# Dads and cads: parental cohabitation and the human sex ratio at birth

Karen Norberg, MD

*Boston University*

*National Bureau of Economic Research, 1050 Massachusetts Ave., Cambridge, MA 02138*

*norberg@nber.org*

Draft: December 30, 2002

Comments gratefully accepted

**Abstract:**

Modern sex allocation theory predicts that parents may be able to vary the sex of their offspring according to the prospects for two-parent care. Using data pooled from four publicly available longitudinal studies, I find that parents who were living with an opposite-sex spouse or partner before the child's birth were significantly more likely to have a male child than parents who were living apart ($X^2$ 21.74, df=1, $p < .0001$). The effect is observable even when the comparisons are made between siblings (odds ratio 1.17, $p < .001$), and even when the sibling comparisons are made before the children's conceptions (odds ratio 1.14, $p < .01$). This "partnership status effect" may be the result of modern reproductive exposures, but a paternal investment effect would fit closely with the predictions of adaptive sex allocation theory.

Dads and cads: parental cohabitation and the human sex ratio at birth

Modern sex allocation theory predicts that parents may respond to a variable environment by varying the sex of their offspring, if the consequences of variations in the environment differ depending on whether the offspring is male or female. (1, 2) Two-parent care is an important, but far from universal, feature of human reproduction; if the fitness consequences of one-parent care were different for sons and daughters, then it might be adaptive for parents to vary the sex of their offspring according to signals forecasting two-parent care. The consequences of one-parent care might be different for sons and for daughters if there is nonlinearity in the effects of parental investment which differs by sex (1), if offspring derive specific benefits from contact with a same-sex parent (3?), or if one sex were more likely than the other to function as "helpers at the nest"(3?). All else being equal, if one-parent care is more of a disadvantage for sons than for daughters, then parents should produce fewer sons under conditions forecasting one-parent care, and more sons under conditions forecasting two-parent care.

Marriage is, by definition, a social institution closely tied to the prospects for two-parent care(5, 6). The present study was motivated by nineteenth century reports of a larger fraction of male children among legitimate births compared to illegitimate births; many of these studies were mentioned by Charles Darwin in his discussion of the sex ratio in *The Descent of Man* (4). However, twentieth century studies no longer find an association between the mother's marital status and offspring sex (7, 8). Paternal investment has received little attention from modern sex allocation theory, perhaps because of this lack of evidence from modern human populations.

There have been many changes in western society since the nineteenth century, and "shotgun" marriage and non-marital cohabitation could have confounded previous efforts to find a paternal investment effect in modern populations. In the present study, I investigate whether parents who were living together before a child's birth were more likely to have a male child than parents who were living apart, whether or not the parents were legally married. Since it is possible that child sex might influence the parent's prenatal behavior, I use observations of the parent's household composition before conception or during the first trimester of pregnancy, when this information is available. The analysis makes use of data pooled from four US surveys: the 1959 National Collaborative Perinatal Project (NCPP), the 1969 National Longitudinal Survey of Young Women, (NLS-YW), the 1979 National Longitudinal Survey of Youth (NLSY), and the 1995 National Survey of Family Growth (NSFG) (9,10,11). Together, these surveys yield a sample of 80,715 children born to 58,104 US families over a span of four decades, from 1959 to 1998. The NLS-YW, the NLSY, and the NSFG were designed to be statistically representative of specific US cohorts, and provide information about the presence or absence of a spouse or unmarried partner in the respondent's household before the child's conception. The earliest survey, the 1959 NCPP, is important to the present study because it predates the legalization of abortion, but it is based on a convenience sample, and household composition is first recorded when the mother registers for prenatal care. About 50% of pregnancies in this sample were registered by 20 weeks of gestation. I consider these four studies separately and together; the pooled sample (with or without the NCPP) is large enough to permit multivariate analyses, and even more importantly, it is large enough to control for many background factors by comparing siblings within the same family. These within-family comparisons estimate the differences in the likelihood of having a male child associated with differences in the responding parent's partnership status before each conception or birth. These comparisons control for the responding parent's genetic endowment, fixed propensities to produce male or female offspring, conditions of rearing, "personality", and other factors that would remain unchanged between pregnancies. However, such comparisons do not control for characteristics of the second parent among half-siblings; nor do they control for household income, or mother's nutritional status, psychological stress, occupational exposures, contraceptive exposures, access to medical care, or other factors which might vary from one pregnancy to another.

Table 1 describes the four samples, and the appendix gives further details. In spite of a decline in the prevalence of marriage, there was a nearly constant association between child sex and parent household composition across the four surveys: there were no significant differences in the fraction of boys among children born to couples who were living together by early pregnancy, but there was a significantly lower fraction of boys among children born to parents who were living apart. This bivariate association was statistically significant in three of the four studies when they were examined separately, although the difference is small. In the pooled sample, there were 51.6% boys among children whose parents were living together, whether married or unmarried, at the first prenatal visit or before the end of the first trimester, and 49.6% boys among children whose parents were living apart, with a $X^2$ value (stratified by study) of 21.74, df=1, $p < .0001$; this corresponds to a stratified odds ratio of 1.081 (95% c.i. 1.046 to 1.117). The association remained significant in the non-NCPP pooled sample when the parent's household was observed before the child's conception (stratified $X^2$ 9.71, df=1, p=.002).

In Table 2, Models 1 and 2 extend this analysis using the pooled sample, with logistic regression models predicting the odds of a male versus female birth, and adjusting for parental age, sex, education, and ethnicity, and child birth order and year of birth. Except for parent's sex, the control variables were selected because they were available in all four surveys, and because they have been associated with offspring sex in other human studies (7, 8). They are presented as linear or dummy variables, but quadratic terms and other transformations did not change the results: partnership status was the strongest predictor of child sex, even when household composition was observed before conception. Model 1 indicates that children whose parents were living together before or during early pregnancy were 6.4% more likely to be male than children whose parents were living apart (odds ratio 1.064, 95% c.i. 1.028 to 1.102, $p < .0001$ CHECK; $Z=3.55$). The partnership status estimates did not differ by study sample, year of child's birth, or parent age, sex, parity, or ethnicity. The estimated odd ratio is 1.059 in Model 2, which is based on the smaller pooled sample (n= 27,216) of households observed before the child's conception.

The estimates were slightly larger in magnitude – although measured with larger standard errors- when the comparisons were made within families. Models 3 and 4 use conditional logistic regression to make comparisons among siblings; in these models, *only* the parent's partnership status predicted child sex. Model 3 indicates that siblings whose parents were living together before or during early pregnancy were 16.7% more likely to be male than siblings whose parents were living apart (odds ratio 1.167, 95% c.i. 1.072 to 1.270, $p < .001$). Again, the estimates change little in Model 4, which uses household observations before the child's conception.

In summary, girls start out in families that differ from the families of boys. The sibling comparisons in model 4 would correspond to a prevalence of about 53% male births among siblings who were conceived when the parents were living together (whether married or unmarried) before the child's conception, compared to a prevalence of about 50% male births among siblings conceived when the parents were living apart. This effect was not explained by paternal bias against the acknowledgment of girls, nor by the parent's ethnicity, age, sex, race, education, or parity, and it was not explained by the fixed genetic, behavioral or environmental background factors that would be shared among full and half siblings within the same family.

It would be hazardous to jump to an adaptive conclusion in the face of modern reproductive exposures. Even if the "cohabitation effect" is not the result of specifically modern exposures, selective miscarriage could be the non-adaptive byproduct of other processes (4, 8). That being said, the within-family results do indicate a cause that is closely tied to reproductive behavior. Although it would be possible for modern reproductive technologies to have previously undetected effects, the partnership status effect is discernable in the early 1960's, well before the legalization of abortion in the US, and it resembles the "legitimacy status" effect widely reported in the nineteenth century, long before the introduction of modern contraception. A paternal investment effect would fit closely with the predictions of sex allocation theory, although much further evidence would be needed to confirm or deny this interpretation.

1. E. L. Charnov, R. L Los-den Hartogh, W.T. Jones and J. van den Assem (1981) "Sex ratio evolution in a variable environment" Nature 289, P 27-33

2. R. L. Trivers, and D. E. Willard (1973): "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring" *Science,* **179**: 90-92

3. D. Hasselquist and B. Kempenaers (2002) “Parental care and adaptive brood sex ratio manipulation in birds” *Phil Trans R. Soc London* 357(1419), 363-372

4. C. Darwin (1871) *The Descent of Man* Random House, New York (chapter 8)

5. G. S. Becker (1974) “A Theory of Marriage: Part II” *Journal of Political Economy* **82** (March/April, Part II) S11-S26

6. R. J. Willis (1999) “A Theory of Out-of-Wedlock Childbearing” *Journal of Political Economy*, 107(6), s33-s64

7. W. H. James (1987)"The human sex ratio, Part 1. A review of the literature" *Human Biology* 59:721-752

8. I. Hardy (2002) *Sex Ratios: Concepts and Research Methods* Cambridge, UK: Cambridge University Press

9. National Archives of the United States, Collaborative Perinatal Project Documentation (1985). Available on the Internet at www.nber.org/docs/cpp/docs/cpp-vol1.pdf

10. U.S. Bureau of Labor Statistics (2001) *NLS Handbook, 2000* US Government Printing Office: 2001-472-553/44953

11. U.S. Department of Health and Human Services (1997) *Public Use Data File Documentation, National Survey of Family Growth, Cycle 5: 1995 User’s Guide* Hyattsville, MD: NCHS

Acknowledgements:

I thank the National Bureau of Economic Research for data and technical support. Don Cox, Dan Feenberg, Claudia Goldin, Peter Gray, Jennifer Jyu, Ronald Lee and Stephan Pischke provided valuable comments and consultation. An earlier version of this paper was presented to the Human Behavior and Evolution Society annual meeting in London, UK, on June 15, 2001.

**Table 1: Partnership Status and Likelihood of a Male Child: Descriptive statistics**

| | NCPP [a] | NLS-YW [b] | NLSY [b] | NSFG [b] | Pooled [b] | Pooled [b] |
|---|---|---|---|---|---|---|
| *N: parents* | 43,721 | 3206 | 8438 | 2739 | 58,104 | 14,383 |
| *N: live births* | 53,515 | 6696 | 17160 | 3360 | 80,715 | 27,216 |
| *Birth year range* | 1959-1966 | 1970-1997 | 1980-1998 | 1991-1995 | 1959-1998 | 1970-1998 |
| Timing of household observation: | First prenatal visit | 6 months or more before birth | 6 months or more before birth | At 6 months before birth | Mixed | Before conception |
| Sampling frame: | low income patients, longitudinal survey | US cohort, longitudinal survey | US cohort, longitudinal survey | US cohort, retrospective survey | All four samples | Excludes NCPP |
| Partnership status: | | | | | | |
| *Married* | 73.3% | 77.6 % | 65.5 % | 63.9 % | 68.7 % | 68.3% |
| *Cohabiting* | 3.2% | 0.5 % | 8.0 % | 18.9 % | 4.9 % | 6.6% |
| *Living apart* | 24.5% | 21.9 % | 26.5 % | 17.2 % | 26.4 % | 25.2% |
| % Male births: | | | | | | |
| *Married* | 51.5% | 53.4 % | 51.6 % | 52.7 % | 51.4 % | 52.2% |
| *Cohabiting* | 51.0% | 46.4 % | 52.8 % | 51.7 % | 51.8 % | 52.4% |
| *Either m. or c.* | 51.1 % | 53.1 % | 52.0 % | 52.6 % | 51.6 % | 52.3% |
| *Living apart* | 49.7% | 51.0 % | 49.5 % | 46.0 % | 49.6 % | 50.0% |
| *Odds ratio:* | 1.061 | 1.083 | 1.087 | 1.205 | 1.081 [c] | 1.091 [c] |
| $X^2$: *(df=1)* | 8.54** (p=.003) | 2.01 p=.16 | 7.64** (p=.006) | 9.38** (p=.002) | 21.74**** (p<.0001) [c] | 9.71*** (p=.002) [c] |

*p <= .05 **p<=.01 ***p<=.005 **** P<=.0001

a: unweighted b: weighted c: stratified by study

NLS-YW: National Longitudinal Survey of Young Women-1969
NLSY: National Longitudinal Survey of Youth-1979
NSFG: National Survey of Family Growth-1995
NCPP: National Collaborative Perinatal Project

*Mixed*: household status is reported six months or more before birth in NSFG, NLSY and NLS-YW; and any time before birth in NCPP.

*Before conception*: household status is reported nine months or more before birth; the NCPP sample is not included

## Table 2: Partnership status and likelihood of a male child: Ordinary and Within-Family Logistic regressions

| | Model 1<br><br>Ordinary | Model 2<br><br>Ordinary | Model 3<br><br>Within-family | Model 4<br><br>Within-family |
|---|---|---|---|---|
| Timing of household observation [a] | Mixed | Before conception | Mixed | Before conception |
| N births | 80715 | 27216 | 80715 | 27216 |
| N families | n/a | n/a | 58104 | 14383 |
| | | | | |
| Parents married or living together | .062***<br>(.017) | .072*<br>(.030) | .154***<br>(.043) | .127**<br>(.050) |
| Parent's age (years) | .0002<br>(.002) | .006<br>(.005) | .006<br>(.030) | .017<br>(.045) |
| Parent's education (years) | .0001<br>(.002) | -.003<br>(.006) | -.010<br>(.006) | .029<br>(.051) |
| Child's birth order | -.003<br>(.005) | -.003<br>(.010) | -.011<br>(.021) | -.012<br>(.029) |
| Child's birth year | -.003<br>(.002) | -.008<br>(.005) | -.005<br>(.030) | -.014<br>(.045) |
| NSFG | .019<br>(.040) | .045<br>(.048) | --- | --- |
| NLSYW | .019<br>(.043) | -.030<br>(.062) | --- | --- |
| NCPP | -.079<br>(.063) | ---- | --- | --- |
| Parent female | -.0002<br>(.031) | .003<br>(.031) | --- | --- |
| Parent Black | -.035 *<br>(.016) | -.013<br>(.032) | --- | --- |
| Parent Hispanic | -.009<br>(.026) | .017<br>(.038) | --- | --- |

*p <= .05 **p<=.01 ***p<=.001 ****p<=.0001

Ordinary logistic regression models use robust standard errors and sampling weight, and within-family models stratify by responding parent's id. Standard errors are in parentheses.

*Mixed*: household status is reported six months or more before birth in NSFG, NLSY and NLS-YW; and any time before birth in NCPP.

*Before conception*: household status is reported nine months or more before birth; the NCPP sample is not included

Dads and Cads: APPENDIX:

*Data sources:*

This study is based on four publicly available data sets which collected information about the presence or absence of unmarried partners as well as marital partners, and which included multiple siblings within the same family. Each of the four data sets has strengths and weaknesses for the present purpose.

The 1969 National Longitudinal Survey of Young Women (NLS-YW) and the 1979 National Longitudinal Survey of Youth (NLSY) belong to a family of nationally representative cohort studies sponsored by the US Bureau of Labor Statistics; subjects are interviewed annually, permitting observation of the parents' marriage or non-marital cohabitation status at the most recent interview before the child's conception. Both samples were drawn to represent the civilian, non-institutionalized US populations in the designated age group at the time of the surveys. Births reported in these surveys were included in the present study if the sample parent was at least 15 years old at the time of the child's birth, if there was at least one interview between 9 and 37 months before the child's birth, and if there were no missing values for child's date of birth, child sex, and parent's education. The NLS-YW began with a cohort of young women between the ages of 14 and 24 in 1968; non-marital cohabitation was still uncommon at the time this survey began, and the NLS-YW did not collect information about unrelated household members in 1975, 1977, and 1980. The NLSY began with a cohort of young men and women between the ages of 14 and 21 in 1979. The NLSY does identify unmarried opposite-sex partners living in the household in every survey year, and distinguishes "partners" from other unrelated household members. Because of the NLS cohort design, children born early in each cohort had younger mothers, on average, than the average child born in the US, and children born later in each cohort had older mothers than the average child born in the US.

In contrast, the 1995 NSFG is a nationally representative, retrospective survey of US women who were between the ages of 15 and 45 when they were interviewed in 1995, and asks directly about month and year of marriages, separations, and divorces, and start and stop dates of non-marital cohabitations. Births reported in the NSFG were included in the present study if the pregnancy ended in a live birth between 1991 and 1995, and if the child's mother was as least 15 years old at the time of the child's birth. The NSFG also provides an explicit estimate of the date of conception for all pregnancies. This survey has been the basis for official US estimates of the prevalence of cohabitation among non-marital births .

Lastly, the National Collaborative Perinatal Project (NCPP) represents a diverse and primarily low-income women registered for prenatal care between 1959 and 1965 at collaborating sites across the US; it is still the largest prospective study linking perinatal risk factors and developmental outcomes ever conducted in the US. The NCPP is important to the present study because the focal pregnancies predate the legalization of abortion in the US, and predate the widespread legalization of contraception for unmarried couples.

*Measures of parents' partnership status:*

For the purposes of the present study, subjects in all four samples could be classified into three mutually exclusive household types: (1) responding parent married, with spouse present in household; (2) responding parent unmarried, with an opposite sex partner living in household; and (3) other. However, the timing of the observations differed across studies.

In both NLS surveys, household composition "before conception" was coded from the sample parent's reported household composition at the most recent interview falling nine months or more before the child's birth, and household composition "before six months" was coded from the most recent interview falling six months or more before the child's birth. The median interval from preconception interview to birth was 16 months. Because households are usually being formed rather than dissolved around the time of conception, information from annual interviews will underestimate the actual proportion of couples who were married or living together around the time of conception. For example, if a couple started a joint household ten months before a child's birth, but if the most recent interview was twelve months before the child's birth, then the couple's status at the time of the child's conception would be misclassified as living apart. This misclassification would tend to bias effects towards the null.

The NSFG collects information about the exact month and year of marriages, separations, and divorces, the start and stop dates of non-marital cohabitations, and the estimated month and year of conception for each reported pregnancy in the focal period from 1991 to 1995. Among the four surveys, the NSFG yields the most sensitive and specific measure of the parent's household status at the time of the child's conception.

In contrast, the NCPP collects its first household information at the mother's first prenatal visit. The timing of the household report is therefore correlated with the quality of the mother's prenatal care; mothers who sought prenatal care earlier in pregnancy may differ from mothers who delayed their first prenatal visits. To avoid selection bias, I make no effort to restrict observations from the NCPP by timing of the first interview, up to the time of the child's birth.

*Other measures:*

The multivariate regression models in Table 2 of the main text also include measures of the age, education, sex, study sample and black or Hispanic ethnicity of the responding parent, and the child's birth order and year of birth. Age, education, birth order, and black ethnicity were selected because they were available in all four surveys, and because they were among the few demographic factors that have been associated with offspring sex in other human studies (12, 13). The NLS-YW did not record Hispanic ethnicity. The parent's age and highest grade completed were measured at the time of the child's birth, and are expressed in years. It would have been possible to report mother's and father's age and education separately, but information about the non-respondent parent is usually missing when that parent does not live in the household, so that missing information is collinear with the classification of household composition; restriction to female respondents, or inclusion of information about the non-responding parent when this information was available, resulted in no significant change in the results that are reported here.

*Statistical methods and sampling weights:*

The ordinary logistic regression models in this study used sampling weights and robust standard errors, which take account of the possible clustering of variation among siblings. Individual weights were adjusted so that each of the four weighted populations was equal to its original sample size. There were no significant differences in point estimates, whether the models were evaluated with or without the use of weights.

*Results:*

*Sample description:*

Appendix Table 1 describes weighted means and frequencies for the key variables in the four studies. Responding parents were divided about equally between mothers and fathers in the NLSY; all of the respondents in the other three samples were female. 51% of the offspring in the pooled sample were males; after appropriate weighting, the NLSYW, NLSY, and NSFG samples included between 12.4% and 14.8% Blacks, and zero to 15.6% Hispanics; however, the NCPP sample was 47.0% Black and 6.5% Hispanic. Mothers in the NCPP were also slightly younger, with less education and larger families, than parents in the other samples.

*Comparisons among measures:*

Appendix Table 2 compares the effects of using different measures and different observation dates for the parent's household composition. Models 1, 2, and 3 compare marriage and cohabitation status as separate indicators. Because knowledge of child sex might influence the occurrence of "shotgun" marriage between conception and birth, these models focus on observations before the child's conception. Models 1 and 2 indicate that married and unmarried couples who are living together before the child's conception have very similar odds of having a male child, compared to parents who are living apart, thus justifying the use of the combined measure in Model 3. The estimates for marriage and non-marital cohabitation are also similar to each other in the NCPP, when household composition is first observed during pregnancy, and they are similar with and without inclusion of other variables in the model.

Models 4, 5, and 6 compare the effects of using different observation dates using the full, pooled sample. Model 4 takes observations at registration in the NCPP, and before conception in the other three samples; Model 5 takes observations at registration in the NCPP, and observations from 6 months or more before the child's birth in the other samples; and Model 6 takes the most recent observation before the child's birth in all four samples. Median observation-to-birth intervals from the NLS samples in these four models were 16, 15, 12, and 4 months, respectively. Overall, the magnitude of effect does not seem very sensitive to differences in the timing of observation, at least within the margins of error of these measures. However, a cutoff at six months (which yielded a median observation interval of about 12 months before the child's birth) produced a slightly better "hit rate" in predicting child sex compared to other cutoffs in the pooled sample.

*Comparisons among samples:*

Appendix Table 3 compares the results of logistic regression models examined separately in the four samples, using observations six months or more before the child's birth (for the NLS and NSFG samples), or at the first prenatal visit (for the NCPP). Of the various cutoffs that could be used, this choice produces the greatest differences in estimates among studies; the coefficients correspond to odds ratios ranging from 1.021 to 1.285. Never the less, the estimates are not significantly different from each other, and formal tests of homogeneity of the effects, including the use of a full set of interaction terms between study sample and partnership status, yield no significant differences between studies. The effect sizes do correspond to the probable sensitivity of the measures in each study. For example, the smallest estimate comes from the NLS-YW; in this study, interviews were not conducted every year, and information about non-marital cohabitation was not collected at all interviews. Likewise, the largest estimate comes from the NSFG; the NSFG provides information about the start and stop dates of both marriages and non-marital cohabitation, allowing the most exact estimate of partnership status at conception and birth.

1. U.S. Bureau of Labor Statistics (2001) *NLS Handbook, 2000* US Government Printing Office: 2001-472-553/44953

2. U.S. Department of Health and Human Services *Public Use Data File Documentation, National Survey of Family Growth, Cycle 5:1995 User's Guide* Hyattsville, MD: NCHS (1997)

3. National Archives of the United States, Collaborative Perinatal Project Documentation (1985). Available on the Internet at www.nber.org/docs/cpp/docs/cpp-vol1.pdf

4. Goldin C. and Katz L. (2000) “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage Decisions” NBER Working Paper No. W7527, National Bureau of Economic Research, Cambridge, MA

5. Blettner, M., Sauerbrei, W., Schlehofer, B., Scheuchenpflug, T., & Freidenreich, C. (1999) “Traditional reviews, meta-analyses and pooled analyses in epidemiology” *Int J Epi* **28: 1-9**

## Appendix Table 1: Sample description

| | NCPP [a] | NLS-YW [a] | NLSY [a] | NSFG [a] | Pooled |
|---|---|---|---|---|---|
| N: parents | 43,721 | 3206 | 8438 | 2739 | 58,104 |
| N: live births | 53,515 | 6696 | 17160 | 3360 | 80,731 |
| Birth year range | 1959-1966 | 1970-1997 | 1980-1998 | 1991-1995 | 1959-1998 |
| *Frequencies:* | | | | | |
| Parent= female | 100.0% | 100.0 % | 52.0% | 100.0 % | 90.0 % |
| Child= male | 50.8% | 52.5 % | 51.0 % | 51.2 % | 51.0 % |
| Black | 47.0% | 12.4 % | 14.8 % | 14.3 % | 40.4 % |
| Hispanic | 6.5% | N/a | 7.9 % | 15.6 % | 9.4 % |
| *Means:* | | | | | |
| Birth year | 1962.6 | 1976.8 | 1988.3 | 1992.6 | 1970.0 |
| Parent's age (yrs) | 24.2 | 27.3 | 27.4 | 26.0 | 24.9 |
| Parent's ed'n (yrs) | 10.8 | 12.8 | 13.2 | 12.6 | 11.4 |
| Birth order | 2.9 | 2.5 | 2.1 | 2.6 | 2.7 |
| Earliest observation used in present study: | First prenatal visit | 9 months or more before birth | 9 months or more before birth | At conception | Mixed |
| *Median interval from observation to birth: (in months)* | | | | | |
| Last observation before conception | N/a | 17.0 months | 15.0 months | 9.1 months | 15.0 months |
| Last observation six months or more before birth | N/a | 14.0 | 12.0 | 6.0 | 12.0 |
| Last observation before birth | 4.0 months | 8.0 | 6.0 | 0.0 | 4.0 |

[a:] NCPP: National Collaborative Perinatal Project
NLS-YW: National Longitudinal Survey of Young Women-1969
NLSY: National Longitudinal Survey of Youth-1979
NSFG: National Survey of Family Growth-1995

N/a = not available

## Appendix Table 2: Partnership Status and Likelihood of a Male Child: Comparisons among measures [a]

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Timing of observation (NLS [b] and NSFG) | 9 months before birth | 9 months before birth | 9 months before birth | 9 months before birth | 6 months before birth | Birth |
| Sample | NLS [b], NSFG | NLS [b], NSFG | NLS [b], NSFG | Whole | Whole | Whole |
| N births | 27216 | 27216 | 27216 | 80715 | 80715 | 80715 |
| | Coefficients (s.e.) | Coefficients (s.e.) | Coefficients (s.e.) | Coefficients (s.e.) | Coefficients (s.e.) | Coefficients (s.e.) |
| Parents married, living together | .078** (.027) | .064* (.031) | | | | |
| Parents unmarried, living together | .100* (.050) | .108* (.052) | | | | |
| Parents married or living together | | ---- | .072* (.030) | .057*** (.017) | .062*** (.017) | 0.058*** (.018) |
| Parent Black | | -.016 (.032) | -.013 (.032) | -.037 * (.016) | -.035 * (.016) | -.036 * (.016) |
| Parent Hispanic | | .017 (.037) | .017 (.037) | -.008 (.026) | -.009 (.026) | -.009 (.026) |
| Parent's age (years) | | .007 (.005) | .006 (.004) | .0002 (.002) | .0002 (.002) | .0003 (.002) |
| Parent's education (years) | | -.003 (.006) | -.003 (.006) | .0001 (.002) | .0001 (.002) | .0001 (.002) |
| Child's birth order | | -.003 (.010) | -.003 (.010) | -.003 (.005) | -.003 (.005) | -.0024 (.005) |
| Child's birth year | | -.008 (.005) | -.008 (.005) | -.003 (.002) | -.003 (.002) | -.003 (.002) |
| NSFG | | .045 (.048) | .045 (.048) | .021 (.040) | .019 (.040) | .019 (.043) |
| NLSYW | | -.029 (.062) | -.030 (.062) | .023 (.044) | .019 (.043) | .027 (.044) |
| NCPP | | ---- | ----- | -.082 (.064) | -.079 (.063) | -.069 (.064) |
| Parent female | | .004 (.032) | .003 (.032) | -.0006 (.031) | -.0002 (.031) | -.0001 (.031) |

*p <= .05 **p<=.01 ***p<=.001 ****p<=.0001

[a] Ordinary logistic regression models use pooled sample, robust standard errors and sampling weight, and within-family models use conditional logistic regressions stratifying by responding parent's id. Standard errors are in parentheses.

[b] "NLS" = NLS-YW and NLSY; see list of study samples in footnote to Appendix Table 1

## Appendix Table 3: Partnership status and likelihood of a male child: Comparisons among samples [a]

| | Pooled sample | NCPP | NLS-YW | NLSY | NSFG [d] |
|---|---|---|---|---|---|
| Timing of household observation [b] | Mixed [c] | First prenatal visit | 6 + months before | 6 + months before | 6 months before |
| N births | 80715 | 53499 | 6696 | 17160 | 3360 |
| Married OR living together | .062***<br>(.017) | .046*<br>(.022) | .021<br>(.067) | .088*<br>(.037) | .251**<br>(.097) |
| Parent = Black | -.035 *<br>(.016) | -.048*<br>(.019) | -.067<br>(.067) | .004<br>(.040) | .119<br>(.093) |
| Parent = Hispanic | -.009<br>(.026) | -.034<br>(.037) | ---- | .022<br>(.042) | .003<br>(.091) |
| Parent's age (years) | .0002<br>(.002) | -.0007<br>(.002) | .007<br>(.011) | .010<br>(.007) | .001<br>(.008) |
| Parent's education (years) | .0001<br>(.002) | .0003<br>(.002) | .017<br>(.014) | -.010<br>(.007) | -.008<br>(.014) |
| Child's birth order | -.003<br>(.005) | -.002<br>(.006) | .011<br>(.017) | -.001<br>(.014) | -.025<br>(.025) |
| Child's birth year | -.003<br>(.002) | -.005<br>(.005) | -.009<br>(.010) | -.011<br>(.007) | -.034<br>(.028) |
| Parent= female | -.0002<br>(.031) | ---- | ---- | .005<br>(.032) | ---- |

*p <= .05 **p<=.01 ***p<=.001 ****p<=.0001

[a] Ordinary logistic regression models use robust standard errors and sampling weight. Standard errors are in parentheses.

[b] "Mixed ": household status is reported at six months before birth in the NSFG; six months or more before birth in NLSY and NLS-YW; and any time before birth in NCPP.

[c] Model 1 also includes dummy variables indicating study; all are non-significant. Model 1 is identical to Model 5 in Appendix Table 2

[d:] NCPP: National Collaborative Perinatal Project
NLS-YW: National Longitudinal Survey of Young Women-1969
NLSY: National Longitudinal Survey of Youth-1979
NSFG: National Survey of Family Growth-1995