Ability bias in the returns to schooling: How large it is and why it matters

Summary

Petr Čala

1 Overview

It has been shown across decades of research that general intelligence is one of the most reliable predictors of person's long term success. Despite the clear consensus on the benefits of education itself, there still exists dispute on what role ability bias plays in this effect. The idea that ignoring a measure of cognitive performance from the Mincer equation (Mincer, 1974) could inflate the returns to schooling has been pioneered in the meta-study of Bowles et al. (2001), and since then, no satisfactory conclusion has been reached in the topic. How much does this ability bias matter? How large are the returns to ability?

I conducted an extensive analysis on this very topic as a part of my master thesis under the supervision of doc. PhDr. Zuzana Havránková, Ph.D. Despite the thesis itself not having been put together yet, I have already synthesized the analysis itself and managed to put together some intriguing results that I would like to present. When it comes to the research goal itself, I hypothesize that the strongest predictor of life success in psychology has an important role in the economic (wage-equation) literature as well. In doing so, I assemble a set of data that should thoroughly capture the decades of wage equation research with 1754 estimates across 115 different studies. Apart from the simple Google Scholar query approach, this data set consists in large part of studies obtained through snowballing, which should help me capture most of the prominent literature the topic has to offer. Furthermore, I offer a robustness check to this analysis in the form of an additional data set across 19 twin studies, which serves to narrow down the effect of education even further.

As far as the technical side of the analysis is concerned, I tackle first the issue of publication bias, which has largely been overlooked in the literature thus far. Furthermore, I explain the variable-level heterogeneity, construct a best-practice estimate, and observe the conomic significance of prominent variables. For methodology, I put together a mix of well-established methods spanning the FAT-PET tests, non-linear methods, Caliper tests, and so forth. To this I add brand new methods such as MAIVE (Irsova et al., 2022), Robust Bayesian Model Averaging (Bartos et al., 2021), and more. I also try to approach the analysis in a robust, replicable style, for which I put together several high-quality R scripts. As an example, one such script allows me to run the Endogenous Kink (Bom & Rachinger 2019) fully in R, for which the code does not exist yet to my awareness.

The results consistently predict the presence of strong publication bias across literature, regardless of the ability bias issue. Contrary to the established notion of returns to an additional year of schooling around 7-8%, the publication bias corrected estimates offer a slightly modest figure, suggesting returns of around 6-7%. When dividing the data into subsamples that take a varying approach to the issue of ability, we find no obvious discrepancy that could be attributed to the ability control.

2 Data Set

I started the data set assembly with a Google Scholar query that efficiently captures the different aspects of the wage equation literature. From the 574 hits, I managed to identify and collect data from 74 studies which sometimes tackled the ability bias issue either head on, other times ignored it completely, and even chose an approach in between these two. Despite the differences in motivation, the structure of these studies remained very consistent. This allowed me to collect these studies as a part of a single data set and single out the effect of ability bias by means of simple subsetting.

To make the data even more robust, I expanded the number of studies by means of literature snowballing. With this, I managed to capture the most prominent studies in the wage equation literature, and believe that little more could be done to improve the data set further. The main effect itself can clearly be defined in the Mincer equation as a percentage return to an additional year of schooling in the Mincer equation. I transformed returns to levels (such as wage increase after attaining a college degree) to this common metric and obtained a directly comparable list of estimates. In total, these ammounted to 1754 estimates across 115 studies, yielding more than 150,000 data points with over 40 variables.

As an additional robustness check, I intend to collect a completely separate data set from 19 twins studies that study the wage equation, in order to narrow down the role of ability even further. As of me writing this summary, I have yet to finalise this data set, but expect to finish the task within a matter of weeks. Some prominent characteristics of the data thus collected can be found in the files appended in the e-mail together with this summary.

3 Publication Bias

My search for publication bias consisted of four main parts. Firstly, I analyzed the data using OLS, Fixed-Effect, Random-Effect, Study and Precision weighted regressions. Afterwards I moved onto non-linear tests, specifically Weighted Average of Adequately Powered (Ioannidis et al., 2017), Top10 (Stanley et al., 2010), Stem-based method (Furukawa, 2019), Selection model (Andrews & Kasy, 2019), Hierarchical Bayes model (Allenby & Rossi, 2006) and the Endogenous Kink model (Bom & Rachinger, 2019). As mentioned in the Overview section, I customized the Endogenous kink method to allow me to run it purely using R, which makes for a direct and more robust comparison of different non-linear methods. I added two methods allowing for endogeneity in the form of an instrumental regression and the p-uniform* method (van Aert & van Assen, 2020). Additionally, I conducted a Caliper test (Gerber & Malhotra, 2008), checked for p-hacking using the p-hacking test proposed by Eliott et al. (2022), and lastly employed the new MAIVE estimator (Irsova et al., 2022). As with the data set summary information, you can find the results of these methods in the appended files.

4 Heterogeneity

I capture the characteristics of data using over 40 variables. These help exlain the estimate and data characteristics, spatial/structural varation, differences in the estimation method, and lastly publication characteristics of studies. The most notable among these all is a dummy variable that can be used to split the data into four categories - studies that directly add ability into their regressions, studies that do not add ability directly, but control for it, studies that do address the issue of ability in their text, but do not control for ability in any way, and lastly, studies that do not mention ability anywhere in the text. None of these categories contains a majority of the number of studies, and thus we can make good esimates of the within-study as well as within-ability-category differences.

When it comes to the importance of other variables, I decided to test these using the Bayesian Model Averaging. This yielded 21 variable groups (and nearly 30 individual variables) with PIP over 0.5, the individual effects of which, as well as their inclusion probabilities across models, can be found in the appended result files. As a robustness check, I add Frequentist model averaging, and construct a best-practice estimate using this specification.

Both standard error and years of schooling have a high Posterior Inclusion Probablity, and a positive sign too. For the former, the positive sign is to be expected from the suggested publication bias from the previous tests, while the latter corresponds to the almost exclusively positive estimates of returns to education. Only one ability variable displays a high PIP, uncontrolled ability, and with a positive sign too. This may partially confirm the suspicion that omission of ability from the Mincer regression may lead to an inflation in the estimate of returns to education.

5 Best-Practice Estimate

To conclude the work, I constructed a subjective best-practice with the results from the BMA and compared these with the best-practice of three other studies from my data set. Generally, the setup of the individual studies would suggest a slighltly higher returns to education coefficient than is the mean of the collected estimates. Lastly, I compute the economic significance of the most important variables identified during the model averaging.

6 Conclusion

I hope my work has helped bring a new insight into the topic of the effect of returns to education and the role that ability bias plays when estimating the Mincer equation. I hope these results and the work I put into obtaining them with the great help of my supervisors and the authors of these methods will serve to contribute to the quality of the conference, be it in any amount possible.