

**Center for Policy Research
Working Paper No. 30**

**HOW WELL CAN WE TRACK COHABITATION USING
THE SIPP? A CONSIDERATION OF DIRECT
AND INFERRED MEASURES**

**Reagan Baughman, Stacy Dickert-Conlin,
and Scott Houser**

**Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu**

October 2000

\$5.00

Up-to-date information about CPR's research projects and other activities is available from our World Wide Web site at www-cpr.maxwell.syr.edu. All recent working papers and Policy Briefs can be read and/or printed from there as well.

CENTER FOR POLICY RESEARCH – Fall 2000

Timothy M. Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret M. Austin
Associate Director,
Budget and Administration

Douglas Holtz-Eakin
Chair, Professor of Economics
Associate Director, Center for Policy Research

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Scott Allard Public Administration
Dan Black Economics
Stacy Dickert-Conlin Economics
William Duncombe Public Administration
Gary Engelhardt Economics
Deborah Freund Public Administration
Vernon Greene Public Administration
Leah Gutierrez Public Administration
Madonna Harrington Meyer Sociology
Christine Himes Sociology
Jacqueline Johnson Sociology
Bernard Jump Public Administration
Duke Kao Economics

Eric Kingson Social Work
Thomas Kniesner Economics
Jeff Kubik Economics
Jerry Miner Economics
John Moran Economics
Jan Ondrich Economics
John Palmer Public Administration
Lori Ploutz-Snyder Health and Physical Education
Grant Reeher Political Science
Stuart Rosenthal Economics
Jodi Sandfort Public Administration
Michael Wasylenko Economics
Assata Zerai Sociology

GRADUATE ASSOCIATES

Reagan Baughman Economics
Robert Bifulco Public Administration
Caroline Bourdeaux Public Administration
Christine Caffrey Sociology
Christopher Cunningham Economics
Tae Ho Eom Public Administration
Seth Giertz Economics
Andrzej Grodner Economics
Rain Henderson Public Administration
Pam Herd Sociology
Lisa Hotchkiss Public Administration
Peter Howe Economics
Benjamin Johns Public Administration

Kwangho Jung Public Administration
James Laditka Public Administration
Xiaoli Liang Economics
Donald Marples Economics
Neddy Matshalaga Sociology
Megan Myers-Hayes Public Administration
Chada Phuengphiniij Economics
Suzanne Plourde Economics
Nora Ranney Public Administration
Mehmet Serkan Tosun Economics
Mark Trembley Public Administration
Rebecca Walker Economics
James Williamson Economics

STAFF

JoAnna Berger Receptionist
Martha Bonney Publications and
Events Coordinator
Karen Cimilluca Librarian/Office Coordinator
Kati Foley Administrative Assistant, LIS
Esther Gray Administrative Secretary
Kitty Nasto Administrative Secretary

Denise Paul Editorial Assistant, *NTJ*
Mary Santy Administrative Secretary
Amy Storfer-Isser Computer Support,
..... Microsim Project
Debbie Tafel Secretary to the Director
Ann Wicks Administrative Secretary
Lobrenzo Wingo Computer Consultant

Abstract

Cohabitation is an alternative to marriage and to living independently for an increasing number of Americans. Despite this fact, research exploring links between living arrangements and economic behavior is limited by a lack of data that explicitly identify cohabiting couples. To aid researchers in using the Survey of Income and Program Participation's (SIPP) rich data for cohabitation issues, our paper considers direct and inferred measures of cohabitation. Our findings suggest that: (1) the best inferred measures in pre-1996 SIPP depends upon a researcher's goals and (2) that the SIPP counts a larger number of cohabiting couples than the widely-used CPS.

I. Introduction

Cohabitation, or living unmarried with a partner of the opposite sex, has become an alternative to marriage and to living independently for an increasing number of Americans. Despite this fact, research that explores links between living arrangements (household structure) and economic status or behavior has been limited by a lack of data that explicitly identify cohabiting couples. The Survey of Income and Program Participation (SIPP) has many merits for addressing cohabitation. However, the SIPP did not explicitly identify cohabiting couples before its 1996 Survey. The goal of this paper is to help researchers use the Survey of Income and Program Participation (SIPP) to analyze issues related to cohabitation. To that end, we consider direct and inferred measures of cohabitation in five panels of the SIPP data. Given the goal, two questions deserve more attention: (1) Why do we care about cohabitation? and (2) Why is the SIPP important in researching cohabitation?

First, cohabitation is an increasingly common living arrangement in the United States, particularly among those of childbearing age. Bumpass and Sweet (1995) find that about one-half the population under the age of 40 has lived with an unmarried partner at some point in their lives. The same study shows that in the past ten years people acceptance of cohabitation as an alternative living arrangement increased among all age groups (Bumpass and Sweet 1995). Ignoring cohabitation when considering economic well-being and ability to pay, for purposes such as calculating tax liability, transfer benefits or poverty, introduces mismeasurement. Using 1990 Census data, Carlson and Danizger (1999) find that the children's poverty rate is understated by 3 percent when ignoring cohabiting couples. Also using 1990 Census data,

Manning and Lichter (1996) find that the poverty rate for single-parent families falls to 31 percent compared to 43 percent when the income of the cohabiting partner is included. Bauman (1999) uses SIPP and CPS data to show the extent to which ignoring cohabiting couples overstates poverty. Ignoring cohabitation also hides inequities. For example, two couples with the same income, but different legal marital status may pay very different taxes (Alm, Dickert-Conlin, and Whittington 1999a) based on that legal marital status.¹ Finally, cohabitation is a margin of behavior that may be affected by economic policy. Alm, Thatcher, and Whittington (1999b), Feenberg and Rosen (1995), Hu (1998) and Moffitt et al. (1998) suggest that some couples choose to cohabit rather than legally marry because of the income tax and welfare penalties for marriage.

Clearly, identifying cohabiting partners will affect measurements of historical trends in living arrangements as well.² In addition to a general need for measurement consistency over time, the time aspect is essential for accurately evaluating changes in family structure in response to recent policies and economic incentives. For example, with the end of welfare as an entitlement following the 1996 Personal Responsibility and Work Opportunity Act, the incentives for living independently diminished. If cohabitation is a likely alternative to living independently, accurate measures of the welfare of women and children depend on accounting for cohabitation. Likewise, the expansion of the Earned Income Tax Credit during the 1990s provides incentives for two-earner couples to cohabit rather than marry (Dickert-Conlin and Houser 2000; Eissa and Hoynes 1999; Ellwood 2000). Having data on family structure before and after these policy changes is essential for analyzing their effects.

Several features make the SIPP an ideal data source for addressing these cohabitation issues. The SIPP consists of a series of eight separate two-to-four year-long nationally

representative panels conducted between 1984 and 1996.³ The SIPP interviews respondents every four months for up to 48 months. Each interview gathers information about the previous four months, resulting in a continuous, monthly record of changes in living arrangements, income, labor force participation, program participation, and other demographics. Because cohabitation is often a short term state (Bumpass and Sweet 1989, find that the median duration of cohabitation is 1.3 years), monthly data enables research into transitions into and out of cohabiting relationships.⁴ Each SIPP panel also includes topical modules, gathered only once or twice during the panel, on wealth, child care, work expenses, taxes, marital and fertility histories (<http://www.sipp.census.gov/sipp/sippov98.htm>). Because of SIPP's rich details on economic well-being, the National Academy of Science recommended that the SIPP replace the Current Population Survey (CPS) as the official data for calculating United States income and poverty statistics (Citro and Michael 1995).

An obstacle to addressing these cohabitation issues in the SIPP is that it did not explicitly identify cohabitators prior to the 1996 panel. Each respondent chose from a list of possible relationships to the reference person (household head), which are explicit for kin or legally married spouses. However, the list did not explicitly identify the relationship of non-relatives to the reference person. In particular, "cohabiting partner" was not an option. Despite this shortcoming, researchers have used the SIPP to address cohabitation, without the ability to directly identify cohabitators (see, for example, London 2000 and Bauman 1999). This problem is not unique to SIPP. In fact, the source used by the Census Bureau to calculate official cohabitation rates, the CPS, also did not explicitly identify cohabitators until 1995 and other researchers have considered these implications (Casper and Cohen 2000 use the CPS and Manning 1995, uses the Census).

To aid researchers in using the SIPP's rich data for cohabitation issues, our paper considers direct and inferred measures of cohabitation. The explicit identification of "unmarried partner" in the 1996 SIPP allows us to gather insight into the challenges of using inferred definitions of cohabitation. We consider the Census Bureau's official CPS-based measure, the POSSLQ, and offer our own measures of inferred cohabitation. We compare our measures to direct measures in the most recent SIPP data and to direct and inferred measures in other data.

Our findings suggest that: (1) researchers using inferred measures in pre-1996 SIPP data would do well to consider their research goal when choosing among several possible inferred cohabitation definitions and (2) estimates of the cohabitating population vary by data source. In particular, the SIPP (whether by direct or inferred measures) counts a significantly larger number of cohabiting couples than the widely-used CPS.

II. Direct and Inferred Measures of Cohabitation in the 1996 Panel

Direct Measure

The SIPP codes all household members by their relationship to the household head. The first panel of the SIPP to explicitly identify cohabitators began in 1996 when the questionnaire offered "unmarried partner" as a choice distinct from roommate or non-relative of the household head. (See the Appendix for the survey question in 1996 and earlier years.) This direct, self-reported measure of cohabitation provides a natural starting point for our analysis. Using a single cross section from January 1996 we calculate the percentage of the adults who report cohabiting.⁵ Throughout the paper, we draw our SIPP samples from the first wave of the panel to avoid attrition biases. Our primary sample includes all adults ages 19 to 65. We use the individual sample weights for the reported "unmarried partners" and their household heads.

By the direct measure, shown in Column (1) of Table 1, we estimate that 4.90 percent of individuals are cohabiting in 1996. Not surprisingly, the rate is highest among younger age groups, 6.03 percent of individuals aged 25 to 44 years cohabit, and lowest among the elderly, only 1.42 percent of individuals aged 60 to 65 years cohabit.

Inferred Measures

Researchers using the 1996 SIPP can directly identify cohabitators. However, for those wishing to use previous SIPP panels, measurement of cohabitation is not straightforward. Our next step is, therefore, to create four inferred definitions of cohabitation in the 1996 SIPP and consider the merits of each relative to the direct measure. There are several obvious demographic characteristics that help to identify *potential* cohabitators within a household. First of all, these individuals must be unmarried (this includes separated, never married, divorced or widowed), of the opposite sex of and unrelated to the head of their household. The household head must also be unmarried. Because the SIPP defines all relationships relative to the household head, unmarried couples in which neither partner heads the household are not identified; therefore, the SIPP will underestimate the number of cohabiting couples in the United States.⁶

The least restrictive method of inferring cohabitation is simply to include anyone who meets each of these four criteria and his or her household head as a cohabiting couple. This method clearly mislabels many individuals who are roommates, boarders or other non-relatives of their household head. However, it also defines the largest possible group of people who *could be* cohabitators.⁷ By this inferred, “Unrestricted” measure, in Column (2) of Table 1, we estimate that 6.99 percent of adults (and 8.22 percent of the 25 to 44 year old subsample) cohabit. Rates follow the same pattern as for the direct measure, peaking in the 25 to 29 year age group.

Tables 2 and 3 provide insight into the sources of error affecting our inferred measure. As a caveat to calling these error rates, we note that direct measure of cohabitation may underestimate true cohabitation rates because of stigma. That is, given the option to explicitly report cohabitation, respondents may be unlikely to do so.⁸ That said, we refer to false positive errors as those when individuals who do not report being cohabitators are labeled as such by an inferred measure. Column (2) of Table 2 shows that 34.2 percent of those identified by the Unrestricted estimate do not report being cohabitators. Column (2) of Table 3 shows that the Unrestricted estimate incorrectly codes older, more educated individuals as cohabitators, relative to the direct measure of cohabiting partners. The Unrestricted measure also incorrectly codes large households without children as cohabiting. These are all characteristics that suggest roommate relationships rather than partnerships. Note also that this definition is likely to code more black individuals as cohabiting than those that report cohabiting.

False negative errors occur when those who report being cohabitators are not labeled as such by an inferred measure. The Unrestricted measure misses only 4 percent of reported cohabitators. Most of these (91 of 105) are part of a same sex couple.⁹ Column (6) of Table 3 shows that the false negative sample has higher education levels and are less likely to have children than the direct measure of opposite sex cohabitators in column (1), which is consistent with findings of same-sex couples by Black et al. 2000. Other papers that investigate direct and inferred definitions of cohabitation (Manning 1995; Casper and Cohen 2000) appear to ignore the possibility that direct measures of cohabitation allow same sex partners to report themselves as such.¹⁰ This source of false positive error will be present in all our inferred definitions. We concentrate on identifying opposite sex cohabiting couples due to the difficulty of inferring estimates of same-sex couples and for comparison with official estimates of cohabitation

partnerships. The remainder of false negatives is couples in which one of the partners reports being married, which are likely to be reporting errors by one of the respondents.

We hypothesize that there are other observable characteristics that we might use to more accurately predict which individuals within this pool are actually the “unmarried partners” of their household heads. To identify the significant characteristics of self-reported cohabitators within the pool of possible ones we defined above, we estimate a logistic regression model. In this model the probability of being part of a couple who reports their status as “unmarried partner” (conditional upon meeting the four criteria listed above) is a function of demographic characteristics such as age, educational attainment, monthly family income, number of children, race and marital status. Other independent variables include the number of adults in the household and the age and education difference between the household head and the possible cohabitor.

Table 4 presents results from this model. Cohabitation rates differ by marital status; those who are widowed are only one-half as likely as the never married to report cohabiting and the divorced are slightly more likely. Women living with unmarried, unrelated male household heads are also significantly more likely to be cohabitators than men living with unmarried, unrelated female household heads. Individuals closer in age to their household head are more likely to report cohabiting, although this is not true for similarity of educational attainment. Finally, the household structure covariates are significantly related to reporting cohabitation. The total number of adults in a household is negatively correlated with the likelihood that it contains a cohabiting couple. At the same time, households with more children are much more likely to contain a cohabiting couple than those without.

Our second inferred estimate of cohabitation uses these regression results to predict cohabitation. The regression allows us to simultaneously consider a number of demographic

characteristics when predicting which individuals in the group of unmarried, unrelated and opposite sex household members might actually be cohabitators.¹¹ From the coefficient estimates in Table 3, we obtain predicted probabilities of cohabitation for each individual. In our pool of possible cohabitators (the least restrictive inferred definition; Column (2) of Table 1), 68 percent report being cohabitators. Our “Regression” inferred definition of cohabitators includes the top 68 percent of the predicted values ($p \geq 0.6395$).

Given the Regression methodology, it is not surprising that the overall rates match the reported rates quite closely - Column (3) of Table 1 shows estimates that 4.81 percent of the total sample and 6.02 percent of the 25 to 44 year old sub-sample cohabit. Because of the regression assumptions, it is also not surprising that both types of errors are equally as large. Column (4) of Table 2 shows that 27.8 percent of Regression-measured cohabitators are false positives and 29.5 percent of reported cohabitators are not recorded as such by the Regression measure.

The regression model suggests at least two criteria that could be used to refine our Unrestricted inferred definition of cohabitation. The first is household structure. Because larger households are much less likely to contain unmarried partners, one might consider restricting household size in an inferred definition. This is the approach taken by the Census Bureau in creating the official, inferred measure of cohabitation, POSSLQ (Persons of the Opposite Sex Sharing Living Quarters). POSSLQ defines cohabiting partnerships as two unrelated individuals of opposite sex living in a household together without *any* other individuals over the age of 15, regardless of their relationships within the household. Compared to our unrestricted measure, this estimate may avoid incorrectly labeling households with roommates and other non-relatives as cohabitators. However, it is also likely to undercount individuals who live with their adult children or other family members.¹²

Column (4) of Table 1 presents cohabitation estimates using the official POSSLQ definition. By this measure, 5.43 percent of the entire adult sample cohabits and the rates by age follow a pattern similar to the direct measure. As expected, because of the more restrictive criteria, the overall POSSLQ rates more closely match the directly measured rates than the Unrestricted measure. Table 3 shows that the rate of false positives is lower (29.1 percent) than for the Unrestricted measure (34.2 percent). However, Table 2 also shows that this estimate does not capture 20.2 percent of the reported cohabitators.

Column (8) of Table 3 confirms that the POSSLQ estimates miss large families with children—most likely because at least one is over age 15. The POSSLQ also misses individuals with less than a high school diploma, most likely because these individual may be living in multiple adult households for income sharing purposes. The POSSLQ method is more likely than other measures to fail to identify black cohabitators. Column (4) of Table 3 shows that, like the Unrestricted estimate, the POSSLQ estimate is likely to include too many individuals with high education and too few individuals with children.

The second restriction suggested by the regression model is age difference between partners. For our final inferred measure of cohabitation, we eliminate from the Unrestricted pool anyone more than ten years older or younger than his or her household head. Ten years is the mean age difference in actual cohabiting couples in the 1996 SIPP. Using this “Age-Restricted” measure, Column (5) of Table 2 shows that an estimated 5.52 percent overall and 6.72 percent of the 25 to 44 year old sub-sample cohabits. As with other inferred definitions, these estimates are higher than the direct measure of married partners in the sample. Column (5) of Table 2 shows that, relative to the Unrestricted sample, this age restriction only slightly lowers the false positive error rate to 31.3 percent from 34.2 percent and greatly increases the false negative error rate to

22.1 from 4.0 percent. Like the three other inferred measures, this estimate includes too many highly educated individuals and omits older individuals who report cohabiting.

Is there a Superior Inferred Measure?

Our results suggest that researchers should be wary of the persistent omission of older individuals in all inferred measures and inclusion of more highly educated individuals, compared to the direct measure. There are also race differences when comparing the inferred measures to the direct measures, suggesting that the living arrangements or reported living arrangements are significantly different between black and white individuals.¹³ With these cautions in mind, researchers looking to infer cohabitation in the SIPP prior to 1996 would do best to choose a definition that fits their research goal.

Taken together, our estimates of rates of cohabitation and demographic characteristics of cohabitators suggest several things. First, the Regression estimates produce overall cohabitation rates that most closely match the direct measures in the 1996 panel of SIPP. However, this is contrived based on knowing what fraction of potential cohabitators is actually cohabiting in that year. Without assuming that the rate of cohabitation among potential cohabitators and that the effect of covariates on the probability of cohabiting are constant over time—two strong assumptions—this inferred cohabitation measure is not particularly useful for measuring historical trends in previous panels of the SIPP.

The POSSLQ definition does the second most accurate job of estimating the direct measure of cohabitation (the job that the Census Bureau primarily intended it to perform). Unlike a regression model, it is simple to implement in datasets without a direct cohabitation measure. On the other hand, the POSSLQ does not accurately describe the demographics of the cohabiting population—primarily because it misses large families with children.

For the purpose of identifying cohabiting couples with children, either an Unrestricted or Age-restricted estimate are preferred. The Unrestricted estimate is the best at capturing the most reported cohabitators. The Age-restricted definition produces the most accurate estimates of cohabitation rates for adults over age 45.

III. Comparisons to Direct and Indirect Measures from other Data Sources

Casper and Cohen (2000) note the need to be cautious when comparing data on cohabitation across different data sources. In this section, we provide researchers with some insight into how SIPP estimates compare to other data.

CPS

We begin by comparing our results to the Casper and Cohen (2000) results for the CPS, the source for official estimates of cohabitation. Although the SIPP sample sizes are typically smaller than the CPS, the SIPP is superior to the CPS in many ways. One significant distinction is that SIPP respondents report income and living arrangements contemporarily, while in the CPS, living arrangements reflect the status at the interview time and income data reflect the previous year.

With respect to identifying cohabitators, the two surveys are comparable because they ask an almost identical survey question (see Appendix) of the same types of people (household heads). In this analysis, we include all SIPP respondents between 15 and 88 (the maximum age) because the CPS data are not age restricted. Table 5 shows that using a direct measure, the CPS estimates that 2,858 thousand couples lived as unmarried partners in 1996, 30 percent fewer than the 4,062 thousand counted in the 1996 SIPP.¹⁴

The first conclusion that we can draw is that even direct measures of cohabitation may vary significantly across data. The most likely explanation is that data is that sampling strategies vary. In the SIPP, interviewers visit households to collect information and assist respondents in identifying their household relationships. In the CPS the first interview is conducted in person, while later ones are often conducted by telephone. If there is stigma attached to cohabitation, relationships within SIPP households may be more accurately reported because of the interviewer's physical presence in the household.

We would like to know if this pattern in the SIPP and CPS is consistent over time. Like the SIPP, the CPS failed to directly identify cohabiting couples in early years (before 1995). To check for consistencies over time, we compare POSSLQ estimates from both data, noting the shortcomings associated with these estimates. From the SIPP data we use Wave 1 cross-sections from the 1990, 1991, 1992 and 1993 panels. For the CPS, we again rely on estimates from Casper and Cohen (2000). Columns (2) and (4) of Table 5 presents these results.

The inferred POSSLQ measures show similar patterns in the growth in cohabitation during the 1990s. However, the SIPP POSSLQ estimates are consistently larger than the CPS estimates, by between 10 and 15 percent. This suggests that there is a persistent difference between the two datasets in the way respondents report household relationships. That is, CPS respondents may be less likely to report the presence of any unrelated individuals in the household and particularly unlikely to report the presence of an unmarried partner.

Census and NSFH

To further test the ability of an inferred SIPP cohabitation measure (POSSLQ) to track cohabitation rates over time, we benchmark it against two other nationally-representative datasets that explicitly identify unmarried partners.

The Census first included a direct measure of “unmarried partners,” relative to the household head, in 1990 (see Appendix). The disadvantage of the Census relative to the SIPP is that it is asked only every ten years and does not have income data contemporaneous with household living arrangements. However, even subsamples of the Census have very large sample sizes: in the 5 Percent Public Use Microdata Sample (PUMS) we have 8.8 million unweighted observations for those over the age of 18. Further, the Census, like the SIPP and the CPS, captures cohabitation only among household heads.

Columns (1) and (2) of Table 6 compare our 1990 SIPP POSSLQ estimates to the 1990 Census. The inferred SIPP cohabitation rate for 1990 overstates the direct rate measured by the Census by approximately 0.5 of a percentage point, or 13 percent. This is in line with the 10 percent difference we found when we compared direct and inferred measures in the 1996 SIPP. The 1990 inferred SIPP estimates also match up relatively well against the Census within age groups. This provides us with some evidence that the inferred measure may do well over time. However, given that the sampling methodology is different in these two samples (Census is primarily by mail), we view this evidence as only suggestive.

Finally, we compare the SIPP POSSLQ measure to a direct measure from the National Survey of Families and Households (NSFH), which was specifically designed to collect information on topics surrounding the family (Bumpass and Sweet 1989). Respondents in the NSFH are randomly chosen adult members of a household, so that cohabitating individuals other than household heads and their partners will be identified as such (for example, a son cohabiting

with his girlfriend while living in his parents' home). For this reason the NSFH provides a more accurate representation of the cohabiting population than can be obtained from the SIPP, CPS or Census. We use 10,008 respondents in the second wave of data (1992-94) because this timeframe most closely matches our recent SIPP data.¹⁵ We identify cohabitators in the NSFH as anyone who responds that a person living with them is their “[L]over/partner” (see Appendix). The NSFH also has the advantage of being a panel—the first wave occurring in 1987 and 1988 and the third wave beginning in the fall of 2000. However, the NSFH has much less detail on sources of income than from the 1992 and 1993 panels of the SIPP.

The NSFH estimates of cohabitation are higher than the SIPP's for all but the oldest age category, 60 to 65 years. For the 19 to 65 sample, NSFH estimates that 5.59 of the population cohabits. This is 22.6 and 16.0 percent higher than the POSSLQ estimates for the 1992 and 1994 SIPPs, respectively. The difference is even greater in the 25 to 44 year age range, where the NSFH estimates find 7.49 percent of the population cohabiting, relative to 5.15 and 5.61 in the 1992 and 1994 SIPPs. We expect higher estimates, than even the SIPP POSSLQ estimates (which, recall, are higher than the direct measures in SIPP), for two reasons. First, the NSFH captures cohabitation among adults other than household heads. Second, the NSFH's goal is to accurately depict family structure, so their interviewing techniques may capture higher levels of cohabitation.

IV. Conclusion

In this data exercise with the 1996 SIPP we develop four inferred estimates of cohabiting relationships and test their effectiveness in both (1) estimating the prevalence of cohabitation in the United States and (2) describing the characteristics of cohabitators. We find that different methods of inferring cohabitation in pre-1996 data perform differently depending upon a

researcher's goal. For tracking the prevalence of cohabitation over time, the Census Bureau POSSLQ measure works well. For better estimates of subpopulations, including couples with children and older adults, we present two inferred cohabitation measures that are better suited to the task.

We also compare SIPP to the CPS, Census and NSFH estimates of cohabitation. By both direct and inferred measures the SIPP counts a larger number of cohabitators than the CPS, which the Census Bureau uses as its official data source for estimating cohabitation and poverty. We believe that this finding reflects the SIPP's more frequent personal interviewing and is incentive to consider using the SIPP to study living arrangements.

Appendix: Survey Questions about Cohabitation

1. The Survey of Income and Program Participation (SIPP)

1996 Variable. Each respondent's relationship to the household reference person is categorized as one of the following:

- Reference person living with relatives
- Reference person living alone or with only non-relatives
- Spouse of reference person
- Child of reference person
- Grandchild of reference person
- Parent of reference person
- Brother/sister of reference person
- Other relative of reference person
- Foster child of reference person
- Unmarried partner of reference person
- Housemate/roommate
- Roomer/boarder
- Other non-relative of reference person

Pre-1996 Variable (Edited).¹⁶ After a data editing process, each respondent's relationship to the household reference person is categorized as one of the following:

- Reference person living with relatives
- Reference person living alone or with only non-relatives
- Spouse of reference person
- Child of reference person
- Other relative of reference person
- Non-relative of reference person but related to others in the household
- Non-relative of reference person and not related to anyone else in the household

2. Current Population Survey (CPS), post-1994

Each respondent categorizes his/her relationship to the household reference person by one of the following:

- Reference person without relatives in the household
- Reference person with relatives in the household
- Spouse
- Child
- Grandchild
- Parent
- Brother/sister
- Other relative of the reference person
- Foster child
- Nonrelative of with own relatives in household
- Nonrelative of without own relatives in household
- Unmarried partner with own relatives in household
- Unmarried partner without own relatives in household
- Housemate/roommate with own relatives in household
- Housemate/roommate without own relatives in household
- Roomer/boarder with own relatives in household
- Roomer/boarder without own relatives in household

3. The 1990 Census (5 Percent Public Use Microdata Extract)

Each respondent categorizes his/her relationship to the household reference person by one of the following:

- Householder
- Husband/Wife
- Son/daughter
- Stepson/stepdaughter
- Brother/sister
- Father/mother
- Grandchild
- Other relative
- Roomer/boarder/foster child
- Housemate/roommate
- Unmarried partner
- Other nonrelative
- Institutionalized Person (excluded from our sample for consistency with the SIPP)
- Other Person in Group Quarters

4. The National Survey of Families and Households (NSFH2), 1992-1994

The respondent is asked to describe his/her relationship to each member of the household. Following are the relationship codes relevant to cohabitation (out of 21 possible codes).

- Husband or wife
- Lover/partner
- Biological child
- Step-child
- Adopted child
- Foster child
- Child of Lover/Partner
- Son- or Daughter- in-Law
- Mother or Father
- Step-parent
- Mother- or Father-in-Law or Partner's parent
- Grandparent
- Brother or Sister
- Step-brother or Step-sister
- Half-brother or Half-sister
- Brother-or-Sister-in-Law
- Grandchild
- Other Relative
- Roommate
- Friend
- Other Non-relative

Endnotes

1. Is it not obvious that two such couples should be treated equally. The extent of income sharing within couples is likely to determine ability to pay (Winkler 1997).
2. In related work, London (1998) shows the undercount of single-mothers when ignoring the possibility that single mothers are not the heads of the family.
3. There were full SIPP panels for 1984, 1985, 1986, 1990, 1991, 1992, 1993 and 1996. Panels were also started in 1988 and 1989, but the Census Bureau terminated both early. The 1996 panel runs through 2000.
4. We do not exploit this advantage in this paper.
5. We use the first wave of data in all SIPP panels discussed in this paper.
6. This problem is common to most other data. SIPP can do somewhat better because relationships between all household members are given in the SIPP Household Relationship Topical Module File. However, this data is only gathered once during each SIPP panel.
7. It is possible for more than one individual per household to be identified as a cohabitor by this method. In these cases, we label as the cohabitor the individual who is closest in age to the household head. In the very small number of cases where the age difference is equal, we do not code anyone in that household as a cohabitor.
8. Bumpass and Sweet (1992) support the possibility of stigma. Using NSFH data, they show that about one-third of those under age 50 and more than one-half of the respondents over age 50 disapprove of cohabitation.
9. There may be odd numbers of individuals because one member of the couple may be outside our 19 to 65 age range.
10. For the purposes of measuring economic well-being, the income sharing among same-sex couples may be similar to opposite sex cohabiting partners. Therefore, this is a group that deserves more attention than we give it. See Black et al. (2000) for more descriptive statistics on same sex couples. Inferred estimates of same-sex partnerships are impossible in our data, due to the prevalence of same sex roommates.
11. We further eliminate from this pool anyone in a household where 2 or more possible cohabitators have been identified. Our process is similar to that used by Casper and Cohen (2000) to create an inferred measures of cohabitation in the CPS.
12. This is the critique of POSSLQ addressed by Casper and Cohen (2000).

13. Manning (1995) finds that there are large errors in race using the 1990 PUMS.
14. Casper and Cohen (2000) find a similar result in a subsample of unmarried women aged 25 to 49.
15. Because of the sampling scheme, only 61 of the respondents in the NSFH2 are under the age of 25. Exactly two are 22 years of age, eight are 23 years of age and 51 and 24 years of age.
16. The pre-1996 panels also contains an unedited version of the household relationship variable which breaks non-relatives into “partner/roommates” and others. However, in previous work (Baughman, Dickert-Conlin, and Houser 1999) we found the unedited version of the variable to be unreliable.

Table 1. Percent of Population Cohabiting: Direct and Inferred Measures 1996 SIPP

Age Group	Direct Measure (1)	Inferred Measures			Age-Restricted (5)
		Unrestricted (2)	Regression (3)	POSSLQ (4)	
19 to 65 (total sample)	4.90	6.99	4.81	5.43	5.52
25 to 44 (subsample)	6.03	8.22	6.02	6.47	6.72
19 to 24	6.41	9.95	7.49	7.95	8.89
25 to 29	8.56	11.77	9.32	10.06	10.75
30 to 34	6.16	8.51	6.66	7.23	7.14
35 to 39	5.53	7.31	4.87	5.33	5.30
40 to 44	4.07	5.61	3.52	3.61	4.09
45 to 49	3.55	5.02	2.80	3.50	3.38
50 to 54	2.88	4.25	2.01	3.08	2.40
55 to 59	2.04	3.43	1.39	2.50	1.82
60 to 65	1.42	2.45	0.98	1.96	1.36

Definitions:

- Direct Measure – An individual who reports being “unmarried partner” to the reference person and his or her reference person.
- Unrestricted – Any unmarried, unrelated individuals of the opposite sex in the same household.
- Regression – Predicted cohabitators based on regression in Table 4. Use predicted value from logit to choose top 68 percent. This is the percent in the Unrestricted estimate who directly report being unmarried partners.
- POSSLQ – Census Bureau Definition. Persons of the Opposite Sex Sharing the Same Living Quarters; restricted to households with only 2 adults over age 15.
- Age Restricted – Same as the Unrestricted, with the age difference between partners in a couple restricted to ten years.

Source: Authors’ calculations using Wave 1 of the 1996 SIPP.

**Table 2. Error Rates: Individuals Misidentified by Inferred Measures
Full Sample, 1996 SIPP**

		Direct Measure (1)	Inferred Measures			
			Unrestricted (2)	Regression (3)	POSSLQ (4)	Age- Restricted (5)
False Positives	<i>Percent of Individuals Identified by Inferred Measure who Are Not Actual Cohabiters</i>	0	34.2	27.8	29.1	31.3
False Negatives	<i>Percent of Actual Cohabiters Not Identified By Inferred Measure</i>	0	4.0	29.5	20.2	22.1

Definitions:

- Direct Measure – An individual who reports being “unmarried partner” to the reference person and his or her reference person.
- Unrestricted – Any unmarried, unrelated individuals of the opposite sex in the same household.
- Regression – Predicted cohabiters based on regression in Table 4. Use predicted value from logit to choose top 68 percent. This is the percent in the Unrestricted estimate who directly report being unmarried partners.
- POSSLQ – Census Bureau Definition. Persons of the Opposite Sex Sharing the Same Living Quarters; restricted to households with only 2 adults over age 15.
- Age Restricted – Same as the Unrestricted, with the age difference between partners in a couple restricted to ten years.

Source: Authors’ calculations using Wave 1 of the 1996 SIPP.

Table 3 Means and Standard Deviations: Demographic Characteristics of Individuals Misidentified by Inferred Measures, Full Sample, 1996 SIPP^a

	Direct Measure (1)	False Positives				False Negatives			
		Unrestricted (2)	Regression (3)	POSSLQ (4)	Age-Restricted (5)	Unrestricted (6)	Regression (7)	POSSLQ (8)	Age-Restricted (9)
N	2,629	1,211	686	809	892	105	745	540	567
Age	34.6 (10.3)	35.3* (11.8)	33.2*** (10.0)	35.1 (11.6)	33.4*** (10.8)	37.9*** (10.0)	38.8*** (10.8)	37.8*** (9.5)	40.0*** (10.4)
Education	12.5 (2.5)	12.6 (2.5)	12.8*** (2.3)	12.8*** (2.4)	12.7*** (2.3)	13.9*** (2.8)	12.7*** (2.8)	12.0*** (2.8)	12.4 (2.8)
Percent less than High School	17.6 (38.1)	13.8*** (34.5)	10.6*** (30.8)	11.6*** (32.0)	11.4*** (31.8)	11.0* (31.4)	19.4 (39.5)	26.2*** (44.0)	21.1* (40.8)
Monthly Income	\$1,772 (\$1,747)	\$1,746 (\$2,871)	\$2,002** (\$3,602)	\$1,782 (\$3,139)	\$1,840 (\$3,204)	\$2,021 (\$1,851)	\$1,841 (\$1,932)	\$1,779 (\$1,599)	\$1,818 (\$2,014)
Percent White	82.7 (37.8)	78.5*** (41.1)	81.0 (39.2)	80.8 (39.4)	78.2*** (41.3)	88.7 (31.8)	78.9** (40.8)	78.0 (41.5)	82.1 (38.4)
Percent Black	13.2 (33.9)	15.9** (36.5)	14.4 (35.2)	14.0 (34.7)	15.7* (36.4)	10.6 (30.9)	13.7 (34.4)	18.5*** (38.9)	14.3 (35.1)
Percent with Children	20.8 (40.6)	16.7** (37.3)	21.1 (40.9)	16.5* (37.1)	17.6** (38.1)	11.8** (32.5)	15.4** (36.2)	29.0*** (45.4)	20.0 (40.1)
Household Size	2.9 (1.2)	3.1*** (1.3)	2.8*** (1.1)	2.5*** (0.9)	3.0 (1.2)	2.7*** (1.4)	3.0*** (1.5)	3.9*** (1.5)	3.0** (1.4)

^aSignificantly different from direct measure value at *** 1% level, ** 5% level, * 10% level.
Source: Authors' tabulations from 1996 Survey of Income and Program Participation.

Table 4 Results of Logistic Regression, Dependent Variable is Direct Measure of Cohabitation, 1996 SIPP^a

Independent Variables	Odds Ratio	Standard Error
Age	1.00	0.005
Education	0.96*	0.022
Monthly Income	1.00	0.001
White	1.38	0.305
Black	1.42	0.365
Divorced	1.26*	0.165
Separated	0.91	0.208
Widowed	0.59*	0.178
Female	1.31**	0.142
Age Difference	0.95***	0.007
Education Difference	1.01	0.028
Number of People in Household	0.47***	0.041
Number of Children of Possible Partner	1.91***	0.262
Number of Children of Household Head	2.71***	0.296
Log Likelihood		-1148.73

^a Sample used for estimation includes 1,953 individuals who meet the criteria of being possible cohabitators; they are (1) unmarried, (2) of the opposite sex of their household head and (3) not related to the household head. We also eliminate from this pool anyone in a household with more than one individual meeting these criteria. Omitted category for race is Other; omitted category for marital status is never married.

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level.

Table 5
 Number of Cohabiting Couples in Thousands
 All Ages

**Table 5. Number of Cohabiting Couples in Thousands,
 All Ages**

	Direct SIPP^a	POSSLQ SIPP^a	Direct CPS^b	POSSLQ CPS^b
	(1)	(2)	(3)	(4)
1996	4,062	4,570	2,858	3,958
1995	NA	NIW1	2,641	3,667
1994	NA	NIW1	NA	3,662
1993	NA	3,931	NA	3,510
1992	NA	3,736	NA	3,308
1991	NA	3,343	NA	3,039
1990	NA	3,303	NA	2,856

NIW1 – Not in Wave 1 of the SIPP Panel

NA – Not available

POSSLQ –Census Definition. Persons of the Opposite Sex Sharing the Same Living Quarters;
 Restricted to households with only 2 adults over age 15.

Sources: ^a Authors' calculations from the various panels of the SIPP. ^b Casper and Cohen
 (2000). There are no age restrictions on the counts in these tables for direct
 comparison to Casper and Cohen.

Table 6. Percent of Population Cohabiting, 1990-1994, Comparison of Inferred SIPP Measures to, Direct Measures from Other Data

Age	1990 Census (1)	1990 SIPP POSSLQ (2)	1992-1994 NSFH (3)	1992 SIPP POSSLQ (4)	1993 SIPP POSSLQ (5)
19 to 65 (sample) ^a	3.59	4.12	5.59	4.56	4.82
25 to 44	5.08	4.82	7.49	5.15	5.61
19 to 24	5.61	6.81	17.27 ^b	7.59	7.12
25 to 29	6.81	6.67	12.48	8.38	8.00
30 to 34	5.42	5.62	7.85	5.37	6.50
35 to 39	4.23	3.79	5.39	3.53	4.02
40 to 44	3.58	2.69	4.21	3.15	3.82
45 to 49	2.89	2.11	3.24	2.31	3.30
50 to 54	2.21	1.98	2.62	2.54	3.08
55 to 59	1.55	1.65	2.75	2.00	1.71
60 to 65	1.05	1.18	0.73	1.78	1.69

^aBecause the maximum age differs in these data, we restrict our sample to 19 to 65 years.

^bBecause of the sampling scheme, only 61 of the respondents in the NSFH2 are under the age of 25. Exactly two are 22 years of age, eight are 23 years of age and 51 and 24 years of age.

Source: Authors' calculations from the 1990 Census 5 Percent Public Use Micro Sample, 1990, 1992 and 1993 SIPPs and the Second Wave of the National Survey of Families and Households.

References

- Alm, James, Stacy Dickert-Conlin, and Leslie Whittington. 1999. "Policy Watch: The Marriage Penalty," *Journal of Economic Perspectives*, 13(3): 193-204.
- Alm, James, Jennifer Thatcher, and Leslie Whittington. 1999. "Shacking up or Shelling Out." Working Paper presented at the National Tax Association Meetings. Atlanta, GA: Georgia State University, November
- Bauman, Kurt J. 1999. "Shifting Family Definitions: The Effect of Cohabitation and Other Nonfamily Household Relationships on Measures of Poverty," *Demography*, 36(3): 315-25.
- Black, Dan, Gary Gates, Seth Sanders, and Lowell Taylor. 2000. "Demographics of the Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources," *Demography*, 37(2): 139-55.
- Bumpass, Larry L. and James A. Sweet. 1995. "Cohabitation, Marriage, and Union Stability: Preliminary Findings from NSFH2," NSFH Working Paper No. 65. <http://www.ssc.wisc.edu/cde/nsfhw/nsfh65.pdf>
- Bumpass, Larry L. and James A. Sweet. 1989. "National Estimates of Cohabitation," *Demography*, 26: 615-625.
- Constance F. Citro and Robert T. Michael (eds.). 1995. *Measuring Poverty: A New Approach*. Washington, D.C. : National Academy Press.
- Carlson, Marcia and Sheldon Danziger. 1999. "Cohabitation and the Measurement of Child Poverty," *Review of Income and Wealth*, 45(2): 179-91.
- Casper, Lynne and Philip N. Cohen. 2000. "How Does POSSLQ Measure Up? Historical Estimates of Cohabitation," *Demography*, 37(2): 237-45.
- Dickert-Conlin, Stacy and Scott Houser. 2000. "EITC and The Female Headship Decision," working paper, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University.
- Eissa, Nada and Hilary Hoynes. 1999. "Good News for Low-Income Families? Tax-Transfer Schemes, and Marriage," working paper, Department of Economics. Berkeley: University of California. http://emlab.berkeley.edu/users/hilary/wp/jcpr_marriage.pdf
- Ellwood, David. Forthcoming. "The Impact of the Earned Income Tax Credit and Social Policy Reforms on Work, Marriage, and Living Arrangements," *National Tax Journal*.
- Feenberg, Daniel and Harvey Rosen. 1995. "Recent Developments in the Marriage Penalty," *National Tax Journal*, 48(1): 91-101.

- Hu, Wei-Yin. 1998. "Welfare, Marriage, and Cohabitation: Experimental Evidence from California," working paper, Department of Economics. Los Angeles: University of California.
- London, Rebecca A. 2000. "The Interaction between Single Mothers' Living Arrangements and Welfare Participation," *Journal of Policy Analysis and Management*, 19(1): 93-117
- London, Rebecca A. 1998. "Trends in Single Mothers' Living Arrangements from 1970 to 1995: Correcting the Current Population Survey," *Demography*, 35(1): 125-31
- Manning, Wendy D. 1995. "Comparisons of Direct and Inferred Measures of Cohabitation," working paper, Population Research Institute. State College, PA: Pennsylvania State University.
- Moffitt, Robert A., Robert Reville, and Anne E. Winkler. 1998. "Beyond Single Mothers: Cohabitation, Marriage, and the U.S. Welfare System," *Demography*, 35(3): 259-78.
- Rindfuss, Ronald R. and Audrey Vandenheuvel. 1990. "Cohabitation: A Precursor to Marrying or an Alternative to Being Single?" *Population and Development Review*, 16(4): 703-726
- U.S. Bureau of the Census. 1997. <http://www.sipp.census.gov/sipp/chap2-3.htm>
- U.S. Bureau of the Census. 1991. Current Population Reports, Series P-20, No. 450. *Marital Status and Living Arrangements: March 1990*. Washington, DC: Government Printing Office, Washington, DC.
- Winkler, Anne E.. 1995. "Does AFDC-UP Encourage Two-Parent Families?" *Journal of Policy Analysis and Management*, 14(1): 4-24
- Winkler, Anne E. 1997. "Economic Decision-Making by Cohabitors: Findings Regarding Income Pooling," *Applied Economics*, 29(8): 1079-90.