Invited reply

Reply to Helle *et al.*

We are grateful for the opportunity to respond to Helle, Lumma and Jokela's (henceforth HLJ; Helle *et al.* 2010) comment on our work. HLJ attribute to us the conclusion that the results in Helle *et al.* (2002) were 'probably false positive owing to smaller sample size'. This is puzzling, for two reasons. First, from a purely statistical point of view, the probability of rejecting a true null (i.e. to obtain a false positive) does not depend on sample size. Second, we never made any statement to this effect. In fact, neither of our two papers uses the phrase 'false positive' or any comparable language. Despite the fact that the conclusion attributed to us by HLJ is inaccurate, the choice of language cuts to the heart of our differences of opinion.

HLJ's comment focuses uniquely on the rejection of a null hypothesis of zero association between sons and maternal longevity. But this confuses statistical significance with what one might call 'demographic significance'. In our papers, we have repeatedly noted that modest effects of sons on maternal longevity cannot be ruled out. HLJ's comment thus sets up a straw man; we are not contesting that sons could have a small negative (or positive) effect on maternal old-age longevity. However, according to HLJ's original estimates (Helle *et al.* 2002) the relative cost of a son was approximately 1 year of post-menopausal longevity in preindustrial humans. No subsequent work on any independent sample has come close to estimating an effect of such a magnitude. The real points of contention are thus: (i) is a one year relative cost plausible? and (ii) what structural interpretation should one give to an empirical relationship between number of sons born and maternal longevity?

In their comment, HLJ augmented their original sample to include 481 women instead of 375 (they also disclose that not all mothers in their original sample were Sami). In addition, they report some analyses where missing data have been imputed, bringing the total sample size to 577. This falls short of the usual standards for a replication since the samples overlap significantly. In this respect, the comment provides little new empirical evidence. A convincing replication would have shown that the pattern documented in Helle *et al.* (2002) holds also among the subset of observations which were previously not analysed. Unfortunately, the authors do not provide this information. HLJ now concede that no subsequent papers have come close to replicating the effect sizes documented in Helle *et al.* (2002). Though HLJ emphasize their disagreements with us, we are struck by this concession.

Concerning the 'structural' interpretation, HLJ's comment does not engage with the points about omitted variable bias and selection bias raised elsewhere in the

literature and in particular in Cesarini *et al.* (2007, 2009). But we do note that the authors have now mostly moved away from causal speculation towards associational language in interpreting their findings and this is a welcome shift. This transition is made without comment, but we would be interested in knowing whether we should infer from this that HLJ now agree with us that the interpretation of the regression coefficients they gave in their 2002 paper was inaccurate. As we have repeatedly emphasized, any statistical inference concerning the absolute effect of sons or daughters on maternal longevity hinges on the assumption that unobserved determinants of maternal longevity are uncorrelated with parity. We have argued that it may be acceptable to assume that the sex composition of children born to a mother is uncorrelated with unobserved determinants of longevity, implying that the relative effect of a giving birth to a son instead of a daughter on maternal longevity can be estimated. It is not reasonable, in our opinion, to assume that fecundity is independent of health, wealth or other factors which may also affect longevity. However, by interpreting their estimated regression coefficients as 'effects' of number of sons on maternal longevity, Helle *et al.* (2002) implicitly made this assumption. We discussed this point in detail (Cesarini *et al.* 2007, 2009) but HLJ's comment does not address the issue. HLJ also write:

'One of the main points emphasized in our original contribution but untested by Cesarini *et al.* (2009) was that children may not have adverse association on [sic!] their mother's lifespan only but also have positive influence on maternal survival, as found among the Finnish Sami where the number of daughters raised to adulthood was positively associated with maternal post-menopausal lifespan' (Helle *et al.* 2002).

This sentence has now been recast in more associational language, unlike the original Helle *et al.* (2002) contribution in which a causal hypothesis was being tested. Yet, the authors still slip into making causal statements ('positive influence'). We therefore consider it misleading to refer to a hypothesis as untested when we have repeatedly emphasized that such hypotheses are *untestable* within the regression framework employed by Helle *et al.* (2002).

HLJ also disagree with our reading of the literature. They note the omission of two potentially relevant works and assert that we misrepresented the findings in three other papers. But these works only further reinforce our conclusions. One omission, McArdle *et al.* (2006), find no association between the sex of offspring and longevity. The other, Carey (2003) reports results suggesting that the number of sons born could be weakly negatively related to post-menopausal longevity. Neither of these omissions presents any evidence in favour of HLJ's original conclusion of a large differential effect of sex of offspring on maternal longevity.

The accompanying comment can be viewed at <http://dx.doi.org/doi:10.1098/rspb.2009.2114>.

Neither do Jasienska *et al.* (2006) whose findings it is suggested that we misrepresented along with those in Van de Putte *et al.* (2004) and Hurt *et al.* (2006). In the words of HLJ: 'Jasienska *et al.* (2006) demonstrated that in rural Polish population mothers with many sons were short-lived but also that daughters born had a similar association with maternal mortality.' The authors also complain that we do not question the findings of Jasienska *et al.* (2006) based on their small sample size. But our paper is a replication of Helle *et al.* (2002), not Jasienska *et al.* (2006), which was only mentioned in passing as part of a literature review. The claim that we misrepresented the findings in Van de Putte *et al.* (2004) and Hurt *et al.* (2006) is also without foundation. Van de Putte *et al.* did not find a statistically significant association between maternal longevity and sex of offspring in their baseline specification, but only in a restricted subset, as described by Cesarini *et al.* (2009). Hurt *et al.* (2006) did indeed find an 'effect' on maternal longevity after controlling for number of surviving sons. But HLJ give no reason for why one should favour a specification with these controls included.

Our interpretation of the anomalous coefficient estimate in Helle *et al.* (2002) is that it is owing to sampling variation, but it is also possible that fundamental differences between the two sample populations explain the discrepancy between our results and those in Helle *et al.* (2002). We have not, however, found any convincing evidence for why this would be the case. More than 90 per cent of the Sami in our sample come from parishes in Northern Lappland situated close to Finland. In fact, one of the parishes in HLJ's sample, Enontekiö, used to belong to the same parish as one of our larger sample populations, Karesuando, prior to the separation of Sweden and Finland in 1809 (Itkonen 1951). Further, there are no striking differences between the two sample populations in terms of the basic demographic characteristics that are available to us. The average time from marriage to first birth in our sample is 0.76 years (2.30 years when excluding mothers who gave birth before their first recorded marriage) compared with 1.34 years for HLJ's sample. The average time between the first and second child was 2.46 years for the women in our data and 2.58 for the women in HLJ's data (we thank Samuli Helle for generously providing this information on very short notice). Of course, this does not imply that there are no differences between our two sample populations. For example, our reading of the anthropological literature and other papers by HLJ's research group suggests that the Sami in HLJ's sample may have relied more heavily on fishing for their subsistence than the Sami in our sample (von Düben 1873; Itkonen 1951; Helle *et al.* 2008).

In our opinion, the most rigorous way to test the hypothesis that sons reduce maternal longevity relative to daughters would be to pool all existing estimates from the literature using a uniform model specification across samples. We would certainly be cooperative in such an effort and are convinced that if such an analysis were conducted, the associations found in HLJ would stand out as an outlier. On a more general note, it seems to us that one of the most serious methodological challenges in contemporary human evolutionary biology, just like in molecular genetics, is how to deal with the challenges of multiple hypothesis testing. Researchers go

to great pains to assemble rich datasets and then use these data to test a large number of hypotheses, of which only a subset are typically reported. Since the large number of variables considered and the model selection procedure are rarely taken into account when *p*-values are computed, interpreting the results from such studies is very difficult. Moreover, reported coefficients will be affected by publication bias. We therefore agree with HLJ's call for more theory driven investigations as it would introduce more discipline and rigour into the way this type of research is conducted. We also believe that journal editors would be well advised to insist more often on replications in independent samples before they accept articles in this area. Otherwise, the risk is great that we will continue to regularly see sensational results being published, only to evaporate under subsequent replication attempts.

David A. Cesarini¹, Erik Lindqvist² and Björn Wallace^{3,*}

¹*Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 02142, USA*

²*Research Institute of Industrial Economics, Stockholm, Sweden*

³*Department of Economics, Stockholm School of Economics, Stockholm, Sweden*

*Author for correspondence (bjorn.wallace@hhs.se).

REFERENCES

- Carey, J. R. 2003 *Longevity: the biology and demography of life span*. Princeton, NJ: Princeton University Press.
- Cesarini, D., Lindqvist, E. & Wallace, B. 2007 Maternal longevity and the sex of offspring in pre-industrial Sweden. *Ann. Hum. Biol.* **34**, 535–546. (doi:10.1080/03014460701517215)
- Cesarini, D., Lindqvist, E. & Wallace, B. 2009 Is there an adverse effect of sons on maternal longevity? *Proc. R. Soc. B* **276**, 2081–2084. (doi:10.1098/rspb.2009.0051)
- Helle, S., Lummaa, V. & Jokela, J. 2002 Sons reduced maternal longevity in preindustrial humans. *Science* **296**, 1085. (doi:10.1126/science.1070106)
- Helle, S., Helama, S. & Jokela, J. 2008 Temperature-related birth sex-ratio bias in historical Sami: warm years bring more sons. *Biol. Lett.* **4**, 60–62. (doi:10.1098/rsbl.2007.0482)
- Helle, S., Lummaa, V. & Jokela, J. 2010 On the number of sons born and shorter lifespan in historical Sami mothers. *Proc. R. Soc. B* **277**, 2909–2911. (doi:10.1098/rspb.2009.2114)
- Hurt, L. S., Ronsmans, C. & Quigley, M. 2006 Does the number of sons born affect long-term mortality of parents? A cohort study of rural Bangladesh. *Proc. R. Soc. B* **273**, 149–155. (doi:10.1098/rspb.2005.3270)
- Itkonen, T. I. 1951 The Lapps of Finland. *S. J. Anthropol.* **7**, 32–68.
- Jasienska, G., Nenko, I. & Jasienski, M. 2006 Daughters increase longevity of fathers, but daughters and sons equally reduce longevity of mothers. *Am. J. Hum. Biol.* **18**, 422–425. (doi:10.1002/ajhb.20497)
- McArdle, P. F. *et al.* 2006 Does having children extend life span? A genealogical study of parity and longevity of the Amish. *J. Gerontol. A Biol. Sci. Med. Sci.* **61**, 190–195.
- Van de Putte, B., Mattijs, K. & Vlietinck, R. 2004 A social component in the negative effect of sons on maternal longevity in pre-industrial humans. *J. Biosoc. Sci.* **36**, 289–297. (doi:10.1017/S0021932003006266)
- von Düben, G. 1873 *Om Lappland och lapparne—Företädesvis de svenske*. Stockholm, Sweden: Nordstedts.