UCD GEARY INSTITUTE DISCUSSION PAPER SERIES

Big and Tall Parents do not Have More Sons

Dr. Kevin Denny

UCD School of Economics & Geary Institute, University College Dublin,
Belfield, Dublin 4
Tel: +353 1 7164632
kevin.denny@ucd.ie

04th February 2008

Geary WP/3/2008

Abstract:

In a 2005 paper Kanezawa proposed a generalisation of the classic Trivers-Willard hypothesis. It was argued that as a result taller and heavier parents should have more sons relative to daughters. Using two British cohort studies, evidence was presented which was partly consistent with the hypothesis. I analyse the relationship between an individual being male and their parents' height and weight using one of the datasets. No evidence of any such relationship is found.

A shorter version of this is forthcoming in the *Journal of Theoretical Biology*

1 Introduction

The paper by Trivers and Willard (1973) on the sex ratio has generated a huge literature. As Carranza (2002) points out "Probably no other case exists in behavioural ecology where a couple of pages have sired so many studies". This research has not necessarily brought clarity and what the theory actually implies, whether it can be tested and whether it applies to humans is disputed (see, for example, Carranza (2002), Brown (2001), Freese and Powell (2001)). Recently Kanazawa (2005) proposes a generalization of the Trivers-Willard hypothesis namely that parents who possess any heritable trait that increase male reproductive success at a greater rate than female reproductive success will have more male offspring. It is proposed that size (height and weight) of the parent is one such trait. Evidence is presented that, it is argued, is partly consistent with the hypothesis, heavier parents have more boys and taller parents have fewer girls controlling for a number of variables.

This note does not question the logic behind this generalization (nor does it endorse it) but shows that analysing the same data somewhat differently leads to very different conclusions. A number of statistical criticisms of the paper and related work by the same author have also been raised by Gelman (2007) and Gelman and Weakliem (2007).

Briefly, Kanazawa (2005) uses the National Child Development Survey (NCDS), a cohort study of individuals born in Great Britain in 1958 and the British Cohort Study, a similar cohort born 1970. As dependent variables he uses, separately, the number of sons and daughters born to cohort members. The independent variables of interest are the cohort member's height and weight. Controls include their years of education, income, sex, whether married and the number of children of the opposite sex.

2 Data analysis

An alternative, and to my mind more direct, test of the theory is to examine the relationship between the sex of the cohort member and the size of each parent (as opposed to the size of the cohort member and the number of sons and daughters that each has). Hence I estimate a model Probability(sex=male) = F(parents' size + controls + ε) using the NCDS data. This is essentially the approach of Almond and Eklund (2007) who focus on socio-economic indicators of parental quality using a large (n=48 *million*) dataset and the results of which generally support the Trivers-Willard hypothesis. Father's and mother's height (measured in centimetres) and their Body Mass Index (BMI, equal to weight in kilograms divided by {height in metres squared}) both measured in 1958 are the independent variables of interest. BMI is the standard measure of weight used in epidemiology. Two other pairs of controls are used: the parents' age at birth and the age at which they left school. For females, in particular, one can think of age as being a decreasing measure of parental quality. Age left education will be positively correlated with income, social status and cognitive ability. Using a categorical measure of father's social class leads to identical conclusions. I use the probit estimator, that is it is assumed that the F(.) function above follows a Normal distribution[1]. Almond and Eklund (2007) use the linear probability model. While this has some undesirable statistical properties (for example predicted probabilities can lie outside the (0,1) range) it typically gives qualitatively similar results to probit and logit.

Table 1 presents three models. The means of the independent variables are in the final column. The dependent variable is a binary variable indicating whether the cohort member is male (1) or female (0), 51.2% of the sample are male. In the first column father's and mother's height and BMI only are included as regressors. Not one of these coefficients is individually

---

[1] See for example Verbeek (2006). As usual, using a logit estimator gives much the same results.

statistically significant nor can one reject the hypothesis that they are jointly insignificant (p value=0.99). To test the theory directly one sided tests (i.e. that the coefficients are positive) could be used however clearly in this case this will lead to the same conclusion as will Bonferroni-type adjustments to the critical p-values as suggested by Gelman (2007).

In column 2, parents' age at birth is included and in column 3 the age at which each parent left full-time education is added. In both cases the result from the first model is unchanged: both parents' height and BMI have no statistically significant effect on the probability of being born male. Since the effect of size might be non-linear (and even non-monotonic) I also experimented with a more flexible functional form by including the square and cube of both height and BMI. In each case one could not reject the hypothesis that the three coefficients were jointly zero. There is no reason to presume that adding other controls is likely to change the results substantively. Note that even if the estimated coefficients were estimated more precisely (i.e. were "statistically significant") they are tiny in magnitude and hence would be of doubtful scientific significance.

3 Conclusion

It is clear from the data that there is no evidence of any relationship between the size of either parent (measured by height and by BMI) and the sex of their offspring. It is not credible that an evolutionary explanation could hold for one generation but not for the one immediately preceding it hence the suggested generalization of Trivers and Willard is not supported. Further mining of the data is unlikely to change to change this conclusion. Why these results seem to differ so starkly from the evidence presented in Kanazawa (2005) is somewhat unclear although the evidence there in favour of the hypothesis seemed weak at best.

References

Almond, D. and Edlund L. (2007) Trivers–Willard at birth and one year: evidence from US natality data 1983–2001, *Proceedings of the Royal Society*, B, 274, 2491–2496

Brown G.R. (2001) Sex-biased investment in nonhuman primates: can Trivers & Willard's theory be tested? *Animal Behaviour*, 61, 683-694

Carranza J. (2002) What did Trivers and Willard really predict ? *Animal Behaviour*, 63, F1-F3

Freese, J. and Powell, B. (2001) Making love out of nothing at all? Null findings and the Trivers-Willard hypothesis, *American Journal of Sociology*, 106, 1776-1788

Kanazawa, S. (2005) Big and tall parents have more sons: further generalizations of the Trivers-Willard hypothesis, *Journal of Theoretical Biology*, 235, 583-590

Gelman A. (2007) Letter to the editor regarding some papers of Dr Satoshi Kanazawa, *Journal of Theoretical Biology*, 245, 597-599

Gelman A. and Weakliem, D. (2007) Of beauty, sex and power: statistical challenges in estimating small effects. Unpublished technical paper, Statistics Department, Columbia University, New York.

Trivers R.L., and Willard D.E. (1973) Natural selection of parental ability to vary the sex ratio of offspring, *Science*, 179, 90-92

Verbeek, M. (2006) *A guide to modern econometrics*, 2nd ed.Wiley, Chichester.

Table 1 : Probit estimates of the probability of being male

	(1)	(2)	(3)	Mean
Father's height (in cm)	0.000	-0.001	0.000	174.6
	(0.100)	(0.310)	(0.080)	
Father's BMI	0.002	0.002	0.001	24.7
	(0.400)	(0.340)	(0.320)	
Mother's height (in cm)	0.000	0.000	0.001	162.1
	(0.010)	(0.140)	(0.230)	
Mother's BMI	0.001	0.002	0.001	23.8
	(0.230)	(0.440)	(0.300)	
Father's age		-0.004	-0.004	30.3
		(1.120)	(1.050)	
Mother's age		-0.001	-0.001	27.4
		(0.130)	(0.002)	
Age father left education			0.002	15.0
			(0.190)	
Age mother left education			-0.019	15.0
			(1.770)	
Constant	0.002	0.149	0.328	
	(0.000)	(0.310)	(0.670)	
χ^2	0.27	0.50	0.26	
	(p=.99)	(p=.974)	(p=.992)	
Pseudo R-squared	0.000	0.004	0.007	

Note: Z statistics (in parentheses) reported. Father's and mother's height, BMI and age are at the time of the birth of the cohort member. Dependent variable =1 if male, =0 if female, n=8249. Data is from the National Child Development Survey, see Kanazawa (2005) for more details. χ^2 is a test for the joint statistical significance of the first four coefficients, d.f.=4.