

# CHAPTER TWO

## Data and Methods

Counting Gay Men and Lesbians  
in Census 2000

Differences between 1990  
and 2000 Counts

Undercount of Same-Sex Unmarried  
Partners in Census 2000

Potential Measurement Error  
within the Same-Sex  
Unmarried-Partner Population

Summary: Who Counts  
and Who Doesn't?

## Data and Methods

Data from Census 2000 are used for the construction of all maps and for the majority of the analyses presented in this book. Of all government data collection processes, none has the weight of this household survey. Census 2000 is the latest decennial census administered by the U.S. Census Bureau for the constitutionally mandated purpose of apportioning congressional seats among the 50 states. In addition to its mandated purpose, the decennial census is also used to determine the appropriate distribution of government funding, draw state legislative districts, evaluate the success of programs, identify populations in need of services, and for a host of other functions (U.S. Census Bureau 2002).

During the census collection period in spring 2000, each housing unit in the 50 states, the District of Columbia, and Puerto Rico received a census questionnaire with six basic questions about each person in the household: name, sex, age, relationship to the householder, Hispanic origin, and race. The householder (the person filling out the census form) was also asked whether the housing unit was rented or owned. These seven questions make up what is commonly called the census "short form."<sup>1</sup>

The specific sources of primary data for the analyses presented in the *Atlas* are as follows:

- Summary File 1 (SF 1): based on short-form questionnaire responses from 100 percent counts of households (including same-sex unmarried partners)
- Summary File 2 (SF 2): based on short-form questionnaire responses from 100 percent counts of households (including same-sex unmarried

- partners), broken down by the race/ethnicity of the householder
- **Summary File 3 (SF 3):** based on long-form questionnaire responses from one in six households
- **Census Special Tabulation:** based on short-form questionnaire responses from 100 percent counts of households as reported in a Census Bureau special tabulation of county-level counts of same-sex unmarried partner households with children under age 18 in the household and counts of individuals within same-sex unmarried partner couples by age

### Counting Gay Men and Lesbians in Census 2000

The census does not ask any questions about sexual orientation, sexual behavior, or sexual attraction, three common ways used to identify gay men and lesbians in surveys (see Laumann, Gagnon, and Michael 1994). Rather, census forms include a number of relationship categories to define how individuals in a household are related to the householder. These fall into two broad categories: related persons (including husband/wife, son/daughter, brother/sister, and so on), and unrelated persons. Since 1990, the Census Bureau has included an "unmarried partner" category to describe an unrelated household member's relationship to the householder.<sup>2</sup> If the householder designates another adult of the same sex as his or her unmarried partner, the household counts as a same-sex unmarried-partner household. The actual question from the census form is shown in figure 2.1.

Research comparing 1990 Census data on same-sex unmarried partners and data from other surveys

Figure 2.1. Census 2000 Household Roster

How is this person related to Person 1?  
 Mark  ONE box.

Husband/wife  
 Natural-born son/daughter  
 Adopted son/daughter  
 Stepson/stepdaughter  
 Brother/sister  
 Father/mother  
 Grandchild  
 Parent-in-law  
 Son-in-law/daughter-in-law  
 Other relative—Print exact relationship.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

If NOT RELATED to Person 1:

Roomer, boarder  
 Housemate, roommate  
 Unmarried partner  
 Foster child  
 Other nonrelative

Source: Census 2000.

provides strong evidence that same-sex unmarried partners counted by the census are by and large gay and lesbian couples (Black et al. 2000).

### Differences between 1990 and 2000 Counts

Same-sex unmarried partners were enumerated, or determined in the official count, differently between the 1990 and 2000 Census data collection processes. The main difference stems from the editing, or data

cleaning, procedures used by the Census Bureau following data collection.

When information about a person or household is missing or inconsistent, the Census Bureau performs a variety of imputations to make the data both consistent and complete. Consistency edits occur when an obvious inconsistency is present in the data, for example if a 4-year-old male is enumerated as the "husband" of a 35-year-old female. Census might edit this record by either altering the age of the male or changing his status to "child," depending on other information provided by the household. Consistency edits are not flagged in the data, and any records of these edits are internal to the Census Bureau.

It is possible that a gay or lesbian couple could consider themselves "married," based on their own interpretation of that social construct, even though at the time the census was collected, no state government officially recognized marital unions between two people of the same sex (although Vermont recognized a category called "civil unions" that provides many of the same rights as those afforded to married couples). Since no federal or state law recognizes same-sex marriages,<sup>3</sup> if a gay or lesbian household designates his or her same-sex partner as a "husband/wife," the Census Bureau could not report these households as same-sex married couples.

During the post-collection data editing for the 1990 Census, the Census Bureau treated most same-sex "married" couples as an inconsistency in the sex variable for the "husband/wife," and usually changed the sex as a consistency edit. This means that in data released by the Bureau the couple was counted as a heterosexual married couple.

Due in part to lobbying by demographers and gay and lesbian interest groups, the Census Bureau changed their post-collection data editing procedures in Census 2000 to treat the issue as an inconsistency

in the relationship to householder rather than in the spouse's sex. That is, the "husband/wife" relationship designation was changed as a consistency edit to an "unmarried partner" relationship. Since the sex variables were not changed, the couple was counted as a same-sex unmarried partner couple.<sup>4</sup>

This change in editing procedures constitutes one factor in the dramatic increases in the counts of same-sex unmarried partners between 1990 and 2000. Large-scale publicity efforts made by national gay and lesbian groups along with increased willingness of gay men and lesbians to identify themselves as partners likely also account for the large increases.

### Undercount of Same-Sex Unmarried Partners in Census 2000

Despite the changes in the Census Bureau's data editing procedures, significant undercount issues remain in Census 2000's count of same-sex unmarried partners. There are several potential reasons for this undercount. Concerns about the confidentiality of their responses may have led many gay and lesbian couples to indicate a status that would not provide evidence of the true nature of their relationship.<sup>5</sup> Other couples may have felt that "unmarried partner" or "husband/wife" does not accurately describe their relationship. A study of the undercount of same-sex unmarried partners in Census 2000 indicates that these were the two most common reasons behind why some gay and lesbian couples chose not to designate themselves as unmarried partners (Badgett and Rogers 2003).<sup>6</sup>

#### *Estimate of the Undercount*

Estimating the size of any potential undercount in the census data is a challenge. A rough estimate of

*The increased willingness of gay men and lesbians to identify themselves as partners likely accounts for the large increase in the counts of same-sex unmarried partners between 1990 and 2000.*

the severity of the undercount of same-sex unmarried partners in Census 2000 can be made by combining estimates of coupling rates among gay men and lesbians and estimates of the prevalence of gay and lesbian individuals in the United States.

Black et al. (2000) measured same-sex coupling rates for two different definitions of "gay" and "lesbian" in the National Health and Social Life Survey (NHLS) data—identifying as gay or lesbian and having had exclusively same-sex sexual partners in the last year. The coupling rate for men ranged from 18.4 percent among those who identified as gay to 28.6 percent among those who had exclusively same-sex sexual partners in the last year. The comparable rates for women were 41.6 percent and 43.8 percent. Clearly, these data suggest that coupling behavior may be quite different for gay men and lesbians. Nonetheless, taking the midpoint of these estimates suggests that 23.5 percent of gay men and 42.7 percent of lesbians are coupled.

Prevalence estimates of the proportion of men and women in the United States who are gay or lesbian are more difficult to obtain. The size of these estimates varies tremendously with how homosexuality is defined (attraction, behavior, or identity). In perhaps the most well-respected recent work on the topic, Laumann et al. (1994) find that 2.8 percent of men and 1.4 percent of women self-identify as homosexual. Chapter 3 explores in greater detail how assumptions of various coupling rates among gay men and lesbians along with different levels of an undercount provide estimates of the size of the total gay and lesbian population. Using an undercount of approximately 25 percent yields estimates of the total gay and lesbian population that are most similar to Laumann's 1994 estimates, although the actual undercount could be higher if Laumann's data also undercount gay men and lesbians.

## Potential Measurement Error within the Same-Sex Unmarried-Partner Population

While the existence of an undercount is quite likely, an equally relevant issue is the possibility that some portion of same-sex unmarried-partner couples might be incorrectly designated as such due to a miscoding of either the "unmarried partner" relationship status or the sex of one of the partners. There are a number of ways a household could be classified in the census data as a same-sex unmarried partner household even though it is not headed by a gay or lesbian couple.

### *Sex Miscoding among Heterosexual Married Couples*

The largest source of measurement error among the same-sex unmarried partner data from Census 2000 is likely a result of sex miscoding errors among heterosexual couples. It can be assumed that some very small fraction of the population makes an error when completing the census form and possibly miscodes a variety of responses, including the sex of the householder or the householder's "husband/wife" or "unmarried partner." Under Census 2000 editing procedures, all these miscoded couples would be included in the counts of same-sex unmarried partners.

Recall that in 1990, married couples that inadvertently checked the wrong sex for the householder or the "husband/wife" most likely had the sex of the "husband/wife" changed. However, in Census 2000, these sex-miscoded couples are now counted as same-sex unmarried partners. Because the ratio between married couples and same-sex couples is so large (roughly 90 to 1), even a small fraction of sex miscoding among married couples adds a sizable fraction of heterosexual married couples to the same-

sex unmarried-partner population, possibly distorting some demographic characteristics, particularly child rearing. While this same error could occur among heterosexual unmarried partners, the smaller ratio between them and same-sex unmarried partners greatly reduces the effects of this form of measurement error on the same-sex couple population.

Ascertaining the extent of sex miscoding is challenging, as the most recent published census study involved the 1970 Census (U.S. Census Bureau 1975). We are grateful to Census Bureau officials, who conducted a detailed internal analysis of Census 2000 data to provide some insights into the possible extent of the measurement error problem among same-sex unmarried-partner couples. Their analyses suggest a sex-miscoding rate no higher than 0.2 percent among heterosexual couples, and even that is likely an upper-bound error rate. Further, they have determined that the sex-miscoding measurement error does not have any significant effect on geographical distribution patterns. As a result, the data used for the construction of all maps and charts in this book concerning geographic distribution of all same-sex couples do not make any corrections for possible measurement error.

However, this form of error could have a much more pronounced effect on data regarding same-sex couples with children under the age of 18, since married couples (the primary source of the measurement error) are much more likely to have children in the home than gay or lesbian couples. These sex-miscoded couples' inadvertent presence among the same-sex unmarried partners could artificially inflate the proportion of same-sex couples with children. The Census Bureau found that 45.6 percent of married couples reported children in their household in Census 2000, compared with 21.8 percent of gay male couples and 32.7 percent of lesbian couples.<sup>7</sup>

This translates to roughly 24.8 million married couples with children and about 162,000 same-sex partners with children; in other words, there are over 150 times as many married couples with children as there are same-sex partners with children. This ratio highlights that even a small sex-miscoding rate of 0.2 percent among the married couples would mean that 49,600 of the 162,000 same-sex unmarried partners with children (or nearly 31 percent) are actually married couples with children.

To at least partially correct for this measurement error, counts in all analyses involving same-sex coupled households with children are adjusted assuming that 0.2 percent of heterosexual couples miscoded the sex of one partner. That is to say, 0.2 percent of married couples with and without children (data on the counts of heterosexual unmarried partners with children were not available) are subtracted from counts of same-sex couples with and without children, respectively. Census Bureau officials concur that the 0.2 percent sex-miscoding rate represents an upper-bound estimate of this error, so calculations for the analyses generate lower-bound estimates of the proportion of same-sex unmarried partners who have children under age 18 present in the household.

### *Other Sources of Measurement Error*

Mistakes in the designation of an unmarried partner could also cause errors. One essentially undetectable form of error (discussed at length in Black et al. 2000) occurs when the person filling out the census form (the householder) does not have a spouse or unmarried partner in the household, but does have a child or other adult in the household living with an unmarried partner. For example, if a female householder classifies the female unmarried partner of her son as an "unmarried partner," then this household

*46 percent of married couples reported children in their household in Census 2000, compared with 22 percent of gay male couples and 33 percent of lesbian couples.*

would be counted as a female same-sex unmarried-partner, or lesbian, household.

While analysis of 1990 Census data suggests that this type of error has negligible effects on the quality of the data at a national level, it could be more common in analyses of certain communities where extended families are more likely to be living in the home and households are larger. For example, Hispanic and American Indian populations are more likely to have extended families living in the home. Communities with large Hispanic and American Indian populations are therefore more susceptible to this type of error because there are proportionally more households where the error could occur. This form of measurement error would have the greatest effect on analyses of same-sex households with children because households most prone to this form of error (those with larger extended families in them) are also more likely to have children in the home. Short of examining the exact family structure within the home (an exercise only possible with more detailed census data), there is no simple correction for this form of measurement error.

Another form of measurement error could be language-based. Confusion may result when respondents fill out a census form not written in their native language or if the census enumerator translations of terms such as "unmarried partner" and "roommate" in other languages, particularly Spanish, do not have the same meanings as the English version. All households in the 50 states and the District of Columbia received English language forms, regardless of the predominant language spoken in either the household or the neighborhood. However, if a census form was not returned, most likely an enumerator who speaks local native languages would have visited the house and assisted the householder in filling out an English-language form.<sup>8</sup> Since 58 percent of Hispanic

households have children under 18 living in them, compared with only 36 percent of all households in the U.S., a measurement error in this population that results in additional mistaken coding of same-sex unmarried-partner households would have a disproportionate effect on statistics involving same-sex couples with children. Unfortunately, we cannot determine the exact nature of any language-related error, nor are we able to estimate the magnitude of such errors.

Reporting figures from regions with larger and more diverse populations helps minimize any disproportionate influence of these uncorrectable measurement-error issues. Therefore, analyses of same-sex couples with children are limited to the nation, states, and large metropolitan areas.

### Summary: Who Counts and Who Doesn't?

Despite these measurement-error issues, Census 2000 data on same-sex partners represent the most comprehensive source of data on gay and lesbian couples living in the United States. That said, many members of the gay, lesbian, bisexual, and transgender (GLBT) community are left out of these data. The largest omission is single gay men and lesbians, who form a majority of the GLBT community. Since only about a quarter of gay men and two-fifths of lesbians are coupled at any given time, over half of all gay and lesbian individuals in the United States are not included in census data counting same-sex couples.

Bisexual and transgender individuals will sporadically appear in the data, but cannot be identified as such. When they are part of a same-sex unmarried partner couple, they are counted as a gay or lesbian couple; however, there is no mechanism to separately identify bisexual or transgender individuals using

census data. Certainly some men and women who are currently living with an unmarried partner of their same sex identify as bisexual, but since there is no census question about sexual orientation self-identification, it is unknown what portion bisexuals make up among the population of same-sex unmarried partners. Likewise, transgender people may also appear in the same-sex couple population if they are living with an unmarried partner and both individuals report the same sex. In this case, they are counted as part of a gay or lesbian couple, although this is likely a very small portion of the same-sex unmarried partner counts.

#### *Endnotes*

1. In addition, one out of every six housing units received a census "long form," which has a much wider array of questions, including questions about individual ancestry, household income, mortgage/rent, length of residence, size of the housing unit, utility bills, and so on.
2. Other relationship categories for unrelated persons are "roomer/boarder," "housemate/roommate," "foster child," and "other nonrelative."
3. In fact, the federal Defense of Marriage Act (DOMA) of 1996 prohibits any federal agency from recognizing a same-sex cou-

ple as being married. The Massachusetts Supreme Judicial Court ruled that the state must begin issuing marriage licenses to same-sex couples in May 2004. Marriage licenses issued to same-sex couples in San Francisco in 2004 will likely face multiple court challenges questioning their validity.

4. For more information about differences between the 1990 and 2000 counts of same-sex unmarried partners, see the Census Bureau's "Technical Note on Same-Sex Unmarried Partner Data From the 1990 and 2000 Censuses," at <http://www.census.gov/population/www/cen2000/samesex.html>.
5. The short form collects the following pieces of identifying information: each household member's name, the exact address of the household, and the household's telephone number.
6. Another potential source of undercount is gay and lesbian couples who do not live together. The short form instructs you not to list "people who live or stay at another place most of the time." But this undercount issue would also apply to married couples who live apart, and we do not see it as a relevant source of undercount.
7. These statistics refer to the presence of children under the age of 18 who are sons or daughters of the householder. For our adjustments, we used statistics on the presence of children under the age of 18, regardless of the children's relationship to the householder.
8. The Spanish census form distributed in Puerto Rico translates "housemate/roommate" as "compañero(a) de casa/compañero(a) de cuarto" and "unmarried partner" as "compañero(a) no casado(a)." If enumerators used the word "compañero(a)" in both categories, this could have created some confusion in distinguishing between the two relationship types.