DSDR | DATA SHARING FOR DEMOGRAPHIC RESEARCH

ICPSR 24384

National Couples Survey, 2005-2006

William R. Grady
Battelle Centers for Public Health Research and Evalulation

John O.G. Billy
Battelle Centers for Public Health Research and Evalulation

Daniel H. Klepinger
Battelle Centers for Public Health Research and Evalulation

Lisa A. Cubbins
Battelle Centers for Public Health Research and Evalulation

Koray Tanfer
Battelle Centers for Public Health Research and Evalulation

Study Methods, Sampling, and Weights

Inter-university Consortium for Political and Social Research
P.O. Box 1248
Ann Arbor, Michigan 48106
www.icpsr.umich.edu

Terms of Use

Married, Cohabitating and Dating Couples: Study Methods

The Married and Cohabitating Couples Study and the Dating Couples Study were jointly fielded for the initial round of data collection in 2005 to 2006 to collect data about contraceptive and disease prevention decisions. Data were collected from each member of each couple so that a determination could be made of how couples jointly decide on the methods they use, how their perception of the risk of sexually transmitted disease influences their decisions, how the decision-making process relates to the consistency and accuracy with which they use that method, and how power differences between the partners affect the way in which they resolve differences in their method preferences and risk perception.

Each study had its own target population which was derived from individuals residing in civilian, noninstitutional housing. Their target populations were restricted to eligible couples defined to be heterosexual couples who currently were potentially at risk for an unintended pregnancy. Specifically, the female partner could not be pregnant or trying to get pregnant and both partners had to assume themselves to be fertile. Because a substantial proportion of women over 35 are married and one or both partners are likely to be sterile, both surveys defined eligible couples to be those where the female partner was age 20 to 35 years old. For the male partner, the only restriction was that he had to be 18 years or older so that he could legally provide informed consent for himself. For the Married and Cohabitating Couples Study, the target population was defined to be those eligible couples currently living together in the same residence as a married or cohabitating couple—there was no requirement that the couple be currently having sexual relations. For the Dating Couples Study, the target population was defined as couples living in separate residences but currently sexually active in a relationship that has lasted for at least one month and where at least one member considered it as their primary relationship.

The sensitive nature of the survey content and the need for simultaneous interviews mandated the use of face-to-face data collection as opposed to telephone interviews. The study was limited to four metropolitan areas: Baltimore, MD; Durham, NC; Seattle, WA; and St Louis, MO. The metropolitan areas studied were defined as the central city plus the county subdivisions immediately adjacent to these cities. These four urban areas were chosen for substantive and pragmatic reasons. On the pragmatic side, Battelle has data collection offices in these four locations, which made the survey cost efficient. On the substantive side, these sites provide diverse populations with respect to culture, race, ethnicity, economic status, and other factors that influence the contraceptive decision-making process.

The sampling frame for this investigation was constructed from the Delivery Sequence File (DSF), a database of residential mailing addresses that the U.S. Postal Service makes available to the general public through license agreements with private companies. Using the DSF to construct the sampling frame allowed us to skip the expensive counting and listing of addresses step typically done to create the frame for face-to-face surveys. Coverage is quite good for metropolitan areas, where most mail is delivered to street addresses. The literature suggests that for urban areas, the DSF file will provide coverage of 95 percent of all residences with the missing units attributable to recent population growth in the locality.[1]

[1] Iannacchione, Vincent G., Jennifer M. Staab, and David T. Redden (2003). "Evaluating the Use of Residential Mailing Addresses in a Metropolitan Household Survey." *Public Opinion Quarterly*, vol. 67, pp. 202-210.

The sample design for this investigation can be characterized as a stratified, probability-proportional-to-size sample of compact clusters. These compact clusters contained 66 adjacent listings from the DSF file, where ZIP+Four zip codes and postal walk sequence codes were used to group addresses so they were physically adjacent. The cluster size of 66 addresses was chosen to ensure that we would obtain 50 occupied housing units on average in each cluster. The sampling began by grouping Census blocks until we obtained segments that were of size 75 or larger based upon our size measure. Using Census-derived data, the segments within each of the four sites were stratified into four strata based upon percent black households. Sample segments were selected with probability proportional to their size measure using systematic sampling.

Each sampled address was visited by an interviewer who verified its residential status and then enumerated the married/cohabitating couples and age-eligible unattached adults living at the sampled address whose members met the study's age requirements. If the household contained multiple couples and/or unattached adults or a mixture of couples and unattached adults, the interviewer consulted a sampling table to randomly select a specific couple or unattached person for eligibility screening based upon the number of couples and the number of unattached persons in the household. This sampling table was custom generated for each sampled address.

For households where a married/cohabitating couple was selected, the interviewer conducted the screening interview with the female partner to confirm she and her partner were age eligible, that they were living together at the sampled address, that she was not pregnant or trying to become pregnant, and that both she and her male partner could be assumed to be fertile. For households where an unattached person was selected, the interviewer confirmed the person's age, that he/she could be assumed to be fertile, that he/she was in a dating relationship with someone of the opposite sex and had sexual relations for at least one month, and that the female partner was not pregnant or trying to become pregnant. If eligible and cooperating, the unattached person then contacted his/her partner and solicited their consent to be contacted by the study interviewer for eligibility screening and study recruitment. To encourage participation in the study, each partner of a sampled married/cohabitating couple was promised an incentive of $40 for completing the interview. Each member of a dating couple was promised $40 for participating and the sampled unattached person was also given $25 for successfully recruiting their partner.

The sample as originally designed was allocated to have approximately equal sample sizes for each of the four urban areas with a modest amount of oversampling used within and across cities to yield the desired sample of completed interviews with eligible married/cohabitating couples and eligible dating couples, with 20% of the married/cohabitating couple interviews and 33% of the dating couple interviews with couples where at least one partner was black. Using available data which was at the national level, we allocated the sample with about 100 segments selected from each urban area, where each segment had 66 addresses sampled for eligibility screening.

In allocating the sample, we first allocated the screeners to each city equally and then allocated them proportionally within strata. This gave exactly equal weights within each study site but did not give the desired allocation of blacks and non-blacks. So we adjusted the sampling rates to get the desired distribution of black couples. Additional sample was added in Durham to boost the overall sample size of completed interviews. This strategy introduced a modest amount of unequal weighting across races within study sites and even less within races within sites.

For the studies supporting the collection of these data, analysis weights were developed that account for the differential probabilities of selection and allow the data for the four sites to be analyzed together. For weight development, we first created analysis weights for each individual site (SITELEVELWT) and then created a combined weight suitable for analyzing the data set as a whole (STUDYLEVELEWT). The analysis weights for each site were constructed by developing sampling weights reflecting the probability of selection of each sampled address and of the couple sampled from that address (if any) and then adjusting these weights to account for nonresponse. The intention of the two couples studies was to analyze the data combined across sites and to have each site have equal impact on the analysis. To facilitate this approach, we created a combined weight by adjusting each site's weights so that they summed to a common population total.

Using the National Couples Survey Data

Data Structure

The dataset includes 2,018 records, with separate records for each individual interviewed. Individuals can be matched into couples on the basis of the **FACESHEETID** variable which has a unique value for each couple.

Variable and Value Labels

The SAS dataset has full variable labels and also contains value labels for all categorical variables. The value labels are contained in the file **formats.sas7bcat**. To use these formats, reference the formats file in the "options" statement in your SAS code. For example:

options pagesize=59 linesize=132 nofmterr fmtsearch=(in.formats);

where "in" is the "libname" for the directory where the format file resides.

Missing Data Codes

Three different missing value codes are defined in the dataset. These are:

.R = "Refused"
.D = "Don't Know"
. = all other missing, including structural missing values (due to valid skips).

Weighting Variables

The dataset includes two weight variables. **SITELEVELWT** contains analysis weights for each individual site (identified in the **SITE** variable). These weights were constructed by developing sampling weights reflecting the probability of selection of each sampled address and of the couple sampled from that address (if any) and then adjusting these weights to account for differential nonresponse. Since the intention of the studies that supported the collection of these data was to analyze the data combined across sites and to have each site have equal impact on the analysis, we also created a combined weight by adjusting each site's weights so that they summed to a common population total. The appropriate weight variable when data from all sites are used is **STUDYLEVELWT**.

Question Wording and Skip Logic Information

In the codebook for the NCS dataset, the question wording on which a variable is based is shown in the **Question/Description** column. Also shown are the conditions defining which respondents were asked the question and, implicitly, who was not asked and have structural missing values on the question. For example, for question **FH04** the question wording shown is: "**Altogether, how many live births have you had?**" Further, below the question wording the following skip information is shown: "**Asked of: women**."