

# CHAPTER 1



## Same-Sex Partnering Data in the 2000 U.S. Census An Overview

In the 1990 decennial census, the U.S. Census Bureau added the “unmarried partner” category to the list of responses a respondent could choose to represent his or her relationship to the householder. The addition of this response, which was used also in the 2000 decennial census, permitted individuals to identify as same-sex unmarried partners, creating both a large and representative data-set with which to study issues of sexual orientation. As we noted in the Introduction, we use these data to undertake the demographic analyses presented in this book. In this chapter, we provide an introduction to the “unmarried partner” response category used in the census. We discuss some of the background leading to its first use in the 1990 decennial census, and cover some of the issues related to the empirical conceptualization of homosexuality and sexuality. Further, we demonstrate the manner in which the unmarried partner household data were categorized in the 2000 Census.

The unmarried partner data from the 2000 Census have certain limitations that should be taken into account when used in analyses of homosexual and heterosexual individuals. In this chapter, we both acknowledge and evaluate the impact of these limitations for research on the demography of sexual orientation. Although there are limitations, the census data nonetheless constitute the best and largest-ever data-set on same-sex and opposite-sex partners, permitting researchers to examine heretofore underexplored issues regarding sexual orientation. In the concluding section of this chapter, we employ these new data to describe some of the characteristics of partnered gay males and lesbians and compare them with married and cohabiting heterosexuals.

## A Short History of the Collection of Census Data on Cohabitation

In the early 1970s social scientists and the general public became increasingly aware of, and interested in, the phenomenon of heterosexual cohabitation, that is, persons of the opposite sex living together. The U.S. Census Bureau was the principal federal agency tasked to provide national data on the increasing numbers of cohabiters. Data from the decennial censuses and the Current Population Surveys (CPS) were used to develop estimates of the numbers of cohabiting adults in the United States.

Until the conduct of the 1990 decennial census and the 1995 CPS, however, there were no census questions that specifically asked respondents if they were nonmarital cohabiters. Consequently, the Census Bureau had to use an indirect approach to obtain cohabitation data. The Bureau defined cohabiting households as nonmarital if they contained “two and only two adults (age 15+) who are unrelated and of the opposite sex” (Casper and Cohen 2000: 237). The Census Bureau used such data to develop nationally representative estimates of cohabiters and their characteristics (Glick and Norton 1977; Glick and Spanier 1980; Glick 1984).

This operational definition of cohabiter, along with the published data, led to the concept of POSSLQ (Partners of the Opposite Sex Sharing Living Quarters), which became a cultural fixture in the 1980s and 1990s. Indeed, Casper and Cohen (2000) note that at least two books were published in the 1980s with POSSLQ in the titles (*There's Nothing That I Wouldn't Do If You Would Be My POSSLQ* [Osgood 1981]; and *Will You Be My POSSLQ?* [Bunting 1987]), and countless newspaper stories and scholarly articles used the POSSLQ data and estimates (Bumpass and Sweet 1989; Bianchi and Spain 1996; Vobejda 1998). The POSSLQ data and the resulting publications demonstrated concretely the increasing prevalence of nonmarital cohabitation in the 1970s and 1980s and helped identify and establish the phenomenon of living together as an emerging national trend. The POSSLQ estimates showed the numbers of cohabiting couples in the United States increasing from 968,000 in 1977 to nearly 2.9 million in 1990 (Casper and Cohen 2000: 239).

But the POSSLQ data were imperfect counts of the numbers of nonmarital opposite sex cohabiters living in marriage-like relationships (Casper and Cohen 2000). For one thing, since the definition was restricted to households with only two adults, it missed cohabiters sharing households with more than two adults. It also incorrectly included households containing adults who, though living together, were not cohabiting in romantic marriage-type relationships, such as roommates, boarders, or those living in other kinds of non-cohabiting relationships. To illustrate, in the 1980 census, one of the ways in which individuals could identify their “relationship to the head of household”

was “roommate.” Persons giving this response had no blood relationship to the householder, but may or may not have had a marriage-like (i.e., an emotional or romantic) relationship with the householder. The POSSLQ approach categorized as cohabiters persons of the opposite sex who were living together as roommates. Consequently, prior to 1990, “couples living outside marriage in marriage-like relationships were not identified separately from (unrelated) individuals living together as roommates” (Black et al. 2000: 140).

Due to both the growing numbers and interest in cohabiters, the Census Bureau recognized the need to provide direct and more precise data on nonmarital cohabitation. Therefore, in the mid-1980s, Donald J. Hernandez, then the Chief of the Marriage and Statistics Branch of the Population Division of the Census Bureau, in consultation with Arthur J. Norton, his supervisor, made the decision to add the response category of “unmarried partner” to the basic question dealing with one’s “relationship to the householder” (Hernandez 2006). This change took effect in the 1990 decennial census and in the 1995 Current Population Survey. This response was also included in the 2000 census and in subsequent Current Population Surveys and American Community Surveys.

The “unmarried partner” response was added to the list of other possible responses (husband, wife, son, grandfather, suitemate, boarder, etc.) to the census question pertaining to one’s “relationship to the householder,” that is, the person in the household designated as “person #1.” Person #1 is typically “the member of the household in whose name the home is owned, being bought or rented” (Barrett 1994: 16). Every person in the household, except for person #1, responds to a question about his or her relationship to person #1. The “unmarried partner” response enables the identification of persons in the household who are unrelated to person #1, but who have a “marriage-like” relationship with person #1. The official definition of an “unmarried partner” is “a person age 15 years and over, who is not related to the householder, who shares living quarters, and who has a close personal relationship with the householder” (U.S. Bureau of the Census 2004). The “relationship to the householder” question used in the 2000 Census has been reproduced in Figure 1.1.

With respect to data on gay males and lesbians, the federal government made no attempt to provide data of any type on this subpopulation prior to 1990. Indeed before the conduct of the 1990 U.S. Census, a large, national-level data-set for the lesbian and gay populations did not exist. The data sets then in existence were limited in scope and were based on smaller samples of the population, for example, the General Social Surveys. In the early 1980s, for instance, Castells and Murphy (1982: 238) wrote that “there is no statistical source that provides information on sexual preferences of residents of specific urban areas. . . . Such an obstacle appears overwhelming to the researcher trying to understand the spatial dynamics of the emerging gay culture.”



As just noted, this all changed in 1990 when the federal government decided that direct data were needed on heterosexual cohabitation and added the “unmarried partner” response to the “relationship” question. Prior to 1990, there was no direct way to use census, CPS data, or both to measure cohabitation; there was no direct way to identify persons in the household who were unrelated to the head of household but who had a marriage-like relationship with the householder. Fortunately, census procedures permitted respondents to check the “unmarried partner” response irrespective of whether their sex was the same as that of the householder. Thus, beginning in 1990, researchers have been able to use the unmarried partner data to obtain information on gay male and lesbian partnering. But what do these data truly convey regarding the identification and enumeration of partnered gay men and lesbians? To address this question we need to discuss issues involved in the conceptualization of sexual orientation.

### The Conceptualization of Sexual Orientation

Most of the social science literature on sexual orientation conceptualizes the phenomenon using two basic perspectives or approaches, or a combination thereof. These two views may be referred to as “essentialism” and “social constructionism” (Laumann et al. 1994: 284). Founded in biology, the essentialist view states that there is an essential characteristic common to all homosexual individuals that is distinct and separate from heterosexual individuals. This common characteristic, or essence, is thought to be based in biology or psychology, and is a fundamental drive or trait that establishes a person’s inclusion into either a homosexual or heterosexual category (Laumann et al. 1994: 285). The essentialist view of homosexuality presumes that a person may be categorized as being or not being homosexual and makes a distinction, often binary, between one who is a homosexual individual and one who is not. Thus, sexual orientation is determined by the definition of two distinct categories.

The social constructionist view of homosexuality counters and critiques the essentialist perspective. Social constructionism argues against the notion of binary categories, that is, that one either is or is not a homosexual individual (Foucault 1978; Butler 1990; Seidman 1996). Instead, this approach argues for a continuum with varying degrees of homosexuality and heterosexuality. Social constructionists point out that homosexual prevalence rates and visibility tend to vary across time and settings, and that the concepts, definitions, and practices of homosexuality are often not the same across context and cultures (Laumann et al. 1994: 285). What in one culture may be defined as “homosexual” may not be defined as such in another culture. For example, an individual may engage in same-sex sexual behavior but not identify him- or herself as a gay male or lesbian. Likewise, one might identify as a homosexual individual but never have experienced same-sex sex. Also, the sexuality definitions and

labels attached to individuals by other persons and by the larger society may be incongruent with how individuals per se self-identify.

When demographers use the census unmarried partner data, they are not necessarily taking an essentialist view. They might have a social constructionist view, or understanding, of sexuality. Nonetheless, the census data only permit sexuality to be measured on one dimension, and without variation. The use of these data involves, by definition, the employment of a clear-cut and straightforward definition of what is a partnered homosexual individual (Black et al. 2000; Smith and Gates 2001; Gates and Ost 2004; Walther and Poston 2004). Consequently, in most of the chapters of this book an essentialist approach, by definition, is applied; the data do not permit any alternative. The census data are essentialist in terms of the way the question has been formulated and the way it can be applied.

However, in other demographic and social science research on homosexuality, the manner in which homosexuality and sexual orientation are measured tends to vary. This is largely due to the different ways sexual orientation has been defined in surveys and conceptualized by researchers (Saewyc et al. 2004). Homosexuality may be defined in terms of sexual behavior, sexual desire (including fantasy), and self-identification (Laumann et al. 1994; Saewyc et al. 2004). In analyses based on data from national surveys, social scientists have used one or more of the above concepts of homosexuality, but particularly those based on self-identification and behavior.

Analyses of homosexuality using data from the General Social Survey (GSS), for instance, usually employ a behavioral definition of homosexuality, such as whether a person's sex partners within a particular timeframe (such as over the past 12 months, or the past five years, or in one's lifetime) have or have not been entirely or predominantly of the same sex as the respondent (Badgett 1995; Berg and Lien 2002; Black et al. 2003). The GSS does not include a question on the self-identification of the respondent's sexual orientation.

Researchers using data from other surveys are able to use a series of different definitions of homosexuality. To date, there are two national surveys conducted in the United States that include questions dealing with sexual behavior, sexual orientation, and sexual desire. One is the National Health and Social Life Survey (NHSL) conducted by Laumann and his associates in 1992 (see *The Social Organization of Sexuality: Sexual Practices in the United States* [1994]). The other is Cycle 6 of the National Survey of Family Growth conducted in 2002 by the National Center for Health Statistics (National Center for Health Statistics, 2004). Because these surveys allow researchers to define homosexuality in various ways, it is possible for their analyses to be more closely attuned to a social constructionist view rather than an essentialist one.

Overall, there are a few fundamental methodological limitations apparent from a review of the literature concerning the conceptualization and measurement of sexual orientation: the lack of common and consistent definitions in

surveys, problems with obtaining sufficiently representative sample sizes, and the lack of sexuality-related questions in large-scale data collections. These limitations may be somewhat related to the purported social stigma attached to homosexuality. This stigma is believed to affect not only the way questionnaires are designed to address or measure sexual orientation, but also the ways in which individuals will respond to survey questions about self-identification, behavior, and desire. For example, some may be reluctant to identify as homosexual or to report homosexual behavior (Laumann et al. 1994: 284).

No doubt there are problems with data on homosexual individuals no matter how the phenomenon is conceptualized and quantified. There are likely methodological limitations and problems inherent in gathering and analyzing data about the gay male and lesbian population, as with any stigmatized minority. Nonetheless, we find that the federal census data on same-sex unmarried partners can be useful if researchers are clear about to whom the data refer. In the next section of this chapter we consider specifically the unmarried partner data from the 2000 U.S. Census and present and discuss a classification of households within which census data on same-sex partnered households are generated. This will enable us to better visualize the makeup and general definition of the populations that are captured by the same-sex unmarried partner data in the 2000 census.

#### Same-Sex Partner Data from the 2000 Census

In 2000, almost 5.5 million unmarried partner households were enumerated in the U.S. Census. These were households in which the couples were living together but were not married. Of these 5.5 million unmarried partner households, almost 600 thousand were same-sex partner households; 301 thousand were male-male households, and 293 thousand were female-female households (see Table 1.1). The same-sex unmarried partner households were located throughout the United States in over 99 percent of all U.S. counties. The largest number (over 85 percent) resided in metropolitan areas (Simmons and O'Connell 2003: 2).

A sorting process exists whereby a household identified in the 2000 Census came to be designated as a same-sex unmarried partner household. By detailing this sorting process, we wish to highlight exactly which households are included and which are not included in the same-sex unmarried partner census data. First, it is important to note that the "relationship to householder" question that produces the same-sex census data is one of seven so-called 100 percent census questions that are asked of all persons who respond to the census. The 100 percent unmarried partner data are available in Table PCT 22 of Summary File 2 of the 2000 Census. This table provides for various levels of geography (e.g., states, counties, census tracts, block groups, etc.) the number

TABLE 1.1. Number of Unmarried Partner and Married Couple Households, by Metropolitan Status and Geographic Region,\* United States, 2000

Household Type	Number	% Metro	% Northeast	% Midwest	% South	% West
Total Households	105,480,101	79.9	19.2	23.4	36.0	21.3
Total Coupled Households	59,969,000	78.7	18.7	23.7	35.9	21.7
Unmarried Partner Households	5,475,768	81.4	19.7	23.0	33.0	24.3
Same-Sex Partners	594,391	85.3	20.1	17.8	35.3	26.9
Male-male	301,026	86.3	19.7	17.3	35.8	27.2
Female-female	293,365	84.2	20.4	18.3	34.8	26.5
Opposite-Sex Partners	4,881,377	80.9	19.6	23.6	32.3	24.0
Married Couple Households	54,493,232	78.5	18.6	23.8	36.2	21.4

\*Includes Puerto Rico Not Shown.

Source: Simmons and O'Connell, 2003, tables 1 and 2.



of households in which person #1 is a male and another male in the household identifies himself as the unmarried partner of person #1; these are known as male-male households. A similar tabulation is provided for the number of female-female households in each geographic area. As mentioned earlier, because the “unmarried partner” response is meant to reflect a “marriage-like” relationship between the two same-sex persons, researchers make the assumption that these data on same-sex households (male-male or female-female) represent households inhabited by partnered gay males, or by partnered lesbians (Black et al. 2000, 2002b; Simmons and O’Connell 2003; Walther and Poston 2004; Gates and Ost 2004). The research we report in Chapter 4 of this book adds further support to this conclusion.

To show how a household ends up being classified as a same-sex household, however, we do not use the 100 percent data from Table PCT 22 of Summary File 2. These are aggregated data and do not permit the statistical manipulation of individual cases. Instead we use data from the five percent Public Use Microdata Sample (PUMS) from the 2000 Census. Following the approach used by Black and colleagues (2000) with 1990 census data, we present in Figure 1.2 a classification of households from the 2000 Census. Although the data shown in Figure 1.2 are drawn from the 5 percent PUMS, we have used sample weights to inflate the numbers to their estimated national levels.

Figure 1.2 shows two boxes with bolded outlines that represent two types of household relationships: those without a “marriage-like” relationship (A), and those with a marriage-like relationship (B). Each of the household relationship boxes is divided into subgroups. The subgroups with bolded outlines, strictly speaking, represent subgroups directly observable in the census data. Subgroups with dotted outlines are not observable in the census data.

The 5 percent PUMS data consist of 4,710,069 households containing single adults or two adults. Inflating these to national levels using their sample weights produces an estimated total number of 94,485,532 households. Excluded in this estimate are all households where the household relationship was imputed; also excluded are households with “multiple-marriage-like relationships.” We begin our classification with these more than 94 million households.

In Figure 1.2 we show that of the more than 94 million households, almost 36 million are households with no marriage-like relationship (A). (Excluded here are households with three or more adults.) The majority of these 36 million households, over 27 million, are single adult households (A.1). These single adult households contain either single gay males and lesbians (A.1.1) or single heterosexuals (A.1.2). But since the census does not contain a question asking about the respondent’s sexual orientation, the numbers of households in boxes A.1.1 and A.1.2 cannot be identified in the census data. This is a shortcoming of the data. Census data do not capture gay men and lesbians who live alone.

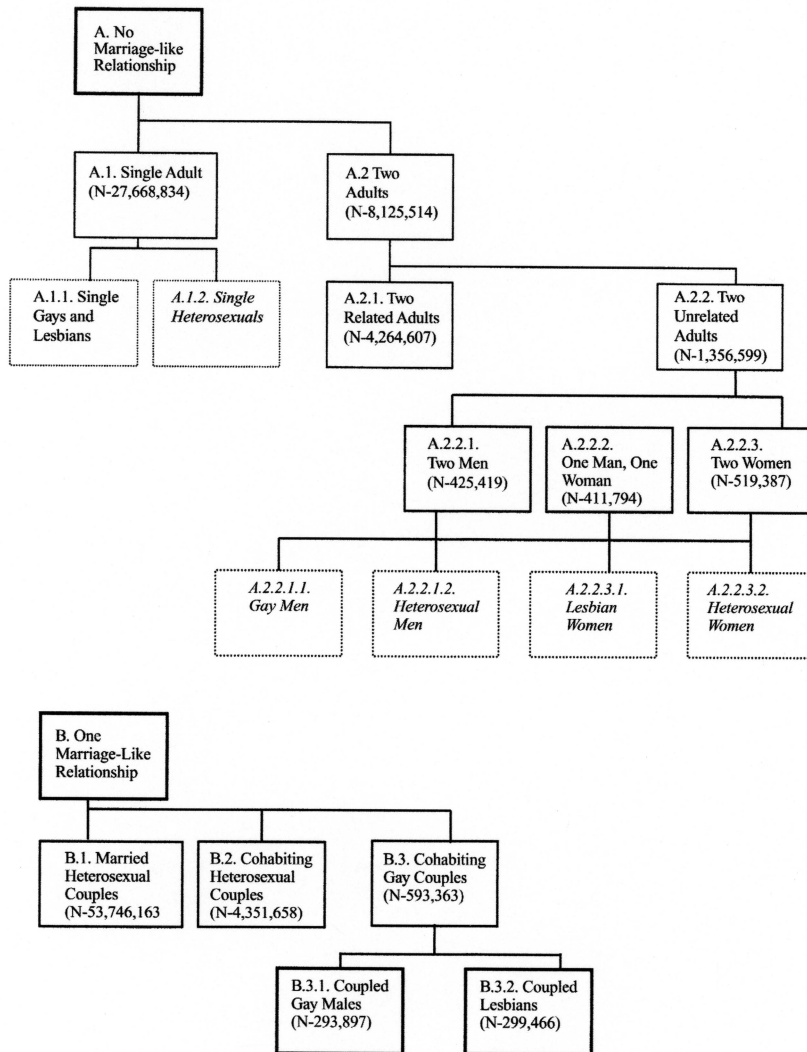


FIGURE 1.2. A Classification of 94,485,532 Households by Marriage-like Relationships for Homosexual and Hetersexual Males and Females: U.S., 2000

The remaining 8 million households are two adult households (A.2). These consist of over 4 million households containing two related adults (A.2.1) and almost 1.4 million households containing two unrelated adults (A.2.2). The box representing households with two related adults (A.2.1) are households with, for instance, two siblings, or with an uncle and a nephew, or with some other combination of two related adults. We do not use this group of households in our analyses.

We are particularly interested in the nearly 1.4 million households containing two unrelated adults (A.2.2). These households may be further subdivided into households with two men (A.2.2.1), one man and one woman (A.2.2.2), and two women (A.2.2.3). These three subgroups contain individuals who are either gay men (A.2.2.1.1), heterosexual men (A.2.2.1.2), lesbians (A.2.2.3.1), or heterosexual women (A.2.2.3.2). But there is no way to make these distinctions with census data because, as already mentioned, the census does not ask about sexual orientation. Thus, gay men or lesbians who are living with other men or women, but not in marriage-like relationships, are also not enumerated in the census.

Of the over 94 million households represented in the classification, almost 59 million are households with one marriage-like relationship (B). These may be subdivided into three subgroups: households with married heterosexual couples (B.1), households with cohabiting heterosexual couples (B.2), and households with cohabiting homosexual couples (B.3). The over 593 thousand cohabiting homosexual couples may be further subdivided into households with cohabiting gay men (B.3.1) and households with cohabiting lesbians (B.3.2). Data on these two groups of same-sex cohabiters shown in boxes B.3.1 and B.3.2 are used in this book's analyses of the social demographic patterns and dynamics of same-sex partners.

### Quality of the Gay Male and Lesbian Partner Data

Although the Census Bureau instructs that the "unmarried partner" category is indicative of a marriage-like relationship, many have questioned whether the respondents understand the implications of this response and, as a result, whether the data truly reflect a homosexual relationship. In the absence of a direct question asking about sexual orientation, a number of concerns have arisen regarding the use of the same-sex unmarried partner data for the purposes of studying issues of homosexuality. Before presenting analyses using these data, therefore, we believe it is important to address these methodological issues in an attempt to appraise the quality of the same-sex partnering census data. First, we ask about the accuracy of the 2000 Census data in portraying the true numbers of partnered gay men and lesbians. Specifically, how well have the 2000 Census data on same-sex partners enumerated the actual numbers of partnered

gay men and lesbians living in the United States in 2000? A second issue concerns the variation across the geographical areas of the United States in the prevalence of same-sex unmarried partners. How valid is this variation? Is there a relationship between this variation and the variation across geographical areas in the true numbers of partnered gay men and lesbians? A third issue concerns the extent to which there could be error in the same-sex partnering census data due to sex miscoding errors. We then conclude this discussion by introducing two reasons that give us further cause to have confidence in the validity of the same-sex partner data.

We first address the validity of the census data on same-sex partners. To do this, we need to know the true numbers of gay men and lesbians living in the U.S. in 2000. There are no such numbers available. However, the numbers may be estimated with data from a national survey that contains sexuality questions dealing with both self-identification and behavior; specifically, Cycle 6 of the National Survey of Family Growth (NSFG) conducted by the U.S. Department of Health and Human Services in 2002 (National Center for Health Statistics 2004). This is a survey of 12,571 persons aged 15–44 in the noninstitutional population. The male and female respondents were asked questions about same-sex behavior and sexual identification. We selected persons self-identifying as gay men or lesbians and those who reported having exclusively same-sex sex partners in the past 12 months. We reasoned that these two characteristics best typify persons who would likely be captured as same-sex unmarried partners in the census data. We combined the two groups and developed an estimate of the percentages of gay men and lesbians in the United States in 2002 self-identifying as homosexual individuals, or engaging in exclusively homosexual behavior in the past year, or both.

Using weighted NSFG data, we determined that 2.55 percent of males may be classified as gay, and that 1.81 percent of females may be classified as lesbian. The upper and lower 95 percent confidence bounds for the males are 3.2 percent and 2.1 percent, and for the females, 2.2 percent and 1.5 percent. These percentage estimates and 95 percent confidence intervals obtained from the 2002 NSFG are remarkably close to male and female homosexuality estimates and confidence intervals obtained from the only other nationally representative survey of the U.S. population that asked the same two sexuality questions, namely, the National Health and Social Life Survey (NHSLs) conducted in 1992 by Laumann and colleagues (Laumann et al. 1994).

As noted, the NSFG gay male and lesbian estimates pertain to persons aged 15 to 44. The U.S. population of males and females aged 15 to 44 counted in the 2000 Census consists of 62,647,145 males and 62,026,997 females. When we multiply these numbers by the NSFG percentages of gay men and lesbians, we obtain estimates of the total numbers of gay men and lesbians in the United States between the ages of 15 and 44, namely, 1,597,502 and

1,122,689, respectively. But how many of these gay men and lesbians are living in committed relationships in the same households?

Gates and Ost (2004: 13) have reviewed several studies to arrive at estimates “that 23.5 percent of gay men and 42.7 percent of lesbians are coupled.” Using their percentage coupling figures, we estimate that in the U.S. in 2000, there were 375,413 gay men in committed relationships living in the same households (that is,  $1,597,502 \times 23.5\%$ ), and 479,388 committed lesbians living in the same households in the United States (or  $1,122,689 \times 42.7\%$ ), all between the ages of 15 and 44.

Using data from the 2000 PUMS and the corresponding person weights, we next determined that Census 2000 enumerated 334,220 same-sex male partners and 345,571 same-sex female partners between the ages of 15 and 44. Comparing these figures with the NSFG-based estimates of the numbers of partnered gay men and lesbians suggests that Census 2000 undercounted 41,193 committed gay men living in the United States, for an undercount of 11.0 percent, and undercounted 133,817 committed lesbian partners, for an undercount of 27.9 percent.

There are many problems with these estimates. For one thing, although the census questionnaire asks about identification, the identity pertains to whether or not one is in an unmarried partnership. We have already noted that the census questionnaire does not include a question asking specifically about the sexual orientation, or the sexual behavior, of the respondents. As have other researchers (Black et al. 2000, 2003; Simmons and O’Connell 2003; Walther and Poston 2004; Gates and Ost 2004), we assume that the census numbers of same-sex male and female partners reflect the numbers of committed gay men and lesbians in the population. Since there are no national-level data available on gay male and lesbian commitment rates and different studies report different estimates, we employed the male and female averages of the various studies developed by Gates and Ost (2004). We acknowledge that these estimates of partnership are, of course, less than ideal.

Despite the problems associated with our population estimates, however, these results are fairly consistent with conclusions reached by other scholars who have also found that committed gay men and lesbians were undercounted in the Census 2000 (Smith and Gates 2001; Badgett and Rogers 2003; Gates and Ost 2004). Indeed, even when widely varied methodologies have been employed to ascertain the validity of the census data, undercount estimates have been surprisingly consistent. In one instance, rather than comparing census results with past nationally representative surveys, researchers at the Institute for Gay and Lesbian Strategic Studies conducted two surveys to determine the use of the unmarried partner response by same-sex couples (Badgett and Rogers 2003). These were nonrepresentative surveys of the gay male and lesbian population; one was conducted of participants attending the gay rights Millennium

March in Washington, DC in April 2000, and the other involved an online survey that included 90 individuals who were in same-sex partnerships at the time of the 2000 census. Both of these samples were likely biased to include an oversampling of individuals identifying as unmarried partners on the census form, since such participants are more likely to be politically active and/or comfortable in disclosing their sexual orientation. Approximately 81 to 84 percent of same-sex partners participating in these surveys either chose the unmarried partner category in the 2000 census, or identified as married and would have been placed in the unmarried partner category by the Census Bureau (Badgett and Rogers 2003). Consequently, the estimated undercount from these surveys for all same-sex partners fell between 16 and 19 percent. This undercount estimate is similar to those presented in our analyses.

We now turn to the second question, that dealing with the validity of the gay and lesbian prevalence indexes across the geographic areas of the United States. How valid is this variation? There are no reliable data available to answer this question because there are no data other than census data “for calculating even the most rudimentary statistics on the [geographic] locations of the gay and lesbian populations” (Black et al. 2000: 149). However, it is possible to examine the face validity of the census-developed geographical distribution data of the partnered gay population by relating its variation with that of the spatial distribution of AIDS deaths. Unfortunately, data on AIDS deaths are only available for large metropolitan areas. Thus, we are only able to compare the variation in AIDS deaths with census-based prevalence rates of the partnered gay population for large metropolitan areas.

We are well aware, and wish to emphasize the point, that AIDS deaths are not restricted to homosexual individuals. Indeed since the 1990s in the United States there have been increasing numbers of heterosexual deaths due to AIDS. But AIDS as a cause of death continues in the United States to be the most prominent for men who have sex with men (as do most gay men) than for the heterosexual population (CDC 2004; Kaiser Family Foundation 2004). One would thus expect that among geographical areas there should be a positive association between the prevalence of gay men and the prevalence of AIDS deaths.

We obtained data on the reported number of AIDS cases for the 12-month period between July 1998 and June 1999 for the 99 metropolitan areas of the United States with populations over 500,000. Similar data are not available for smaller metropolitan areas, or for nonmetropolitan areas. We first examine this relationship in a relative way by correlating the rates of partnered gays (per 1,000 unmarried males) with the rate of AIDS cases per 100,000 persons in the area. The correlation between the two rates across the 99 metropolitan areas is .52. The relationship is positive and strong.

The actual number of reported AIDS cases and the actual number of male-male households among the 99 metropolitan areas are highly skewed so we used

their natural logs in a second examination of the relationship. The correlation between the logged values of number of AIDS deaths and number of male-male households is 0.86, indicating a strong positive correlation. Similar comparisons conducted with census and AIDS data for 1990 produced comparable high positive correlations (Black et al. 2000; Walther and Poston 2004).

These tests increase our confidence in the quality of the same-sex partner data obtained in the 2000 decennial census, particularly the validity of the geographical distribution of these data in large metropolitan areas. As already noted, there are no similar data on AIDS deaths for the smaller metropolitan areas or for nonmetropolitan areas. However, the fact that the two variances are so closely related in the larger metropolitan areas gives us reason to believe that it is likely that the variances for other geographical areas would also be closely related.

In addition, there are no similar data available for examining the face validity of the partnered lesbian data from the 2000 Census. We do show in Chapter 2 of this book, however, that the partnered gay male rates and partnered lesbian rates are themselves highly and positively related (also see, Black et al. 2000; and Walther and Poston 2004). This provides some indication of the face validity of the lesbian data. Specifically, if the AIDS death rates support the face validity of the partnered gay male data, and the partnered gay male rates are highly related to the lesbian rates, then logic suggests that the partnered lesbian data also have face validity.

The third issue to be addressed is the degree to which there could be error in the same-sex partner data, perhaps due to individuals miscoding their sex. In the 1990 census, if a same-sex couple indicated that their relationship was that of married, postcollection census editing treated this as an inconsistency, and “usually changed the sex as a consistency edit. This means that in data [for 1990] released by the Bureau the couple was coded as a heterosexual married couple” (Gates and Ost 2004: 12). The Bureau changed this postcollection editing decision in the 2000 Census to treat it “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 [as they were in 1990], the couple was counted as a same-sex unmarried partner couple” (Gates and Ost 2004: 12).

In the 2000 U.S. Census, there was a notable increase in the total number of individuals classified as same-sex unmarried partners, with almost 600,000 couples reported as same-sex unmarried partner couples in 2000 compared to 145,130 in 1990. The quadrupling of couples identifying as same-sex unmarried partners led researchers to speculate about the cause of the “increase.” Indeed, Black and colleagues (2002a) have cautioned that some of these couples might actually be heterosexual couples, misclassified by the Census Bureau as same-sex partners in an attempt to rectify contradictions between individuals’ selected sex and marital status. If this is the case, they note that researchers

should try to adjust for misclassified heterosexuals when using the Census data to study the homosexual population.

Because the federal government does not recognize marriage between two individuals of the same sex, the Census Bureau does not accept responses where the householder (person #1) identifies another individual of the same sex as a spouse (Fields and Clark 1999). Rather, in 2000 the Bureau accepted the sex indicated by the respondents, but reclassified the couple as unmarried partners (U.S. Census Bureau 2001). Further, same-sex unmarried partners who selected "married" on the marital status question were also reallocated by the Bureau into a category other than married. Assuming that these individuals are same-sex couples who wish to indicate a marital relationship, this strategy maintains the integrity of the responses since the unmarried partner category is designed to capture marriage-like relationships. Black and his colleagues (2002a), however, observe that a side effect of this allocation process is that married heterosexual couples who misreport the sex of one spouse will be reclassified as same-sex unmarried partners.

In the several geographical-based analyses of gay male and lesbian partnering that we present in this book, the miscoding problem is not an issue. Because the sex miscoding measurement error is very small, it does not appear to have "any significant effect on geographical distribution patterns" (Gates and Ost 2004: 14). In analyses employing the Public Use Microdata Samples, however, any sizable sex miscoding could prove to be problematic.

Black and his associates created models to test the extent of this measurement error; they also presented a method for recalculating reliable estimates using the census data. Briefly stated, to obtain the portion of same-sex unmarried partners who are actually misclassified heterosexuals, they used a figure from a 1975 Census Bureau study that indicated that the error rate for sex miscoding was less than .002 for each observation; they estimated that the error rate for miscoding one's own sex and one's partner's sex would be between .003 and .004. They then engaged in two exercises to estimate the numbers of misclassified heterosexuals: one assumes that the average number of children for homosexual households is the same for those with both allocated and nonallocated marital status, and the second assumes that the rate of sex miscoding is the same for both married and unmarried heterosexual partners.

Based on these analyses, Black and associates (2002a) concluded that between 30 and 35 percent of all same-sex unmarried partner couples are actually misallocated married heterosexual couples. Their assumption in the first exercise, however, may be problematic. Fields and Clark (1999) found in Census test studies that same-sex unmarried partners who self-identify as married couples have different characteristics than those who do not so identify. In fact, they found that the presence of children in the household (the characteristic employed by Black and associates in their analysis) was an area in which these households were particularly likely to differ, with same-sex households with children being seven times more likely to have identified as "married." Thus, the



assumption by Black and colleagues that allocated and nonallocated households are similar may be problematic. They attempt to adjust for this weakness in their second exercise by allowing the average characteristics to vary among allocated and nonallocated households, and assuming that the rate of sex miscoding is the same for married and unmarried heterosexual partners. They note, however, that doing so results in increased sensitivity to the assumptions that they make regarding the rate of sex miscoding because in the first exercise they assume the value of sex miscoding only for the smaller, unmarried group.

Consequently, the analyses of Black and his associates may well contain problematic assumptions, which could result in a biased estimate of the number of misclassified heterosexuals in the 2000 Census. Nonetheless, their cautions concerning the possible existence of sizable misclassification errors need to be taken into account when conducting analyses with these data.

However, research undertaken at the U.S. Bureau of the Census indicates that the number of persons mistakenly included in the same-sex unmarried partner data because of sex miscoding is offset in part by persons mistakenly excluded because of sex miscoding. O'Connell and Gooding (2006, 2007) examined the first names of opposite-sex couples (married or unmarried) in the 2004 test census of New York and compared their names with their reported sex. They found, for instance that "98 percent of the people with the name of 'Elizabeth' . . . reported that they were female, compared with 79 percent of people with the name of 'Morgan' and 75 percent of people with the name of 'Leslie.' Some respondents with these names may have mis-marked their response in the sex item as male while others may, in reality, be male and not female" (O'Connell and Gooding 2006: 5). They then presented various approaches using "name" responses instead of "sex" responses for editing sex responses. Using the most conservative "name" approach, namely, that in which "99 percent of people with that name were of the opposite sex" (O'Connell and Gooding 2006: 5) in the census data, they found that there is as much of a gain in the number of same-sex persons based on their names as there is a loss based on their sex miscoding.

They showed that "using first names in an impartial and systematic way to invalidate reported sex responses will yield more same-sex couples than originally reported" in the census data (O'Connell and Gooding 2006: 5); indeed the number gained is near the number lost; and the characteristics of the two groups are similar. O'Connell and Gooding concluded that the inclusion of persons in the same-sex counts due to sex misclassification errors is not as serious an issue as believed by Black and his associates.

Finally, two other points may be made that further increase our confidence in the validity of the same-sex partner data from the 2000 Census. One pertains to the national "Make Your Family Count" publicity campaign that was initiated, sponsored, and conducted by gay male and lesbian communities prior to the conduct of Census 2000. Spearheaded by the Institute for Gay and Lesbian

Strategic Studies and the Policy Institute of the National Gay and Lesbian Task Force, the campaign encouraged gay male and lesbian couples to mark the “unmarried partner” category in order to be counted in the census (Bradford et al. 2002; McManus 2003; Badgett and Rogers 2003). In the months of January through March of 2000, gay and lesbian organizations and communities publicized the 2000 Census via the Internet, newspapers, and mailing lists to make their constituents aware that Census 2000 was about to be conducted. Furthermore, they encouraged gay men and lesbians in partnered relationships to fill out the census questionnaire and to be sure to use the “unmarried partner” response when answering the question on “relationship to the householder.” Although we do not know the complete effects of the campaign, it has been credited by some as helping to increase the numbers of same-sex unmarried partner respondents fourfold in 2000 from the 1990 Census (Bradford et al. 2002).

A second point concerns the fact that the actual numbers and rates based on the same-sex partner census data for the census tract neighborhoods of many cities and metropolitan areas of the United States have been shown in many contexts to be large and high in exactly those neighborhood areas “known” to be gay and lesbian enclaves; and the opposite has been shown to be true for neighborhoods known as heterosexual areas. Chapters 2 and 4 in this book present data along these lines. To illustrate, the Castro District in San Francisco is well known and cited in the literature as being a prominent, if not the preeminent, gay male enclave in the United States (Abrahamson 1996; Murray 1992). According to the census data on same-sex unmarried partners, it does indeed have a very high concentration of male unmarried partners, as well as female unmarried partners (Gates and Ost 2004).

Similarly, a district in Oakland, California, known by many to be a lesbian enclave (see chapter 4) reveals a high concentration of female unmarried partners in the census data (Zamora 2004; Gates and Ost 2004). Another well-known gay enclave in the Southwestern United States, the Montrose District of Houston, also shows a very high concentration of male unmarried partners according to the 2000 Census data. Conversely, other areas of Houston, such as Kingwood and Sugarland, other areas of San Francisco, such as the Sunset and the Parkside, and many other areas in other cities also known to be heterosexual neighborhoods, report very low numbers of same-sex unmarried partners in the 2000 Census. This correspondence between the spatial distribution of same-sex unmarried partners and areas known to be gay enclaves, or known to be predominantly heterosexual, provides additional evidence about the validity of these data.

In this chapter, we have evaluated the general quality of the same-sex partner data from the 2000 census in several different ways. Despite the shortcomings of these data, we agree with Black and his associates (2000) who conducted similar analyses of the 1990 data that the census data on same-sex partners are not the product of measurement error and that, indeed, the bulk of the same-sex couples enumerated in the census data are same-sex partners.

Having examined the validity of the census data, we now turn to a general description of unmarried same-sex partners and their households, as revealed by these data. We compare them on several characteristics with married heterosexuals and with cohabiting heterosexuals to obtain a better understanding of this population.

### Characteristics of Same-Sex Unmarried Partners and Their Households

In later chapters of this book we undertake analyses examining several demographic questions pertaining to sexual orientation. Before doing so, we provide here an introduction to some of the basic characteristics of individuals who identified as same-sex unmarried partners on the 2000 Census. What percentage of racial minorities selected this category? What is the average income of individuals in this group? What is the average age of an unmarried partner? And how do these characteristics, and others, compare with those of individuals in other couple-types, specifically married and heterosexual unmarried partners? Some of the characteristics of same-sex partners and their households are presented in this section. They will provide insights into certain information on homosexuality that can be gleaned from the 2000 Census, as well as offer a foundation for the analyses to follow in later chapters of this book.

We noted previously that Census 2000 enumerated over 105 million households. Almost 60 million were households inhabited by couples, of which over 54 million were married couples. This leaves almost 5.5 million unmarried partner households. Of these, over 595 thousand were same-sex unmarried partner households (see Table 1.1). Thus, 1 in 10 of the unmarried partner households captured in Census 2000 were same-sex unmarried partner households; of these, 301 thousand were male-male and 293 thousand were female-female households.

Almost 80 percent of all households, 79 percent of coupled households, and 81 percent of unmarried partner households, were located in metropolitan areas (Table 1.1). Same-sex partner households, in comparison, had a larger metropolitan presence, with over 85 percent in metropolitan areas. This is the largest metropolitan presence of the several categories of coupled households. Also shown in Table 1.1 is the slightly different distribution by geographic region of same-sex partner households compared to other households, particularly married-couple households, with slightly more in the Northeast and West and less in the Midwest and South.

The state of California had more unmarried partner households (12 percent of the total) than any state in the nation. The percentage of same-sex unmarried partner households in California was 16 percent, also the highest in the country (Simmons and O'Connell 2003: Table 1.2). The highest percentages of same-sex unmarried households were in cities on the West and East

Coasts. San Francisco had the highest percentage of same-sex unmarried households compared to all households (2.7%), followed by Fort Lauderdale (2.1%), Seattle (1.9%), Oakland (1.8%), and Berkeley (1.8%). Only one of the top ten cities in this group was in the Midwest, namely, Minneapolis (1.6%) (Simmons and O'Connell 2003: Table 1.3). The metropolitan and regional differences examined in the preceding paragraphs receive more attention in Chapter 2 of this book.

We now examine characteristics of same-sex unmarried partners and compare them with heterosexuals who are cohabiting and with heterosexuals who are married. Here we use data from the 5 percent Public Use Microdata Samples (PUMS) from the 2000 Census. We undertake these analyses first for males, then for females. In the first analysis we compare characteristics of males who are same-sex partners with males who are cohabiting with females and with males who are married to females. The second analysis is undertaken in a similar way for females.

Since most of the analyses involve socioeconomic comparisons of the groups, we introduce several constraints. To be included in the analyses reported here, we required that the males and females were in the labor force with a job and earning at least \$1,000 in 1999. We also use statistical sample adjustment methods (Stata Corp 2005) that introduce survey adjustment estimators to adjust our analyses according to the population weights assigned in the 5 percent PUMS.

Table 1.2 compares the characteristics of male same-sex unmarried partners, with cohabiting male heterosexuals and married male heterosexuals. The top panel of the table examines mean values for the three groups. On average, same-sex male partners reported annual earnings in 1999 of over \$40,000, compared to almost \$32,000 for male cohabiters and over \$48,000 for males who were married. These earnings differences among labor force participants in the three groups require more attention. Chapter 8 in this book analyzes earnings differences between homosexual and heterosexual males and females.

We next examine mean occupational status scores for males in the three groups. The occupational status score is a score assigned to persons in each detailed census occupation based on the median earnings for that occupation. The occupational status score is meant to represent the material rewards accruing to persons in different occupations where the higher the value of the score, the higher the status (Nam and Boyd 2004). Married males have slightly more occupational status than same-sex partnered males, and both have more status than male cohabiters. Occupational differences between same-sex partners and heterosexual partners are examined in further detail in chapter 9 of this book.

Although married men have higher occupational status than same-sex partnered males, almost two-thirds of the same-sex male partners have a college degree or higher compared to less than 60 percent of married males and 47 percent of male cohabiters (Table 1.2).