‘Natural Selection Across Three Generations of Americans’: response to reviewers

Dear Professor Medland,

Thank you for the opportunity to revise our manuscript. Thank you also to the reviewers. We respond to their points below.

One major change is that I have asked Tobias Edwards to be a co-author. I discovered that he was working on similar topics, so we merged our efforts. One effect is just that we have more polygenic scores to analyse. The original results are not much changed, which is encouraging. Another is that we look in more detail at two extra generations, using children and grandparents of the respondents. We’ve given the paper a title which reflects that. Lastly, we have more careful and accurate estimates of effect sizes of “true” and/or “chip-wide” scores.

I hope this is OK. I think it has appreciably improved the paper.

With best wishes,

David Hugh-Jones

Reviewer Comments:

# Reviewer 1

*Thanks for the opportunity to read the manuscript ‘Natural selection in the Health and Retirement Study’. This study is of great interest to the behavior genetics and population genetics community and beyond. It demonstrates ongoing natural selection in a descriptive manner and partly confirms ideas about the trade-off between childbearing and work. I have a few comments.*

*First, it would be good to state in the abstract the expectations about the tested interactions. Otherwise it reads as though interactions with low income etc have been tested arbitrarily.*

*Second, in the introduction, it is implied that people with larger human capital work more, but it remains unclear why. Only in the discussion section, it is elaborated on the idea of substitution effect, that larger human capital leads to larger income loss when work is reduced for alternative activities. Note that this phenomenon is also known as opportunity costs of larger human capital. This should be explained early and explicitly as in absolute terms, one might argue that larger human capital increases net income and therefore might lead to working less and having more children with the same income.*

*In a similar vein, it would be very welcome to reiterate reason why the expected effect is stronger amongst low income etc people.*

We now detail the theory more in the abstract, and we’ve described the substitution effects and income effects straight away in the intro. We’ve tried to keep the language as simple and intuitive as possible.

*I was wondering why the authors control for PCAs in the PGI analyses. As stated by the authors, this study is not a causal study but one of correlations. The consequence of a correlation (natural selection) should be independent of its causes.*

I think for two reasons. First, we are ultimately interested in selection on the “true PGI”. Stripping out PCs is an admittedly imperfect way to control for confounding population stratification. Second, PGIs are typically residualized in this way, so by examining the effect of these we’re following the “industry standard” and learning about scores in the form they are actually used.

To simplify interpretation, we’ve decided to residualise the PGIs on the principal components and then standardize them, instead of control for them.

*In general, the authors state that they are less interested in significance but in effect sizes. First, this is very generous. Whilst I find this acceptable, it would be good to read one or two more sentences on why the authors think this is legitimate in this case. Second, I believe such a descriptive approach might be useful studying the patterns of natural selection, however, when it comes to the interactions, the author state that effects ‘Appear larger for unmarried’, for example. It remains unclear to me when an effect is considered larger or not, which becomes important as the author differentiates between interactions which are real and those which are not. What are the criteria here.*

First, one could argue that the literal null hypothesis is not really plausible here. If *any* PGI is under selection in the population (clearly true), then some underlying variants must also be under selection. Since all PGIs are just weighted sums of underlying variants, it will then be a knife-edge case for any other PGI to have exactly zero correlation with fertility. What matters more is whether effect sizes are big, small or trivial.

More specifically, we’re interested in the pattern of selection effects across different PGIs, and how it correlates with the PGI’s relationship to income. So we want to include PGIs with small selection effects, which we predict to have a low absolute correlation to income. Given the sample size, these PGIs are very likely to have insignificant individual selection effects, but we shouldn’t ignore them in the cross-PGI analyses.

When we examine relationships across scores, we indeed test for significance, always by bootstrapping the sample.

*The authors state that the correlation between the PGI and fertility quantifies natural selection. However, to my knowledge it is the covariance following the Robert-Price Identity: Morrissey, M. B., L. E. B. Kruuk, and Alastair J. Wilson. “The danger of applying the breeder’s equation in observational studies of natural populations.” Journal of evolutionary biology 23.11 (2010): 2277-2288.*

We quantify natural selection with the regression coefficient of the PGI on relative lifetime reproductive success. At times in the original manuscript this was wrongly referred to as a correlation but thus has now been updated. The polygenic scores are standardized to have a standard deviation of 1. Under a univariate model, the regression coefficient is the covariance divided by the variance in the independent variable. The variance is standardized to 1, so the regression coefficient is equal to the covariance and thus the two are interchangeable. A brief discussion of the issue is now given in the Data section.

In the previous manuscript we used a multivariate model, controlling for genetic principal components. We now no longer control for the PCs. Now we residualise the polygenic scores on the PCs before standardizing the polygenic scores. This ensures our regression estimates are exactly the covariance between our polygenic scores and RLRS.

*The authors might want to discuss the expected selection in terms of relevance. They quantify it but they do not discuss these quantifications. This is very important shaping our expectation of ongoing natural selection. It might also be very interesting to know more about whether those trends follow the secular trends in phenotypes.*

We have now elaborated on the error corrected results and added a final paragraph to the Discussion section on this. We estimate how selection against intelligence will affect the phenotype, finding the result is not especially large. In particular we discuss PGIs for health phenotypes, as a corrective to the existing literature which is perhaps a bit over-focused on cognition and IQ.

We briefly mention that genetic change in traits is often opposite to the phenotypic change. This highlights that environmental change is likely to be of greater importance for phenotypic trends than genetic change.

*Finally, the statement ‘in the UK Biobank sample, correlations of PGS coefficients between younger and older parents are negative (-0.319), and for 42 out of 58 PGS, coefficients have opposite signs. However, in bootstraps, these statistics are not significantly different from 0 and 29 respectively.’ was not entirely clear to me. What does it mean that 42 PGS had opposite sign effects? Does that mean for old versus younger parents’ coefficients were very different, even in the opposite direction? How does this relate to the first part of the sentence that coefficients were negative. Also, why has this been tested against 0 and 29?*

The earlier claim is not “coefficients are negative overall” but “coefficients among younger parents are negatively correlated with coefficients among older parents”. We are dealing with correlations of correlations: this can get confusing.

Negatively correlated, and more strongly, opposite-signed, effects between young and old parents support the theory that income effects dominate for older parents, while substitution effects dominate for younger parents. To be a bit more explicit: if a PGS simply doesn’t correlate with human capital, we expect it to not be selected on among either group of parents. If a PGS is correlated with human capital, that will lead to income and substitution effects for those with more human capital. If the substitution effect dominates for younger parents then the PGS will have a negative correlation with fertility; if the income effect dominates for older parents, then the PGS will have a positive correlation with fertility.

The test against 29 had the following logic: if all effects were really zero, and estimates were pure noise with mean 0, then we’d expect opposite-signed and same-signed “effects” to be equally likely; so, 29 is half of the 58 scores.

However, thinking about this issue a bit more made us change our tests more generally. The original situation was that we tested most pairs of coefficients across sample subgroups by comparing their absolute size. Coefficients for older and younger parents had similar absolute sizes, but we noticed they were opposite-signed and hence ran the tests above. This is not very satisfactory: it’s too like special pleading, and it also suggests the original tests weren’t adequate.

We now improve our tests in table 1 as follows. We take the null hypothesis to be “selection coefficients are the same in the different subgroups, for all PGIs”. Under the null, the true difference between selection coefficients is zero; the difference in our sample (and in bootstrapped subsamples) should then not be significantly different from zero. Our alternative hypothesis is that, for at least some scores, selection coefficients against human capital are smaller among the richer/more educated/older/married group. That is, if the selection coefficient among the poorer and richer groups are , and is the sign of the PGI’s correlation with human capital, then

Under the null, we’d expect the above equation to hold in the sample for just 50% of the scores. If it holds for significantly more than 50% of the scores (using bootstrapped C.I.s) then we can reject the null in favour of the alternative.

We also try a simpler approach. Using the PGIs with the largest and most precisely estimated selection effects in the whole sample, we estimate selection effects in different subgroups and test for a significant difference between them via a simple interaction term. We report these results too.

# Reviewer 2

*The author sought to examine natural selection as indicated by polygenic scores in a contemporary representative sample of the U.S. and found that the results partially support the economic theory of fertility as a result of natural selection among both white and black individuals. The authors’ hypotheses are based primarily on the economic theory of fertility, which emphasizes a trade-off between the number of children and income: i) higher income makes raising children costlier, but ii) higher income makes having more children more affordable. As far as I understand the results of the paper, the results support prediction 1, but are more ambiguous for prediction 2?*

*However, in my opinion the theoretical database, data and analysis are adequate, the problem I have is that the MS is a bit confusing. It should be made clearer why an analysis was done and what the results mean in terms of rejecting or accepting the theoretical assumptions. The possible link to biological selection should also be explained in more detail, as this has been done in the case of education, income and reproduction to some extent on the basis of comparable data (Kong et al. 2017, Beauchamp 2016, Fieder & Huber 2002) and also in many other studies.*

We have added a sentence explicitly addressing our motivation in the introduction. We have also clarified the biological theory in the Data section. We provide an explanation of the intuition for our approach and how it is established in biology, through the Robertson-Price identity, to estimate the selection differential.

In terms of accepting or rejecting our theory, our paragraph in the discussion was:

The results here support prediction 1 but are more ambiguous for prediction 2. PGS coefficients on RLRS also appear larger for low-income groups and unmarried respondents. But there is little evidence for larger coefficients among people with lower education, or younger parents. This may be due to the low sample size. But in the UK, the between-group differences were large; differences that big would surely have been visible here. The theory can accommodate this non-result, if preferences for income are less curved in the US for whatever reason. But note that *any* theory with a negative relationship between education and fertility will give prediction 1. In this sense, results here are less supportive of the economic theory specifically.

We don’t think these results give a simple thumbs up or down to the theory, but we’re open to suggestions that would help clarify what we say.

*The division into a “black” and a “white” is necessary due to technical reasons, but further complicates the results/discussion. Presumably the results for the larger sample (I think in the case of the HRS this are the whites) should be presented in the main text, and the results for the smaller population moved into a supplement?*

Done. Indeed, the black sample is really too small to say much.

*Also, there seem to be many associations of PGS with RLRS (Figure 4), such as ADHD, diabetes, etc., which I think are interesting findings, even if of low significance, worth at least a brief discussion.*

We agree. We now discuss some health-related traits in the conclusion:

Whilst past research has focused on [cognitive] traits, we find selection differentials are often of similar magnitudes for health related traits. The selection differential for self-rated health is greater than those of education and cognitive performance. The most significant, positively selected trait was ADHD, which was also found in the UK (Hugh-Jones and Abdellaoui 2022). Future research should study the health and medical implications of natural selection, in addition to its social implications.

*What is the data base for the mean number of children of people born in the same year - is pooling appropriate, as the means may vary considerably according to ethnicity?*

We thought about this, and now agree that it is better to use mean number of children of whites and of blacks separately for each ethnic group:

‘RLRS is calculated separately by ethnicity. This is not ideal, because it treats the ethnicities as separate breeding populations. But the alternative of calculating pooled RLRS would effectively be estimating natural selection in the whole US population by treating whites as representative. We therefore focus on the white US population, with the caveat that results from this “one data point” may not replicate in other ethnicities or countries.’

Hugh-Jones, David, and Abdel Abdellaoui. 2022. “Human Capital Mediates Natural Selection in Contemporary Humans.” *Behavior Genetics* 52 (4-5): 205–34.