

Homophobic Innumeracy?  
Estimating the Size of the Gay and Lesbian Population

Michael D. Martinez  
Kenneth D. Wald  
Stephen C. Craig

Department of Political Science  
University of Florida  
P. O. Box 117325  
Gainesville, FL 32611-7325  
martinez@ufl.edu  
kenwald@polisci.ufl.edu  
sccraig@ufl.edu

**Abstract:** Previous research has shown that mass perceptions about the sizes of minority populations are influenced by sociodemographic, threat, and context variables. This paper extends the analysis to a population group that has not previously been considered from a national perspective, that is, gays and lesbians. Our analysis of a statewide survey of Florida residents in 2002 shows that people in sociodemographic groups associated with low levels of political knowledge were more likely to report higher estimates of the gay population. Threat variables and objective context had relatively little impact, but estimates also were higher among individuals who reported personal contact with gays and lesbians.

Paper presented for delivery at the Annual Meetings of the American Political Science Association, Philadelphia, PA, August 31-September 3, 2006. © American Political Science Association

Long assailed for its lack of political literacy (Berelson, Lazarsfeld, and McPhee 1954; Converse 1975; Delli Carpini and Keeter 1996), the American public has also been charged with statistical innumeracy. According to numerous surveys conducted over the years, many citizens consistently and systematically misperceive the relative size of certain minority groups in the population; in particular, there is a tendency for people to overestimate the proportion of African Americans, Latinos, and Jews who live in the United States. Prior research suggests that such inflation is the product of three forces: low levels of political knowledge, a misreading of environmental cues, and perceived threat. Of these, political knowledge (more specifically, the sociodemographic correlates of knowledge) has generally emerged as the strongest influence on accurate estimates about the presence of minorities in the population.

The present study extends the analysis to a population group that has not previously been considered from a national perspective, that is, gays and lesbians. Unfortunately, estimating the gay population at *any* level presents a number of challenges beyond those associated with ethnic and racial groups – challenges that make it difficult even for politically well-informed and attentive citizens to produce accurate percentages. First, due to a lack of official or authoritative information about the size of the gay population, there simply are no trustworthy baseline data available upon which one might anchor personal estimates.<sup>1</sup> Moreover, unlike race or ethnicity, which may be visible to the naked eye, sexual orientation is not apparent on the surface and many gay people choose to remain deeply closeted because of the stigma that still attaches to homosexuality in the United States. Third, the kind of contextual cues that might signal a strong gay presence to insiders (including bars, bookstores, and other institutions that cater to gays) may not be recognized as such by heterosexuals, thereby depriving individuals of the kind of useful information that helps them to gauge the presence of other minorities. Finally, gays are stigmatized much more intensely than the racial and ethnic groups examined in earlier research. Notwithstanding the changing social standing of gays and their gradual success in winning civil

rights protection (Brewer 2003), it remains the case that the majority of the population still bears a negative orientation toward them (Yang 1997; 2001). On the "feeling thermometers" that measure affect toward various groups in society, for example, gays continue to define the bottom of the scale rankings, well below the average scores of blacks, Hispanics, and groups used in previous innumeracy research.<sup>2</sup>

These differences between gays and other minorities, chiefly racial and ethnic groups, raise the possibility that different dynamics may come into play when people are asked to estimate the size of the gay population. The distinction here is more than a curiosity because public perceptions and misperceptions about minorities are not only durable, they also are potent influences on policy preferences (Kuklinski et al. 2000). Accordingly, the analysis that follows should help to clarify the sources of resistance to, and support for, gay-related issues on the public agenda.

### **Background**

When asked in surveys to estimate the size of various minority groups, individuals routinely inflate the official figure by several orders of magnitude. For example, a majority of respondents in the 1991 ANES Pilot Study believed that African Americans constituted more than 30% of the national population (vs. 12% as revealed by the U.S. Census); almost four in ten thought that Jews – whose numbers have been reliably pegged at about 2-3% – actually made up a fifth or more of the population (Nadeau, Niemi, and Levine 1993: 334). Similarly, a 1995 survey revealed that non-Hispanic whites believed that blacks, Hispanics, and Asians collectively added up to half the population, an estimate that doubled the actual proportion of these minorities in the United States (Thernstrom 2002: 14-15). A more recent such survey shows even more extreme misperceptions: "Americans believe there are three times as many African Americans, almost three times as many Hispanics, and fully six times as many Asians in

the U.S. as there actually are," and overestimate the Jewish population by a factor of ten (Theiss-Morse 2003: 4-5). The tendency to overestimate minority population size is greater when the estimate is for the United States as a whole, but is also evident when respondents estimate minority population sizes in their local communities (Alba, Rumbaut, and Marotz 2005: 908-9).

This statistical innumeracy does not appear to be distributed randomly, but rather varies as a function of at least three sets of influences.<sup>3</sup> For estimates of blacks and Hispanics, several studies have documented the strong impact of *social background factors* such as age, education, race, gender, and political interest (Nadeau, Niemi, and Levine 1993; Alba, Rumbaut, and Marotz 2005). The correlation of these variables with accuracy in estimating black and Hispanic population size likely reflects the influence of political information and attentiveness;<sup>4</sup> that is, people with deeper stores of information should be more capable of looking beyond popular stereotypes that are fed by the visible but misleading presence of minorities in entertainment and news. Yet it is equally possible that sociodemographic variables represent another quality that might also affect statistical numeracy: tolerance (Overby and Barth 2002). Whatever the force behind the relationship, however, the strong ties between social background factors and estimates of other minorities means that this set of variables constitutes the default model in our own analysis.

Estimates of minority population also appear to be shaped by *contextual perceptions*. This finding is consistent with evidence that respondents generally draw on salient personal experience and unrepresentative anecdotes when responding to information questions in opinion surveys (Nadeau and Niemi 1995: 327). Two kinds of environmental cues have been found to influence population estimates of group size: geographical concentration of minority groups and personal contact with members of those groups. Individuals appear to generalize estimates about black and Hispanic population size in the United States from the density of these groups in their own communities (Nadeau, Niemi, and Levine 1993; Sigelman and Niemi 2001, Alba, Rumbaut,

and Marotz 2005). A higher level of personal contact with members of racial and ethnic minority groups similarly leads to higher estimates of minority population size in the nation as a whole (Alba, Rumbaut, and Marotz 2005). Population density and personal contact can be understood, respectively, as objective and subjective measures of the local context.

A third factor that influences population estimates of minorities is *perceived threat*: There is a tendency to overstate the population size of minority groups regarded as threatening. When a group is perceived as dangerous and hostile, individuals sometimes monitor their environment intensely and "see" the threat in exaggerated terms (Allport 1954; Gallagher 2003; Nadeau and Niemi 1995: 327-8; Whaley and Link 1998). Of course, the reverse is equally plausible because people may derive a sense of threat from nearby groups that are highly concentrated (Fossett and Kiecolt 1989). Nonetheless, perceived threat has emerged as a significant predictor of overestimates of minority populations.

All these factors may influence citizens' estimates of the gay and lesbian populations, albeit in a somewhat different manner than they affect estimates of racial and ethnic minorities. If we assume that media coverage and popular entertainment over-represent the gay population (Gallagher 2003), then it is reasonable to anticipate that people with higher levels of political knowledge will be less apt to rely on the sometimes misleading impressions gleaned from those sources. Thus, we expect that people with more education and greater political interest will provide smaller (which is to say, more accurate) estimates of the gay population. However, contextual information might be less useful in estimating gays than it is in assessing the density of racial and ethnic groups. As noted above, the environmental cues that help individuals judge the size of the local black and Hispanic populations are less readily apparent in the case of gays and lesbians; hence, there are less "objective" data to assist individuals in forming their guesses about the size of the gay-lesbian population. The failure of Jewish density to predict accurate estimates of the Jewish population, a consequence of the relative invisibility of Jews, may hold

for gays as well (Nadeau, Niemi, and Levine 1993: 340). On the other hand, personal contact might be even more powerful in stimulating overestimates of gay population because of the strong evidence that knowing gays and lesbians personally correlates with more positive views.<sup>5</sup>

The threat variables might also work in a distinctive manner for gays and lesbians. Based on data about racial and ethnic groups, we assume that negative affect prompts individuals to overstate the size of groups they dislike and fear. There is a twist, however, that could produce precisely the opposite pattern when gays and lesbians are the target population. Many opponents of gay rights insist that homosexuality is a chosen lifestyle rather than an immutable social trait and, as a result, that gays and lesbians have the option of changing their behavior. If this is true, then we might expect that people who are predisposed to dislike gays, such as religious traditionalists, will also be prone to minimizing the size of the gay population. If nobody is truly gay (meaning gay by nature) then, by definition, there is virtually no gay population. Hence, threat may produce denial in the case of gays and lesbians rather than the overestimation routinely induced by racial and ethnic minorities.

Our interpretation here is close to the conclusion offered by Overby and Barth (2006) in the published study that, so far as we have been able to determine, is most similar to the present inquiry. Using data from a national survey that asked respondents what percentage of gays or lesbians are present in their own communities, the authors found little evidence to indicate that threat variables prompted inflated estimates. Based on that finding, Overby and Barth concluded that individuals with hostile or homophobic views don't believe that gays are threatening because they do not constitute a "real" category.

We want to stress that our study differs methodologically from Overby and Barth (2006) in several ways. First, and in our view least important, is the fact that their analysis was based on a national survey while ours involved a statewide (Florida) sample.<sup>6</sup> Second, our approach follows most prior research in that respondents were asked to provide a general rather than a

specifically local estimate of the gay population. Third, the two studies test different models. Whereas Overby and Barth lacked measures of several likely influences on citizens' estimates (and thus had to make critical assumptions about those unmeasured influences), our survey was designed so as to permit direct tests of the predictor variables identified in the literature dealing with political innumeracy. Taking each of these factors into account, it is clear that our research is not an exact replication of the work done by Overby and Barth. Nevertheless, a comparison of the findings provided by the two studies should not only prove interesting in its own right, but also suggest new directions that scholars might wish to take in the future.

### **Data and Methodology**

The core analysis presented here is based on a telephone poll of 601 Florida residents conducted in June, 2002, by the *Florida Voter* survey organization. The sampling frame was drawn using a random-digit-dialing (RDD) procedure; if needed, four callbacks were made to each working number in an effort to obtain a completed interview. Interviewers asked to speak with "the youngest male, eighteen years of age or older, who is now at home." When no male was available, they inquired about "the youngest female, eighteen years or older, who is now at home" (see Crespi 1988; Gaziano 2005). A total of seventy questions were included in the survey, and the mean time for completion was twenty-four minutes. The margin of error was plus or minus 4.1 percentage points.<sup>7</sup>

Much of the survey was devoted to questions asking respondents for their attitudes about gays and gay rights, including a battery that measured ambivalence toward gay rights issues (Craig et al. 2005). Other items tapped egalitarianism, traditional marriage values, traditional lifestyle values, moral judgment, moral traditionalism, party identification, ideology, religiosity, and demographics. Near the end, the question was posed, "What percentage of the overall population would you estimate is either gay or lesbian? (If you had to make a guess, what

percentage would you say?)" A substantial number of people (107, or 17.8%) either could not or would not venture a guess, and estimates offered by the remaining 494 respondents varied quite a lot, as shown in Figure 1. Unlike previous research on popular estimates of other minority populations, we have no baseline "objective" assessment against which to measure the accuracy of these figures. What we know, however, is that the modal response (10%) and the median (12%) were close to the 10% figure popularized by the Kinsey studies in the 1950s and sometimes advanced by gay and lesbian activists. The mean (17.4%) exceeded both that estimate and the 2-5% figure derived from the most authoritative surveys of sexual orientation (Badgett 1998: 9-11). Roughly one-quarter of our respondents believed that gays and lesbians account for 5% of the population or less, while another quarter estimated 25% or more (including one individual who, perhaps facetiously, weighed in at 90%).<sup>8</sup> It is possible that the emphasis given in our survey to questions about gays and gay rights primed respondents to give higher estimates than they might have provided in another context. A separate survey of 502 Florida households<sup>9</sup> conducted by the Florida Survey Research Center in June, 2006 provides some reassuring evidence on that point, however. At the conclusion of a monthly statewide consumer confidence survey, respondents were asked to assess percentages of blacks, Jews, and "gays and lesbians" for the country as a whole. Average estimates of the size of the gay population were only slightly lower (mode and median of 10, mean of 15.7) than those observed in our 2002 survey.<sup>10</sup>

Figure 1 about here

As discussed earlier, we hypothesize that respondents' estimates of the size of the gay and lesbian population are related to demographics (education, race, gender, age, income) and other personal factors (political interest, partisanship) that serve as surrogates for political knowledge or "numeracy," perceived threat (measured here in terms of religious denomination, practice, guidance, and belief, as well as moral judgment and traditional values), context (referring to the prevalence and visibility of the local gay population), and individual-level contact with gays and

lesbians. Most of our demographic and background variables were based upon respondent self-report in the survey. *Educational attainment* and *age* were scored according to number of years, while *race/ethnicity* and *income* allowed people to place themselves in the appropriate category. *Political interest* was captured by a single question asking how often the individual "follows government and public affairs," and *partisanship* by the standard American National Election Studies "direction" (Republican, Democrat, Independent) question.<sup>11</sup> *Gender* was recorded by the interviewer without asking, unless it was necessary to do so.

Several questions dealing with religion (representing perceived threat) were included in the survey. Religious *denomination* was measured with a single item that allowed us to classify respondents as either traditional Protestants, Evangelical Christians (including Protestant Fundamentalists, reflecting the overlap in political outlooks between these two groups more than the theological distinction between evangelicalism and fundamentalism), Catholics (including the Greek Orthodox, for much the same reason), Jewish, affiliated with another religion, or not affiliated. In addition, we asked about respondents' *interpretation of the Bible* (whether it is the word of God, to be taken literally) and *frequency of attendance* at religious services. Our measure of *religious guidance* is based on answers to two questions: Respondents who did not regard "religion to be an important part" of their lives were coded in the lowest category, while all others were scored according to their self-assessment of the amount of guidance religion provides them in daily living.

We also asked respondents whether they believed that "homosexuals are that way because they choose to be," or that "people do not choose to be homosexual, they are born that way." The threat hypothesis leads us to expect that estimates of the gay population should be higher among people who say that gays are gay by choice, though Overby and Barth (2006: 206) found higher estimates at the local level among those who believed that sexual preference was biologically determined.

In addition, Weisberg (2005; also see Stoker 1987) draws our attention to the differences between *moral judgment* and *moral traditionalism*. Moral judgment, a potent influence on policy preference in several areas, involves acceptance or rejection of others' nontraditional lifestyle choices, particularly sexual behaviors, and is measured here by an index ( $\alpha = .86$ )<sup>12</sup> involving questions about the morality of sexual relations between consenting adults of the same sex, between unmarried heterosexuals, between unmarried teens, and between a married person and someone other than his or her spouse, as well as respondent assessments about whether homosexual sex is "just plain wrong" and "disgusting" (separate questions for male homosexuals and lesbians; see Appendix). Moral traditionalism, on the other hand, has to do with one's feelings about traditional family and social structures. Within this broader value dimension, we distinguish between *traditional lifestyle values* and *traditional marriage values* (see Craig, Kane, and Martinez 2002; Craig et al. 2005). The former is measured with an index ( $\alpha = .50$ ) based on agreement or disagreement with the following statements: (1) The world is always changing and we should adjust our view of moral behavior to those changes. (2) This country would have many fewer problems if there were more emphasis on traditional family ties. (3) The newer lifestyles are contributing to the breakdown of our society. Our index of traditional marriage values ( $\alpha = .66$ ) is based on agreement or disagreement with the following: (1) All in all, family life suffers when the woman has a full time job. (2) It is more important for a wife to help her husband's career than to have one herself. (3) A husband's job is to earn money; a wife's job is to look after the home and family. The threat hypothesis leads us to expect higher estimates of the gay population from individuals who are morally judgmental and who possess more traditional values regarding lifestyle and marriage.

We supplemented our survey responses with contextual data from the county where the respondent resides, following the theory that people who live in communities where gays and lesbians are more prevalent and prominent might generalize from their local estimates to the

general population. The first set of indicators represents objective measures of gay presence, which have been used successfully in studies of gay rights policy adoptions (Wald, Button, and Rienzo 1996). Percent *unmarried partner households* at the county level was calculated using 2000 census data; scores range from a low of 0.34 in Sumter (a mostly rural county in North Central Florida) to a high of 1.59 in Monroe (comprising the Florida Keys). A dummy variable was included for *gay ordinance*, which has a value of one for the six counties that have either a wide-ranging law prohibiting discrimination based on sexual orientation or are home to one or more localities with a similar law (van der Meide 2000). For each county, we also coded the number of *information lines* (including entries for GLBT organizations of any sort, such as telephone switchboards), *gay bars* (including gay clubs), *gay bookstores* (excluding general bookstores with GLBT sections), *gay shops* (those that cater explicitly to GLBT, not including dealers in erotica), *local gay publications*, and *local gay service organizations* (see *Damron Men's Travel Guide* 2002). In addition, our pollster (see note 7) provided us with a classification of the respondent's community as *urban* (central city within a Metropolitan Statistical Area), *suburban* (any portion of an MSA county not in the central city), or *rural* (a non-MSA county), based on the area code and prefix of the respondent's telephone.<sup>13</sup> In general, we expect to see higher estimates of the gay population from respondents in urban counties, and from those living in counties with more visible gay institutions.

Finally, our survey included two questions intended to capture one's personal experience with homosexuals. Respondents were asked, "Do you know anyone personally who is either gay or lesbian?" (481 of 601 said they did.) And the last question in the survey asked respondents to identify "the term that best describes yourself – heterosexual, or straight; homosexual, gay, or (female respondents only) lesbian; or bisexual." (568 self-identified as heterosexual, sixteen as gay, lesbian, or bisexual, four as "other" or "don't know," and only thirteen refused to answer the question.) Our expectation is that higher estimates will be provided by respondents who have had

*personal contact* with homosexuals, and from those who are themselves gay, lesbian, or bisexual.

The question posed by our research is fairly simple: What factors (knowledge, threat, context, or personal contact) contribute to higher estimates of the size of the gay and lesbian population? Despite its simplicity, we faced several important methodological issues in trying to provide a credible answer. First, as is common in survey research, many people responded "don't know" (or, less often, declined to express any opinion at all) on one or more items of theoretical interest to us. Indeed, given the sensitivity of the topics covered in this particular survey, it is not surprising that fewer than half the respondents (289 of 601) provided valid answers to all of the questions used to measure our independent variables. Although listwise deletion is the most commonly used method of dealing with missing data, King et al. (2001) have cautioned that such a strategy can lead to biased estimates in multivariate analyses, even where missing data are much less prevalent than in this particular survey. To address that problem, we estimated our multivariate models using multiple imputation of the missing values on explanatory variables. Results of the multivariate analyses presented below are pooled estimates based on separate analyses of five imputed datasets.<sup>14</sup>

Second, the substantial number of people with missing data on the dependent variable posed a different problem. While many of the 107 respondents who answered "don't know" to our question about the size of the gay and lesbian population (even after being pressed for a guess) may have been sincerely admitting that they hadn't thought about the issue, others might have refused an answer or used the "don't know" option to conceal what they perceived as a socially undesirable answer (Berinsky 1999). If factors that affect the likelihood of responding to the question are correlated with the responses themselves, listwise deletion of missing cases on the dependent variable would likely result in biased estimates of coefficients in multivariate analyses. We address this concern by using the Heckman procedure,<sup>15</sup> which first estimates the

probability of responding to the survey question using a probit model, and then includes the inverse Mills Ratio (essentially, the hazard of selection) in the second-stage linear model of the outcome variable.

Third, while we are fortunate to have multiple indicators for several concepts that might help explain variation in estimates of the size of the gay and lesbian population, the correlations that exist between many of our explanatory variables could easily mask the independent effects of any one variable in a multivariate analysis. To minimize that risk, we will present bivariate and multivariate results for each group of variables, and then retain substantively and statistically significant variables for the estimation of a comprehensive model.

## **Results**

### Demographic/Background Variables

Political knowledge (numeracy) is regarded as the baseline predictor of estimated population size against which other models are evaluated. Following Nadeau, Niemi, and Levine (1993), several demographic and personal background variables were employed as proxies for such knowledge; that is, we expected to observe lower, hence more accurate, estimates of the gay and lesbian population in groups whose members tend to be more knowledgeable (men, the well educated, older people, the affluent, whites, those interested in politics, and partisans). Table 1 shows mean estimates of the size of the gay and lesbian population, as well as the proportion of each group that provided a valid response to the question.<sup>16</sup>

Table 1 about here

Most variables commonly associated with political knowledge show the predicted relationship. In contradistinction to Overby and Barth (2006), we find that the average estimate offered by men is nearly four points lower than the average for women, respondents with more formal education tended to provide lower estimates than those who self-reported fewer years of

education, and whites' estimates of the size of the gay and lesbian population were smaller, on average, than those offered by minorities (most notably, blacks and Hispanics). Other results from our survey are more consistent with Overby and Barth: older respondents and those who reported paying more attention to politics also gave lower estimates, and (excluding the handful of people with household family incomes under \$10,000) income is negatively associated with the estimated size of the gay population. The one exception to our expectation is partisanship. We anticipated finding higher levels of political knowledge and, consequently, lower estimates of the gay and lesbian population among partisans (both Republicans and Democrats) relative to Independents. But while Republicans did provide the lowest average estimates, they were not significantly lower than those set forth by Independents – and the latter estimates were, in turn, three points *lower* than those by Democrats (again, consistent with Overby and Barth 2006). Nevertheless, with the lone exception of partisanship, these patterns are very much in line with the hypothesis that groups possessing more political knowledge will tend to give lower estimates of the size of the gay and lesbian population.<sup>17</sup>

Further, as shown in the last column of Table 1, most of the traits that are negatively related to the estimated size of the gay population also are associated with higher response rates on that particular question. Do the bivariate patterns hold when we correct for this possible selection bias by estimating a multivariate model using the Heckman two-stage procedure?<sup>18</sup> In Table 2, we report the results of a second-stage OLS model that provides unbiased estimates of the effects of demographic and other background variables on estimates of the gay and lesbian population.

Table 2 about here

The multivariate analysis generally confirms the bivariate findings in Table 1. Once again, most demographic categories that are characterized by higher levels of political knowledge are also associated with lower estimates of the size of the gay and lesbian population. Men's

estimates were, on average, 4.5 points lower than those provided by women, *ceteris paribus*. Whites' estimates were lower than those provided by blacks and Hispanics (by 6.7 points and 4.3 points, respectively), while those with more schooling also tended to provide lower estimates; interestingly, the effect of four years of formal education was slightly less than that of gender and being Hispanic, and substantially less than the effect of being black. Age was significant as well, with the effect of six years of life experience being comparable to that of a single year of education. The effect of political interest, income, and partisanship were trivial, controlling for other personal traits. On the whole, however, the effect of background variables on estimates of the size of the gay and lesbian population are notable, in keeping with similar analyses involving other minority populations.

### Threat Variables

We also anticipate that people who feel that their religious beliefs and values are threatened by "the gay lifestyle" will be more inclined to "fear" that gays and lesbians compose a larger share of the population. Table 3 displays the mean estimates among groups defined by religion (denomination, practice, guidance, belief) and by value orientation (moral judgment, traditional marriage values, traditional lifestyle values). On the whole, support for the threat hypothesis is fairly limited. As predicted, Evangelicals offered higher estimates than did other Protestants, Catholics, Jews, and non-believers, and Biblical literalists tended to think that there were more gays and lesbians than did non-literalists. However, the estimates of those who attend religious services weekly were not much different than those who never attend, nor were there substantial differences between estimates of individuals for whom religion provides a great deal of guidance in their daily lives and those for whom religion is not particularly important. Also, in contrast to Overby and Barth (2006: 206), respondents who said that "homosexuals are that way because they choose to be" gave higher estimates than others – but the differences are small.

And those who scored high on the moral judgment scale or who espoused traditional lifestyle values actually offered somewhat *lower* estimates than did other respondents; there was no discernible (or monotonic) pattern on the traditional marriage values variable.

Table 3 about here

Table 4 shows results of the multivariate analysis using imputed data for missing values and controlling for the likelihood of responding to the question, and the story is similar. Among our measures of religious denomination, practice, guidance, and belief, only the coefficient for the Biblical literalism dummy variable is significant at conventional levels. Controlling for other variables in the model, Biblical literalists' estimates were about eight points higher, on average, than were the estimates offered by those who ascribed authorship of the Bible to men. Evangelicals' estimates were about four points higher than those of non-believers, *ceteris paribus*, but the coefficient is not significant ( $p < .23$ ). Among the values variables, only the coefficient for traditional lifestyles values is significant at conventional levels, and, as with the bivariate analysis in Table 4, individuals with more traditional values saw *fewer* gays and lesbians. On the whole, these results provide, at best, weak support for the threat hypothesis.

Table 4 about here

### Contextual Variables

Next, we examined the possibility that people might draw inferences about the overall size of the homosexual population from the size and visibility of gay populations in their own communities, and from their personal contact with gays and lesbians. Using objective indicators of gay density, the top half of Table 5 shows that respondents living in counties with more visible gay populations (as indicated by the number of gay bars, shops, bookstores, and publications) tended to provide slightly higher estimates than those from other Florida counties, though differences between the two groups are generally small. For example, the mean estimate

of people in counties with at least thirteen gay bars was three points higher than that offered by residents of counties with no gay bars. In addition, census estimates of the proportion of same-sex households in a county have no discernible relationship at all with respondents' estimates of the size of the gay population at the bivariate level. The one contextual variable that seems to have the greatest effect is the simple categorization of areas as urban, suburban, or rural, with urban residents' estimates being noticeably larger than those proposed by rural and suburban residents. The Heckman multivariate regression analysis (with imputed values for missing data) in the top portion of Table 6 shows similar results. Keeping in mind that there is a fair amount of multicollinearity among the explanatory variables in this model, gay bars and urban residence stand out as the contextual variables that have a positive relationship with respondents' estimates of the size of the gay population. Unexpectedly, percent same-sex households in a county is negatively related to the size of the estimate, *ceteris paribus* and allowing for a generous standard of statistical significance ( $p = .12$ ).

Tables 5 and 6 about here

### Contact Variables

In contrast to the generally weak relationships observed with our threat and objective contextual variables, personal contacts with gays and lesbians appears to have a significant effect on estimates of the size of the gay population. The bottom portion of Table 5 shows that the majority of respondents who said that they personally know someone who is gay, and especially the small minority of respondents who identified themselves as gay, lesbian, or bisexual, had significantly higher estimates than did heterosexuals who professed not to know any gays themselves. The Heckman multivariate analysis in the bottom half of Table 6 indicates that knowing someone who is gay increases one's estimate of the gay population by about three

points ( $p < .07$ ), and gay identification contributes another ten points. Contact with gays seems to have a direct effect on one's perception of the size of the gay population.

### Comprehensive Models

We present comprehensive Heckman second-stage models in Table 7. Recognizing the potential havoc that multicollinearity might wreak on a "kitchen sink" model, our strategy was to exclude variables with coefficients that had  $p$  values greater than .25 in the multivariate analyses presented above. Although .25 is a generous standard, we wanted to guard against the possibility of omitting any variables that might emerge as significant in a more fully specified model. In fact, results from the first model tested in Table 7 clearly support some of our earlier findings while undermining others.

Table 7 about here

The independent effects of education, age, gender, and ethnicity remain powerful in the comprehensive model, underscoring the impact that these demographic variables have in shaping people's estimates of the size of any minority population. To some extent, then, public estimates of the gay population mirror patterns that have been observed for other minorities: higher estimates can be expected from individuals who are less educated, younger, nonwhite, and female. At the same time, however, estimation of a comprehensive model attenuates the effects of some other variables. The initial multivariate analysis of threat variables (Table 4) indicated that Biblical literalists gave estimates about 8 points higher than those provided by citizens who doubt the divine authorship of the Bible, but about two-thirds of that effect can be accounted for by the inclusion of demographic, contextual, and contact variables in the comprehensive model. (The significance level of the regression coefficient also exceeds conventional standards.) Similarly, the magnitude of the effect of Evangelicalism is also diminished, while coefficients for traditional lifestyle values and traditional marriage values continue to have unexpected

negative signs (and are not statistically significant). In our initial analysis of contextual variables (Table 6), estimates were higher among urban residents and those living in counties with larger numbers of gay bars. In the comprehensive model, the effect of urban (as opposed to suburban) residence becomes statistically trivial, though the gay-bars coefficient is still meaningful at a modestly generous ( $p < .10$ ) standard. Finally, the magnitude of the contact variables' effects are diminished slightly.

We estimated the second model in Table 7 by dropping traditional marriage values and the urban and rural dummy variables that had high probability values ( $p > .25$ ) in Model 1. For the most part, the story remains the same, though the significance levels for some coefficients (specifically, traditional lifestyles and gay bars) give us a little more confidence that their effects are real. Taken together, the bulk of the evidence suggests that most of the variance in the estimates of the size of the gay and lesbian population that we can account for is explained by simple demographic variables, as prior research has found with respect to estimates of other minority populations. We cannot be certain whether this finding reflects greater awareness among certain groups of the social and political contours of the nation, or perhaps their grasp of the simple mathematical concept of a percentage, but demographics form strong baseline effects in estimates of the size of the minority populations. Against that baseline, the effects of threat variables appear modest (and perplexing, in the case of the negative relationship with traditional lifestyles values). There is some evidence that the visibility of the local gay population (measured by the number of gay bars) heightens perceptions of the size of the gay population, and yet the best measure we have of the that population's actual size (county-level data on the proportion of same-sex households) is *negatively* related to estimates of the overall size of the gay population. Personal contact does matter, but its effects are small in comparison to those of demographic variables.

## Discussion

There are some important similarities between the public's estimates regarding gays and lesbians and their beliefs about the size of other minority populations. First and foremost, many people aim high (often very high). While the lack of consensus on definition combined with questionable measurement preclude any truly objective assessments, the subjective estimates provided by a substantial portion of the mass public clearly exceed those that we have seen among advocacy groups with the most to gain from basic knowledge about the incidence of homosexuality in the United States (Badgett 1998). Second, the people who aim the highest generally belong to the same demographic groups (women, less educated, young, black or Hispanic) that tend to provide higher estimates of other minority populations. We therefore agree with researchers (Nadeau, Niemi, and Levine 1993) who maintain that these demographic correlates point to political knowledge as the underlying but unmeasured variable explaining much of the variation in our respondents' estimates. People with higher levels of education usually have higher levels of factual knowledge about politics (Nadeau and Niemi 1995), and some of the same processes that lead them to know which party holds a majority in the House of Representatives also help them form realistic impressions about minority populations.<sup>19</sup>

However, unlike other minorities, our respondents' estimates of the size of the gay and lesbian population are only very weakly related to local environmental context. As important as gay bars, media outlets, and service organizations may be to the life of the gay community, they either remain out of sight to the majority straight community or fail to convey very much information about the prevalence of gays. Other scholars have found that the perceived size of minority groups nationally are strongly related to the perceived size of local minorities (Alba, Rumbaut, and Marotz 2005), but we did not see any links between "objective" cues at the county level and estimates of the gay population. Indeed, our most objective indicator (proportion of same-sex households as determined by the Census Bureau) was *negatively* related to individuals'

estimates of the size of the gay population. Such results amplify Overby and Barth's (2006) suggestion that perceptions of local populations are themselves endogenous. To the extent that context matters, the relevant context seems to be the person's circle of contacts, as people who know gays personally tended to report higher estimates.

Finally, if religious traditionalists and citizens who possess traditional moral values feel threatened by homosexuals, it is not manifest by magnifications of the perceived size of the gay and lesbian population. Our specific findings differ from Overby and Barth's (2006) conclusion that beliefs about homosexuality being a matter of choice help some people to minimize the threat, but there is a slight tendency for respondents with the most traditional values to report lower estimates of the size of the gay population. As we suggested at the outset, threat variables may operate differently in the case of gays and lesbians; specifically, people who have a negative affect toward gays may manifest that orientation by denying gayness as anything other than a (perverse) lifestyle choice – a persistent motif in the rhetoric of the anti-gay movement. Thus, paradoxically, people who see the gay "lifestyle" as a threat may respond by understating, rather than exaggerating, the proportion of gays in the population. If so, threat affects the perception of gay population but in the opposite way that it does for other racial and ethnic groups. This intriguing hypothesis warrants further study.

We hope that this contribution to understanding the sources of people's beliefs about the sizes of minority populations will help spur future research on the effects of those beliefs on public policy preferences. Public opinion is shaped, in part, by domain-specific knowledge (Althaus 1998, Kuklinski *et al.* 2000), and there is evidence that perceptions of group size have independent effects on some related public policy preferences (Alba, Rumbaut, and Marotz 2005). In our own data, for example, perceived size of the gay population (logged) is modestly correlated with the personal importance attached to issues of homosexuality and gay rights ( $r = .12, p < .02$ ) and with agreement that "laws are sometimes needed in order to protect the basic

rights of homosexuals" ( $r = .11, p < .02$ ). Although these bivariate relationships are weak, they suggest that innumeracy is related to both issue salience and preferences; paradoxically, however, the strongest support for gay rights legislation apparently comes from those who hold the *least* plausible beliefs about the size of the gay population. Surely, these bivariate relationships do not tell the whole story, and it is time for the scholarly community to determine when mass beliefs about minority population sizes can move from the left-hand side to the right-hand side of the equation.

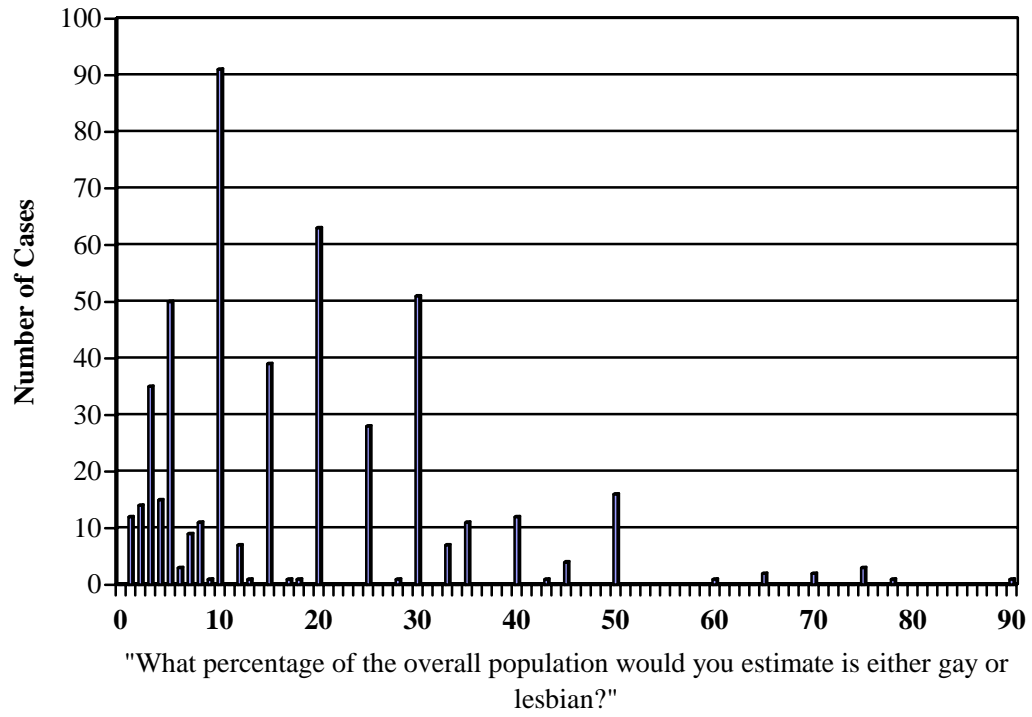
**Figure 1: Estimated Percent Gay**

Table 1  
 Estimates of Size of Gay Population and Percent Responding: Demographic Groups

	Number of valid responses	Mean estimate	Standard Deviation	Percent Responding
<u>Gender</u>				
Male	240	15.5	14.3	85.4
Female	254	19.3	14.7	79.4
<u>Years of Education</u>				
1 thru 12	148	22.3	16.8	77.5
13 thru 15	153	17.5	13.4	83.2
16	103	15.1	14.4	85.8
17 and up	89	12.2	9.8	86.4
<u>Age</u>				
18 thru 29	94	24.0	17.5	93.1
30 thru 39	83	21.0	15.7	87.4
40 thru 49	104	17.5	13.1	88.1
50 thru 59	97	15.0	12.9	81.5
60 thru 69	54	12.7	12.1	74.0
70 thru 79	31	12.2	8.3	64.6
80 thru 97	23	8.8	10.4	67.6
<u>Follow politics</u>				
Hardly at all	33	24.2	16.6	73.3
Now and then	49	21.6	15.4	77.8
Some	148	17.1	12.9	85.5
Most	263	15.9	14.6	82.7
<u>Race</u>				
White	341	14.7	11.8	80.6
Black	78	25.7	19.2	87.6
Hispanic	59	21.9	16.8	93.7
Other/DK/Refused	16	19.4	16.1	61.5
<u>Income</u>				
Less than \$10,000	19	16.5	12.4	82.6
\$10,000 to \$30,000	121	22.4	17.2	85.2
\$30,000 to \$50,000	108	16.7	13.9	78.8
\$50,000 to \$70,000	98	16.6	13.7	89.9
Over \$70,000	102	15.1	12.7	91.9
<u>Party Identification</u>				
Republican	158	15.9	13.8	85.4
Democrat	157	19.8	15.7	84.0
Independent	145	16.6	13.8	79.2
Other	17	20.2	16.9	85.0

Source: Survey of Florida residents, June 2002

Table 2  
 Heckman Second Stage Regression of Gay Population Estimate on Demographic Variables

	Coefficient	Standard Error	t	Significance
(Intercept)	38.915	5.575	6.980	0.000
Years of Education	-1.008	0.268	-3.758	0.000
Female	4.558	1.332	3.423	0.001
Age in Years	-0.167	0.068	-2.442	0.015
Follow Politics	-0.689	0.828	-0.832	0.406
Black	6.728	1.861	3.616	0.000
Hispanic	4.361	2.209	1.974	0.049
Other Race	2.534	4.172	0.608	0.544
Income category	-0.241	0.582	-0.414	0.679
Democrat	1.615	1.664	0.971	0.332
Republican	0.064	1.755	0.037	0.971
Other Party	2.710	3.451	0.785	0.432
Inverse Mills Ratio	-1.898	8.078	-0.235	0.814
Average Rho	-0.144			
Average R <sup>2</sup>	0.195			
Number of cases	494			

Source: Survey of Florida residents, June 2002

Table 3  
 Estimates of Size of Gay Population and Percent Responding: Threat Variables

	Number of valid responses	Mean estimate	Standard Deviation	Percent Responding
<u>Religious Denomination</u>				
Trad. Protestant	163	16.3	13.9	82.3
Evangelical	78	22.9	17.2	86.7
Catholic	110	15.3	12.3	78.0
Jew	23	13.3	9.7	88.5
Other	76	19.6	17.3	81.7
Atheist/agnostic	37	16.5	12.0	84.1
<u>Religious Attendance</u>				
Never	60	16.8	12.8	80.0
Few times a year	149	17.2	14.1	88.7
Once or twice a month	91	19.0	15.6	83.5
Every week	182	17.4	15.4	78.1
<u>Religious Guidance</u>				
Not important	99	17.5	16.3	83.9
Some	87	18.0	13.6	85.3
Quite a bit	89	15.3	13.4	88.1
Great deal	198	18.4	14.9	78.6
<u>View of the Bible</u>				
Written by men	83	15.1	12.8	89.2
Not literal word of God	219	15.9	13.4	84.9
Literal word of God	167	20.6	16.3	78.0
<u>Beliefs about homosexuals</u>				
Born that way	201	16.8	13.5	84.5
Mixed/In between	39	16.1	16.0	76.5
Choose to be gay	220	18.6	15.3	86.3
<u>Moral Judgment</u>				
Low	135	19.0	14.7	86.0
Middle	133	17.2	15.8	88.1
High	144	17.0	13.7	80.0
<u>Traditional Lifestyle Values</u>				
Low	135	18.3	14.3	86.5
Middle	163	17.3	13.4	84.0
High	159	15.2	13.2	79.9

(continued next page)

Table 3 (continued)

Traditional Marriage Values

Low	273	17.8	14.7	83.0
Middle	113	18.7	14.4	83.1
High	92	14.4	12.3	82.1

Source: Survey of Florida residents, June 2002

Table 4  
 Heckman Second Stage Regression of Gay Population Estimate on Threat Variables

	Coefficient	Standard Error	t	Significance
(Intercept)	25.400	3.284	7.734	0.000
Athiest/agnostic	omitted			
Traditional Protestant	0.175	3.083	0.057	0.955
Evangelical Protestant	4.043	3.353	1.206	0.228
Catholic	-0.570	3.167	-0.180	0.857
Jew	-2.784	3.898	-0.714	0.475
Other Religion	1.507	3.06	0.492	0.623
Bible written by men	omitted			
Bible non-literal, written by God	2.774	2.073	1.338	0.181
Bible literal, written by God	8.314	2.404	3.458	0.001
Never attend religious services	omitted			
Attend few times a year	-0.376	2.285	-0.165	0.869
Attend once or twice a month	2.327	2.62	0.888	0.375
Attend every week	-0.770	2.558	-0.301	0.764
No religious guidance	omitted			
Some religious guidance	-0.329	2.319	-0.142	0.887
Quite a bit of religious guidance	-2.254	2.358	-0.956	0.340
Great deal of religious guidance	0.336	2.276	0.148	0.883
Homosexuals born that way	omitted			
Mixed/In between	-0.948	2.394	-0.396	0.692
Homosexuals choose to be gay	1.168	1.674	0.698	0.486
Moral Judgment scale	-0.096	0.117	-0.821	0.412
Traditional Lifestyles scale	-0.545	0.247	-2.208	0.028
Traditional Marriage scale	-0.311	0.195	-1.59	0.112
Inverse Mills Ratio	-8.693	4.071	-2.135	0.033
Average Rho	-0.591			
Average R <sup>2</sup>	0.105			
Number of Cases	494			

Source: Survey of Florida residents, June 2002

Table 5  
 Estimates of Size of Gay Population and Percent Responding: Contextual Variables

	Number of valid responses	Mean estimate	Standard Deviation	Percent Responding
Percent same sex households in county				
Less than 0.55	173	18.5	15.6	86.1
Between 0.55 and 0.70	163	16.4	14.5	81.9
More than 0.70	158	17.4	13.6	78.6
Gay bars in the county				
None	98	15.6	14.7	81.7
1 thru 7	133	17.3	14.1	78.2
8 thru 12	60	17.2	12.9	81.1
13 thru 25	203	18.6	15.3	85.7
Gay rights ordinance in the county (or a city within the county)				
No	306	17.4	14.5	80.7
Yes	188	17.6	14.8	84.7
Gay information lines				
0	173	16.5	14.2	79.4
1	203	17.7	14.3	82.5
2 thru 4	118	18.3	15.8	86.1
Gay Publications				
0	230	16.6	14.3	80.7
1	61	17.0	13.2	77.2
2	30	18.3	14.7	90.9
3	173	18.6	15.5	84.8
Gay shops in County				
0	282	17.0	14.6	79.9
1	108	17.0	14.8	85.0
2	104	19.1	14.5	86.0
Gay Bookstores in County				
0	344	17.0	14.3	80.0
1	150	18.4	15.3	87.7
Place				
Urban	102	20.1	16.8	81.6
Suburban	358	16.9	13.7	82.5
Rural	34	15.3	16.0	81.0
Know someone who is gay				
No	81	13.7	15.0	77.1
Yes	406	18.1	14.4	84.4
Self-identification				
Hetero/straight	470	17.0	14.0	82.7
Gay/Lesbian/Bisexual	16	26.8	16.5	100.0

Table 6a

## Heckman Second Stage Regression of Gay Population Estimate on Contextual Variables

	Coefficient	Standard Error	t	Significance
(Intercept)	24.664	4.265	5.783	0.000
Percent Same Sex households	-12.480	8.044	-1.552	0.121
Gay Ordinance	-2.221	2.051	-1.083	0.279
Gay Information Lines	-0.219	1.358	-0.161	0.872
Gay Bars	0.303	0.169	1.786	0.075
Gay Services	1.002	0.904	1.109	0.268
Urban	2.870	1.638	1.752	0.080
Rural	-0.643	2.684	-0.240	0.811
Inverse Mills Ratio	-8.012	4.060	-1.973	0.049
Average Rho	-0.531			
Average R <sup>2</sup>	0.033			
Number of Cases	494			

Table 6b

## Heckman Second Stage Regression of Gay Population Estimate on Contact Variables

	Coefficient	Standard Error	t	Significance
(Intercept)	16.158	2.090	7.732	0.000
Know someone who is gay	3.264	1.806	1.808	0.071
Gay/Lesbian/Bisexual	11.434	4.385	2.608	0.009
Inverse Mills Ratio	-6.506	3.837	-1.696	0.090
Average Rho	-0.440			
Average R <sup>2</sup>	0.041			
Number of Cases	494			

Source: Survey of Florida residents, June 2002

Table 7

## Heckman Second Stage Regression of Gay Population Estimate: Comprehensive Model

	Model 1				Model 2			
	Coefficient	Standard Error	t	Significance	Coefficient	Standard Error	t	Significance
(Intercept)	42.939	5.862	7.326	0.000	42.617	5.779	7.375	0.000
Years of Education	-1.052	0.250	-4.211	0.000	-1.040	0.248	-4.186	0.000
Female	4.118	1.323	3.112	0.002	4.262	1.318	3.234	0.001
Age in Years	-0.128	0.060	-2.116	0.035	-0.134	0.060	-2.220	0.027
Black	6.023	1.833	3.285	0.001	6.146	1.832	3.354	0.001
Hispanic	5.173	2.141	2.416	0.016	4.919	2.140	2.298	0.022
Other Race	2.420	4.051	0.597	0.550	2.788	4.051	0.688	0.492
Evangelical Protestant	2.434	1.781	1.367	0.172	2.424	1.785	1.358	0.175
Bible non-literal, written by God	0.540	1.761	0.306	0.759	0.464	1.775	0.261	0.794
Bible literal, written by God	2.562	2.362	1.085	0.279	2.240	2.372	0.944	0.345
Traditional Lifestyles	-0.319	0.218	-1.468	0.143	-0.407	0.205	-1.989	0.047
Traditional Marriage	-0.205	0.180	-1.141	0.254				
Percent Same Sex households	-13.855	6.877	-2.015	0.044	-14.702	6.846	-2.148	0.032
Gay Bars	0.241	0.139	1.726	0.085	0.261	0.138	1.886	0.060
Urban	1.248	1.488	0.839	0.402				
Rural	-1.618	2.463	-0.657	0.512				
Know someone who is gay	2.634	1.706	1.544	0.123	2.861	1.700	1.683	0.093
Gay/Lesbian/Bisexual	8.862	4.474	1.981	0.048	8.798	4.389	2.004	0.045
Inverse Mills Ratio	-0.999	7.186	-0.139	0.889	-0.720	7.221	-0.100	0.921
Average Rho	-0.078				-0.056			
Average R <sup>2</sup>	0.233				0.229			
Number of Cases	494				494			

Source: Survey of Florida residents, June 2002

**Notes**

1. Apart from the intrinsic difficulty of measuring sexual orientation, the stigma of homosexuality is likely to discourage individuals from volunteering such information in surveys or censuses. In 1990, the U.S. Census did introduce a category for households headed jointly by same-sex partners. For many reasons, however, the resulting measure should not be regarded as a comprehensive estimate of the gay population, though it may still be useful for our purposes as we use it to contrast counties by the population density of gays and lesbians.

2. We do not address whether gays face more or less hostility than other minorities. Nor do we overlook the possibility that racial and ethnic prejudice persists in different form than traditionally measured and that, as a result, feeling thermometers and other indicators of affect may be heavily influenced by social desirability. Our argument is simply that anti-gay sentiment is widespread and relatively open, perhaps producing different dynamics when respondents are asked to estimate population size.

3. Alba, Rumbaut, and Marotz (2005) reserve the term innumeracy for those who have difficulty translating perceptions into proportions. The average respondent in the 2000 General Social Survey, for example, believed that the five largest racial-ethnic groups added up to 150% of the U.S. population. Alba and colleagues distinguished between such errors and misestimation of the *relative* size of different groups, recommending ratios (say, estimated white percentage to estimated black percentage) as a better way of determining misperception.

4. According to Gallagher (2003), individuals who are made aware of the disparity between their estimates and the true percentages often attribute it to the impact of local television news. In depth interviews, they claimed that the amount of broadcast attention received by various minorities suggested a much larger population than official data revealed. Paradoxically,

paying attention to local news is commonly perceived as an indicator of political attentiveness that should theoretically lead to more accurate information.

5. Overby and Barth (2002) pointed out that the context-support relationship could be due to endogeneity; that is, individuals who are tolerant of gays may be more likely to have contact because they are not hostile to gays and lesbians, *and* the latter may be more open about their sexuality with people who are believed to be sympathetic.

6. Although no state is a perfect microcosm of the entire country, Florida – from its small towns and rural areas in the north, to the retirement communities in the southeast, to the I-4 Corridor and Disney World in between – is quite diverse in sociodemographic terms, and its political behavior in recent elections has closely mirrored the national pattern.

7. Each respondent was further qualified to ensure that he or she was at least eighteen years of age and a permanent resident of the state. Using American Association for Public Opinion Research's final disposition standards (AAPOR 2000; also see <http://www.aapor.org/pdfs/standarddefs2004.pdf>), the response rate was 36%. Additional information regarding the survey can be obtained from *Florida Voter* directly (954-584-0204), or from the Graduate Program in Political Campaigning in the Political Science Department at the University of Florida (352-392-0262).

8. We were concerned that such outlandishly wild estimates might pose a threat to our general findings. Ultimately, though, we were satisfied that they did not from two replications of the final comprehensive model, presented in Table 7 below. In the first replication, the dependent variable is the natural log of the respondent's estimate of the size of the gay population. In the second replication, we deleted ten cases in which respondents estimated that more than half the population is gay. The substantive findings were not affected by either change in specification.

9. The sampling frame for the UFSRC survey was Florida households with at least one adult resident; the margin of error is plus or minus 4.5 points. Additional information can be obtained directly from the authors.

10. Respondents in 2006 resembled those observed in other surveys, in that they offered fairly high estimates of the size of the African-American (mode and median of 30, mean of 31) and Jewish populations (mode of 10, median 15, mean 18.1) living in the nation as a whole (see Nadeau, Niemi, and Levine 1993).

11. Exact wordings for most questions used in this analysis are provided in the appendix.

12. The alpha statistics reported here are means computed after multiple imputations of missing values (a procedure discussed in greater detail below).

13. The list vendor created the urban/suburban/rural specifications based on census-tract codes for the plurality of telephone numbers in a given area code and exchange.

14. We used the MICE ("multiple imputation using chained equations") routine in the R language. MICE and similar algorithms operate by replacing missing values with a random draw from a distribution estimated from a maximum likelihood function based on other variables in the dataset. (A general discussion can be found in King et al. 2001, and the MICE package is explained in detail by Van Buuren and Oudshoorn 1999.) We drew five imputed values for each missing value which, when combined with the observed non-missing data, form five replicate datasets. Indices were constructed after the multiple imputations were completed.

15. We used the "heckit" routine in the micEcon package in R to estimate the Heckman model. Overby and Barth (2006) were cognizant of the same selection problem, noting that a significantly higher proportion (43.6%) of their respondents said they did not know enough to provide an estimate or simply refused to do so. Obviously, we cannot say whether their higher non-response rates were due to the local focus of their question, to our encouraging respondents

who initially demurred to make a guess, or to some other factor or combination of factors.

16. In bivariate analyses, means and standard deviations are computed from the actual (non-imputed) data. Respondents who declined to classify themselves on the explanatory variables are therefore excluded (pairwise), but are included in the multivariate analyses using the multiple imputation procedure described above in note 14.

17. Even so, the mean estimates provided by high knowledge groups (men, older, college educated, affluent, politically interested, white) are still high compared to the most authoritative surveys of sexual orientation (Badgett 1998: 9-11).

18. The first stage of the Heckman procedure is a probit model of the likelihood of offering a response. Identification of the Heckman model depends on the inclusion of at least one variable in the first-stage (selection) model that is excluded from the second-stage (outcome) model, so we include (along with personal traits) the number of attempts to reach the respondent, as well as a measure of one's interpretation of the Bible (see the appendix). Missing values for the explanatory variables were imputed five times, and the results in Table 2 reflect the pooled estimates of five replicate datasets.

19. We also recognize these patterns could reflect varying levels of simple "numeracy" among demographic groups. That is, people with higher levels of education may have a firmer grasp of the simple mathematical concept of percentages that might help them to recognize, for example, that mutually exclusive and exhaustive groups' shares of the population will necessarily sum to 100%.

### References

- Alba, Richard , Rubén G. Rumbaut, and Karen Marotz. 2005. "A Distorted Nation: Perceptions of Racial/Ethnic Group Sizes and Attitudes Toward Immigrants and Other Minorities." *Social Forces* 84: 901-919.
- Allport, Gordon W. 1954. *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.
- Althaus, Scott L. 1998. "Information Effects in Collective Preferences." *American Political Science Review* 92: 545-558.
- American Association for Public Opinion Research. 2000. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, MI: AAPOR.
- Badgett, M. V. Lee. 1998. *Income Inflation: The Myth of Affluence Among Gay, Lesbian, and Bisexual Americans*. New York: The Policy Institute of the National Gay and Lesbian Task Force (accessed on July 31, 2006 at <http://www.thetaskforce.org/downloads/income.pdf>).
- Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.
- Berinsky, Adam J. 1999. "The Two Faces of Public Opinion." *American Journal of Political Science* 43: 1209-1230.
- Brewer, Paul R. 2003. "The Shifting Foundations of Public Opinion about Gay Rights." *Journal of Politics* 65: 1208-1220.
- Converse, Philip E. 1975. "Public Opinion and Voting Behavior." Pp. 75-169 in *Handbook of Political Science*, vol. 4, eds. Fred I. Greenstein and Nelson W. Polsby. Reading, MA: Addison-Wesley.
- Craig, Stephen C., James G. Kane, and Michael D. Martinez. 2002. "Sometimes You Feel Like a Nut, Sometimes You Don't: Citizens' Ambivalence about Abortion," *Political Psychology* 23: 285-301.

- Craig, Stephen C., Michael D. Martinez, James G. Kane, and Jason Gainous. 2005. "Core Values, Value Conflict, and Citizens' Ambivalence about Gay Rights." *Political Research Quarterly* 58: 5-17.
- Crespi, Irving. 1988. *Pre-Election Polling: Sources of Accuracy and Error*. New York: Russell Sage Foundation.
- Damron Men's Travel Guide*. 2002. San Francisco: Damron Company.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven: Yale University Press.
- Fossett, Mark A., and K. Jill Kiecolt. 1989. "The Relative Size of Minority Populations and White Racial Attitudes." *Social Science Quarterly* 70: 820-835.
- Gallagher, Charles A. 2003. "Miscounting Race: Explaining Whites' Misperceptions of Racial Group Size." *Sociological Perspectives* 46: 381-396.
- Gaziano, Cecilie. 2005. "Comparative Analysis Of Within-Household Respondent Selection Techniques." *Public Opinion Quarterly* 69: 124-157.
- King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95: 49-69.
- Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *Journal of Politics* 62: 790-816.
- Nadeau, Richard and Richard G. Niemi. 1995. "Educated Guesses: The Process of Answering Factual Knowledge Questions in Surveys." *Public Opinion Quarterly* 59: 323-346.
- Nadeau, Richard, Richard G. Niemi, and Jeffrey Levine. 1993. "Innumeracy About Minority Populations." *Public Opinion Quarterly* 57: 332-347.

- Overby, L. Marvin and Jay Barth. 2002. "Contact, Community Context, and Public Attitudes Toward Gay Men and Lesbians." *Polity* 34: 433-456.
- Overby, L. Marvin, and Jay Barth. 2006. "Numeracy About Minority Populations: Americans' Estimations of *Local* Gay Population Size." *Polity* 38: 194-210.
- Sigelman, Lee, and Richard G. Niemi. 2001. "Innumeracy About Minority Populations: African Americans and Whites Compared." *Public Opinion Quarterly* 65: 86-94.
- Stoker, Laura L. 1987. "Morality and Politics: Conduct and Control. A Report on New Items in the 1987 National Election Pilot Study." ANES Pilot Study Report, No. nes002273 (accessed on July 13, 2006 at <http://www.umich.edu/~nes/>).
- Theiss-Morse, Elizabeth. 2003. "Characterizations and Consequences: How Americans Envision the American People." Paper presented to the 2003 Annual Meetings of the Midwest Political Science Association, Chicago, IL.
- Thernstrom, Stephan. 2002. "The Demography of Racial and Ethnic Groups." Pp. 13-36 in *Beyond the Color line: New Perspectives on Race and Ethnicity in America*, eds. Abigail Thernstrom and Stephan Thernstrom. Stanford, CA: Hoover Institution Press.
- Van Buuren, Stef, and Karin Oudshoorn. 1999. *Flexible Multivariate Imputation by MICE*. Leiden, Netherlands: TNO Preventie en Gezondheid.
- van der Meide, Wayne. 2000. *Legislating Equality: A Review of Laws Affecting Gay, Lesbian, and Transgendered People in the United States*. New York: Policy Institute of the National Gay and Lesbian Task Force.
- Wald, Kenneth D., James W. Button, and Barbara A. Rienzo. 1996. "The Politics of Gay Rights in American Communities: Explaining Antidiscrimination Ordinances and Policies." *American Journal of Political Science* 40: 1152-1178.

- Weisberg, Herbert F. 2005. "The Structure and Effects of Moral Predispositions in Contemporary American Politics." *Journal of Politics* 67: 646-668.
- Whaley, A. L., and B. G. Link. 1998. "Racial Categorization and Stereotype-Based Judgments About Homeless People." *Journal of Applied Social Psychology* 28: 189-205.
- Yang, Alan S. 1997. "Trends: Attitudes toward Homosexuality." *Public Opinion Quarterly* 61: 477-507.
- Yang, Alan S. 2001. *The 2000 National Elections Study and Gay and Lesbian Rights: Support for Equality Grows*. Washington, DC: National Gay and Lesbian Task Force Foundation (accessed on July 30, 2006 at <http://www.thetaskforce.org/downloads/NES2000.pdf>).

## **Appendix**

Question wordings for most of the variables employed in our analysis are provided below.

Estimate of Gay Population. What percentage of the overall population would you estimate is either gay or lesbian? (If you had to make a guess, about what percentage would you say?)

Political Interest. Some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say you follow what's going on in government and public affairs most of the time, some of the time, only now and then, or hardly at all?

Party Identification. Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?

Religious Denomination. Is your religious affiliation Traditional Protestant, Evangelical Christian, Catholic, Jewish, or something else?

Religious Attendance. How often do you attend religious services – every week, once or twice a month, a few times a year, or never?

Religious Guidance. Do you consider religion to be an important part of your life, or not? (If yes) Would you say your religion provides some guidance in your day-to-day living, quite a bit of guidance, or a great deal of guidance in your day-to-day life?

View of the Bible. Which of these statements comes closest to describing your feelings about the Bible: One, the Bible is the actual word of God and is to be taken literally, word for word; or, Two, the Bible is the word of God but not everything in it should be taken literally, word for word; or Three, the Bible is a book written by men and is not the word of God.

Beliefs about Homosexuals. Now I'm going to read two statements and ask you to tell me which one comes closer to your own opinion. You might agree to some extent with both, but we want to know which one is closer to your views: (a) Homosexuals are that way because they choose to

be; or (b) People do not choose to be homosexual, they are born that way. Do you feel strongly or not so strongly about that?

Moral Judgment. (1) Now I'm going to read two statements and ask you to tell me which one comes closer to your own opinion. You might agree to some extent with both, but we want to know which one is closer to your views: (a) Sexual relations between two adults of the same sex is morally wrong; or (b) There is nothing immoral about sexual relations between two consenting adults of the same sex. Do you feel strongly or not so strongly about that? (*Answers are coded from 1 to 5, with those who offer a "mixed" response placed in the middle.*) (2) There's been a lot of discussion about the way morals and attitudes about sex are changing in this country. If a man and a woman have sexual relations before marriage, do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all? (3) What if they are in their early teens, say 14 to 16 years old? In that case, do you think sexual relations before marriage are always wrong, almost always wrong, wrong only sometimes, or not wrong at all? (4) And finally, what is your opinion about a married person having sexual relations with someone other than the marriage partner – is it always wrong, almost always wrong, wrong only sometimes, or not wrong at all? (*Answers to items 5-8 range from strongly agree to strongly disagree.*) (5) Sex between two men is just plain wrong. (6) Sex between two women is just plain wrong. (7) I think male homosexuals are disgusting. (8) I think female homosexuals, or lesbians, are disgusting.

Traditional Lifestyle Values. Strongly agree to strongly disagree (5-point scale) with the following: (1) The world is always changing and we should adjust our view of moral behavior to those changes. (2) This country would have many fewer problems if there were more emphasis on traditional family ties. (3) The newer lifestyles are contributing to the breakdown of our society.

Traditional Marriage Values. Strongly agree to strongly disagree (5-point scale) with the following: (1) All in all, family life suffers when the woman has a full time job. (2) It is more important for a wife to help her husband's career than to have one herself. (3) A husband's job is to earn money; a wife's job is to look after the home and family. The threat hypothesis leads us to expect higher estimates of the gay population from individuals who are morally judgmental and who possess more traditional values regarding lifestyle and marriage.

Personal Contact. Do you know anyone personally who is either gay or lesbian? (Yes or no)

Sexual Preference. I'm going to read a list of terms people sometimes use to describe themselves: (a) heterosexual, or straight; (b) homosexual, gay, or (*female respondents only*) lesbian; and (c) bisexual. Now, as I read the list again, please stop me when I get to the term that best describes how you think of yourself.