

Human Sex Ratios and Sex Distribution in Sibships of Size 2

WILLIAM D. STANSFIELD¹ AND MATTHEW A. CARLTON²

Abstract We previously analyzed data from the U.S. National Health Interview Survey (NHIS, 1998 to 2002) on families with two biological children (10 years of age and younger) and found that the distribution of families with two boys, two girls, and one boy + one girl did not statistically conform to a binomial distribution regardless of the boy/girl sex ratio used. Using the best estimate of the sex ratio from the data, we found that there were significantly more families with opposite-sex siblings than families with same-sex siblings. No biological mechanism could explain these results at the time. In the present study we conducted an analysis of the first two children in sibships of size 3 from the same data source and found that there are significantly more same-sex sibships than unlike-sex sibships. Combining the two sets of data for the first two children produced observed numbers in close agreement with the expected numbers. A hypothesis of parental choice (family planning) appears to be strongly supported as an explanation for the discrepancies in the two sets of data individually. For example, parents who have a boy and a girl (either order) as their first two children are more likely to stop having children (“stopping rule”) than are parents whose first two children are of the same sex.

In our earlier paper (Carlton and Stansfield 2005), we analyzed data in sibships (also called families) of size 2 from the 1998–2002 U.S. National Health Interview Survey (NHIS). Using all 50,936 children, 10 years of age and younger, in the NHIS data set (the youngest age cohort available), we calculated that 51.38% were boys (B) and 48.62% were girls (G); the B/G ratio was 1.0566, or about 106 boys for every 100 girls. A standard large-sample z test of the hypothesis $\Pr(B) = \Pr(G) = 0.5$ gave a test statistic of $z = 5.50$ ($p < 0.0001$). A chi-square goodness-of-fit test (Table 1) was highly significant, leading us to conclude that the number of boys in two-child families did not follow any binomial distribution, even one with a sex imbalance. The binomial model, using the sex ratio in this sample, predicted that considerably more same-sex sibships would appear than were presented in the NHIS data; GG sibships contributed more to the chi-square statistic than BB sibships did. A test for independence of siblings’ sexes led us to conclude that the sexes of second children were not independent of the sexes of their older siblings.

Table 1. Observed Counts, Expected Counts, and Chi-Square Contributions for a Binomial Goodness-of-Fit Test Using the NHIS Data Set for the First Two Children in Sibships of Size 2

	<i>Number of Boys</i>		
	<i>0</i>	<i>1</i>	<i>2</i>
Observed	5,844	13,079	6,545
Expected	6,021.32	12,724.35	6,722.32
χ^2	5.222	9.885	4.677

After having read our paper in *The American Statistician*, Philip Hougaard (Biostatistics Department, H. Lundbeck A/S, Valby, Denmark) contacted us by e-mail. He was surprised that there were not more same-sex sibships than unlike-sex sibships. We intimated in our paper that if parents were more likely to stop having children after they had one child of each sex than after they had two boys or two girls, then there would be proportionately fewer same-sex sibships in two-child families than in the first two births of three-child families. But at that time we had no data on how often this type of family planning limited family size or how strongly this nonbiological effect, if present, influenced our analysis. Hougaard suggested that we analyze the first two children in all families with at least two siblings. Our purposes in writing this paper are (1) to report the results of analyzing the first two children in sibships of size 3, (2) to compare the results with those reported in our previous paper, and (3) to discuss the results of following Hougaard’s suggestion.

Materials and Methods

We returned to the 1998–2002 NHIS and extracted data for all biological children, 10 years of age and younger (the youngest age cohort available), in three-child families. Using only the first two children in this data set, we calculated the B/G ratio. This ratio was then used to calculate the expected numbers for a binomial distribution of three types of sibships using only the first two children therein: BB, (BG + GB), and GG. A chi-square statistic was calculated as usual. The probability that the deviation of the expected number from the observed number is due to chance was determined.

Results

The relative frequency of boys represented by the first two children in 7,541 sibships of size 3 is 0.5176. Using this value, we found that a chi-square test of binomial distribution (Table 2) produced a much larger, highly significant statistic ($\chi^2 = 73.90, p \ll 0.0001$) than that produced in our study of sibships of exactly size 2. These data do not conform to any binomial distribution with a single B/G ratio of any kind, as we reported for sibships of only 2 in our previous paper. However,

Table 2. Observed Counts, Expected Counts, and Chi-Square Contributions for a Binomial Goodness-of-Fit Test Using the NHIS Data Set for the First Two Children in Sibships of Size 3

	<i>Number of Boys</i>		
	<i>0</i>	<i>1</i>	<i>2</i>
Observed	1,941	3,393	2,207
Expected	1,754.60	3,765.81	2,020.60
χ^2	19.80	36.90	17.20

now more same-sex sibships are reported than expected (in keeping with Hougaard’s understanding of the literature), with GG sibships contributing slightly more to the chi-square statistic than BB sibships.

Discussion

The fact that the two sets of data give essentially opposite results coincides with the theory of parental choice and/or birth control. If parents with one boy and one girl are more likely to stop having children than those with two children of the same sex, we should see a higher number of BG and GB two-child families than independence would predict; indeed, this is what we found earlier (Carlton and Stansfield 2005). Concordantly, if parents with two boys or two girls are more likely to have another child, then we should see higher than expected counts of BB and GG pairs among the first two children in three-child families; again, the data (Table 2) bear this out.

According to Anjani Chandra (personal communication, 2004), the NHIS is neither a census nor a mandatory data collection effort. Rather, it is a voluntary household-based survey yielding nationally representative statistics on the topics covered. The NHIS does not obtain information on children who do not reside in the home when the interview takes place, and this could conceivably affect the proportions of sexes. The NHIS also fails to note the incidence of identical twins, half-siblings, stillbirths, and other biological phenomena. Finally, the NHIS data are not stratified by potentially confounding variables, such as health status, paternal age, or socioeconomic status (Lazarus 2002). Despite these statistical concerns, we believe that it is reasonable to treat these 33,009 families as if they are a simple random sample of all two- and three-child U.S. families. In particular, because relatively few young families (i.e., parents with children 10 years of age or younger) have more than three children, these 33,009 families might reasonably represent a cross-section of all U.S. families currently with young children. We therefore combined these two sets of families from the NHIS data. The results of this approach are summarized in Table 3.

The expected counts are based on the binomial model and the maximum-likelihood estimate from the first two children in these families: relative frequency

Table 3. Observed Counts, Expected Counts, and Chi-Square Contributions for a Binomial Goodness-of-Fit Test Using the NHIS Data Sets for Two-Child and Three-Child Families^a

	<i>Number of Boys</i>		
	<i>0</i>	<i>1</i>	<i>2</i>
Two-child families	5,844	13,079	6,545
Three-child families	1,941	3,393	2,207
Total observed	7,785	16,472	8,752
χ^2	0.0108	0.0204	0.0096

a. Only the first two children in a three-child sibship pertain to the data.

of boys = $[2(8,752) + 16,472]/2(33,009) = 0.5146$. The observed counts are almost identical to the expected counts; in particular, the chi-square statistic is a very small 0.0408, with a correspondingly very high probability (> 99%) that the difference between the observed and expected values could occur by chance alone. That is, when we do not subdivide the data into two-child and three-child families, the combined data for the first two children conform closely to a binomial distribution. In particular, the apparent validity of this binomial model implies that (1) sexes of successive children are independent and (2) a single sex ratio governs the probability of a male child, regardless of whether this is the parents' first or second child. It seems that the lack-of-independence evidence, mentioned in our previous paper, was just a byproduct of not including three-child families. The parental choice explanation seems plausible in light of this analysis, whereas the data do not support any theory suggesting that the sex of child 1 can biologically affect the likely sex of child 2.

Although it was not known to us when we first analyzed two-child families in our 2005 paper, the family stopping rule and related phenomena have been known and reported in the literature for at least four decades, as evidenced by some of the following examples. If parents with two children are more likely to stop having children after having one boy and one girl, such stopping rules can reduce the average sibship size in the population but should not affect the overall sex ratio of the population. Families in which the first two children are GB (girl oldest) are least likely to increase in size, whereas families in which the first two children are of the same sex are most likely to grow. This pattern seems to continue as the sibship size increases (Edwards 1966). In addition,

As long as the sex of each birth is independently determined, no degree of birth control, even if directly related to the desire for male or female offspring, can change the overall population sex ratio. Family planning in relation to sex can, however, affect the distribution of the sequences of births, and thus the apparent correlation between successive births. (Cavalli-Sforza and Bodmer 1971, p. 661)

Using data from women interviewed in the 1965 and 1976 U.S. National Fertility Studies (NFS) and the 1976 U.S. National Survey of Family Growth (NSFG), Sloan and Lee (1983) found that the sex of women's previous children has an effect on their subsequent fertility (family size) intentions.

An extensive review of the literature on the human sex ratio up to 1987 was conducted by W. H. James (1987a). However, in this review James was mainly concerned with variables and biological mechanisms that might influence the sex ratio rather than the family structure of two- and three-child sibships. In his next paper (James 1987b), James found that the Caucasian secondary sex ratio (at birth, without malformations) in the population he studied averaged 0.514 with a standard deviation of 0.05. (We note that the frequency of males reported by James to three decimal places is equal to the value in our data; see Table 3.)

Maternal age seems to be of little (if any) effect, whereas it is well established that the sex ratio declines with increasing paternal age and is correlated with declining male androgen levels. Yamaguchi and Ferguson (1995) reported that the "sex composition effect of children born [in two-child and three-child families] is larger for highly educated women than for those with lower education attainment and for women in younger cohorts than for those in older cohorts" (p. 272).

Grant (1998) cited the stopping effect twice in his book. Some parents may prefer boys, others may prefer girls, but "it is hard to find evidence to support either pattern" (p. 183). However, on the next page Grant says that demographers have found good evidence for son preference in Western cultures, and in the last 10 to 15 years, evidence exists of a stronger than ever trend toward son preference in Asian cultures.

Lazarus (2002) does not mention "stopping effect" in his book but does say that many societies exhibit a negative relationship between wealth (status) and fertility (family size) that may be due to "birth control." Pollard and Morgan (2002) examined the strength of the stopping rule on both the fertility behavior and intentions of parents over several decades using (in part) data from the NSFG and concluded that changes in the societal sex system are expected to weaken this pronatalist (family planning) effect in recent periods. "Consistent with this expectation, there has been some attenuation of the effect of sex composition of previous children on the third birth, suggesting declining salience of children's gender for parents" (Pollard and Morgan 2002, p. 600).

None of these or other references we have consulted provide data and chi-square analyses of a binomial distribution of the combined sexes of the first two children in sibships of size 2 and 3 that support the hypothesis of family planning, birth control, or a stopping effect as we have reported here.

Literature Cited

- Carlton, M. A., and W. D. Stansfield. 2005. Making babies by the flip of a coin. *Am. Stat.* 59(2):180–182.
- Cavalli-Sforza, L. L., and W. F. Bodmer. 1971. *The Genetics of Human Populations*. San Francisco: W. H. Freeman.
- Edwards, A. W. F. 1966. Sex-ratio data analyzed independently of family limitation. *Ann. Hum. Genet.* 29:337–347.
- Grant, V. J. 1998. *Maternal Personality, Evolutionality, and the Sex Ratio: Do Mothers Control the Sex of the Infant?* London: Routledge.
- James, W. H. 1987a. The human sex ratio. 1. A review of the literature. *Hum. Biol.* 59(5):721–752.
- James, W. H. 1987b. The human sex ratio. 2. A hypothesis and a program of research. *Hum. Biol.* 59(6):873–900.
- Lazarus, J. 2002. Human sex ratios: Adaptations and mechanisms, problems and prospects. In *Sex Ratios: Concepts and Research Methods*, I. C. W. Hardy, ed. Cambridge, England: Cambridge University Press, 287–311.
- Pollard, M. S., and S. P. Morgan. 2002. Emerging parental gender indifference? Sex composition of children and the third birth. *Am. Sociol. Rev.* 76(4):600–613.
- Sloan, D. M., and C.-F. Lee. 1983. Sex of previous children and intentions for further births in the United States, 1965–1976. *Demogr.* 20(3):353–367.
- Yamaguchi, K., and L. R. Ferguson. 1995. The stopping and spacing of childbirths and their birth-history predictors: Rational-choice theory and event-history analysis. *Am. Sociol. Rev.* 60(2):272–298.