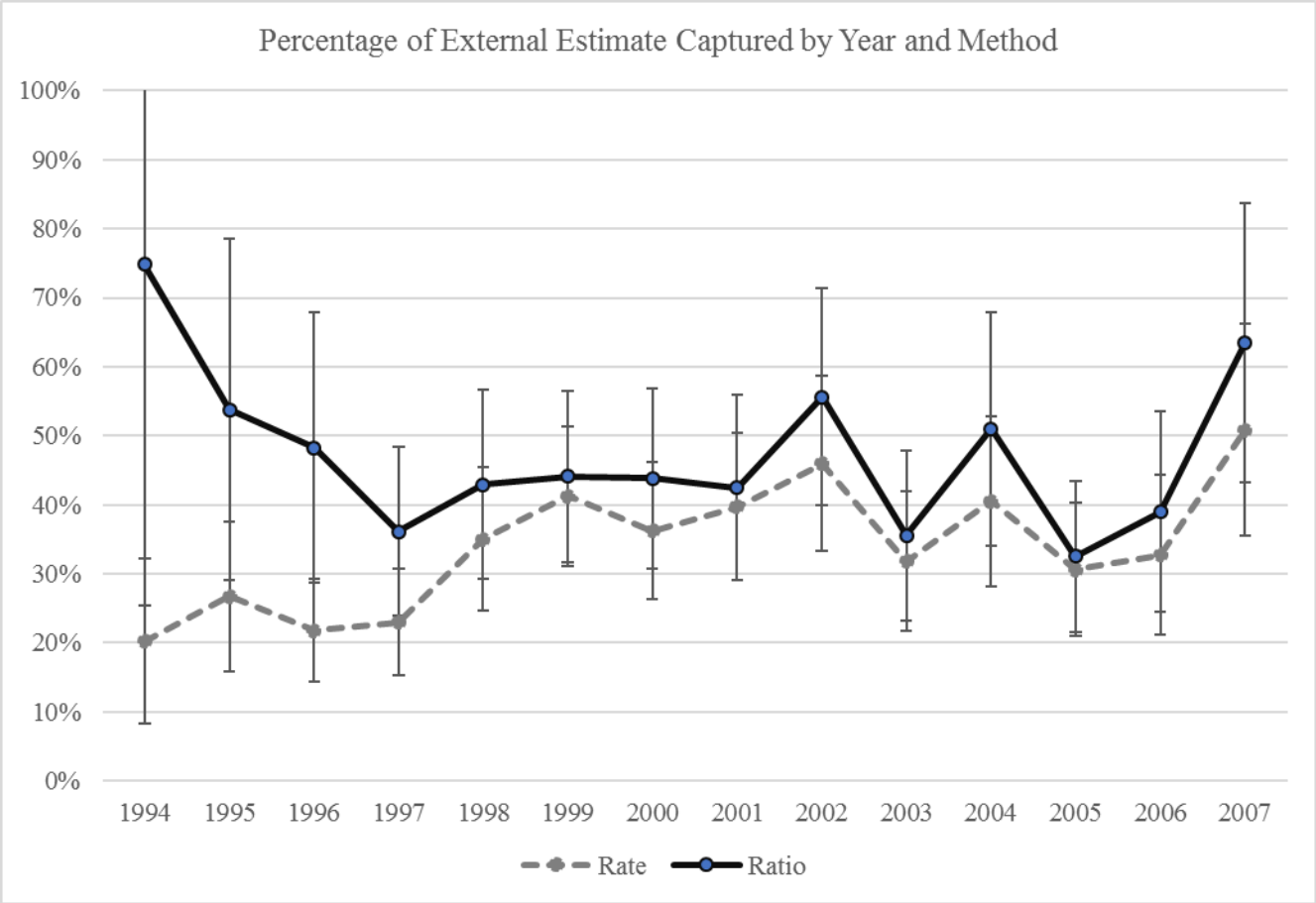
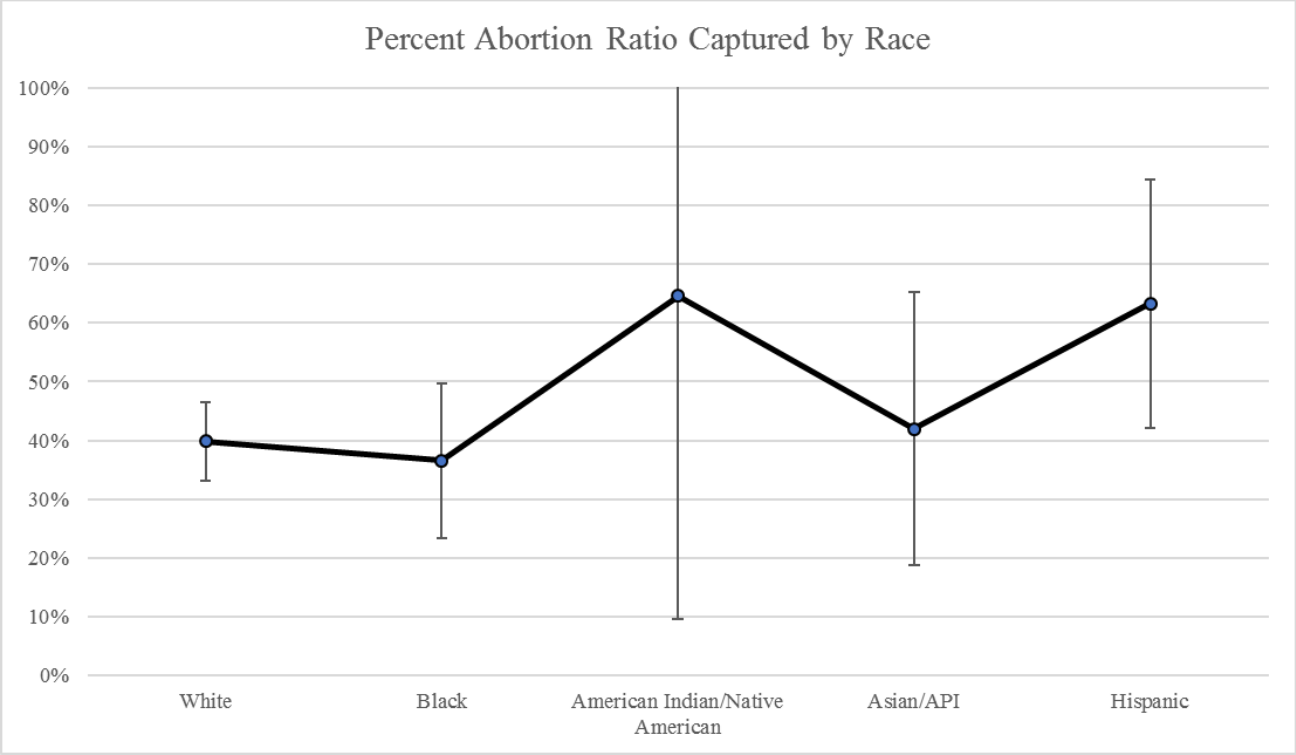


FIGURE 2: PERCENTAGE OF EXTERNAL ESTIMATES CAPTURED BY ADD HEALTH ACROSS METHODS, BY YEAR



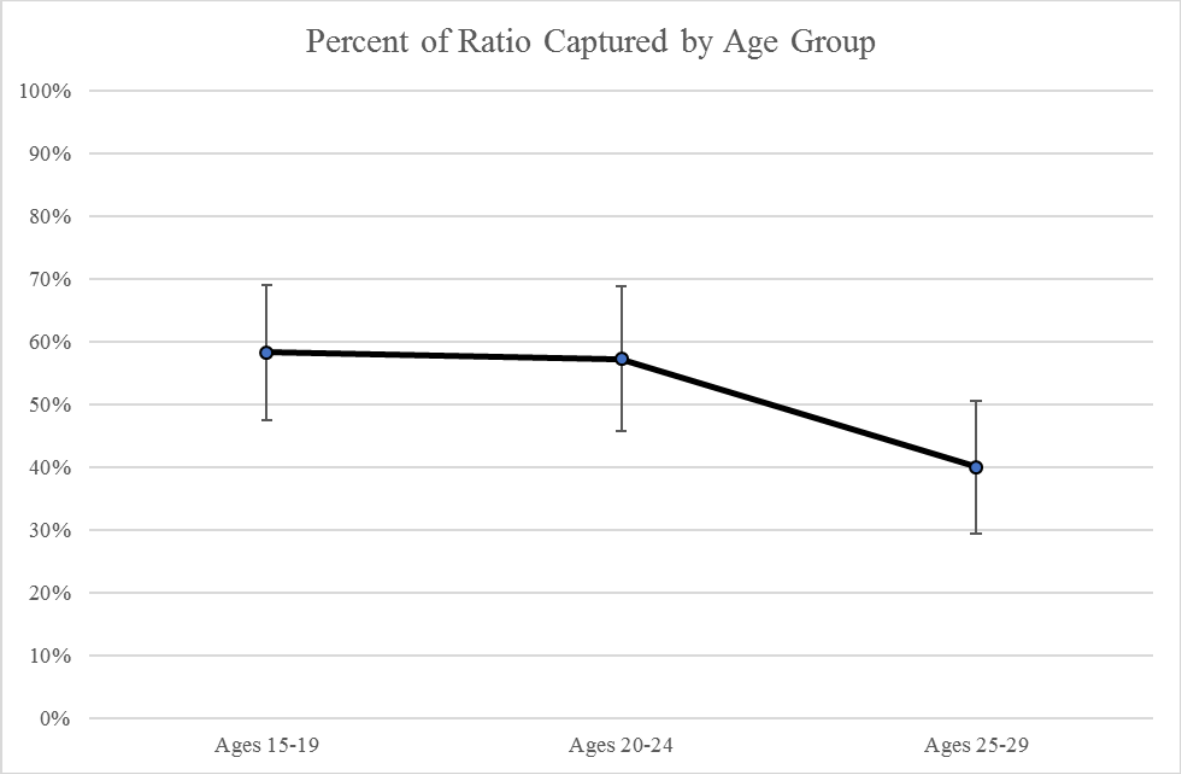
Note:
Bars denote 95% confidence intervals
95% Confidence Interval Bars capped at 100%, though bars may exceed this value (see Tables 4 and 5 for further details).
Source for birth data: National Center for Health Statistics (1994-2007)

FIGURE 3: PERCENTAGE OF EXTERNAL ABORTION RATIO CAPTURED BY ADD HEALTH, BY RACE



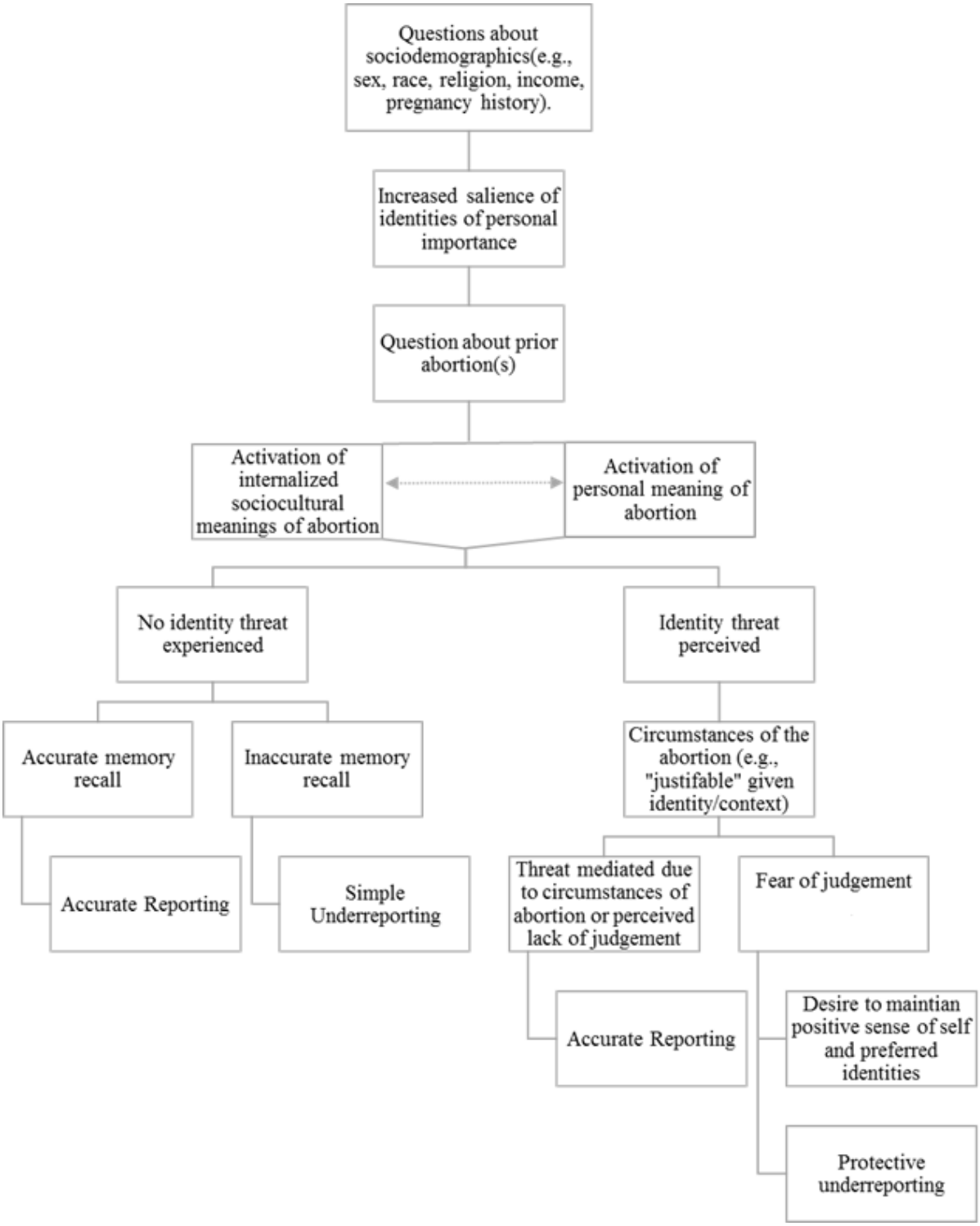
Notes: Bars denote 95% confidence intervals

FIGURE 4: PERCENTAGE OF EXTERNAL ABORTION RATIO CAPTURED BY ADD HEALTH, BY AGE



Notes: Bars denote 95% confidence intervals

FIGURE 5: REVISED THEORETICAL FRAMEWORK



APPENDIX 1: ALTERNATIVE INCOME CODING SCHEMES

The coding scheme used in the body of the paper is referred to as the "best guess" coding method. The first alternative coding prioritized temporal sequencing. In this coding scheme, abortions that occurred between 2005 and 2008 used income reported at wave 4, abortions that occurred between 2000 and 2004 used income reported in wave 3, abortions that occurred between 1996 and 1999 used personal income reported at wave 2, and abortions that occurred earlier than 1996 used wave 1 data.

In the second alternative coding scheme parents reported income was used as a proxy for income among women with early life abortions; abortions that occurred before 1999 used income data reported by parents unless this data was missing. If these data were missing the personal income reported in wave 1 and 2 was used following coding scheme 1. Parental income was measured using the following question from the Wave 1 parental questionnaire: In addition, in Wave 1 data collection a parental survey was administered that asked " In Wave 3 participants were asked, "Including all the income sources you reported above, what was your total personal income before taxes in {2000/2001}? Please include all of the income sources you identified in the previous question," Participants gave their income in whole dollars.

Figure A below shows the distribution of the raw data by income group across coding methods. The percentage captured by group changes significantly in some cases. Tables A-C below show the differing results for the consistency analyses when using different coding schemes. Although the underlying distribution varies somewhat, no substantial differences are observed in the analyses as a result of the coding scheme used.

Figure A: Raw distribution of women in each income category at time of abortion by coding method b

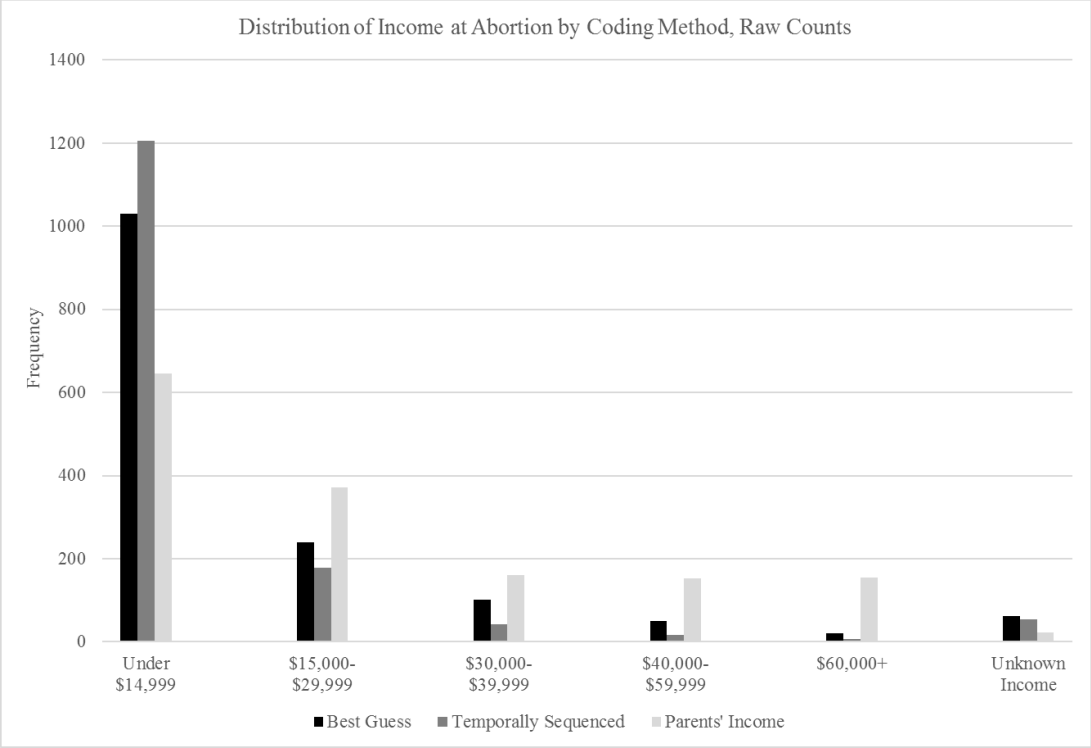


Table A: Descriptive statistics by coding method for consistency analyses (N=1248)

	Mean (*SD*)	Range
ABORTION SPECIFIC CHARACTERISTICS		
Income at Abortion (Best-Guess Coding)	$10,117.79 (*13647.44)*	$0-$86,000
Income at Abortion (Temporal Sequence Coding)	$6,813.17 (*9461.32)*	$0-$61,000
Income at Abortion (Parents' Income)	$8943.96 (*14086.52)*	$0-$86,000

Table B: Bivariate consistency results by coding method
(Number of Abortions=1248, Number of unique women=829, Number of Schools=122)

	Best Guess	Temporal Sequencing	Parents Income
Income at Abortion (IHS Transformed)	1.27***	1.24***	1.26***
	(0.04)	(0.03)	(0.04)
Constant	0.07***	0.11***	0.07***
	(0.02)	(0.03)	(0.02)
Individual Effect (Estimate)	0.27	1.43	1.47
	(0.31)	(0.23)	(0.24)
School Effect (Estimate)	1.47	0.40	0.27
	(0.24)	(0.22)	(0.31)
Log Likelihood	-749.90	-758.06	-749.90

*p<.05, **p<.01, ***p<.001
NOTES:
Exponentiated coefficients presented, standard errors in parentheses
LR-tests comparing the multilevel model to the marginal logistic model were all significant at the .001-level.

Table C: Full multivariate multilevel consistency of reporting model by income coding used
(Number of Abortions=1248, Number of unique women=829, Number of Schools=122)

	Best Guess	Temporal Sequence	Parents Income
INDIVIDUAL CHARACTERISTICS			
Race (Reference: White)			
Black	2.05	2.12	2.06
	(1.05)	(1.07)	(1.06)
Hispanic	1.57	1.57	1.58
	(0.92)	(0.90)	(0.92)
Other/Multiracial	1.61	1.70	1.61
	(1.29)	(1.37)	(1.30)
Abortion Attitude Scale	0.74*	0.74*	0.74*
	(0.10)	(0.09)	(0.09)
Ever on Welfare Before 18 (Reference: No)			
Yes	0.52	0.52	0.52
	(0.39)	(0.39)	(0.41)
ABORTION-SPECIFIC CHARACTERISTICS			
Age at Abortion	7.12***	6.85***	7.08***
	(2.11)	(1.99)	(2.11)
Income at Abortion (IHS Transformed)	.98	1.02	0.99
	(0.04)	(0.05)	(0.05)
Religion at Abortion (Reference: Protestant)			
Catholic	0.49	0.48	0.49
	(0.26)	(0.26)	(0.26)
Other	1.39	0.80	1.38
	(0.57)	(0.57)	(0.90)
None	1.39	1.31	1.60
	(0.81)	(0.64)	(0.81)
Marital Status at Abortion (Reference: Unmarried)			
Married	0.72	0.69	0.72
	(0.44)	(0.42)	(0.44)
Time Since Abortion (Reference: 1 year or less)			
More than 1-year past	0.001***	0.001***	0.001***
	(.002)	(.002)	(.002)
Constant	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)
Individual Effect (Estimate)	1.19	1.16	1.17
	(0.56)	(0.56)	(0.57)
School Effect (Estimate)	2.46	2.44	2.49
	(0.47)	(0.46)	(0.46)
Log Likelihood	-258.80	-259.41	-258.88

*p<.05, **p<.01, ***p<.001

NOTES:

Exponentiated coefficients presented, standard errors in parentheses

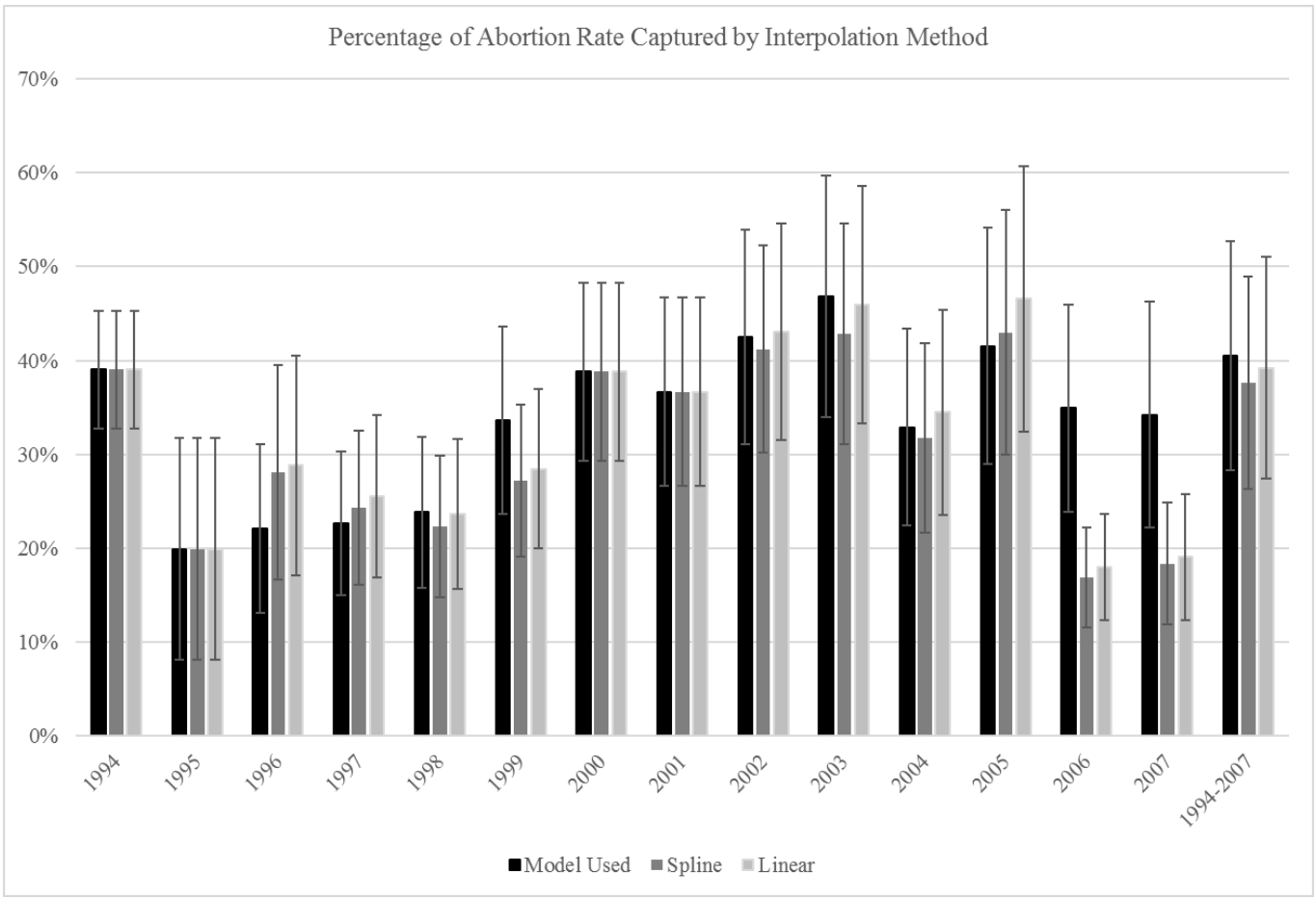
LR-tests comparing the multilevel model to the marginal logistic model were all significant at the .001-level.

APPENDIX 2: ADD HEALTH AGE MATCHING INFORMATION

Year	Ages Covered in Add Health	Add Health Age Range Used in Analyses	Age Range Used for External Estimates	Notes
1994	11-21	15-21	15-21	
1995	12-22	15-22	15-22	Add Health women aged 22 were included because Guttmacher Institute data was available for all ages in this year.
1996	13-23	15-24	15-24	CDC age-data was used, as a result, the external estimate includes an additional year not covered by the Add Health survey in this year.
1997	14-24	15-24	15-24	
1998	15-25	15-24	15-24	Add Health women aged 25 were excluded from this year.
1999	16-26	16-24	16-24	Add Health women aged 25 and 26 were excluded from this year. Year by year data for teens was used to exclude 15-year-olds.
2000	17-27	17-27	17-27	All ages were included Guttmacher Institute data was available for all ages in this year.
2001	18-28	18-28	18-28	All ages were included Guttmacher Institute data was available for all ages in this year.
2002	19-29	19-29	19-29	Year by year data available on teens was used to include 19-year-olds.
2003	20-30	20-29	20-29	Add Health women aged 30 were excluded from this year.
2004	21-31	21-29	20-29	Add Health women aged 30 and 31 were excluded from this year.
2005	22-32	22-29	20-29	Add Health women 30 and over were excluded from this year.
2006	23-33	25-29	25-29	Add Health women aged 23 and 24 were excluded from this year.
2007	24-34	25-29	25-29	Add Health women aged 24 were excluded from this year.

NOTE: Add Health women under 15 and over 29 were excluded from all analyses

APPENDIX 3: ALTERNATIVE AGE INTERPOLATIONS



Brackets represent 95% Confidence Intervals

APPENDIX 4: ADDITIONAL INFORMATION RELATED TO THE POISSON ANALYSES

TABLE A: INFORMATION ON THE ADDITIONAL VARIABLES USED IN THE COUNT AND EXCESS ZERO POISSON ANALYSES

	Add Health Question	Coding Notes
Ever Worked Full Time	Asked in Wave 4: "Have you ever worked full time at least 35 hours a week at a paying job while you were not primarily a student? Do not include summer work." Responses: Yes, No	
Sexual Orientation	Asked in Wave 4: Please choose the description that best fits how you think about yourself. Responses:100% heterosexual (straight); mostly heterosexual (straight), but somewhat attracted to people of your own sex; bisexual, that is, attracted to men and women equally; mostly homosexual (gay), but somewhat attracted to people of the opposite sex; 100% homosexual (gay); not sexually attracted to either males or females.	Individuals who responded 100% heterosexual and mostly heterosexual were coded as heterosexual and all other sexual orientations were collapsed.

TABLE B: COUNT POISSON ANALYSES, WEIGHTED (n=7375)

	Estimate	Linearized Standard Error[1]
COUNT		
Race (*Ref: White)*		
Black	0.00	0.69
Other	0.37	1.01
Hispanic	-1.64	2.12
Age at Wave 4 (*Ref 24-29)*		
30-34	0.09	0.89
Number of Marriages	0.23	0.39
Religion at Wave 4(*Ref: Protestant*)		
Catholic	-1.66	8.20
Other	-2.06	7.68
None	-1.48	6.66
Education (*Ref: Less than College)*		
College/More than College	-0.91	2.04
Ever Worked Full Time (*Ref: Never Worked Full Time)*		
Have Worked Full Time	0.03	0.45
Sexual Orientation (*Ref:* Not Heterosexual)		
Heterosexual	0.38	0.42
Number of Births 1994-2007	-0.04	0.07
Exposure ln(t)	1.00	
Constant	-0.42	9.96
Number of observations	7375	
Log pseudolikelihood	-4618330.3	

^p<.10 *p<.05, **p<.01, ***p<.001

1. The linearized standard error used by Stata 14's svy estimation suite is the Huber-White sandwich estimator.

TABLE C: EXCESS ZERO POISSON ANALYSES, WEIGHTED (n=7375)

	Estimate	Linearized Standard Error[1]
EXCESS ZEROS		
Race (*Ref: White)*		
Black	-0.11	0.22
Other	-1.85	1.16
Hispanic	-0.13	0.57
Age at Wave 4 (*Ref 24-29)*		
30-34	-0.45	0.31
Number of Marriages	0.34^	0.19
Religion at Wave 4(*Ref: Protestant*)		
Catholic	-0.22	0.31
Other	0.21	0.27
None	-0.80^	0.41
Education (*Ref: Less than College)*		
College/More than College	-0.13	0.31
Ever Worked Full Time (*Ref: Never Worked Full Time)*		
Have Worked Full Time	-0.59	0.52
Sexual Orientation (*Ref:* Not Heterosexual)		
Heterosexual	-0.10	0.91
Number of Births 1994-2007	-0.03	0.09
Constant	1.65^	0.85
Number of observations	7375	
Log pseudolikelihood	-4618330.3	

^p<.10 *p<.05, **p<.01, ***p<.001

1. The linearized standard error used by Stata 14's svy estimation suite is the Huber-White sandwich estimator.

REFERENCES

Astbury-Ward, Edna, Odette Parry, and Ros Carnwell. 2012. “Stigma, Abortion, and Disclosure—Findings from a Qualitative Study.” *The Journal of Sexual Medicine* 9(12):3137–47.

Avalos, Lisa. 1999. “Hindsight and the Abortion Experience: What Abortion Means to Women Years Later.” *Gender Issues* 17(2):35–57.

Belmi, Peter, Rodolfo Cortes Barragan, Margaret A. Neale, and Geoffrey L. Cohen. 2015. “Threats to Social Identity Can Trigger Social Deviance.” *Personality and Social Psychology Bulletin* 146167215569493.

Biggs, M.Antonia, Ushma D. Upadhyay, Charles E. Mcculloch, and Diana G. Foster. 2016. “Women’s Mental Health and Well-Being 5 Years After Receiving or Being Denied an Abortion: A Prospective, Longitudinal Cohort Study.” *JAMA Psychiatry*. Retrieved December 16, 2016 (http://jamanetwork.com/journals/jamapsychiatry/fullarticle/2592320).

Bird, Sheryl Thorburn and Laura M. Bogart. 2003. “Birth Control Conspiracy Beliefs, Perceived Discrimination, and Contraception among African Americans: An Exploratory Study.” *Journal of Health Psychology* 8(2):263–76.

Bock, Jane D. 2000. “Doing The Right Thing?: Single Mothers by Choice and the Struggle for Legitimacy.” *Gender & Society* 14(1):62–86.

Charles, Vignetta E., Chelsea B. Polis, Srinivas K. Sridhara, and Robert W. Blum. 2008. “Abortion and Long-Term Mental Health Outcomes: A Systematic Review of the Evidence.” *Contraception* 78(6):436–50.

Chumlea, William Cameron et al. 2003. “Age at Menarche and Racial Comparisons in US Girls.” *Pediatrics* 111(1):110–13.

Cockrill, Kate, Ushma D. Upadhyay, Janet Turan, and Diana Greene Foster. 2013. “The Stigma of Having an Abortion: Development of a Scale and Characteristics of Women Experiencing Abortion Stigma.” *Perspectives on Sexual and Reproductive Health* 45(2):79–88.

Coleman, Priscilla K. 2011. “Abortion and Mental Health: Quantitative Synthesis and Analysis of Research Published 1995–2009.” *The British Journal of Psychiatry* 199(3):180–86.

Coleman, Priscilla K., Catherine T. Coyle, Martha Shuping, and Vincent M. Rue. 2009. “Induced Abortion and Anxiety, Mood, and Substance Abuse Disorders: Isolating the Effects of Abortion in the National Comorbidity Survey.” *Journal of Psychiatric Research* 43(8):770–76.

Davis, Liane V. And Jan L. Hagen. 1996. “Stereotypes and Stigma: What’s Changed for Welfare Mothers.” *Affilia* 11(3):319–37.

Edin, Kathryn and Maria Kefalas. 2011. *Promises I Can Keep: Why Poor Women Put Motherhood Before Marriage, With a New Preface*. University of California Press.

Ellison, Marcia A. 2003. “Authoritative Knowledge and Single Women’s Unintentional Pregnancies, Abortions, Adoption, and Single Motherhood: Social Stigma and Structural Violence.” *Medical Anthropology Quarterly* 17(3):322–47.

Fine, Cordelia. 2010. *Delusions of Gender: How Our Minds, Society, and Neurosexism Create Difference*. New York: W.W. Norton & Company, Inc. Retrieved April 8, 2016 (https://books-google-com.libproxy.lib.unc.edu/books/about/Delusions_of_Gender_How_Our_Minds_Societ.html?Id=s2ztdax83ymc).

Finer, Lawrence B. And Stanley K. Henshaw. 2003. "Abortion Incidence and Services in the United States in 2000." *Perspectives on Sexual and Reproductive Health* 35(1):6–15.

Fu, Haishan, Jacqueline E. Darroch, Stanley K. Henshaw, and Elizabeth Kolb. 1998. "Measuring the Extent of Abortion Underreporting in the 1995 National Survey of Family Growth." *Family Planning Perspectives* 30(3):128–38.

Goffman, Erving. 1963. *Stigma: Notes on the Management of Spoiled Identity*. Simon and Schuster.

Guttmacher Institute. 2011. "Are You IN THE KNOW: Abortion Providers in the United States." Retrieved (https://www.guttmacher.org/in-the-know/abortion-providers.html).

Halfmann, Drew Thomas. 2011. "Recognizing Medicalization and Demedicalization: Discourses, Practices, and Identities." *Health:* 1363459311403947.

Hancock, Ange-Marie. 2003. "Contemporary Welfare Reform and the Public Identity of the 'Welfare Queen.'" *Race, Gender & Class* 10(1):31–59.

Harris, Kathleen Mullan. 2009. The National Longitudinal Study of Adolescent to Adult Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007-2009 [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. DOI: 10.3886/ICPSR27021.v9

Harris, K. M. Et al. 2009. "The National Longitudinal Study of Adolescent to Adult Health: Research Design [WWW Document]." Retrieved April 26, 2016 (http://www.cpc.unc.edu/projects/addhealth).

Henshaw, S.K., *1994-1995 National Survey of Contraception Use Among Women Having Abortions.* (Data Set O6 (Carley, M.L., Kelley, M.S., & Lang, E.L., Archivists). [machine-readable data file and documentation]. New York, NY: The Alan Guttmacher Institute (Producer). Los Altos, CA: Sociometrics Corporation, Data Archive on Adolescent Pregnancy and Pregnancy Prevention (Producer & Distributer)

Hertz, Rosanna. 2006. *Single by Chance, Mothers by Choice : How Women Are Choosing Parenthood without Marriage and Creating the New American Family: How Women Are Choosing Parenthood without Marriage and Creating the New American Family*. Oxford University Press, USA.

Hobcraft, John, Jane Menken, and Samuel Preston. 1982. "Age, Period, and Cohort Effects in Demography: A Review." *Population Index* 48(1):4–43.

Huntington, Dale, Barbara Mensch, and Nahid Toubia. 1993. "A New Approach to Eliciting Information about Induced Abortion." *Studies in Family Planning* 24(2):120–24.

Jagannathan, Radha. 2001. “Relying on Surveys to Understand Abortion Behavior: Some Cautionary Evidence.” *American Journal of Public Health* 91(11):1825–31.

Jones, Elise F. And Jacqueline Darroch Forrest. 1992. “Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988.” *Demography* 29(1):113–26.

Jones, R.K. (2008). *2008 Survey of Abortion Patients*. [machine-readable data file and documentation]. New York, NY: The Alan Guttmacher Institute (Producer)

Jones, R.K., Darroch, J.E., and Henshaw, S.K. (2004). *2000-01 National Survey of Contraception Use Among women Having Abortions.* (Data Set R3, mckean, E.A., J.J.). [machine-readable data file and documentation]. New York, NY: The Alan Guttmacher Institute (Producer). Los Altos, CA: Sociometrics Corporation, Data Archive on Adolescent Pregnancy and Pregnancy Prevention (Producer & Distributer)

Jones, Rachel K. And Megan L. Kavanaugh. 2011. “Changes in Abortion Rates Between 2000 and 2008 and Lifetime Incidence of Abortion.” *Obstetrics & Gynecology* 117(6):1358–66.

Jones, Rachel K. And Kathryn Kost. 2007. “Underreporting of Induced and Spontaneous Abortion in the United States: An Analysis of the 2002 National Survey of Family Growth.” *Studies in Family Planning* 38(3):187–97.

Kumar, Anuradha, Leila Hessini, and Ellen M. H. Mitchell. 2009. “Conceptualising Abortion Stigma.” *Culture, Health & Sexuality* 11(6):625–39.

Lara, Diana, Jennifer Strickler, Claudia Díaz Olavarrieta, and Charlotte Ellertson. 2004. “Measuring Induced Abortion in Mexico A Comparison of Four Methodologies.” *Sociological Methods & Research* 32(4):529–58.

Lehrer, Evelyn L. 2004. “The Role of Religion in Union formation: An Economic Perspective.” *Population Research and Policy Review* 23(2):161–85.

Logel, Christine et al. 2009. “Interacting with Sexist Men Triggers Social Identity Threat among Female Engineers.” *Journal of Personality and Social Psychology* 96(6):1089–1103.

Loretta J. Ross. 1993. “African American Women and Abortion: 1800-1970.” In *Theorizing Black Feminisms: The Visionary Pragmatism of Black Women*. London: Routledge.

Lupton, Joseph and James Smith P. 2003. “Marriage, Assets, and Savings.” In *Marriage and the Economy: Theory and Evidence from Advanced Industrial Societies*, edited by S. Grossbard-Shechtman. Cambridge University Press.

Mcmahon, Martha. 1995. *Engendering Motherhood: Identity and Self-Transformation in Women’s Lives*. Guilford Press.

Nash, Elizabeth, Rachel Benson Gold, Gwendolyn Rathbun, and Yana Vierboom. 2015. “Laws Affecting Reproductive Health and Rights: 2014 State Policy Review.” *Guttmacher Institute*. Retrieved January 22, 2017 (https://www.guttmacher.org/laws-affecting-reproductive-health-and-rights-2014-state-policy-review).

National Center for Health Statistics and The National Bureau of Economic Research. 2016. “NCHS’ Vital Statistics Natality Birth Data.” Retrieved December 14, 2016 (http://www.nber.org/data/vital-statistics-natality-data.html).

Newport, Frank, David W. Moore, and Lydia Saad. 1999. “Long-Term Gallup Poll Trends: A Portrait of American Public Opinion Through the Century.” *Gallup.com*. Retrieved January 10, 2017 (http://www.gallup.com/poll/3400/longterm-Gallup-Poll-Trends-Portrait-American-Public-Opinion.aspx).

O’Brien, George Dennis. 2010. *The Church and Abortion: A Catholic Dissent*. Rowman & Littlefield Publishers.

Pazol, Karen, Andreea A. Creanga, and Suzanne B. Zane. 2012. “Trends in Use of Medical Abortion in the United States: Reanalysis of Surveillance Data from the Centers for Disease Control and Prevention, 2001–2008.” *Contraception* 86(6):746–51.

Pew Research Center. 2016. “Public Opinion on Abortion.” *Public Opinon on Abortion*. Retrieved April 28, 2016 (http://www.pewforum.org/2016/04/08/public-opinion-on-abortion-2/).

Preston, S. H., P. Heuveline, and M. Guillot. 2001. *Demography — Measuring and Modelling Population Processes*. Malden, MA: Blackwell Publishing.

Rasch, V., H. Muhammad, E. Urassa, and S. Bergström. 2000. “Self-Reports of Induced Abortion: An Empathetic Setting Can Improve the Quality of Data.” *American Journal of Public Health* 90(7):1141–44.

Reese, Ellen. 2001. “The Politics of Motherhood: The Restriction of Poor Mothers‘ Welfare Rights in the United States, 1949–1960.” *Social Politics: International Studies in Gender, State & Society* 8(1):65–112.

Ross, Catherine E., John Mirowsky, and Karen Goldsteen. 1990. “The Impact of the Family on Health: The Decade in Review.” *Journal of Marriage and the Family; Minneapolis* 52(4):1059.

Russo, Nancy Felipe. 1976. “The Motherhood Mandate.” *Journal of Social Issues* 32(3):143–53.

Scheepers, Daan and Naomi Ellemers. 2005. “When the Pressure Is up: The Assessment of Social Identity Threat in Low and High Status Groups.” *Journal of Experimental Social Psychology* 41(2):192–200.

Smith, Lauren B., Nancy E. Adler, and Jeanne M. Tschann. 1999. “Underreporting Sensitive Behaviors: The Case of Young Women’s Willingness to Report Abortion.” *Health Psychology* 18(1):37–43.

Snow, David A. And Leon Anderson. 1987. “Identity Work Among the Homeless: The Verbal Construction and Avowal of Personal Identities.” *American Journal of Sociology* 92(6):1336–71.

Steele, Claude M., Steven J. Spencer, and Joshua Aronson. 2002. “Contending with Group Image: The Psychology of Stereotype and Social Identity Threat.” Pp. 379–440 in, vol. 34, edited by B.-A. In E. S. Psychology. Academic Press. Retrieved February 21, 2016 (http://www.sciencedirect.com/science/article/pii/S0065260102800090).

Tennekoon, Vidhura. 2017. “Counting Unreported Abortions: A Binomial-Thinned Zero-Inflated Poisson Model.” *Demographic Research* 36(2):41–72.

Udry, J. R., M. Gaughan, P. J. Schwingl, and B. J. Vandenberg. 1996. “A Medical Record Linkage Analysis of Abortion Underreporting.” *Family Planning Perspectives* 28(5):228–31.

Uecker, Jeremy E. And Charles E. Stokes. 2008. “Early Marriage in the United States.” *Journal of Marriage and Family* 70(4):835–46.

Umberson, Debra. 1992. “Gender, Marital Status and the Social Control of Health Behavior.” *Social Science & Medicine* 34(8):907–17.

United States Census Bureau. 2016. “International Data Base: Mid-Year Population by Single Year Age Groups - Custom Region - United States.” Retrieved January 2, 2017 (https://www.census.gov/population/international/data/idb/region.php?N=%20Results%20&T=15&A=separate&RT=0&Y=1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,2005,2006,2007&R=-1&C=US).

U.S. Department of Health and Human Services. 2007. “2007 HHS Poverty Guidelines.” *ASPE*. Retrieved January 24, 2017 (https://aspe.hhs.gov/2007-hhs-poverty-guidelines).

Waite, Linda J. And Evelyn L. Lehrer. 2003. “The Benefits from Marriage and Religion in the United States: A Comparative Analysis.” *Population and Development Review* 29(2):255–75.

Yang, Yang and Kenneth C. Land. 2013. *Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications*. CRC Press.