Adam Altmejd

EDUCATION & REPLICATION

STOCKHOLM SCHOOL OF ECONOMICS

Education & Replication

Essays on the Determinants of College Choice and the Predictability of Lab Replications

Adam Altmejd

Dissertation for the Degree of Doctor of Philosophy, Ph.D.,
in Economics
Stockholm School of Economics, 2018

Education & Replication

ISBN 978-91-7731-086-0 (printed)
ISBN 978-91-7731-087-7 (pdf)

This book was typeset by the author using LaTeX.

Front cover photo: © Hedda Selder

Back cover photo: ARCTISTIC/Photo: Nicklas Gustafsson

Printed by: BrandFactory, Gothenburg, 2018

Keywords: college choice, sibling spillovers, peer effects, returns to fields of study, returns to education, lab experiment replicability, replication prediction.

Foreword

This volume is the result of a research project carried out at the Department of Economics at the Stockholm School of Economics (SSE).

This volume is submitted as a doctoral thesis at SSE. In keeping with the policies of SSE, the author has been entirely free to conduct and present his research in the manner of his choosing as an expression of his own ideas.

SSE is grateful for the financial support provided by the Jan Wallander and Tom Hedelius Foundation which has made it possible to carry out the project.

Göran Lindqvist

Director of Research
Stockholm School of Economics

Tore Ellingsen

Professor and Head of the
Department of Economics
Stockholm School of Economics

Acknowledgements

I thank my supervisor Tore Ellingsen. He has guided me masterfully over these five years, and steered me clear of many pitfalls. I am so grateful for all his encouraging words and insightful comments.

I am thankful to Magnus Johannesson and Anna Dreber Almenberg, who were always supportive and invited me to collaborate on a fruitful and inspiring research endeavor, and to Colin Camerer for a very rewarding collaboration. I thank the Jan Wallander and Tom Hedelius Foundation for financial support.

Among fellow students, I am so happy to have Siri Isaksson, whose energy and commitment has always been a source of motivation, encouragement, and great enjoyment. Thank you also to Niklas A., Eskil F., Karl H., Emma H., Hannes M., Dany K., Elin M., Thomas S., Bengt S., and Erik Ö. for being such fantastic colleagues. I am also grateful to Aron Vallinder, who spurred my intellectual curiosity all throughout my youth. I fondly remember our promise in high school to never go into academia, and am glad that we both broke it.

The support I received from my family has been invaluable. I thank my parents, Aleksy and Frieda, who never doubted my ability, not even for a second, my sister Nina, who did not follow my education choices but has an intellectual clarity that is a great source of inspiration to me, and my uncle Josef, who knew I would get a doctorate long before I did.

I am fortunate enough to share my life with Hedda, to whom I cannot express enough gratitude. Thank you for all the engaging conversations, for making me rethink my often erroneous preconceptions, and most of all, for your unrelenting support and trust in me. I love you.

Finally, Dagny, thank you for sharing your last pug years with me, and for not snoring to loudly when I smuggled you into the office.

Stockholm, September 10, 2018
Adam Altmejd

Contents

Introduction

This doctoral thesis is a collection of three essays. Each paper is a self-contained and independent research article.

The thesis begins with an essay titled "Sibling Influence on College Choice", where I investigate how preferences for different kinds of college education are affected by the education experience of siblings. In the subsequent chapter, "Relative Returns to Swedish College Fields", I measure returns to different study orientations in Sweden. The third and last paper is called "Predicting Replication" and introduces a simple tool for predicting the outcome of laboratory experiment replications.

Below, I give short summaries of each paper. I encourage the interested reader to also explore the introductions of each paper. These provide slightly longer overviews, without too much technical detail.

* * *

In the first paper, I study how preferences for college are influenced by the education experience of siblings. It answers the question; If your sibling studies business at Stockholm School of Economics, are you more likely to apply there? The answer is yes, up to 50% as likely!

It is well known that children with educated parents are more likely to go to university. Sociologists often stress the importance of familial influence, citing the lack of role models as one important reason for the persistent inequality in schooling (Ceja, 2006; Kaczynski, 2011). But how exactly does the familial transmission of education preferences work? Is it just that people are from either college-going families or not, or do the experiences of family members matter for what one decides to study?

Finding the right college education is not easy, especially without enough information. College applicants have been shown to not know enough about either the costs or the benefits of different degrees to make good decisions (French and Oreopoulos, 2017). Could an older sibling help fill this information gap?

To study this question I collected a large data set of admissions to Swedish universities. By looking at applicants who have scores exactly at the cutoff, where

a lottery determines if they receive an offer to a program or not, I can isolate the effect of a sibling's admission. Around 20% of sibling pairs decide to study at the same university. But most chose to do so because they have correlated preferences. They would go there even if the other sibling was not admitted. I isolate the spillover effect and see that the likelihood that an individual applies to a specific school increases by about five percentage points (or 25%) if their sibling has already been admitted.

By studying different groups of siblings separately, I identify a number of interesting patterns. Siblings are five times more likely to follow each other into the same institution compared to the same field of study. The effect is stronger when both siblings are male but does not vary with the education level of the parents, or with the popularity of the program. I conclude that the patterns that I observe indicate that siblings follow each other mostly because it is convenient.

* * *

It has been said that all economist try to measure the returns to education at some point in their career. The second paper of the thesis is my attempt. It is a methodological replication of Kirkebøen et al. (2016), where I estimate the economic returns to different fields of study in Sweden.

Sweden has the lowest college premium in the developed world (OECD, 2017). At the age of 30, the difference in yearly earnings between individuals with a college education and those without is only $6,000. However, among university degree holders, the variability between fields of study can be almost three times as large.

It is complicated to properly separate the effect of education from the fact that the most productive students attend the best schools. In Sweden individuals with humanities degrees earn $22,000 on average, while those with degrees in the best paid fields have wages that are twice as high. Is this because the returns to humanities are low, or are humanities students just less productive? By studying applicants at the margin, I am able to control for this selection bias. I find large differences in returns between fields.

Medicine, Engineering and Business often yield premia of over $10,000 per year compared to other fields. Humanities, on the other hand, frequently has a large enough negative payoff that the applicant ends up earning less than those who do not go to university at all.

When compared to other countries, Swedish applicants are more likely to apply to fields where they can expect a negative payoff. With the lowest education premia in the developed world, even the poor returns to humanities are not that discouraging. It is possible that Swedish students simply care less about

future earnings when choosing what to study. Instead, they might prioritize the consumption value of college and study topics that are fun or interesting, rather than tailored to the demands of the labor market.

* * *

When a study is replicated, an independent researcher recreates it to evaluate the original findings. Replications are a cornerstone of scientific progress. "Non-reproducible single occurrences are of no significance to science" as Popper famously wrote (Popper, 2005, 16.22).

But replications are also time-consuming, expensive, and not very rewarding. In a world with scarce scientific resources we should think carefully about which studies to replicate and, like we do in the last chapter of this thesis, improve institutions to curtail the costs.

Together with a group of co-authors, I develop a tool to help researchers predict if laboratory experiments replicate in the third essay. When the tool is provided with data from an experimental study it gives the user a probability that a replication will be successful. The model is surprisingly accurate, it produces forecasts that are as accurate as the predictions of experts. With the help of the algorithm, a researcher can make better informed decisions about which papers to replicate first. Famous papers with low predicted chance of replication seem like a good start.

In the future, journals may manage replications and a model like the one presented here could be implemented in the peer review process. Before publication, each paper would stand a small chance of being subject to a replication. To make the process more cost-effective, the likelihood that a specific paper is chosen for replication could be made dependent on the probability predicted by a model like ours.

* * *

The three essays in this thesis share one important theme. While the last one tackles the question of replications head on, the other two apply reproducible research methods to empirical investigations.

In recent years, social science has witnessed a reproducibility crisis (Open Science Collaboration, 2015). When researchers attempted to replicate famous findings, many turned out to be false positives — random patterns in the data that had been misinterpreted as systematic. In the wake of the crisis, psychologists greatly improved their research practices to minimize the risk of more false discoveries in the future (Simmons et al., 2011). The requirement to register analysis

plans before experiments are conducted has been one effective way to reduce the "researcher degrees of freedom" in analysis that could lead to p-hacking[1].

How can we be sure that empirical studies to not face the same problems? Compared to experimental researchers, empiricists have considerably more leeway in their design and analysis choices. Among a million regressions, many will turn out to be interesting and statistically significant. The obvious safeguard against data mining is to use pre analysis plans. However, without the control over data generation that experiments provide, the need to commit to a certain analysis before exploring the data has been viewed as too restrictive. To evaluate this claim, I registered analysis plans for the two first essays in this thesis (Altmejd, 2017).

An empirical pre analysis plan will never accurately predict all contingencies. But the plan does not need to be perfect to be of use. Indeed, I was forced to deviate from my plan on certain occasions. But it also significantly reduced the degrees of freedom at my disposal, ensuring that the choices that I made were not influenced by the structure of the data.

The two first papers of this thesis constitute "methodological experiments", evaluating if pre analysis plans are at all usable for empirical research. My conclusion is in the affirmative. The benefits from committing to a route through what Gelman and Loken (2013) call the "garden of forking paths", far outweighs the costs.

[1] P-values express the probability to identify an effect when there in fact is none. If one runs a large number of tests, some will seem significant by pure chance. When not all tests are reported, p-values are therefore meaningless.

References

Altmejd, A. (2017). *Pre Analysis Plan: Effects of Education Choice*. URL: https://osf.io/rj6t7/.

Ceja, M. (2006). "Understanding the Role of Parents and Siblings as Information Sources in the College Choice Process of Chicana Students." In: *Journal of College Student Development* 47.1, pp. 87–104.

French, R. and P. Oreopoulos (2017). "Behavioral Barriers Transitioning to College." In: *Labour Economics*.

Gelman, A. and E. Loken (2013). "The Garden of Forking Paths: Why Multiple Comparisons Can Be a Problem, Even When There Is No 'Fishing Expedition' or 'p-Hacking' and the Research Hypothesis Was Posited Ahead of Time."

Kaczynski, K. M. (2011). *Exploring the Influence of Siblings and Their Relationships on the College Choice Process*. PhD Thesis. The George Washington University.

Kirkebøen, L. J., E. Leuven, and M. Mogstad (2016). "Field of Study, Earnings, and Self-Selection." In: *The Quarterly Journal of Economics* 131.3, pp. 1057–1111.

OECD (2017). *Education at a Glance 2017*.

Open Science Collaboration (2015). "Estimating the Reproducibility of Psychological Science." In: *Science* 349.6251.

Popper, K. (2005). *The Logic of Scientific Discovery*. Routledge.

Simmons, J. P., L. D. Nelson, and U. Simonsohn (2011). "False-Positive Psychology Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." In: *Psychological Science* 22.11, pp. 1359–1366.

EDUCATION & REPLICATION

ESSAYS ON THE DETERMINANTS OF COLLEGE CHOICE AND THE PREDICTABILITY OF LAB EXPERIMENTS

This Ph.D. thesis is a collection of three research articles.

In "Sibling Influence on College Choice" the author studies how preferences for unviversity schooling are affected by the education experience of siblings. An individual is considerably more likely to apply to a specific program if their sibling studies it. The effect is driven by siblings having a preference for going to the same institution. It is twice as strong when both siblings are male, but does not change with neither parental education nor the popularity of the program. A possible explanation is that siblings follow each other out of convenience.

"Relative Returns to Swedish College Fields" is a paper about the economic returns to different college orientations. Even though Sweden has the lowest average college premium in the developed world, differences between fields can be large. Medicine and engineering have returns of over $10,000 per year compared to other fields. Humanities, on the other hand, has a large enough negative payoff that the degree holder likely earns less than those who do not go to university at all. Interestingly, applicants do not seem to care much about low returns when chosing what to study, but instead prioritize non-pecuniary benefits of college education.

"Predicting Replication" covers a completely different topic. In the paper, the author team design a simple machine learning algorithmic that predicts the outcomes of laboratory experiment replications in Psychology and Economics. The model is very accurate, on par with the forecasts of experts. It could be used to make better decisions about which studies to replicate, and thus increase cost-effectiveness of replication efforts.

Adam Altmejd holds a B.Sc. in Business an M.Sc. in Economics from Stockholm School of Economics. His main research interests are Behavioral Economics, Empirical Microeconomics and Household Finance.

ISBN 978-91-7731-086-0

DOCTORAL DISSERTATION IN ECONOMICS
STOCKHOLM SCHOOL OF ECONOMICS, SWEDEN 2018

STOCKHOLM SCHOOL OF ECONOMICS