

On the Mathematics of Fraternal Birth Order Effect and the Genetics of Homosexuality

Cite this article as: Tanya Khovanova, On the Mathematics of Fraternal Birth Order Effect and the Genetics of Homosexuality, Archives of Sexual Behavior <https://doi.org/10.1007/s10508-019-01573-1>

This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.springer.com/aam-terms-v1>

Author accepted manuscript

On the Mathematics of Fraternal Birth Order Effect and
the Genetics of Homosexuality

Tanya Khovanova¹

¹MIT, Cambridge, MA. USA, email: tanyakh@yahoo.com

Author accepted manuscript

Abstract

Mathematicians have always been attracted to the field of genetics. The mathematical aspects of research on homosexuality are especially interesting. Certain studies show that male homosexuality may have a genetic component that is correlated with female fertility. Other studies show the existence of the fraternal birth order effect, that is the correlation of homosexuality with the number of older brothers.

This paper is devoted to the mathematical aspects of how these two phenomena are interconnected. In particular, we show that the fraternal birth order effect implies a correlation between homosexuality and maternal fecundity. Vice versa, we show that the correlation between homosexuality and female fecundity implies the increase of the probability of the younger brothers being homosexual.

Keywords: Fraternal birth order effect, male homosexuality, fecundity, genetics.

1 Background

According to the study by Blanchard and Bogaert [3] (1996): “[E]ach additional older brother increased the odds of [male] homosexuality by 34%.” (see also Blanchard [1] (2004) and Bogaert [4] (2006), Bogaert, Skorska, Wang, Gabrie., MacNeil, Hoffarth, VanderLaan, Zucker, and Blanchard [5] (2018), and a recent survey by Blanchard [2] (2018)). The current explanation is that carrying a boy to term changes their mother’s uterine environment. Male fetuses produce H-Y antigens which may be responsible for this environmental change for future fetuses.

The research into a genetic component of male gayness shows that there might be some genes in the X chromosome that influence male homosexuality. It also shows that the same genes might be responsible for increased fertility in females (see Ciani, Cermelli, & Zanzotto [6] (2008) and Iemmola, & Ciani [7] (2008)).

In this paper we compare two mathematical models. In these mathematical models we disregard girls for the sake of clarity and simplicity.

The first mathematical model of *Fraternal Birth Order Effect* (FBOE), which we denote FBOE-model, assumes that each next-born son becomes homosexual with increased probability. This probability is independent of any other factor.

The second mathematical model of *Female Fecundity* (FF), which we denote FF-model, assumes that a son becomes homosexual with probability depending on the total number of children and nothing else.

We show mathematically how FBOE-model implies correlation with family size and FF-model implies correlation with birth order. That means these two models are mathematically intertwined.

We also propose the *Brother Effect*. Brothers share a lot of the same genes. It is not surprising that brothers are more probable to share traits. With respect to homosexuality, we call the correlation that homosexuals are more probable to have a homosexual brother than a non-homosexual the Brother Effect. The existence of genes that increase predisposition to homosexuality implies the Brother Effect. The connection between FBOE-model and the Brother Effect is more complicated.

We also discuss how to separate FBOE and FF in the data.

Section 2 contains extreme mathematical examples that amplify the results of this paper. Section 3 shows how FBOE-model implies the correlation with family size. Section 4 shows how FF-model implies the correlation with birth order. In Section 5 we discuss the connection between FBOE-model and the Brother Effect. In Section 6 we discuss how to separate the birth order from the family size.

2 Extreme Examples

First consider extreme theoretical examples. In the first two examples, suppose mothers only give birth to sons and only to one or two sons.

First Extreme example. This is an extreme variation of FBOE-model. Suppose the

first son has a zero probability of being gay (which means that first sons are never gay) and the second son has probability one of being gay (which means he is always gay). Then all mothers of one son will have a straight son. All mothers with two sons will have one gay and one straight son. Homosexuals appear only in two-son families and never in one-son families. Therefore, FBOE-model implies the correlation with family size.

Second Extreme example. This is an extreme variation of FF-model. Suppose mothers with one son have probability zero of having a gay son. Suppose mothers with two sons have two homosexual sons with probability one. The first born is sometimes gay and sometimes straight, but the second son is always gay. Hence, it is more probable that the second son is gay. Therefore, FF-model implies the correlation with birth order.

These extreme variations of FBOE-model and FF-model show that these two models are intertwined.

The next two sections explain this in more detail.

3 FBOE-model and the family size

Let us build a model with variables for numbers that correspond to FBOE-model. In this simple model we assume that the probability of a child being gay depends only on birth order and nothing else.

FBOE-model. Let us assume that mothers have either one or two boys. Let a be the probability of a woman having one boy, and correspondingly, $1 - a$ of having two boys. Suppose N is the total number of women in consideration. Suppose p_1 is the probability that the first boy is homosexual and p_2 is the probability that the second boy is homosexual. The fraternal birth order effect means that $p_2 > p_1$.

Now we produce the results of such a model.

Let us first estimate the total number of boys T :

$$T = aN + 2(1 - a)N = (2 - a)N.$$

The number of homosexuals in the one-son families is expected to be ap_1N . The expected number of homosexual first-born sons in two-son families is $(1 - a)p_1N$ and the expected number of homosexual second-born sons is $(1 - a)p_2N$. The total expected number of homosexuals H is the sum:

$$H = p_1N + (1 - a)p_2N.$$

The probability that a randomly chosen boy is a homosexual is

$$\frac{H}{T} = \frac{p_1 + (1 - a)p_2}{2 - a}.$$

Let us see what happens with fecundity. Suppose we pick a mother randomly, then pick her son randomly. If there is only one son, then he is the one we have to pick. The probability that we pick a gay son, given that we picked the mother with one child is p_1 . The probability that we pick a gay son, given that we picked the mother of two children is $(p_1 + p_2)/2 > p_1$. This difference is the source of the correlation with fecundity.

To calculate this properly we need to choose a boy randomly and find the average fertility of the mother. The formula is given by the following equation:

$$\frac{\# \text{ number of single sons} + 2 \cdot \# \text{ number of non-single sons}}{\# \text{ number of sons}}. \tag{1}$$

First we calculate average maternal fertility per boy:

For a randomly chosen boy (including both homosexual and non-homosexual boys), there are aN mothers of one son and $(1 - a)N$ mothers of two sons. Hence, a mother of a randomly chosen boy has on average

$$\frac{aN + 2 \cdot 2(1 - a)N}{(2 - a)N}$$

children, which is equal to

$$\frac{a + 4 - 4a}{2 - a} = 2 - \frac{a}{2 - a}.$$

Let us see what happens with homosexual boys. We have ap_1N expected gay boys from one-son families and $(1 - a)(p_1 + p_2)N$ expected gay boys from two-son families. Now we plug this into the Eq. (1) where we replace a randomly chosen boy with a gay boy to get:

$$\begin{aligned} & \frac{ap_1N + 2(1-a)(p_1+p_2)N}{ap_1N + (1-a)(p_1+p_2)N} \\ &= \frac{ap_1 + 2(1-a)(p_1+p_2)}{ap_1 + (1-a)(p_1+p_2)} \\ &= \frac{2(ap_1 + (1-a)(p_1+p_2)) - ap_1}{ap_1 + p_1 - ap_1 + p_2 - ap_2} \\ &= 2 - \frac{ap_1}{p_1 + p_2 - ap_2}. \end{aligned}$$

If we denote by c the ratio p_2/p_1 , then the average maternal fertility per gay boy is

$$2 - \frac{a}{1 + (1-a)c}.$$

As $c > 1$, then $1 + (1 - a)c > 2 - a$. Therefore, $2 - \frac{a}{1+(1-a)c} > 2 - \frac{a}{2-a}$. It follows that the average maternal fertility per gay boy is greater than the overall average maternal fertility.

It is useful to note, that if $c = 1$, then there is no correlation with the family size; that is, the average fertility is the same for randomly chosen boys and gay boys. This is the expected result. Indeed, $c = 1$ means homosexuality does not depend on the birth order and is assigned completely randomly.

The impact of FBOE-model on the correlation with the family size is stronger if we consider larger families. Suppose $p_1 < p_2 < p_3 < \dots$ are the probabilities of the first, second, and so on child being gay, correspondingly. Then the average probability, x_k , of being gay per child in a k -son family is

$$x_k = \frac{p_1 + p_2 + \dots + p_k}{k}.$$

When we add larger numbers to the average, the average increases. Thus, $x_j > x_i$, when $j > i$.

Among boys with many older brothers, there is a larger proportion of homosexuals. Thus, they contribute more to the calculation of average fecundity. Hence, if we add to our model the possibility of more than two boys where each next boy has a higher probability of being homosexual, the correlation will be more impressive.

The results show that FBOE-model implies correlation with the family size. The female fecundity correlation with male homosexuality was shown not only for mothers, but also for maternal aunts and grandmothers [6, 7]. This means, the fecundity results as a whole are not threatened by my examples. We will describe in Section 6 how to work with the data to mathematically separate birth order and family size.

4 FF-model implies birth order correlation

Let us build a mathematical model with variables instead of fixed numbers that correspond to the correlation of homosexuality with female fecundity. In this simple model we assume that the probability of a child being gay depends only on the family size and nothing else.

FF-model. Let us assume that mothers have either one or two boys. Let a be the probability of a woman having one boy, and correspondingly, $1 - a$ of having two boys. Suppose N is the total number of women in consideration. Suppose q_1 is the probability that a boy in a one-son family is homosexual and q_2 is the probability that a boy in a two-son family is homosexual. We assume that $q_2 > q_1$ to support the correlation of female fecundity with homosexuality.

Here are the results of such a model. In our notation we use index f for first sons and s for second sons.

Let us see what happens with birth order. We start with first sons. The total number of first sons T_f is N :

$$T_f = N.$$

The number of homosexuals in one-son families is expected to be aq_1N . The number of homosexual first sons in two-son families is expected to be $(1-a)q_2N$. The total number of first sons that are homosexual, H_f , is expected to be:

$$H_f = aq_1N + (1-a)q_2N.$$

The probability that the first-born is homosexual is

$$\frac{H_f}{T_f} = aq_1 + (1-a)q_2.$$

Now we do the same for the second-born sons. The expected total number of them T_s is:

$$T_s = (1-a)N.$$

The expected number of homosexuals among them H_s is:

$$H_s = (1-a)q_2N.$$

The probability that the second-born son is homosexual is

$$\frac{H_s}{T_s} = q_2.$$

The final mathematical step needs to show that the probability that the first born is homosexual is less than the probability that the second born is homosexual. It follows from the fact that $q_1 < q_2$. Indeed:

$$\frac{H_f}{T_f} = aq_1 + (1-a)q_2 < aq_2 + (1-a)q_2 = q_2 = \frac{H_s}{T_s}.$$

If we denote by c the ratio q_2/q_1 , then the ratio of increase of the proportion of the

homosexuals in later children, that is $\frac{H_s}{T_s}$ divided by $\frac{H_f}{T_f}$ is:

$$\frac{q_2}{aq_1 + (1-a)q_2} = \frac{c}{a + (1-a)c}.$$

It is useful to note, that if $c = 1$, then there is no correlation with the birth order: the first-born sons and second-born sons are homosexuals with the same probability.

The impact of FF-model on the birth order is stronger if we consider larger families. Suppose $q_1 < q_2 < q_3 < \dots$ are the probabilities of sons being homosexual in families of size 1, 2, and so on, respectively. Then the average probability y_i of a child number i being gay depends on the distribution of family sizes. Suppose the number of families of size m is N_m , then we can calculate y_i as:

$$y_i = \frac{q_i N_i + q_{i+1} N_{i+1} + q_{i+2} N_{i+2} + \dots}{N_i + N_{i+1} + N_{i+2} + \dots}.$$

We can show that $y_j > y_i$, when $j > i$. Let us denote $Q_1 = q_i N_i + q_{i+1} N_{i+1} + q_{i+2} N_{i+2} + \dots + q_{j-1} N_{j-1}$ and $Q_2 = q_j N_j + q_{j+1} N_{j+1} + q_{j+2} N_{j+2} + \dots$. Further, let us denote $M_1 = N_i + N_{i+1} + N_{i+2} + \dots + N_{j-1}$ and $M_2 = N_j + N_{j+1} + N_{j+2} + \dots$. Then $y_i = \frac{Q_1 + Q_2}{M_1 + M_2}$ and

$$y_j = \frac{Q_2}{M_2}.$$

The important observation is that

$$\frac{Q_1}{M_1} \leq q_{i-1} < q_i < \frac{Q_2}{M_2}.$$

Therefore,

$$Q_1 M_2 < Q_2 M_1.$$

It follows that

$$Q_1 M_2 + Q_2 M_2 < Q_2 M_1 + Q_2 M_2.$$

This implies

$$y_i = \frac{Q_1 + Q_2}{M_1 + M_2} < \frac{Q_2}{M_2} = y_j.$$

The results show that FF implies correlation with the birth order. The fraternal birth order effect was shown only for brothers and not for sisters. This means, the FBOE as a whole is not threatened by my examples. We will describe in Section 6 how to separate birth order and family size mathematically in the data.

5 Brothers

Siblings share a lot of genetic material. Not surprisingly they have a lot of common traits. If a trait is genetic, then the probability that a sibling has it is higher than the probability that a randomly chosen person has it. Very often the fact that siblings share traits with higher probability than random people share traits serves as a confirmation that the trait is genetic. That means the existence of a homosexual gene would imply the higher probability that a gay person has a gay brother than the probability that a randomly chosen person has a gay brother. We call this correlation the Brother Effect.

Is there a mathematical way to connect the birth order with the Brother Effect? The answer: it is complicated.

Let us look at how the fraternal birth order effect influences the probability that a gay boy has a gay brother. The probability that a gay person has a gay brother depends on the number of boys in the family. If a boy does not have brothers he cannot have a gay brother. If a boy has a million brothers, then with extremely high probability at least one of them will be gay.

Here are two extreme mathematical examples where we assume that mothers have only one or three sons.

Third extreme example. This is an extreme variation of FBOE-model. Suppose the first son has zero probability of being gay and the second and third sons have probability

one of being gay. That is, $p_1 = 0$ and $p_2 = p_3 = 1$. All gay boys in this model have a gay brother, while a randomly chosen boy sometimes has one and sometimes does not. With these particular probabilities FBOE-model implies the Brother Effect.

Fourth extreme example. This is an extreme variation of FBOE-model. Suppose the first and second sons have zero probability of being gay and the third son has probability one of being gay. That is, $p_1 = p_2 = 0$ and $p_3 = 1$. No gay boy in this model has a gay brother, while some randomly chosen boys have one. With these particular probabilities FBOE-model contradicts the Brother Effect.

It follows that depending on the actual numbers FBOE might or might not imply the Brother Effect.

6 Separating birth order and female fecundity

Our simplistic models in Sections 3 and 4 showed that FBOE-model and FF-model imply each other. That means birth order and female fecundity are intertwined in the data. It is important to separate these two different models. To do it we need to fix some variables.

Method 1. To show how the birth order works independently of female fecundity, we need to fix the family size. Suppose we consider only families of size 2. Then the fertility does not play a role. In this case, according to the fraternal birth order effect, the second son is gay with higher probability than the first son. The corresponding probabilities derived from real data should confirm FBOE-model without interference of FF-model.

Method 2. To show how the female fecundity works independently of the fraternal birth order effect, we need to consider only the first sons. Then the FBOE-model does not play a role. In this case, according to FF-model, the first son in a larger family is gay with higher probability than the first son in a smaller family. The corresponding probabilities derived from real data should confirm FF-model without FBOE-model.

Consider the theoretical discussion of families of size one and two in Sections 3 and 4.

Here is the joint mathematical model, which we call FBOE-FF-model.

FBOE-FF-model. Let us consider only the case of women who have one or two boys. Let a be the probability of a woman having one boy, and correspondingly, $1 - a$ of having two boys. Suppose N is the total number of women in consideration. Suppose p_{11} is the probability that the first boy in a one-son family is homosexual, p_{12} is the probability that the first boy in a two-son family is homosexual, and p_{22} is the probability that the second boy in a two-son family is homosexual. The female fecundity means $p_{12} > p_{11}$. The ratio $\frac{p_{12}}{p_{11}}$ shows a contribution of FF independent of FBOE. The fraternal birth order effect means $p_{12} > p_{22}$. The ratio $\frac{p_{12}}{p_{22}}$ shows the contribution of FBOE independent of FF.

7 Conclusion

We showed mathematically that FBOE and FF are interrelated in the data: FBOE implies FF, and FF implies FBOE. We proposed data analysis that will separate these two contributions. To show the impact of FBOE without interference of FF, one has to look at families of the same size. To show the impact of FF without interference of FBOE one has to look at the first born sons.

References

- [1] Blanchard, R. (2004). Quantitative and theoretical analyses of the relation between older brothers and homosexuality in men. *J Theor Biol.* 230, 173–187. doi:10.1016/j.jtbi.2004.04.021
- [2] Blanchard, R. (2018). Fraternal Birth Order, Family Size, and Male Homosexuality: Meta-Analysis of Studies Spanning 25 Years. *Arch. Sex. Behav.*, 47, 1–15. doi: 10.1007/s10508-017-1007-4

-
- [3] Blanchard, R., & Bogaert, A. F. (1996). Homosexuality in Men and Number of Older Brothers. *Am J Psychiatry* 153, 27–31. doi: 10.1176/ajp.153.1.27
- [4] Bogaert A. F. (2006). Biological versus nonbiological older brothers and men's sexual orientation. *PNAS*, 103(28), 10771–10774. doi: 10.1073/pnas.0511152103
- [5] Bogaert, A. F., Skorska, M. N., Wang, C., Gabrie, J., MacNeil, A. J., Hoffarth, M. R., VanderLaan, D. P., Zucker, K. J., and Blanchard, R. Male homosexuality and maternal immune responsivity to the Y-linked protein NLGN4Y. (2018) *Proc Natl Acad Sci USA*, 115, 302–306. doi: 10.1073/pnas.1705895114
- [6] Ciani, A. C., Cermelli, P. & Zanzotto, G. (2008). Sexually Antagonistic Selection in Human Male Homosexuality. *PLoS ONE*, 3(6):e2282. doi: 10.1371/journal.pone.0002282
- [7] Iemmola, F. & Ciani, A. C. (2009). New Evidence of Genetic Factors Influencing Sexual Orientation in Men: Female Fecundity Increase in the Maternal Line. *Arch Sex Behav*. 38(3), 393–9. doi: 10.1007/s10508-008-9381-6
- [8] Khovanova, T. (2008) Fraternal Birth Order Threatens Research into the Genetics of Homosexuality, *Tanya Khovanova's Math Blog*. Retrieved from <http://blog.tanya-khovanova.com/?p=40>