

WILLIAM D. STANSFIELD¹ AND MATTHEW A. CARLTON²

Abstract In two-child families containing at least one boy, the expected probability that such a family has two boys is $1/3$, provided that the boy/girl (B/G) ratio is 1.0 and the population to which they belong has a binomial distribution of BB, (BG + GB), and GG families. It is commonly known that in most human populations the sex ratio at birth (i.e., the ratio of the number of boys to the number of girls) is greater than 1.0. Teachers and textbook writers seldom discuss the more realistic expected distributions in populations where the sex ratio is greater than 1.0. We present data from two federal surveys with sex ratios greater than 1.0 and find that the observed proportions of two boys in families of size 2 with at least one boy range from 0.3335 to 0.3941. It has been reported in the literature that the probability (p) of a male birth is subject to both within-sibship variation (Poisson variation), for which our data are suggestive, and possibly also between-sibship variation (Lexis variation). These deviations (biases) from the assumptions of a simple binomial distribution are involved in the calculation of values of p and standard 95% confidence intervals, thereby foiling attempts to make reliable statistical inferences from the data. Analysis of the data is also complicated by family planning that falsifies the assumption of randomness in the binomial gender distribution model. Families of size 2 (and their sex composition) are often discussed in a wider context. Overpopulation in some parts of the world has caused mass starvation and threatens to do the same worldwide unless the birth rate drops to agriculturally sustainable levels. Even if every woman of fertile age has only two children on average from now on, the world's population is predicted to continue growing toward 9 billion people by 2050. Other sociological problems are bound to follow. Although the birth rate in China has recently dropped, the average age of the population has risen, so that by 2035 it is projected that for each person over age 65 there will be just three working-age people. Furthermore, China's one-child policy has already led to a sex imbalance where there is a large excess of men for whom marriage and parentage is denied.

¹Biological Sciences Department, California Polytechnic State University, San Luis Obispo, CA 93407.

²Statistics Department, California Polytechnic State University, San Luis Obispo, CA 93407.

Human Biology, February 2009, v. 81, no. 1, pp. 3–11.

Copyright © 2009 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: BEHAVIORAL GENETICS, BIRTH CONTROL, FAMILY PLANNING, FAMILY SIZE, FAMILY STRUCTURE, PROBABILITY THEORY, SEX RATIO, STOPPING RULE, TWO-BOY PROBLEM, NATIONAL HEALTH INTERVIEW SURVEY (NHIS), MARILYN VOS SAVANT, WILL LASSEK.

Marilyn vos Savant (listed in the Guinness Book of World Records Hall of Fame for highest IQ) is a columnist for *Parade* magazine. *Parade* is distributed by more than 400 Sunday newspapers and has a circulation of 32 million and a readership of 71 million. This is why we consider her column to be the major source of disseminating what we call “the most widely publicized gender problem in human genetics.” In this paper we aim to explain what this problem is exactly. We present and analyze data on sex distributions in two-child families from two federal surveys to provide an answer to this problem. We also reveal some of the biases that might be present in these surveys and analyses, discuss how unequal male/female ratios in a population can evolve rapidly and create various societal problems, and evaluate plans to limit population growth by restricting family size to no more than two children.

Materials and Methods

In her column of October 19, 1997, vos Savant posed the following problem. “A woman and a man (unrelated) each have two children. At least one of the woman’s children is a boy, and the man’s older child is a boy. Do the chances that the woman has two boys equal the chances that the man has two boys?” (Vos Savant 1997). The history of this two-boy problem and vos Savant’s solution to it can be found on the Internet at the Wikipedia URL given in the Literature Cited. Vos Savant avers that “the chances the woman has two boys are about 1 in 3 and that the chances the man has two boys are about 1 in 2.” These are the well-known, correct solutions to this math question, provided that we accept two underlying assumptions: (1) the sexes of successive children are independent and (2) the probability of a male birth is 0.5 on every birth, or, equivalently, that the sex ratio is a constant 1.0 from birth to birth throughout the population.

In other words, the number of boys in two births follows a binomial distribution with success probability 0.5. Edwards (1960) discusses some types of variation on this traditional mathematical model. Some of vos Savant’s readers doubted that her 1/3 solution was correct. So she solicited data from her women readers “with two children (no more), at least one of which is a boy (either child or both of them).” She obtained 17,946 responses by letters and e-mails. Without reporting the sex ratio in the sample, she states that about 35.9% of respondents (“about 1 in 3”) said they have two boys. However, she then says, “Given that there are about 106 boys born for every 100 girls, the actual percentage in the population would be closer to 33.9%.” This is actually an incorrect adjustment, as noted by Carlton and Stansfield (2005). Vos Savant seems to have simply divided 35.9% by the boy/girl (B/G) ratio 106/100 to obtain 33.9% to one decimal place. The correct formula for the adjusted probability of two boys, given at least one boy and still assuming a binomial distribution, is

$$P(\text{BB} \mid \geq 1\text{B}) = \frac{r}{r+2}, \quad (1)$$

Table 1. NHIS Data on 42,888 Families of Exactly Two Children for 1987–1993

<i>Sex Sequence</i>	<i>Observed Number of Families</i>	<i>Observed Percentage of Sample</i>	<i>Expected Percentage of Sample if $r = 1.05$</i>
Girl-girl (GG)	9,523	22.204	23.795
Boy-girl (BG)	11,118	25.923	24.985
Girl-boy (GB)	10,913	25.445	24.985
Boy-boy (BB)	11,334	26.427	26.235

r = population human sex ratio (boys to girls).

where r is the assumed sex ratio. Applying this formula and assuming $r = 1.06$ (as suggested by vos Savant) produces the adjusted probability of 34.64%.

In that same column in *Parade*, vos Savant was asked if she could get Census Bureau data for comparison. She cited an answer from Will Lassek (then at the Center for Health Data and currently at the University of Pittsburgh): “The U.S. Census Bureau interviews a random sample of families each year for the National Health Interview Survey” (NHIS). These data on 42,888 families with exactly two children were for the years 1987 to 1993 (Table 1). Lassek noted in his response to vos Savant that the number of families with two girls was smaller than expected.

We decided to see how well the NHIS data conform to a binomial model using the best estimate of the sex ratio from the data, approximately 109/100. We analyzed Lassek’s data using a chi-square test, and, as anticipated, we found it to be highly significantly different from any binomial distribution (neither independence nor constant probability was plausible), but we did not write a paper on this. Instead, we asked the NHIS if they could verify Lassek’s data or send us more recent data. We received only data from a more recent NHIS for the period 1998–2002. We used the data set containing two-child families with the youngest cohort of biological children (10 years of age or younger). We analyzed this new data and found that it also was highly deviant from a binomial distribution, and even though we did not come to any conclusion about why the test was significant at the time, we published our findings in *The American Statistician* (Carlton and Stansfield 2005). Later, we analyzed additional data from the 1998–2002 NHIS and were able to provide evidence to support a hypothesis of parental choice (stopping rule) as one of the main reasons that the data did not conform to a binomial distribution. These findings were published in *Human Biology* (Stansfield and Carlton 2007).

Results

None of Lassek’s correspondence with vos Savant nor any of his data from the NHIS 1987–1993 data set appeared in either of our two previous papers. Lassek is quoted by vos Savant as having said, “Limiting the sample to families with two children (as you did in your survey) introduces some biases; note, for example, the smaller-than-expected percentage (22%) of two-children families with two girls. Including the first two children of any family with at least two children would

Table 2. Five Estimates of Two Boys in Two-Child Families with at Least One Boy (Pertinent Families) from Three Sources

<i>Data Source</i>	<i>All Families</i>	<i>All Families with Two Boys</i>	<i>Pertinent Families</i>	<i>Raw Probability</i>	<i>Adjusted Probability</i> ^a
vos Savant, if B/G = 106/100	No data	~6,443	17,946	0.3590	0.3464
NHIS, 1987–1993; B/G = 108.82/100	42,888	11,334	33,365	0.3397	0.3524
NHIS, 1998–2002; two-child, B/G = 105.66/100	25,468	6,545	19,624	0.3335	0.3457
NHIS, 1998–2002; first two in three-child family, B/G = 107.30/100	7,541	2,207	5,600	0.3941	0.3492
NHIS, 1998–2002; first two in (2 + 3) child families, B/G = 106.01/100	33,009	8,752	25,224	0.3470	0.3465

a. Adjustments made using estimated sex ratio from the sample.

probably give figures closer to those predicted.” He did not state what those predicted figures would be, but we will see the consequences of that bias in what follows.

Lassek appears to have calculated an estimate of about 34.0% BB families from the 33,365 families of size 2 in his sample with at least one boy (Table 2). According to vos Savant, Lassek wrote, “Close enough, I’d say.” He did not state the expected percentage to which his data was “close enough”; if we assume that he meant close enough to 1/3 as opposed to 1/2, then we would agree. Because Lassek did not report the sex ratio in the entire NHIS 1987–1993 survey of 342,018 families, the sex ratio had to be derived from the 42,888 of those families containing only two children. Lassek (personal communication, 2008) correctly questioned the validity of using our estimate of the sex ratio of 1.09 for calculating the adjusted percentage of two boys in families of size 2 with at least one boy in his data. According to him, “If you are basing that on the sample of two-child families, you are calculating an expected value from a sample which is already biased by the same thing you are investigating. Because of the stopping rule, the two-child family sample has an excess of boys. It seems to me that you should use the sex ratio at birth, which is probably close to 1.05.”

If the assumed population sex ratio is about 1.05 for all five data sets in Table 2 (Matthews and Hamilton 2005), then the adjusted probability (and answer to vos Savant’s problem)—still assuming an underlying binomial model—becomes 0.3443:

$$P(\text{BB} \mid \geq 1\text{B}) = \frac{1.05}{3.05} = 0.3443. \quad (2)$$

Note that the probability formula $r/(r + 2)$ is monotone in r , which in turn increases with the proportion of boys in the data used to estimate r . Thus any sample

with a disproportionate number of boys (such as the NHIS 1987–1993 data) will result in an overestimate of r and of the probability $P(\text{BB} \mid \geq 1 \text{ B})$.

The more recent 1998–2002 data from the NHIS for families of size 2 with at least one boy gave an adjusted probability of 34.57% (Table 2). The largest sample size in the three data sets from the 1998–2002 NHIS survey was the one that combined families of size 2 plus the first two children of families of size 3 with at least one boy. This group had the lowest absolute deviation between raw and adjusted estimates (0.0014, or 0.14%) of any of the four NHIS data sets. This agrees with Lassek’s prediction: “Including the first two children of any family with at least two children would probably give figures closer to those predicted.” In our previous paper (Stansfield and Carlton 2007), we reasoned that combining families of size 2 with the first two children in families of size 3 (NHIS 1998–2002 data set) would produce a data set least likely to be biased by family choice and the stopping rule. This assumption is supported in that paper because that group conforms closely to the expected numbers of BB, (BG + GB), and GG families in a binomial distribution with $r = 1.06$, and it is deemed more likely to resemble the population of the first two children in all families with at least two children than any of the other data sets in Table 2.

We note that the B/G ratio in the 1987–1993 NHIS data is the highest of any of the data sets in Table 2, even though those data contained the largest sample size. However, the National Center for Health Statistics reports that the B/G ratios at birth from 1940 to 2002 ranged from a high of 105.9/100 in 1946 to a low of 104.6/100 in 1991 (<http://www.cdc.gov/nchs/pressroom/05facts/moreboys.htm>). Because the B/G ratio of Lassek’s data (1.0882) is so highly biased, it seems more logical to use the B/G ratio of approximately 105/100, as suggested by Lassek, to calculate the expected value (of two boys in two-child families with at least one boy) for his 1987–1993 data from the NHIS. Lassek continues, “The sex ratio at birth is determined by biology, not the future makeup of the population, so it is not really valid to base it on any [sample] other than newborns. A much larger excess of boys are actually conceived, but they are more likely to abort or be stillborn, so the ratio of 1.05 is actually much smaller than the conception difference.” Lassek’s points are well taken. Assuming a binomial distribution and a B/G ratio = 1.05, the expected percentages are reported in Table 1.

It should be understood that the deviation between raw and adjusted probability estimates in Table 2 is not a metric for any hypothesis test. At best, these deviations show how a sample’s estimate for the fraction of BB sibships in two-child families with at least one boy changes as the sex ratio deviates from 1.0. If the raw vs. adjusted difference is large, this indicates that (1) the sex ratio in the sample is quite different from 1.0 and/or (2) the sample is relatively small.

Discussion

Sex ratios can be influenced by numerous factors (Edwards 1966; Jacobsen et al. 1999; James 1976, 1985, 1987a, 1987b, 1995, 1996, 2001, 2008; McWhirter

1956; Pickles et al. 1982). Edmond Malinvaud (1955) studied almost 4 million births in France for 1946–1950 and concluded that the probability estimate p (measured as a percentage) of a pregnancy producing a boy is fairly well fitted by the linear relationship $p = 51.45 + 0.3n - 0.5m$, where n is the number of preexisting boys and m is the number of preexisting girls in the sibship. In families in this population with the first child being a boy, the probability of the second child being a boy is 51.75%; with the first child being a girl, the probability of the second child being a boy is 50.95%. The implication is that successive sexes are not independent, although this contradicts the findings of our earlier paper (Stansfield and Carlton 2007). According to James (1976), the probability of a male birth (p) declines with birth order within individual sibships (Poisson variation). This trend might be present in Table 1, where the sex ratio for all families is 1.0882, whereas the sex ratio for the firstborn is 1.09886. In the presence of such within-sibship variation, data would not be expected to conform to a binomial distribution, even in the absence of parental choice. James (personal communication, 2008) notes that the probability of a male birth is also affected by between-sibship variation (Lexis variation). “The evidence for both is overwhelming in regard to mammalian, including human, sex ratio at birth.” Because a mixture of Lexis and Poisson variation may mimic a binomial distribution (the two effects can mask each other; see James 2000), it is invalid to infer from a seemingly binomial distribution that the probability of a male birth is equal at all trials. In particular, the assumption that all couples have the same probability of male births (homogeneity) is invalid.

Our original purpose in acquiring the NHIS data was not to demonstrate the effect of family planning but to compare the numbers of BB, (BG + GB), and GG sibships in families of size 2 with the numbers expected under the assumption of a binomial distribution using the sex ratio from the NHIS data. The difference between observed and expected numbers was highly significant and suggested to us that the data are highly biased and that family planning and the stopping rule were highly likely factors. However, to properly document a stopping rule, we would require a different approach with proper conditional probabilities (progression ratios conditional on the number of boys and girls already born), perhaps akin to the equation developed by Malinvaud (1955).

It has been suggested to us (by one of the anonymous reviewers of this paper) that because the average sex ratio in the U.S. population during 1987–2002 was approximately 1.05, the expected proportion of BB families among all families of size 2 with at least one boy is 34.4%. This value is lower than the observed estimate vos Savant reported (35.9%) but not much less than the adjusted expected estimate of 34.6% (assuming $r = 106/100$ for her data). We must remember that vos Savant’s data are not a proper random sample of the population. Her data were obtained from volunteers by letters and e-mails, whereas the NHIS data were collected through initiated contact by trained NHIS personnel who properly randomly sampled the population to obtain data truly representative of the U.S. population. One of our main purposes in the present report is to show that the expected probability of producing two boys in families of two children with at least one boy is

not exactly 33% but somewhat more because the sex ratios in the data are greater than 1.0. In any event, the answer to vos Savant's question is much closer to 1/3 than to 1/2.

It should be pointed out that the NHIS does not obtain information on children who do not reside in the home when the interview takes place. Thus older children who have moved out or enrolled in boarding school are not included, nor are any children who died before the interview. This might bias the sex ratio if girls are more likely than boys to live outside the family. Because the youngest age cohort available from the NHIS reports families with children 10 years of age or younger, we consider that this "migratory effect," if present, would most likely be smaller than in cohorts of families with children older than 10 years.

More recently on the Internet, Snell and Peterson (2007) raised the following questions: "Does the 95% confidence interval for the probability obtained from Marilyn's survey include $p = 1/3$? Answer the same for Will Lassek's data." Our response follows. If the survey sample is drawn at random from the target population, a standard 95% confidence interval (CI) for a proportion is constructed by the formula

$$CI = p \pm 1.96 \left[\frac{p(1-p)}{n} \right]^{1/2}, \quad (3)$$

where p denotes the sample proportion (e.g., of BB families among families of size 2 with at least one boy) and n denotes the sample size. This formula is valid only if the sample taken can be regarded as a simple random sample from the population of interest; vos Savant's sample is not. In fact, none of the NHIS data cited here represent a simple random sample from any well-defined population. Therefore no statistically valid confidence intervals can be calculated for any of these data sets. In Table 2 all the adjusted estimates are greater than 1/3, or 0.3333 (range = 0.3457–0.3524), as expected for B/G ratios greater than 1.0000. The observed value (0.3335) in the two-child families of the 1998–2002 NHIS data is very close to 1/3 despite a B/G ratio of 105.66.

In conclusion, in our attempts to analyze vos Savant's problem using data sets from her and from two NHI surveys, we found that the data contained so many possible sources of bias and deviations from binomial distributions that they were essentially refractive to standard statistical procedures. We have not seen any data sets for human sex distributions in families of size 2 from other investigators that have attempted an analysis of the two-boy problem comparable to ours. Perhaps such attempts have been made but were terminated, without publication, when it became obvious that normal statistical tests could not be used to make rigorous inferences (such as 95% confidence intervals). It is little wonder that these kinds of data and attempts to analyze them are missing from standard biology textbooks. Writers of such texts and teachers would naturally not want to mislead their readers or students by oversimplifications. So if the percentage of BB, (BG + GB), and GG sibships is presented as conforming to a binomial distribution, it should be

emphasized that this is not a model to which human populations commonly conform. The assumptions underlying all such models and possible sources of bias in the data should be carefully examined before drawing conclusions from them. Our analyses of vos Savant's two-boy problem reveals some of the complexities that are seldom encountered in textbooks that discuss only idealized models rather than actual data. We certainly do our graduate students no favor by sheltering them from such realities.

In a much broader sense, it seems to us that our research on the sex distribution in two-child families has direct connection to what many people believe to be one of the world's most pressing problems, namely, overpopulation and various methods by which population growth can be halted or reversed. Even if every woman now of fertile age has only two children on average, the UN predicts that the world's population will still be growing toward 9 billion people by 2050 (MacKenzie 2008). Reproductive rates are now rapidly declining in some countries, and if this trend continues, population size is predicted to decrease after 2050. The only ways to slow it further are an even more rapid decrease in birth rates or a massive rise in mortality (*New Scientist* editor 2008). China's one-child policy has been draconian and has already led to such a gender imbalance that, if continued, it will lead to far too many men for a restricted pool of women. At one time, the B/G ratio at birth in China was as high as 1.3. In parts of China, family size is still restricted to one child. However, some exceptions are now allowed; for example, couples who are both only children can have a second child. The birth rate has actually dropped (evolved) lately in China, partly because of sex-selective abortion. This drop raises the average age of the population so that by 2035 it is projected that for each person older than 65, there will be just three working-age people, compared to ten today (Nowak 2007). Social upheavals are bound to follow in the wake of these demographic changes. These, and many other problems linked to family size and population structure, are of great importance to physicians, anthropologists, social scientists, and politicians.

In the August 26, 2007, issue of *Parade*, a reader posed the following question to vos Savant. "Say that a country permits families to have only one child. However, if the firstborn is a girl, parents are allowed to have one more child if they wish to try for a boy. Over time, will this policy produce more males, more females, or neither?" Marilyn responds as follows. "The policy will limit population growth, but it has no effect on nature, which determines the gender balance of babies: Boys and girls will be produced in the usual proportion. However, fewer parents will have two children, so fewer families will include children of both sexes" (Vos Savant 2007).

Acknowledgment We are indebted to C. M. Andreoli for translating Malinvaud (1955) from French into English.

Received 18 August 2008; revision accepted for publication 12 March 2009.

Literature Cited

- Carlton, M. A., and W. D. Stansfield. 2005. Making babies by the flip of a coin? *Am. Statistician* 59:180–182.
- Edwards, A. W. F. 1960. The meaning of binomial distribution. *Nature* 186:1074.
- Edwards, A. W. F. 1966. Sex-ratio data analyzed independently of family limitation. *Ann. Hum. Genet.* 29:337–346.
- Jacobsen, R., H. Moller, and A. Mouritsen. 1999. Natural variation in the human sex ratio. *Hum. Reprod.* 14:3120–3125.
- James, W. H. 1976. Sex ratio and the sex composition of the existing sibs. *Ann. Hum. Genet.* 38(3):371–378.
- James, W. H. 1985. Sex ratio, dominance status, and maternal hormone levels at the time of conception. *J. Theor. Biol.* 114:505–510.
- James, W. H. 1987a. The human sex ratio. 1. A review of the literature. *Hum. Biol.* 59:721–752.
- James, W. H. 1987b. The human sex ratio. 2. A hypothesis and a program of research. *Hum. Biol.* 59:873–900.
- James, W. H. 1995. What stabilizes the sex ratio? *Ann. Hum. Genet.* 59:243–249.
- James, W. H. 2000. The variation of the probability of a son within and across couples. *Hum. Reprod.* 15(5):1184–1188.
- James, W. H. 2001. The causes of the variation of the probability of a son within couples. *Hum. Reprod.* 16(3):596–597.
- James, W. H. 2008. Evidence that mammalian sex ratios are partially controlled by parental hormone levels around the time of conception. *J. Endocrinol.* 198:3–15.
- Mackenzie, D. 2008. The population paradox. *New Scientist* 260(2683):20.
- Malinvaud, E. 1955. Relations entre la composition des familles et le taux de masculinité. *J. Soc. Statist. Paris* 96:49–64.
- Matthews, T. J., and B. E. Hamilton. 2005. *Trend Analysis of the Sex Ratio at Birth in the United States*. National Vital Statistics Reports, v. 53, no. 20. Hyattsville, MD: National Center for Health Statistics.
- McWhirter, K. G. 1956. Control of sex ratio in mammals. *Nature* 178:870–871.
- New Scientist* editor. 2008. [Editor's response to a letter suggesting family planning as a part of the solution to slow the growth of population.] *New Scientist* 199(2663):20.
- Nowak, R. 2007. Hope I get rich before I get old. *New Scientist* 196:62–63.
- Pickles, A. R., R. Crouchley, and R. B. Davies. 1982. New methods for the analysis of sex ratio data independent of the effects of family limitation. *Ann. Hum. Genet.* 46:75–81.
- Snell J. L., and B. Peterson. 2007. [Ask Marilyn discussion questions]. Available from http://www.dartmouth.edu/~chance/chance_news/recent_news/chance_news_6.12.html#two-boy%20problem. Accessed May 18, 2008.
- Stansfield, W. D., and M. A. Carlton. 2007. Human sex ratios and sex distribution in sibships of size 2. *Hum. Biol.* 79:255–260.
- Vos Savant, M. 1997. Ask Marilyn [column]. *Parade* (October 19), 8–9.
- Vos Savant, M. (2007). Ask Marilyn [column]. *Parade* (August 26), 16.
- Wikipedia. n.d. Marilyn vos Savant. Available from http://en.wikipedia.org/wiki/Marilyn_vos_Savant. Accessed May 18, 2008.