Base-rate neglect and birth order dependent academic effort in Michael Sandel's Harvard Justice class

Antony Millner*[1] and Raphael Calel[1]

[1]Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, London, UK.

July 2010

Abstract

This short article highlights a misleading statistical inference in Michael Sandel's recently published 'Justice' (Sandel, 2009). Sandel conducted an informal survey among his Harvard students, and discovered that 75-80% of them are first born children. He uses this data to infer that the academic effort a child exerts depends on birth order, with first born children exhibiting high effort levels, and thus being over-represented at Harvard. However Sandel makes this inference without reference to the 'base-rate' - the proportion of first born children amongst the offspring of mothers of Harvard students. Such faulty reasoning about conditional probabilities is widely documented by psychologists. We present a simple model that shows that the birth order effect needs to be unreasonably large in order for it to explain Sandel's data. A more plausible explanation is that mothers of Harvard students have significantly fewer children than the national average.

*Email: antony.millner@lse.ac.uk

Michael Sandel's 'Justice' (Sandel, 2009) is a rewarding and accessible account of political philosophy, based on a course he has taught at Harvard for over two decades. It cites an interesting 'unscientific' (his words) survey he has conducted with his students in order to demonstrate John Rawls' critique of feudal, libertarian, and meritocratic conceptions of distributive justice. Rawls suggests that all three of these theories base just distributive shares on morally arbitrary chance endowments – whether of birth, property, or genetically inherited natural abilities – and that only his difference principle, the idea that inequality is permissible only if it benefits the worst off in society, avoids this arbitrariness. A potential critique of Rawls' position is that it neglects the causal relationship between effort and achievement, however Rawls argues that even effort may be a product of favourable, and arbitrary, individual circumstances. In order to support this point, Sandel asks his Harvard students to raise their hands if they are the first born children in their families. According to Sandel, 'About 75-80 percent raise their hands. The result has been the same every time I have taken the poll.'

Sandel seems to be inferring that the fact that so many Harvard undergraduates are first born suggests that birth order has a strong effect on academic effort (assuming it has no effect on innate academic abilities). However this reasoning is a classic case of base-rate neglect, a phenomenon widely studied in the psychology literature (Bar-Hillel, 1980). Whether 75-80% is a high or low number depends on the proportion of first born children amongst the offspring of mothers of Harvard students, a statistic that Sandel does not make reference to. Put another way, Sandel estimates $\Pr(1^{\text{st}} \text{ born}|\text{Harvard})$, when he is really interested in how $\Pr(\text{Harvard}|n^{\text{th}} \text{ born})$ depends on n. The

former quantity depends on the relative frequencies of n^{th} born children amongst mothers with children at Harvard, the 'base-rate', and is thus not identifiable with the latter.

To get a handle on what Sandel is really inferring from his data, begin by noting that 41% of the children born in the US in 1991 (likely the birth year of many of his current students) were first born (Martin et al., 2009). There is thus a big gap between the national base-rate and Sandel's sample. If we assume that this gap is exclusively due to a birth order effect, how large does this effect need to be to explain the data? In order to answer this question, consider the following simple model:

Given a population of mothers, consider the set of *all* their childern. Let F be the subset of first born children in this set, and let H be the subset of Harvard students in this set. Let $p(F|H)$ be the probability of being a first born child, given that you are a Harvard student, and $p(H|F)$ be the probability of being at Harvard, given that you are a first born child. Then Bayes' formula tells us that

$$p(F|H) = \frac{p(H|F)p(F)}{p(H|F)p(F) + p(H|F^c)(1 - p(F))} \tag{1}$$

where $p(H|F^c)$ is the probability of going to Harvard given that you are *not* a first born child (F^c is the complement of F), and $p(F)$ is the unconditional probability of being a first born child.

We now transform this expression into more meaningful variables. Consider the quantity $p(F)$. In any population of children born to N mothers, there will be exactly N first born children (since each mother has at least

one child). Let the fertility rate amongst mothers of Harvard students be λ – then the total number of children these mothers give birth to is just λN. Thus $p(F)$, the proportion of first born children in this population, is given by

$$p(F) = \frac{N}{N\lambda} = 1/\lambda. \tag{2}$$

Substituting (2) into (1), dividing the numerator and the denominator of (1) through by $p(F)p(H|F^c)$, and defining

$$r := \frac{p(H|F)}{p(H|F^c)}$$

we have that

$$p(F|H) = \frac{r}{r + \lambda - 1}. \tag{3}$$

This expression relates the proportion of Harvard first borns to the strength of the birth order effect, measured by r, and the fertility rate amongst mothers of Harvard students, measured by λ. When *only* the birth order effect is present, the fertility rate of Harvard mothers must be the same as that in the general population of mothers, i.e. approximately $1/0.41 = 2.44$. If this is the case, equation (3) requires that first born children be 4.3-5.8 times more likely to attend Harvard than later born children to be consistent with Sandel's data. Thus Sandel's reasoning relies on a very large birth order effect indeed.

Equation (3) however offers an alternative explanation for the large gap between the proportion of first borns in Sandel's class and that in the general population – perhaps the fertility rate amongst mothers of Harvard students

is lower than the national average? Figure 1 plots the complete set of values of r and λ that are consistent with Sandel's estimates of $p(F|H) \in [0.75, 0.8]$. From the graph we see that the data can also be explained without recourse to a birth order effect if the fertility rate of mothers with children at Harvard is 1.25-1.33. There is also a continuum of intermediate possibilities in which the birth order effect and exceptional fertility rates both play partial explanatory roles.

Given the information that Sandel presents, it is impossible to tell which parameter values best fit his data. Nevertheless, it seems clear that if we wish the birth order effect to be the sole explanatory variable, it needs to be unreasonably large. Moreover, given the well-known negative correlation between fertility rates and womens' income and education levels (Jones and Tertilt, 2008), two variables that are also likely to be correlated with their chances of having a child at Harvard, the low fertility rates needed to explain the data are not implausible. Thus at the very least, Sandel's inference that the high proportion of first born children in his class is evidence for a strong birth order effect is tenuous, and requires further data to be substantiated.

While Sandel's philosophical pedagogy is a model of clarity, his reasoning about statistical evidence is (no doubt unintentionally) misleading. Surely there is as much need to foster careful thought about empirical evidence amongst Harvard undergraduates as there is about questions of morality? After all, normative judgements are statements about the desirability of factual states – we probably need to have the facts straight first before we can decide whether they are 'good'!

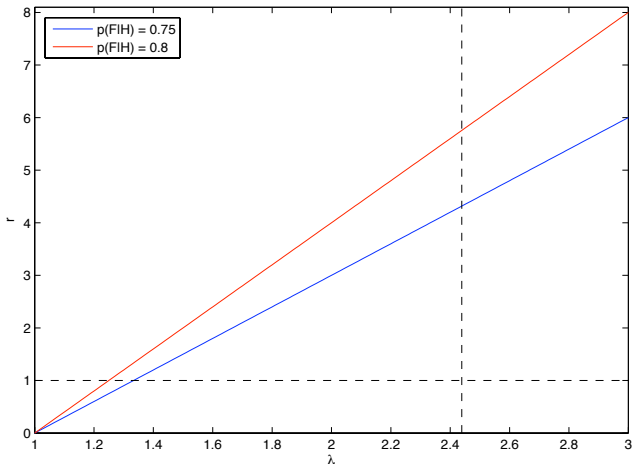

Figure 1: Values of r and λ consistent with Sandel's estimate of the proportion of first born children in his class. The horizontal dashed line at $r = 1$ corresponds to the case in which his data are entirely explained by a low fertility rate amongst mothers of Harvard students, while the vertical dashed line at $\lambda = 2.44$ indicates the case in which his data are entirely explained by the birth order effect.

References

Bar-Hillel, Maya. 1980. "The base-rate fallacy in probability judgments." *Acta Psychologica* 44(3):211–233.

Jones, L.E. and M. Tertilt. 2008. An economic history of fertility in the U.S.: 1826-1960. In *Frontiers of Family Economics, Vol. 1*, ed. P. Rupert. Emerald Press pp. 165–230.

Martin, Joyce A., Brady E. Hamilton, Paul D. Sutton, Stephanie J. Ventura, Fay Menacker, Saharon Kirmeyer and T.J. Mathews. 2009. "National Vital Statistics Reports." 57(7).

Sandel, Michael. 2009. *Justice*. Farrar, Straus and Giroux.