DISSERTATION SERIES 2022:5

Essays on the economics of education and health

Iman Dadgar

IFAU INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

For dissertations, quality is ensured in the traditional way, through the academic review procedure at Uppsala /Stockholm University. This dissertation has not been reviewed by IFAU.

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Friday 13 May 2022 at 13.00 in Nordenskiöldsalen,Geovetenskapens hus, Svante Arrhenius väg 12.

Swedish Institute for Social Research 107

Stockholm University

Essays on the economics of education and health

Iman Dadgar

Doctoral Thesis in Economics at Stockholm University, Sweden 2022

Essays on the economics of education and health

Iman Dadgar

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Friday 13 May 2022 at 13.00 in Nordenskiöldsalen, Geovetenskapens hus, Svante Arrhenius väg 12.

Abstract

Study I: This paper investigates the effect of the academic ordinal rank position of Swedish grade 9 students relative to their school peers on future educational achievement and adult earnings. The results show evidence of a positive impact of being more highly ranked in the class, and the effects are concentrated to the top and the bottom of the ordinal rank distribution. High-ability students from low-income families gained the most from having a higher ordinal rank in grade 9. The results contrast with US findings, which suggest a similar impact across the rank distribution.

Study II: This paper studies the effect of a reform that increased school-level autonomy in determining how to allocate time between different subjects in Sweden. It evaluates the impact of the reform using registry data in a Difference-in-Differences framework. The results suggest that students' educational outcomes, including the subsequent choice of educational track, were not affected by the reform. However, there are some indications that students in large schools and students from low socioeconomic households may have benefited from the reform.

Study III: Research suggests that increases in gross domestic product (GDP) lead to increases in traffic deaths plausibly due to the increased road traffic induced by an expanding economy. However, there also seems to exist a long-term effect of economic growth that is manifested in improved traffic safety and reduced rates of traffic deaths. Previous studies focus on either the short-term, procyclical effect, or the long-term, protective effect. The aim of the present study is to estimate the short-term and long-term effects jointly in order to assess the net impact of GDP on traffic mortality. We performed error correction modelling to estimate the short-term and long-term effects of GDP on the traffic death rates. The estimates from the error correction modelling for the entire study period suggested that a one-unit increase (US$1000) in GDP/capita yields an instantaneous short-term increase in the traffic death rate by 0.58 ($p<0.001$), and a long-term decrease equal to −1.59 ($p<0.001$). However, period-specific analyses revealed a structural break implying that the procyclical effect outweighs the protective effect in the period prior to 1976, whereas the reverse is true for the period 1976–2011.

Study IV: Unemployment might affect several risk factors of cardiovascular disease (CVD), which is the leading cause of death globally. The characterization of the relation between these two phenomena is thus of great significance from a public-health perspective. The main aim of this study was to estimate the association between the unemployment rate and mortality from CVD and from coronary heart disease (CHD). We used time-series data for 32 countries spanning the period 1960–2015. We applied two alternative modelling strategies: (a) error correction modelling, provided that the data were co-integrated; and (b) first-difference modelling in the absence of co-integration. Separate models were estimated for each of five welfare state regimes with different levels of unemployment protection. We also performed country-specific ARIMA-analyses. Because the data did not prove to be co-integrated, we applied first-difference modelling. Our findings, based on data from predominantly affluent countries, suggest that heart-disease mortality does not respond to economic fluctuations.

Keywords: *education, ordinal rank, decentralization, timetable, accident mortality, heart-disease mortality, unemployment, GDP.*

Stockholm 2022
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-203003

ISBN 978-91-7911-830-3
ISBN 978-91-7911-831-0
ISSN 0283-8222

Stockholm University

Department of Economics

Stockholm University, 106 91 Stockholm

ESSAYS ON THE ECONOMICS OF EDUCATION AND HEALTH

Iman Dadgar

Stockholm University

Essays on the economics of education and health

Iman Dadgar

ISBN print 978-91-7911-830-3
ISBN PDF 978-91-7911-831-0
ISSN 0283-8222

Printed in Sweden by Universitetsservice US-AB, Stockholm 2022

To my parents, Hossein
and Shahin.

Contents

Acknowledgment 2

Abstract 5

Svensk sammanfattning 7

Introduction 9
- School autonomy, student grade rank, and educational achievement 10
  - The effect of ordinal rank in school on educational achievement and income in Sweden 10
  - School autonomy and subject-specific timetables 11
- Economic fluctuation and different causes of death 11
  - Short-term and long-term effects of GDP on traffic deaths 11
  - Is there a link between cardiovascular mortality and economic fluctuations? 12
- References 13

Acknowledgment

Starting and completing Ph.D. programs can either result in success (=1) or fail (=0). The question is: which factor contributes to making a Ph.D. journey a success? For me, the probability to start and succussed was about zero unless many variables – exogenously - pushed it to 1. I was born in a middle-class family; I was in the middle part of the ability distribution in the middle of the class and always got the middle grades. Therefore, I am fortunate that I was randomly assigned to people who helped me finish this journey.

First and foremost, a warm thanks to Karin Edmark, my main supervisor, for her support and advice. Karin, I have learned a lot from you, especially from how you address critical comments on papers and your clever solutions. Our continuous research discussions have truly been a source of comfort during my Ph.D. studies. You always had planned ahead and knew the next steps. Matthew Lindquist, my co-advisor, you have a visionary view on research, which I hope has transmitted to me. Thank you, Matthew, for being a source of inspiration and encouragement. You have always helped me to envision my long-term research paths. Thanks to both of my supervisors, I could not wish for better ones.

I started at SOFI by being a research assistant to Thor Nordström, I was fortunate to co-author three papers with him. Totto, thank you for your trust, endless support, and great dinners with Thelmo! It has always been my pleasure to work for/with you.

I also want to thank my master thesis supervisor, Hans Grönqvist, for his encouragement, which led me to SOFI. Anders Stenberg, thank you for patiently reading my papers and providing constructive and detailed feedback for last year. I am also grateful to Jonas Vlachos, the discussant in my final seminar, for providing helpful comments.

I feel privileged to have been in a dynamic and constructive research environment like the one at SOFI. Malin, I am glad that we were fellow Ph.D. colleagues – our chats about research, life, and kids have been a source of comfort during these years! Christine, thanks for your openness and idea exchange. Roza, thanks for being generous with your time and advice. Orsa and Marie-Pascal, thank you for helping and encouraging me to apply for various jobs-last year. Jose, you have brought the Spanish warmth to SOFI.

SOFI would not have been the same without the guidance and support of its administrative staff, which have helped (and shielded) me from various administrative matters. Maria Mårtensson, thanks for all the letters to Migrationverket. Tara Nabavi, for your structure, Anne Jenson, for managing courses and study plans. Katarina Hagelin, Irma Muñoz, Elma Sose, Julio Lundborg, and Daniel Rossetti, thank you for your help and patience!

I would like to express my gratitude to Yassaman Rostamian (@yassamanrostamian) for the great illustration of my thesis cover.

I appreciate the funding from Handelsbanken and IFAU. The Handelsbanken scholarship allowed me to go to Oxford University as an exchange. This was a unique chance for me to visit a great institution and get inspired by Oxford's leading researchers' interaction. I also would like to thank Prof. Brian Nolan for hosting me during my time at Oxford.

When I came to Sweden ten years ago, I only knew one person, but now Sweden is my second home, and I am surrounded by lovely friends.

Roujman, it is impossible to thank you in a couple of sentences. I am lucky that our paths crossed, and I got to know you. You have supported me during all these years, from the first presentation I had at the master's program to the last steps of finalizing my dissertation. We have spent lots of time together: laughing, improving each other's ideas, and discussing everything from how "good" research should be conducted to which is the tastiest candy bar on campus! I cannot imagine finishing this journey without you. Thank you for being my closest friend.

Amir and Fatemeh, we have known each other from the beginning of my journey in Sweden; your kindness is endless, and the world would have been a better place if there were more people like you. Mahsa & MR. Asadi and the crew, for the love, support, and "Bahman koochik"! Amin & Behnaz being with you make winter nights brighter and more pleasant; thank you for being such amazing friends and for "Eshgh o Haal"!Omid, I have learned so much from you, and I am proud to have you as "Amo." Hoda and Adonis, we are going to have more excitement next year!

Justine and Nico, thank you for being true friends through the years, for all the great times we spent together, and for many more! Amin & Ghazal, thank you for all your kindness and for always being great friends. Mehdi, I have known you since high school, and you have not been changed a bit all these years! Thanks for your "Ma'refat"!

My old-economist friend, Javad, for being a true companion since our BSc, you were always a great source of will and motivation. I am looking forward to starting working on various projects together. Mohammad, thanks for all the musical exchanges and discussions; I hope we see each other more in the near future. All my old friends, Behnam, Inoor, Ali M, Ali H, Reza, Sajedeh, Parisima, Ehsan, for the great time we spent together.

My mother in-low, Farideh, thank you for always being kind and caring! Alireza, for all the memories and the times we spent together.

My dear "Khanom Darabi," my mom, you were the symbol of a real teacher in our family; we always had discussions about the education system, teachers' rights, and improving students' grades. I miss you a lot. I hope you were here. Ali, thanks for being a perfect "big brother" to me; you were always available

when I needed you. You are a great source of energy in our family. Laya, Artin, Liane and Ali, I missed you lot and I hope see you more.

Pouneh, you are not just a sister but my oldest friend. We shared a lot of critical moments in life together. You are the source of inspiration and a role model to me! Baba Janam, thanks for your unconditional love and support. You are the kindest person I know; I wish I could be more like you.

Nava, since you came into my life, I deeply understood that nothing matters more than your smiles. You are a source of joy, fun, and love. Thanks for making our life colorful and beautiful.

Idayi, how can I even thank you? In every single step, from the first day we met at SBU in Iran until now that I am graduating from SU, you were always supportive, kind, patient, and most importantly, you always believed in me. To me, you are the definition of love. To you, Ida;

مصلحت نیست که از پرده برون افتد راز ورنه در مجلس رندان خبری نیست که نیست

(حافظ)

Iman Dadgar
Stockholm, April 2022

Abstract

Study I: This paper investigates the effect of the academic ordinal rank position of Swedish grade 9 students relative to their school peers on future educational achievement and adult earnings. The results show evidence of a positive impact of being more highly ranked in the class, and the effects are concentrated to the top and the bottom of the ordinal rank distribution. High-ability students from low-income families gained the most from having a higher ordinal rank in grade 9. The results contrast with US findings, which suggest a similar impact across the rank distribution.

Study II: This paper studies the effect of a reform that increased school-level autonomy in determining how to allocate time between different subjects in Sweden. It evaluates the impact of the reform using registry data in a difference-in-differences framework. The results suggest that students' educational outcomes, including the subsequent choice of educational track, were not affected by the reform. However, there are some indications that students in large schools and students from low socioeconomic households may have benefited from the reform.

Study III: The aim of the present study is to estimate the short-term and long-term effects jointly in order to assess the net impact of GDP on traffic mortality. We performed error correction modeling to estimate the short-term and long-term effects of GDP on traffic death rates. We used time-series data for 18 countries spanning the period 1960–2011. The estimates suggested that a one-unit increase (US$1000) in GDP/capita yields an instantaneous short-term increase in the traffic death rate by 0.58 and a long-term decrease equal to −1.59. However, period-specific analyses revealed a structural break implying that the procyclical effect outweighs the protective effect in the period prior to 1976, whereas the reverse is true for the period 1976–2011.

Study IV: The main aim of this study was to estimate the association between the unemployment rate and mortality from cardiovascular disease (CVD) and from coronary heart disease (CHD). Additional aims were (a) to assess whether the associations are modified by the degree of unemployment protection, (b) to determine the impact of GDP on heart-disease mortality, and (c) to assess the impact of the Great Recession in this context. We used time-

series data for 32 countries spanning the period 1960–2015. The estimated effect of unemployment and GDP on CVD as well as CHD was statistically insignificant across age and sex groups and across the various welfare state regimes. An interaction term capturing the possible excess effect of unemployment during the Great Recession was also statistically insignificant.

Svensk sammanfattning

Studie 1
Studien undersöker om svenska elevers rangordning i förhållande till skolkamraterna, med avseende på deras studieprestationer, påverkar deras utbildning och inkomster senare i livet. Resultaten tyder på att det har en positiv effekt att vara högre rankad i klassen, och att effekterna är koncentrerade till toppen och botten av den fördelningen. Studenter med höga betyg från låginkomstfamiljer påverkas särskilt starkt av att vara högre upp i rangordningen. Resultaten kontrasterar mot tidigare studier från USA, där man funnit en likartad effekt i hela fördelningen.

Studie 2
Studien studerar effekten av en decentraliseringsreform som ökade svenska skolors inflytande över fördelningen av undervisningstid mellan skolämnen. Reformens effekter undersöks genom att applicera difference-in-difference-metoden på ett detaljerat registerdatamaterial. Resultaten tyder på att reformen inte hade någon effekt på att elevernas utbildningsresultat eller val av utbildningsval i genomsnitt. Det finns dock indikationer på att elever i stora skolor och elever från familjer med lägre socioekonomisk status kan ha gynnats av reformen.

Studie 3
Syftet med studien är att skatta de kortsiktiga och långsiktiga effekterna av BNP på trafikdödlighet. Vi använde tidsseriedata för 18 länder som täckte perioden 1960–2011. Resultaten av den ekonometriska analysen tyder på att en ökning av BNP/capita motsvarande 1000 USD ger en omedelbar kortsiktig ökning av trafikdödlighetsraten (per 100 000) med 0,58, och en långsiktig minskning lika med -1,59. Periodspecifika analyser visade dock på ett strukturellt skifte, som innebar att den procykliska effekten uppväger den skyddande effekten under perioden före 1976, medan det omvända gäller för perioden 1976–2011.

Studie **4**
Huvudsyftet med denna studie var att skatta sambandet mellan arbetslöshet och dödlighet i hjärt-kärlsjukdom och kranskärlssjukdom. Ytterligare syften var (a) att undersöka om sambandet varierar beroende på utformningen av arbetslöshetsskyddet, (b) att skatta BNP:s inverkan på dödligheten i hjärt-kärl-

sjukdomar och kranskärlssjukdom, och (c) att skatta effekten av finanskrisen 2008 i detta sammanhang. Vi använde tidsseriedata för 32 länder som täcker perioden 1960–2015. Den skattade effekten av arbetslöshet och BNP på dödlighet i hjärt-kärlsjukdom och kranskärlssjukdom var statistiskt insignifikant; detta gällde för olika ålders- och könsgrupper liksom för olika system för arbetslöshetsskydd. En interaktionsterm avsedd att skatta effekten av arbetslöshet under lågkonjunkturen 2008 var också statistiskt insignifikant.

Introduction

This thesis uses insights and methods from Economics to study two aspects central to society's welfare: health and education. The first part of the thesis includes two single-authored papers on the topic of the Economics of Education, while the second part of the thesis includes two co-authored papers (together with Thor Norström) on the relationship between macroeconomic fluctuations and mortality.

Two fundamental yardsticks of any welfare state are to what extent and at what level of quality it provides education and healthcare to its citizen. The share of GDP allocated to these areas indicates how high these two aspects are placed on the political agenda. In the year 2018, almost 11% of Gross Domestic Product (GDP) was spent on healthcare in Sweden and 7% on education. [1]

The main objective of an economist is to investigate how resources are allocated and what are the consequences of these allocations for individuals, groups, and society as a whole. The aim of this thesis is to contribute to such analyses in several aspects: The first study analyzes the impact that the ordinal rank of students in school, in terms of academic achievement, has on their long-term educational and labor market outcomes. Although ordinal rank has received some attention in previous literature (Jonsson and Mood 2008, Booij, Leuven et al. 2017), I contribute by applying a sophisticated empirical strategy (Murphy and Weinhardt 2020, Denning, Murphy et al. 2021) to Swedish register data. In the second study, I investigate the long-term effects of a decentralization policy, implemented in the year 2000, that gave Swedish schools more authority over the timetable. Surprisingly, very little research has been performed on this policy, and the studies that do exist are mostly of a qualitative nature (Nyroos, Rönnberg et al. 2004, Rönnberg 2007). The second study of this thesis aims to quantitatively evaluate if this decentralization policy has had any effect on pupils' long-term outcomes, such as their Grade Point Average (GPA), the probability of attending the STEM field, and years of education. Third, previous research on health and macroeconomic fluctuations has tended to either focus on the short-term (procyclical effect) or the long-term (protective effects) between GDP growth and traffic deaths. In the third study, we combine both of these effects in the same model in order to assess the net impact of GDP on traffic mortality. The

[1] Source: World Bank

fourth study continues this line of research by investigating how cardiovascular mortality responds to changes in unemployment.

School autonomy, student grade rank, and educational achievement

In modern society, students spend almost one-third of their lives in schools, and it is vital to have an efficient and equal education system. Hanushek (2020) listed four essential factors of the education production function: family attributes, school resources, teacher quality, and the quality of peers. This part of the thesis focused on two empirical questions about school management and peers' effects on students. The first paper discusses the impact of the ordinal educational performance rank of the students within schools on their long-term outcomes, and the second paper investigates how higher school autonomy affects students' educational achievement.

The effect of ordinal rank in school on educational achievement and income in Sweden

This paper investigates the effect of the academic ordinal rank position of Swedish grade nine students relative to their school peers on future educational achievement and adult earnings. The ordinal rank effect is a version of the peer effect, but instead of focusing on the average effect of peers, it focuses on students' positions in the ability distribution. A lower rank in the class could impact student self-confidence and social status, leading to lower efforts in the future. It is also possible that being surrounded by high-performance students (and having lower rank) increases students' motivation(Black, Devereux et al. 2013) and gives them better networks in the future. There are also effects of parents and teachers who can react to the low ordinal rank in different ways (Lavy, Paserman et al. 2012).

To empirically analyse the effect of ordinal rank on student outcomes, it is essential to control for all confounding factors that affect outcomes from other channels. The model isolates the rank effect by conditioning on student ability, school-cohort fixed effects, and school types.

The results show evidence of a positive impact of being more highly ranked in the class, and the effects are concentrated to the top and the bottom of the ordinal rank distribution. These results contrast with US findings, which suggest a similar impact across the rank distribution (Denning, Murphy et al. 2021). For example, increasing the rank from the 75th percentile to the top of the class increases the probability of attending the STEM field by almost 6%. The ordinal rank had an extreme effect on students with very low ability: going from the bottom to the second percentile increased the probability of finishing upper secondary school by almost 25%. The paper also finds heterogeneous effects of ordinal rank: high-ability students from low-income households and immigrants gained the most from having a higher ordinal rank in grade nine. The paper also investigated if those students are more prone to comparing themselves to students of the same gender. The girls' rank among girls strongly affects attending in the STEM track, years of education, and income rank for girls at the top of the class. In the left part of the distribution, the effect of boys' position among boys is more substantial among low achieving boys. Here, a boy's lower rank hurts income, finishing upper secondary school, and attending vocational track.

School autonomy and subject-specific timetables

This paper studies the effect of a reform that increased school-level autonomy in determining how to allocate time between different subjects in Sweden. The reform took place in 900 of Sweden's approximately 3000 primary and lower secondary schools in the early 2000s. This study evaluates the impact of the reform using registry data in a difference-in-differences framework. The results suggest that students' educational outcomes, including the subsequent choice of academic track, were not affected by the reform. However, there are some indications that students in large schools, and students from low socioeconomic households may have benefited from the reform.

Economic fluctuation and different causes of death

This part contains two papers. The first paper investigates the short-term and long-term effects of GDP growth on traffic accident mortality. The second paper studies the effect of unemployment on heart disease mortality.

Short-term and long-term effects of GDP on traffic deaths

Previous research suggests that increases in the GDP lead to increases in traffic deaths plausibly due to the increased road traffic induced by an expanding economy (Neumayer 2004, Ruhm 2015, He 2016). However, there also seems to exist a long-term effect of economic growth that is manifested in improved traffic safety and reduced rates of traffic deaths (Van Beeck, Borsboom et al. 2000, Yannis, Papadimitriou et al. 2014). Previous studies focus on either the short-term, procyclical effect, or the long-term, protective effect. The aim of the present study is to estimate the short-term and long-term effects jointly in order to assess the net impact of GDP on traffic mortality.

We extracted traffic death rates for the period 1960–2011 from the WHO Mortality Database for 18 OECD countries. Data on GDP/capita were obtained from the Maddison Project. We performed error correction modeling to estimate GDP's short-term and long-term effects on traffic death rates. The estimates from the error correction modeling for the entire study period suggested that a one-unit increase (US$1000) in GDP/capita yields an immediate short-term increase in the traffic death rate by 0.58 ($p<0.001$) and a long-term decrease equal to −1.59 ($p<0.001$). However, period-specific analyses revealed a structural break implying that the procyclical effect outweighs the protective effect in the period prior to 1976, whereas the reverse is true for the period 1976–2011.

Is there a link between cardiovascular mortality and economic fluctuations?

Unemployment might affect several risk factors of Cardiovascular Disease (CVD) (Catalano, Goldman-Mellor et al. 2011), which is the leading cause of death globally. The characterization of the relation between these two phenomena is thus of great significance from a public-health perspective. The main aim of this study was to estimate the association between the unemployment rate and mortality from CVD and from Coronary Heart Disease (CHD). Additional aims were (a) to assess whether the associations are modified by the degree of unemployment protection, (b) to determine the impact of GDP on heart-disease mortality, and (c) to assess the impact of the Great Recession in this context. We used time-series data for 32 countries spanning the period 1960–2015. We applied two alternative modeling strategies: (a) error correction modeling, provided that the data were co-integrated, and (b) first-difference modeling in the absence of co-integration. Separate models were estimated for each of the five welfare state regimes with different levels of unemployment protection. We also performed country-specific ARIMA-analyses. Because the data did not prove to be co-integrated, we applied first-difference modeling. The estimated effect of unemployment and GDP on CVD as well as CHD was statistically insignificant across age and sex groups and across the various welfare state regimes. An interaction term capturing the possible excess effect of unemployment during the Great Recession was also statistically insignificant. Our findings, based on data from predominantly affluent countries, suggest that heart-disease mortality does not respond to economic fluctuations.

References

Black, S. E., et al. (2013). "Under pressure? The effect of peers on outcomes of young adults." Journal of Labor Economics **31**(1): 119-153.

Booij, A. S., et al. (2017). "Ability peer effects in university: Evidence from a randomized experiment." The Review of Economic Studies **84**(2): 547-578.

Catalano, R., et al. (2011). "The health effects of economic decline." Annual review of public health **32**: 431-450.

Denning, J. T., et al. (2021). "Class rank and long-run outcomes." The Review of Economics and Statistics: 1-45.

Hanushek, E. A. (2020). Education production functions. The economics of education, Elsevier**:** 161-170.

He, M. M. (2016). "Driving through the Great Recession: Why does motor vehicle fatality decrease when the economy slows down?" Social science & medicine **155**: 1-11.

Jonsson, J. O. and C. Mood (2008). "Choice by contrast in Swedish schools: How peers' achievement affects educational choice." Social forces **87**(2): 741-765.

Lavy, V., et al. (2012). "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom." The Economic Journal **122**(559): 208-237.

Murphy, R. and F. Weinhardt (2020). "Top of the class: The importance of ordinal rank." The Review of Economic Studies **87**(6): 2777-2826.

Neumayer, E. (2004). "Recessions lower (some) mortality rates:: evidence from Germany." Social science & medicine **58**(6): 1037-1047.

Nyroos, M., et al. (2004). "A Matter of Timing: time use, freedom and influence in school from a pupil perspective." European Educational Research Journal **3**(4): 743-758.

Rönnberg, L. (2007). "The Swedish experiment with localised control of time schedules: Policy problem representations." Scandinavian Journal of Educational Research **51**(2): 119-139.

Ruhm, C. J. (2015). "Recessions, healthy no more?" Journal of health economics **42**: 17-28.

Van Beeck, E. F., et al. (2000). "Economic development and traffic accident mortality in the industrialized world, 1962–1990." International journal of epidemiology **29**(3): 503-509.

Yannis, G., et al. (2014). "Effect of GDP changes on road traffic fatalities." Safety science **63**: 42-49.

I

The Effect of Ordinal Rank in School on Educational Achievement and Income in Sweden*

*I would like to thank Karin Edmark, Matthew Lindquist, Jonas Vlachos, Anders Stenberg, Erik Lindqvist, Malin Tallås Ahlzén and Roujman Shahbazian for their comments and suggestions. I would also like to thank the seminar participants at the Swedish Institute for Social Research (SOFI) and seminar participants at Institute for Evaluation of Labor Market and Education Policy (IFAU) . Financial support from Handelsbankens forskningstiftelser and from the IFAU is gratefully acknowledged.

1 Introduction

This paper investigates an aspect of the peer effects literature, which until recently, has received relatively little attention. Specifically, this paper examines the importance of the ordinal rank of a student in terms of academic achievement. Using plausibly random variation in the ordinal rank within schools over time, I estimate the causal impact of a student's ordinal rank on their medium- and long-run educational and labor market outcomes. These include the likelihood of attending a STEM-related field in upper secondary school, finishing upper secondary school, years of education at age 33, and average income between the ages 30 and 33. The analysis uses registry data from several cohorts of the full population of Swedish 9$^{\text{th}}$ graders.

How might the ordinal rank affect students' outcomes today and in their future? One potential mechanism is that being one of the lowest-performing students in the classroom may decrease a child's self-confidence (Denning et al. 2021), social status, and mental health (Kiessling & Norris 2020). These negative side effects may lower students' performance today and induce them to lower their effort in the future (Denning et al. 2021). On the other hand, being surrounded by higher-performing students could improve a student's drive and motivation (Black et al. 2013, Booij et al. 2017, Carrell et al. 2018, Carrell et al. 2009). Furthermore, the ordinal ranks of students may impact how their teachers interact with them and the resources made available to them (Lavy, Paserman & Schlosser 2012). For example, being a lower-ranked student could result in more attention and help from the teacher. Alternatively, this could give the teacher a negative signal, which could lower the teacher's motivation and expectations concerning that student. Being among the top may induce teachers to provide students with more and better learning opportunities and materials.

The ordinal rank model is a version of the peer-effects mode (see e.g.Sund 2009, Lavy, Silva & Weinhardt 2012, Ammermueller & Pischke 2009, Hanushek et al. 2003), but it builds on distinct mechanisms compared to the average peer-effect model. Moreover, the omission of ordinal rank from the analyses of peer effects means that the peer effect analysis may be suffering from an omitted variable bias (Bertoni & Nisticò 2019). Ribas et al. (2020) analysis of peer effects supports this notion: being low ranked in the class harms students, but this effect is partly mitigated by a positive effect of the average peer ability. In other words, the impact of average peer ability and the student's ordinal rank work in opposite directions, potentially canceling each other if both are not considered.

The ordinal rank literature is closely related to the literature that investigates "The big fish little pond effect." The consensus within this literature is that students perform better by being the best in the class or worse by being worse in the class (Marsh 1987, Marsh et al. 2007). This paper contributes to a recent growing number of economics studies that estimate the impact of ordinal rank on student outcomes, such as grade point average (Denning et al. 2021, Cicala et al. 2018), finishing high school (Elsner & Isphording 2017), personality traits (Pagani et al. 2021), major-specific choices at university (Elsner et al. 2021), the likelihood of smoking, drinking alcohol, and having unprotected sex (Elsner & Isphording 2018), mental health (Kiessling & Norris 2020), the likelihood of committing a violent act in school (Comi et al. 2021) and labor market outcomes (Denning et al. 2021). The common denominator for these studies is that students' ordinal rank positively affects their outcomes.

The empirical strategy adopted in this paper builds on Denning et al. 2021. [1]. I condition on a student's percentile rank in the national ability distribution to isolate the effects of ordinal rank from the effects of student ability. This together with school-cohort fixed effect and additional school type control variables result in a specification where there is no remaining systematic correlation between the ordinal rank and observed pre-determined student characteristics. This finding is taken as evidence that the

[1]Denning et al. 2021 use registry data from Texas to analyze the effect of ordinal rank in grade 3 on individuals' outcomes in their mid-20s.

ordinal rank-effect that I estimate reflects the causal impact of ordinal rank and not the effect of other (potentially unobservable) student differences. [2]

This study contributes to the literature mentioned above in several ways. First, my access to Swedish registry data for full cohorts of students during the 1990s allows me to evaluate the impact of the ordinal rank on long term outcomes – i.e., when the students are in their 30s. This is an important benefit, as the data provide outcome measures that more accurately measure a student's long-term labor market outcomes, a measurement that previous studies have ignored. For example, I measure income by averaging over the observed annual earnings between age 30 and 33. As a comparison, Denning et al. 2021 measure income at age 24. Second, I study not only the linear but also the non-linear effect of ordinal rank on long-term outcomes. My results show that this is important as the estimated effect varies considerably across the distribution. In addition, I strengthen the empirical strategy by performing many non-linear balancing tests with respect to pre-determined student background characteristics. The non-linear balancing test is critical since the main effect appears in the non-linear model.[3] Third, the detailed registry data allow me to carry out interesting heterogeneity analyses with respect to gender and family background. This is an important aspect because the effects of the ordinal rank prove to vary substantially across groups of students.[4] Finally, this is, to the

[2]I studied the ordinal rank during the final year of compulsory schooling. I argue that this year is important for two reasons. First, compulsory school students spend a lot of time together in well-defined groups and therefore clearly understand their ordinal position within their group. Second, 9th grade is critical for Swedish students since they choose which upper secondary school they want to attend and what course of study they would like to follow. If ordinal rank affects these choices, then the ordinal rank will probably have consequences for long-term outcomes as well.

[3]To my knowledge, other studies do not perform any non-linear balance tests.

[4] Denning et al. 2021 do not provide detailed information on important parental characteristics such as income and educational information. They use proxies such as information about free lunch and parents' race/ethnicity, which only to a limited extend captures background characteristics and tap into parental resources.

best of my knowledge, the first study on the non-linear ordinal rank impact in Sweden. [5]

The results show that the ordinal rank positively impacts many (but not all) of the outcomes studied in the paper. Years of education completed by age 33 is positively affected, whereas the probability of graduating from upper secondary school is not. The probability of attending a STEM field in upper secondary school increased for students, but only for those in the top quartile of the ability distribution (where most STEM students are drawn from). Income and income rank between age 30 and 33 both increase relative to ordinal rank, but the effect size varies across the distribution of student ability. Only those in the tails of the ability distribution benefit from experiencing a higher ordinal rank. Interestingly, the gains by high ability students are even larger among those from low-income families. A heterogeneity analysis by gender shows that the ordinal rank has a beneficial effect for girls in the top of the class distribution. This effect is larger when the ordinal rank is measured only in the girls.

The results were subjected to several robustness checks. First, I found that the pattern of the results is the same in both small and large schools, although the magnitude is smaller in small schools. Second, mathematics and Swedish grade were added to the model, but this did not change the results.

The rest of the paper is organized as follows. In the next section, I present a short description of the data. Section 3 describes the empirical strategy in more detail. In section 4, I examine tests of the internal validity of my identification strategy. The main results are presented in section5. Results on heterogeneous effects are presented in section 6. Section 7 demonstrates the robustness of my findings, and section 8 discusses the conclusions.

[5] The study by Facchinello 2020, is the most closely related Swedish Study as it uses data from the Swedish Evaluation Through Follow-up Longitudinal Study (which includes a 10% sample of students). He found that although enrolling in a better class helps young students perform better on some tests, it lowers their desire to sign up for more challenging courses. He also found that this effect remains even in upper secondary schools and leads to a lower GPA.

2 Data

This section contains a short description of compulsory school in Sweden and describes the data variables used in the study. Table 1 shows the descriptive statistics of the variables. Appendix B provides more information about the detailed descriptive statistics and the data sources.

2.1 Compulsory school in Sweden

During the period under study (1990-1997), the Swedish education system was based on nine years of compulsory school, voluntary upper secondary school (three years), and voluntary university or college. Students entered the compulsory school the year they turned seven: elementary school (grade 1 to grade 6) and lower secondary school (grade 7 to 9). Almost all students finished compulsory school (Halldén 2008, Stanfors 2000) and continued to upper secondary education. During the last year of compulsory school, students chose among several vocational and academic upper secondary school tracks. Admission to these educational tracks is based on their 9th grade final grade point average (GPA).

Since 1992, students also could choose schools outside of their residential area (Edmark et al. 2014). This was implemented almost at the same time as the introduction of private schools.[6] Grading in compulsory school was done by teachers based on exams and other qualifications of the students during the school year.

2.2 Rank, ability and school types

As will be shown in section 3, there are two important variables in the analysis: the students' ordinal rank in their school by cohort and their ability relative to their nation-

[6]The school choice reform was implemented in 1992. Before the reform, students could only enroll in schools near where they live.

wide cohort. This section describes how these variables are measured. Both variables are based on the final grades at the end of compulsory school (grade 9). The final grade is the student's GPA and each subject was graded from 1 (lowest) to 5 (highest). At the time of the study, the grading system was norm based – i.e., the students were graded such that their cohort followed a normal distribution, centered around 3. For core subjects, standardized exams guided teachers how to assign grades, although no central enforcement mechanism was in place to ensure a normal distribution. The teachers in each topic were responsible for grading the students based on exams, attendance, and the overall performance of the students.

The student's rank within their school and cohort is calculated as $R_{isc} = \frac{n_{isc}-1}{N_{sc}-1}$, where n_{isc} is the GPA rank of the student within the school and cohort, and N_{sc} is the number of students in that school and cohort. Thus, the rank variable is a number between 0 and 1. I also create 20 dummy variables corresponding to the 20 ordinal ranks of the students by sorting them within the school-cohort from the lowest to the highest GPA and dividing them into 20 equal groups. The number 20 was chosen since it allows me to include all schools, even those with as few as 20 pupils. In this way, each dummy covers 5% of the students from the lowest GPA rank to the highest GPA rank in each school and cohort. I call this the "rank measure".

To calculate the student's ability in relation to the nationwide cohort, I used 9$^{\text{th}}$ grade GPA and ranked all students in the country cohort. The GPA is scaled from 1 to 5 with one decimal, summing to a 50-level ability rank from the lowest GPA to the highest GPA. I call this the "ability measure."

Interestingly, both ability and rank measurements are based on the same 9$^{\text{th}}$ grade GPA. Although the ability distribution nationwide is unique for all students in a cohort and approximately follows the normal distribution, the school rank is driven by the cohort differences in school grade distributions. In addition to GPA, in some cases, I added the Mathematics and Swedish grades separately to the regression model to capture student ability more precisely.

As will be explained in the methods section, the empirical model also included a set of dummy variables for school type. These school types were constructed based on the characteristics of the GPA distribution in the schools in terms of the variance and mean. More precisely, the mean and variances of GPA for each school cohort were calculated and sorted from the lowest to the highest in four and ten groups, respectively. Table B.2 in the appendix shows the descriptive statistics for each of these school types.

2.3 Outcome variables

I constructed four long-term outcomes all measured when the students were 33 years old. The first is the years of study, which is calculated as the total number of years in education based on the students' highest level of completed education. The years of study could potentially relate to the rank as top-ranked students might be encouraged to continue to higher levels of education, whereas low-ranked students might not consider higher levels of education. During the study period, almost all students finished compulsory education. On average, the students attended school for 13.2 years, and almost 25% of students had some post-secondary degree.

The second related outcome variable is the probability of finishing upper secondary school, which translates to at least 12 years of schooling. It is reasonable that the probability of finishing upper secondary schools is mostly relevant for students in the low ranks, whereas most students in the high ranks probably have a low risk for dropping out irrespective of marginal changes in the rank position. In our data, almost 86% of students completed upper secondary education at the age of 33.

The third outcome variable measures the field of study that the students were in enrolled in the first year of upper secondary school. This category is defined as two types of education: STEM, which is arguably the most challenging field academically, and vocational training. As seen in Table1, almost 22% were enrolled in a STEM-related track and 37% of the students were enrolled in a vocational track. Previous studies have shown that having a higher ordinal rank in school is associated with choosing a more challenging educational track (Denning et al. 2021, Murphy & Weinhardt 2020, Facchinello 2020), which motivated the study of the academic STEM field separately. A detailed description of these statistics is found in Table B.1 in the appendix.

The last outcome is the average income of the students at the age of 30–33.[7] The rank could potentially affect income in different ways. First, through the intermediate variable such as years of education and the field of education that is explored with other outcome variables. Second, through a direct impact on students' beliefs, motivations and attitudes. I use two different income variables; one is the average of the log of earnings from 30-33 and the other is the rank of the same variable. To calculate the income rank, income of the cohort is sorted from the lowest to the highest and divided into 100 equal-sized groups.

2.4 Control variables

I use individual, family, and demographic background variables to control for student background and study whether the rank impact is heterogeneous with respect to student gender, parental education, or income. All variables were measured when students were in grade 9, – i.e., when the rank variable was measured.

To measure parent income, I added the father and mother income to get the household's income in logarithmic form. The household income was sorted from the highest to the lowest in the four groups. The same procedure was implemented for parent education: parents' years of education were added and ranked from highest to the lowest in four levels. To capture the immigration status, a dummy variable was used to measure whether students were born outside of Sweden. I also controlled for the age of immi-

[7] It is vital to average over several years, since previous research has shown that income information is volatile.

gration since this is an important indicator for educational attainment and other relevant outcomes. The student's birth year and month were also used to capture the effect of age of starting school. Previous studies have shown that the background characteristics and educational outcomes vary systematically due to the month of birth.

3 Empirical strategy

This paper estimates the causal effect of students' ordinal ranks on their medium- and long-run educational and labor market outcomes. Ordinal rank is determined by a student's position in the distribution of final grades in their school for grade 9. I compared the outcomes of equivalent students who had the same innate academic abilities, but who obtained different ordinal ranks by chance. In this setting, "chance" is generated by the fact that students go to different schools have somewhat different student grade distributions. The key is that some of the variation in the distribution of grades across schools is purely random and generates random differences in ordinal ranks among otherwise equivalent students.

Figure 1 illustrates this variation. A student with median ability (= 25 on the x-axis) can have a rank as low as 7 or as high as 13. This figure shows that students with the same ability can be assigned different ordinal ranks depending on which school and cohort they are enrolled in. For this method to have sufficient power, many schools and cohorts are needed. Following Denning et al. 2021, I model outcome y of student i in school s and cohort c as follows:

$$y_{isc} = g(R_{isc}) + A(a_{isc}) + S_{sc} + \gamma\,\mathbf{X_{ist}} + \varepsilon_{isc}, \tag{1}$$

In Equation 1, $g(.)$ is a function of ordinal rank, R_{isc}, $A(.)$ is a function of student ability, a_{isc}, S_{sc} is a school-by-cohort fixed effect, and $\mathbf{X_{ist}}$ is a set of control variables that includes student characteristics. The ordinal rank function $g(.)$ can take two forms.

The first is a linear function of rank λR_{isc}. The linear rank variable is calculated as $R_{isc} = \frac{n_{isc}-1}{N_{sc}-1}$, where n_{isc} is the GPA rank of the student within his or her school and cohort, and N_{sc} is the number of students in that school and cohort. Thus, the linear rank variable is a number between 0 and 1. The second functional form for ordinal rank is a set of dummy variables 1–20. [8] Each dummy covers 5% of the students from the lowest GPA rank to the highest GPA rank in each school and cohort. Rank 10 is the reference category that will be dropped in the model. Thus, the non-linear function g(.) is $\sum_{r=1,r\neq 10}^{20} \beta_r I_r$, and the main estimates of interest are β_r. These estimates tell us how much the outcome changes when students belong to rank r as compared to rank 10.

To obtain an unbiased estimate of the ordinal rank, it is critical to control for student ability since there is a high correlation between rank and ability and since ability will likely have a direct impact on the outcomes of interested. As discussed in section 2, I use a student i's rank in the entire country by cohort distribution of GPA as a proxy for student i's ability, a_{isc}. This "national rank" is calculated based on the same measure used to determine the school-by-cohort rank, namely GPA at the end of grade 9. To calculate national rank, all students in each cohort are sorted from the lowest to the highest GPA. Then, national ranks between 1 and 50 are constructed such that each rank covers 2% of the student-by-cohort population. [9]

Using national rank as an ability measurement may, however, produce biased estimates as there is a potential correlation between a student's rank and the ability distribution in the student's school.[10] This correlation is illustrated in Figure2, which shows two schools with the same mean but one with high variance (school A) and the other

[8] The number 20 was chosen because it allows me to include all schools, even those with as few as 20 students.

[9] I chose 50 ranks because 50 rank covers all possible GPA positions, since ninth grade GPA ranges from 1.0 to 5.0 in the data and is only reported with one decimal place, i.e. 4.5 is a valid GPA, while 4.56 is not.

[10] This bias is discussed at length in Booij et al. (2017) and Denning et al. (2021).

with low variance (school B). In this example, students with a GPA of 4.0 (above the average mean) in high variance schools have lower ordinal ranks than students with the same GPA in low variance schools. In contrast, students with a GPA of 2.0 (below average, dash line) in high variance schools have lower ordinal ranks than students in low variance schools. This means that a student's ordinal rank is not only related to his or her own ability but also to the variance of grades in his or her school.

Thus, we risk mixing the causal effect of ordinal rank with other factors related to a school's grade distribution if we are not careful. For example, if some types of parents choose schools based on their child's ability and on the grade distributions observed at a particular set of schools, then school-level differences in grade distributions together with student sorting based on observable variables could lead to omitted variable bias in the main model. To solve this problem, the function A(.) in model 1 includes student ability and school type, where school type is defined by the mean grade level of the school and by the variance of grades in that school. I created a set of dummy variables that control for school grade distribution type interacted with the student's ability such that:

$$A(a_{isc}) = \sum_{d=1}^{D} \sum_{a=1}^{50} \beta_{ad} I_d I_{isc}. \tag{2}$$

In equation 2, I_{isc} is the dummy variable that equals one if students i belong to rank r (from 1–50) in the country among cohort c. School type, d, ranges from 1 to D. School types are defined based on the mean and the variance of the school's grade distribution. To define these types, the annual school grade means and variances were sorted from the lowest to highest, and five categories of schools were defined: 1) quartile of mean; 2) quartile of variance; 3) decile of the mean; 4) decile variance; and 5) quartile of mean and quartile of variance. [11] Based on these alternatives, a different number of dummies in the model were used: alternatives 1 and 2 – 200 dummies (4*50); alternative 3 and 4

[11] The descriptive statistics of school types are shown in table B.2.

– 500 dummies (50*10); and alternative 5 – 800 dummies (16 * 50). To choose between these alternatives, I ran a set of linear and non-linear balancing tests.

In model 1, S_{sc} is the school- and cohort-fixed effect, so the rank- and school-effect are assumed to have a separate additive effect (see Denning et al. 2021 page 8). Different control variables were also used in the model: gender, immigration status, parent education (four levels), and parent income (four levels), as previously described in the data section. Adding the model components together results in the following estimation equation:

$$y_{isc} = \sum_{r=1, r\neq 10}^{20} \beta_r I_r + \sum_{d=1}^{D} \sum_{a=1}^{50} \beta_{ad} I_d I_a + S_{sc} + \gamma \mathbf{X_{ist}} + \varepsilon_{isc}. \tag{3}$$

3.1 Identifying variation

As discussed in the previous section, different numbers of dummies can be used in the estimation equation to control for ability and school grade distribution type. One question is, after adding all of these dummies, is there still a sufficient amount of identifying variation left in the model to affect outcomes? Keep in mind that the identification strategy uses differences in the grading distributions across schools that is conditionally random. If we control for too much – so that the residual grade distributions across schools are all equal – then the ordinal ranks of students with the same ability will be identical and we will no longer be able to estimate the effect of ordinal rank. So we need to categorize schools so that grade distributions are ”sufficiently” similar, but not identical. One can frame this problem as a trade-off between omitted variable bias (when we do not control enough for the grade distribution types of schools) and a failure to capture the full effect of rank after controlling away too much of the variation in grade distributions across schools. One method for choosing the correct set of control dummies is to see how using different specifications affects a balance test where we estimate the correlations between observed student characteristics and ordinal rank. This is done in section 4.

Here, I investigate how much variation is left after controlling for school grade distributions and student ability. Figure 3 shows the relationship between student ability (student country rank) and student ordinal rank in four types of schools: low mean and low variance; low mean and high variance; high mean and low variance; and high mean and high variance. Ordinal ranks vary at each ability level in all four school types.

To check the size of the variation after controlling for all factors in the equation, we can see that the residual in the equation 4 varies across ability distribution.

$$R_{isc} = \sum_{d=1}^{16} \sum_{a=1}^{50} \beta_{ad} I_d I_a + S_{sc} + \gamma \, \mathbf{X_{ist}} + \varepsilon_{isc} \tag{4}$$

The standard deviation of the ε_{isc} at each point in the ability distribution indicates how rank varies across the ability distribution after controlling for other factors. Figure 4 shows the standard deviation of ordinal rank and the residual of ordinal rank plotted across the ability distribution. The standard deviation of the ordinal rank is higher on average and has an inverted U-shape. When the residual of regression 4 is used, the variation change is stable across the ability distribution (on average, 0.7 standard deviation). This is the variation used in the main model. That is, on average, a student can get 0.7 ordinal ranks higher or lower in any ability point. This variation corresponds to 3.5% changes in their position in the class.

3.2 Measuring ability in grade 9

Ordinal rank is based on the student's final GPA in compulsory school. This is measured in the spring of grade 9, when students are 15–16 years old. Previous studies have used rank measured at younger ages. Using the rank measured at older age has both advantages and disadvantages. As older students understand their relative position to other students in the school and understand the importance of rank, they understand that they will soon have to compete for slots in upper secondary tracks. This finding is in line with Elsner et al. 2021, which predicts that older students will place more importance on rank than younger students.[12]

However, ordinal rank measured in grade 9 may already incorporate the compound effects of a student's ordinal rank at earlier grades. That is, ordinal rank at grade 9 may underestimate the actual – or compound – rank effect.

3.3 Teacher grading

As we discussed in the data section, using a teacher-graded GPA as a measure of student ability can have both advantages and disadvantages. On the one hand, it could capture multi-dimensional measures of the human capital; on the other hand, it could increase measurement error in the ability measure. This section addresses the downside of teacher grading in two cases.

A first concern is that the grading generosity differs between teachers. First, suppose that "grading generosity" is at the school level and that it affects all students in a school and cohort similarly. In this situation, the composition of teachers in a school and year is captured by the inclusion of the fixed effects of school-by-cohort in the regression model, which alleviates this concern. However, if this generosity varies across classroom within the same cohort and school, it may give rise to a biased estimate of the ordinal rank effect. Students in different classes may face teachers with different grading behaviors. At the same time, this concern is likely mitigated by the fact that students in grade 9 tend to meet many teachers who teach specific subjects, which lowers the chance that one class of students meets a specific subset of teachers having their own grading standards. Later, as a robustness check, this issue is addressed by looking at the effects found in a sample of small schools. These schools have fewer teachers, and all students are most likely graded by the same teachers, which should mitigate this

[12]This is more relevant in the Swedish context because students are assigned their first grades in grade 9

problem.

Another concern is that the ability distribution in the school may itself influence what grade teachers assign. The grading system in place at the time was designed so that grades are relative to the standard normal distribution in the cohort nationwide. That is, in principle, the teachers should not let the school-level distribution influence their grade setting. In other words, it was in principle possible to give all students the highest grade if all students were considered top students in relation to the nationwide distribution. However, teachers may still have been affected by the ability distribution of the students in the school. The more likely scenario is that the ability of a student would result in slightly lower grades if the student were the worst in the class rather in the tier of the school-level cohort distribution. Similarly, being the very best in the class (again, conditional on ability) may contribute to being assigned slightly higher grades.

Two crucial facts should be noted in relation to this issue. First, the empirical model partially controls for the school-level (observed) ability distribution by including the A(.) function. Second, the case outlined above (i.e., the within school ability distribution contributes to a wider set of grades being assigned) will most likely result in a downward bias of the ordinal rank effect on later academic outcomes as being ranked low in the school is correlated with a lower GPA, given academic ability. The empirical analysis includes dummy variables for the nationwide rank of GPA, which means that the study is based on comparing students with equal GPA but different school level ranks. If low school rank is correlated with lower GPA (given actual ability), then the lower-ranked students will have "worse" GPA than they really have, so they will be grouped with "worse students" when compared with students with the same GPA but with higher school level rank. That is, a low ordinal rank's negative impact on later outcomes will be counteracted by the fact that the "comparison students" are worse in absolute ability. For top students, however, the grade will be an overstatement of absolute ability (if having a high rank in the school leads to getting higher grades), and this will mitigate the impact of the school-level rank when the student is compared with students with the same GPA but with lower school level rank and higher absolute ability.

4 Balance test

Before estimating the regression according to equation 3, this section provides an informal test of whether the controls added to the regression equation are sufficient to control for omitted variable bias. This is done by testing whether different pre-determined student characteristics are uncorrelated with the ordinal rank when different sets of controls are added to the model. If this is the case, I can reasonably assume that unobserved student characteristics are also likely to be uncorrelated with the ordinal rank conditional on the included controls. The model is, therefore, likely to produce the causal effect of ordinal rank. I do this in two formats: a linear balance check and a non-linear check.

4.1 Linear balance test

For the linear balance test, I run equation 3 and replace the outcome variable with different characteristics of the students. Following Denning et al. 2021, I use five alternative specifications. If the background characteristics are not related to the ordinal rank, it is reasonable that unobserved characteristics are also not associated with rank. Therefore, these regressions provide information on which specification can be used to produce estimates on the impact of the ordinal rank that are plausibly unaffected by unobserved variable bias.

The results presented in Table 2 show the estimates of the ordinal rank coefficient for the following dummy outcome variables: being male; having high- and low-educated parents; having high- and low-income parents; and immigrant status of students. Panel A shows the estimates when only the ability measurement is used as a control variable in the model without interaction with school types. Panel B shows the estimates when the quartiles of school mean GPA are interacted with ability measurement (four school types interacted with 50 ability types). Panel C allows for deciles of school mean GPA ($10 \times 50 = 500$ dummies), panel D quartiles of school variance in GPA (200 dummies), and panel E deciles of school variance in GPA (500 dummies). The last panel uses quartiles of GPA variance and the quartiles of GPA mean (800 dummies). This is the preferred specification that Denning et al. 2021 used in their model. The table also reveals that no coefficient is statistically significantly different from zero in the last two specifications – i.e., the student's background characteristics are not correlated with the linear rank.

4.2 Nonlinear balance check

As ordinal rank may have a non-linear impact, I wanted to test whether the pre-determined features are uncorrelated with the rank in a non-linear format. More specifically, in model 3, outcome variable were replaced with student characteristics. As was the case for Table 2, statistically insignificant coefficients provide support for the notion that the ordinal rank is conditional on the included controls (i.e., not correlated with student characteristics) and that the specification will yield estimates of the causal effect of the rank.

Figure 5 shows the estimates of each of the characteristics for three specifications.Three specifications were used: no controls (green); control for ability (orange); and control for ability and 16 school types (blue). The last two correspond to specifications (a) and (f) (Table 2). In each figure, the x-axes show the student's ordinal rank in school, where each rank is covered 5% of students, and the y-axes show the coefficient estimates of the corresponding rank. Note that the scale of figures is not the same since the dependent variable in each has different scales.

The figure illustrates that the correlation between ranks and characteristics is high when school type and ability are not controlled. When the model controls for the ability, most of the estimates move toward zero. In the last specification, when the 16 types of schools are interacted with the student's ability (800 dummies), the correlation is zero across student's ordinal rank in almost all cases. That is, using these controls in the model takes care of all observable characteristics of the students. I take this as suggesting that it likely also controls for all unobservable student characteristics and therefore will continue with this specification.

5 Results

This section presents estimates of the effects of ordinal rank on future educational and labor-market outcomes. Two types of estimators are considered, linear and non-linear. For the non-linear effect, where separate coefficients are estimated for each ventile of the school by cohort rank distribution, the results are presented visually because of the large number of coefficients.

5.1 Linear effect of ordinal rank

The estimates of ordinal rank are presented in the panels of Table 3. The estimates of school rank in each row show how the outcomes change when a student's rank increases by one rank in the school. Note that a one rank increase is equal to an increase of 5% the position of the students because the ordinal rank variable is defined as the ventiles of the school by cohort distribution. The effect size is calculated by dividing the estimates by the average of the outcome variable. Each column from (a) to (f) corresponds to specification (A) to (F) defined in the data section, where each specification number of dummies is used to capture ability and school types. Specification (f) is the preferred specification.

For most of the estimates in Table 3, the magnitude of the effect is small although the school rank significantly affects the outcomes. For example, the estimate for the income rank in the specification (f) is 0.24. When a student's position in the school randomly increases by 5%, the income rank rises by 0.24 (of 100). The effect size is less than half a percent. For other outcomes, the size of the effects is also very small.

In terms of the specification, the results show that the estimates in all cases are smaller when the model controls for school variances – column (d) and (e) – compared to when controlling for school mean – column (b) and (c) – or only controlling for ability – column (a). Our preferred specification is a specification (f) where both school variances and school means interact with the ability.

Table 3 reveals that the effect size is small when the ordinal rank is used in the linear format. As discussed in the empirical strategy section, this might be because the rank effect is not homogeneous across all ranks. If the effect is zero in some parts of the rank distribution, the overall estimate of the rank effect will decrease. This is addressed by estimating the impact of ordinal rank in a non-linear format.

5.2 Non-linear effect of ordinal rank

Figure 6 shows the main results of the non-linear effect of ordinal rank on different outcomes. The figure shows the result of equation 3; each dot is the estimated coefficient of each of the 20 ventiles and the corresponding 95% confidence intervals. Students in the 10^{th} ventile (approximately in the middle of the school by cohort distribution) are in a reference category; that is, the coefficient of the ith ventile shows the impact of being ranked in the ith ventile in the school by cohort distribution relative to being in 10^{th} ventile. Note that the scales for the figures are not the same since the scales for the outcome variables are not the same.

The field choices for upper secondary school are in panels (a) and (b). Panel (a) shows that the ordinal position of the students has a positive and significant effect on attending the STEM field only for students with a high rank. Moving from rank 15 (75 percentile) to the top of the class (100 percentile) increases the probability of attending the STEM field by almost 6% compared to rank 10. This is the most immediate effect of the rank since students choose their upper secondary field of study when they finish grade 9. Attending a vocational track has an inverted-U shaped relation with the ordinal rank impact: the likelihood that a student attends the vocational track increases with the rank when students have a low rank (below 25 percentile) and decreases when students have a high rank (above 75 percentile).

Next, we explored whether the ordinal rank affects the probability of finishing upper secondary school and the years of schooling. Panel (c) shows that ordinal rank influences how many years of education a student completes by the age of 33 for the low-rank interval (below 15 percentile) and the top rank interval (above 80 percentile). In contrast, for the interval in-between, the coefficients are statistically indistinguishable from zero. That is, students with very high (and low) ability and with a higher rank were more likely to continue their studies. For students in the middle of the rank distribution, the ordinal rank had no effect on the years of education compared to the reference category.

Panel (d) illustrates that ordinal rank had an extreme effect on students with very low ability: rank increased from one to two for the probability of finishing upper secondary school by almost 25% (from 0.04 to 0.15 in panel d, rank 1 to 2). For rank 3 through 11, ordinal rank had no effect; for rank 12 and above, ordinal rank had a very small effect.

Panel (e) and (f) show that ordinal rank positively affects average income between age 30 and 33. This holds both when the income log (e) and income rank (f) are used. The pattern is almost similar for both and suggests that the slope of the effect is high for students in low rank and lower for students in the higher rank. For example, the income rank of the students increased one rank when the students' ordinal rank randomly increased from one to two in the school, whereas if the students' rank in the school increased from 19 to 20, the income rank increased by 0.5. [13]

[13] part of the effect might be due to that getting higher income rank in the high end of the income

6 Heterogeneity effect

In this section, I analyze if the ordinal rank has varying effects on students with different backgrounds. The heterogeneity study is limited to the non-linear specification and is carried out for the following four background characteristics: gender, immigration status, parental income, and parental education. All results are presented in Figures A.1 - A.5. Figure A.1 shows that females react slightly more than males to an increase in the ordinal rank in terms of the outcomes for income and years of education. The other outcomes almost follow the same pattern.

Next, I investigated the hypothesis that students are more prone to comparing themselves to students of the same gender. In other words, does the ordinal rank measured among the girls in the class have a stronger impact on the outcomes of girls than the ordinal rank measured among the boys? To test this, model 3 was run for girls and boys separately, but the rank in their gender group was used as the explanatory variable rather than the ordinal rank in their school. Figure A.2 shows the estimates of girl's rank among girls (left panel) and estimates of boy's rank among boys (right panel) for different outcomes. Each dot marks the effect of the rank on the outcomes compared to the student in rank 10.

Figure A.2 shows that the girls' rank among girls has a strong effect on enrolling in the STEM track, years of education, and income rank for girls at the top of the class. The pattern of the estimates of rank among girls is like the pattern of the estimates in Figure A.1 (a girl's rank among all students). The only difference is that the effect on attending the vocational track is almost insignificant for girls. On the other hand, the effect of boy's rank among boys is stronger among low achieving boys, in the left part of the distribution. Here, a boy's lower rank has a negative effect on income, finishing upper secondary school, and attending vocational track. The pattern is almost the same as the effect when measured among all students. The probability of finishing upper

distribution is harder than getting a higher rank in low-income levels.

secondary school is one example: being in the lowest 5% of the class decreases the chance of finishing upper secondary school by almost 8%. This number is 3% for girls among girls. That is, it is more harmful for boys than girls to be in the low rank in the class.

Figure A.3 illustrates that ordinal rank has a higher effect on the future income of immigrant students when they are in the top ranks of the school. In contrast, for the native group, higher ordinal rank has a more positive effect in the lower parts of the rank distribution. As seen in panel (f), the ordinal rank does not affect the likelihood of immigrant students of attending a vocational track, whereas the effect of ordinal rank on attending a vocational track for native-born students is an inverse U shape.

Figure A.4 and A.5 show the differences between students with high and low parental income and education. For the income heterogeneity of parents, the differences can be seen in the student's income and the probability of finishing upper secondary education.). The impact of school rank on the income was strong, particularly among top-ranked students with low-income parents. An increase in the school rank from 18 to 20 (top 90% to 100%) increased the income rank from almost two to five, an effect size of almost 6% (three increases with an average 50). Looking at the effect on finishing upper secondary school, low-income students also benefited from higher rank if they were among the top 15% of the schools.

7 Robustness check

Whereas the empirical strategy section discusses the validity of the estimators, this section carries out a set of robustness checks that addresses each of these concerns.

7.1 School size

The most preferred setting is to have class-level data and capture the student's ordinal rank in the classroom. To investigate how this might affect our estimates, the models were run in large and small schools separately.[14] Analyzing the results for different school size also helps observe teacher grading. In a small school, the number of teachers in each topic is limited and, in most cases, students have the same teacher in all grades, so almost all students are graded by the same teachers.

On the other hand, there is an advantage to looking at large schools with respect to measurement error. In a large school, there are more students in each ordinal rank bin, and more students with the same ordinal rank have similar abilities. This can decrease the biases from ability measurement because it covers more students with the same ability. Figure A.6 shows the estimates of small and large schools. Almost in all outcomes, the estimates for large schools were larger than for a small school for students at top of the class (above 75%). Denning et al. (2021) found similar pattern.

7.2 Ability measurement

To measure the ability of the students, I used national student rank in the country. The Swedish, English, and Mathematics grades were added to the model to check whether the results are robust with other ability measures. The grade is added with four dummies that indicate the grade from 1 to 5. Since the Mathematics and English grades were offered in two levels, each level's dummies were used in the model. The results are shown in figure A.7 and illustrate that adding Mathematics and Swedish grade separately in the model leads to a slightly smaller effect in the top ranks. In other words, adding grades in Mathematics, Swedish, and English reduced the ordinal rank effect on the outcomes somewhat.

[14]The large schools define whether the school size (captured by the number of students) is more than the mean school size in each year.

8 Conclusion

This paper studies the causal impact of a students' ordinal rank in their school and cohort during grade 9 on their subsequent educational and labor market outcomes. The study builds on the methodology used in Denning et al. 2021 but adds a set of non-linear balancing tests that prove to be important and informative when choosing the proper specification of the model. The main finding of the study is that a higher ordinal rank has a positive impact on later life outcomes although only for those in the tails of the ability distribution. This stands in contrast to results from the U.S., where Denning et al. (2021) report rank effects across the entire ability distribution.

In Sweden, moving up five rank positions to the top of the class increases the probability of attending a STEM track in upper-secondary school by almost 6%. Ordinal rank also positively affects the number of years of education measured at age 33, but only for pupils below the 15th percentile or above the 80th percentile. A rank effect on finishing upper secondary school is only estimated for those at the very bottom of the rank distribution; moving up by one ordinal rank in this part of the distribution increased the probability of finishing upper-secondary school by almost 3%. The effect of ordinal rank on income between the ages 30 and 33 is modest but larger for high-ability students from low-income families. Overall, girls tend to gain more than boys from an increase in ordinal rank, especially high ability girls near the top of their class. The results also suggest that the beneficial effect of ordinal rank for girls at top of the class is mostly driven by the rank when measured among girls. In other words, the girls seem to compare themselves with the other girls rather than with the entire class. There is also some evidence that suggests that the ordinal rank is relatively more important for students from immigrant families and low-income families.

Compared to the U.S. literature, I found that the ordinal rank had smaller effects. This difference might be because this study measures the ordinal rank in grade nine[15]

[15] In the Swedish context at the study period, students got their grades when they were in grade nine,

rather than in grade 3 as in the U.S. study and is based on the teacher-graded GPA rather than on test scores. Both these aspects imply that the effects of estimated ordinal rank found in this paper are lower bounds for the true impact.

The policy implications of ordinal rank effects are unclear as the understanding of the underlying mechanisms is limited. Are rank effects due to how schools allocate resources, including teachers' time, attention, and goodwill? Does a higher rank boost confidence or act as a motivator for future effort? Are the effects of rank symmetric? We would like to harness the potentially beneficial effects of ordinal rank – especially for those from families from low socio-economic status – without harming others. At the very least, educators should be made more aware of the disadvantages associated with a low ordinal rank and work towards reducing the degree of the disadvantages associated with worse outcomes in adulthood.

and there was no other formal evaluation before this stage.

References

Ammermueller, A. & Pischke, J.-S. (2009), 'Peer effects in european primary schools: Evidence from the progress in international reading literacy study', *Journal of Labor Economics* **27**(3), 315–348.

Bertoni, M. & Nisticò, R. (2019), 'Ordinal rank and peer composition: Two sides of the same coin?'.

Black, S. E., Devereux, P. J. & Salvanes, K. G. (2013), 'Under pressure? the effect of peers on outcomes of young adults', *Journal of Labor Economics* **31**(1), 119–153.

Booij, A. S., Leuven, E. & Oosterbeek, H. (2017), 'Ability peer effects in university: Evidence from a randomized experiment', *The review of economic studies* **84**(2), 547–578.

Carrell, S. E., Fullerton, R. L. & West, J. E. (2009), 'Does your cohort matter? measuring peer effects in college achievement', *Journal of Labor Economics* **27**(3), 439–464.

Carrell, S. E., Hoekstra, M. & Kuka, E. (2018), 'The long-run effects of disruptive peers', *American Economic Review* **108**(11), 3377–3415.

Cicala, S., Fryer, R. G. & Spenkuch, J. L. (2018), 'Self-selection and comparative advantage in social interactions', *Journal of the European Economic Association* **16**(4), 983–1020.

Comi, S., Origo, F., Pagani, L. & Tonello, M. (2021), 'Last and furious: Relative position and school violence', *Journal of Economic Behavior & Organization* **188**, 736–756.

Denning, J. T., Murphy, R. & Weinhardt, F. (2021), 'Class Rank and Long-Run Outcomes', *The Review of Economics and Statistics* pp. 1–45.
URL: *https://doi.org/10.1162/rest_a_01125*

Edmark, K., Frölich, M. & Wondratschek, V. (2014), 'Sweden's school choice reform and equality of opportunity', *Labour Economics* **30**, 129–142.

Elsner, B. & Isphording, I. E. (2017), 'A big fish in a small pond: Ability rank and human capital investment', *Journal of Labor Economics* **35**(3), 787–828.

Elsner, B. & Isphording, I. E. (2018), 'Rank, sex, drugs, and crime', *Journal of Human Resources* **53**(2), 356–381.

Elsner, B., Isphording, I. E. & Zölitz, U. (2021), 'Achievement rank affects performance and major choices in college', *The Economic Journal* **131**(640), 3182–3206.

Facchinello, L. (2020), 'Peer effects in education: When beliefs matter', *Available at SSRN 2966549* .

Halldén, K. (2008), 'The swedish educational system and classifying education using the isced-97', *The International Standard Classification of Education (Isced-97): An evaluation of content and criterion validity in* **15**.

Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003), 'Does peer ability affect student achievement?', *Journal of applied econometrics* **18**(5), 527–544.

Kiessling, L. & Norris, J. (2020), 'The long-run effects of peers on mental health', *MPI Collective Goods Discussion Paper* (2020/12).

Lavy, V., Paserman, M. D. & Schlosser, A. (2012), 'Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom', *The Economic Journal* **122**(559), 208–237.

Lavy, V., Silva, O. & Weinhardt, F. (2012), 'The good, the bad, and the average: Evidence on ability peer effects in schools', *Journal of Labor Economics* **30**(2), 367–414.

Marsh, H. W. (1987), 'The big-fish-little-pond effect on academic self-concept.', *Journal of educational psychology* **79**(3), 280.

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J. & Köller, O. (2007), 'The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation', *American Educational Research Journal* **44**(3), 631–669.

Murphy, R. & Weinhardt, F. (2020), 'Top of the class: The importance of ordinal rank', *The Review of Economic Studies* **87**(6), 2777–2826.

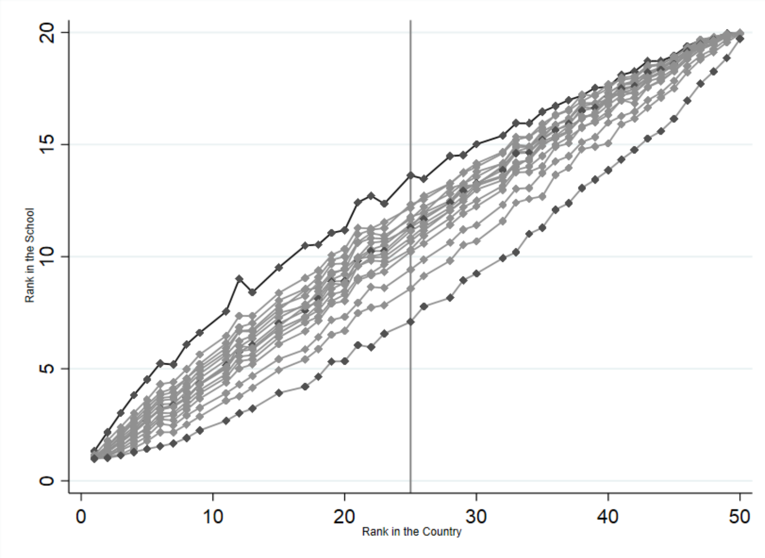
Pagani, L., Comi, S. & Origo, F. (2021), 'The effect of school rank on personality traits', *Journal of Human Resources* **56**(4), 1187–1225.

Ribas, R. P., Sampaio, B. & Trevisan, G. (2020), 'Short-and long-term effects of class assignment: Evidence from a flagship university in brazil', *Labour Economics* **64**, 101835.

Stanfors, M. (2000), 'Säkert och sakta. en historisk översikt över kvinnor i naturvetenskaplig och teknisk utbildning.', *Högskoleverkets rapportserie* .

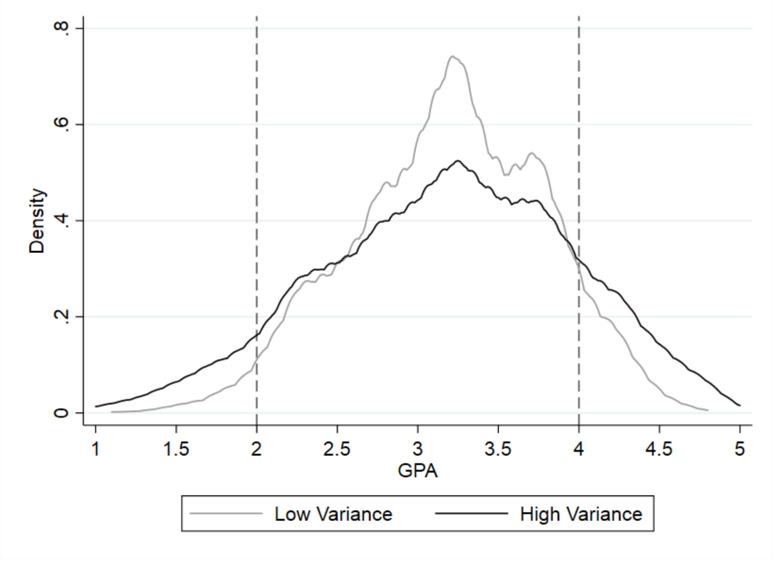
Sund, K. (2009), 'Estimating peer effects in swedish high school using school, teacher, and student fixed effects', *Economics of Education Review* **28**(3), 329–336.

Figure (1) Relationship between ability and school rank



Note: The figure shows the relationship between the ordinal rank of students and ability. The ability is measured according to the student's grade 9 GPA rank in the country, ranging between 1 and 50. Each dot marks the mean of school rank for each ability point, and the lines represent the types of school, from low- to high-mean in 20 groups

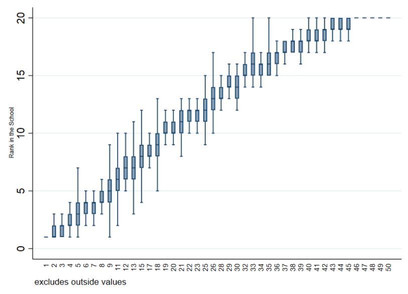
Figure (2) Two schools with the same GPA mean and high and low variance



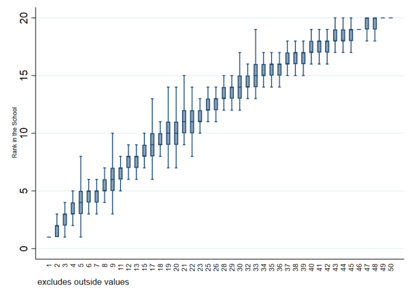
Note: This figure shows the grade distribution in two schools with the same mean GPA but with two variances: high and low. The black line is the distribution of high variance schools.

Figure (3) Relationship between ability and school rank in different types of schools

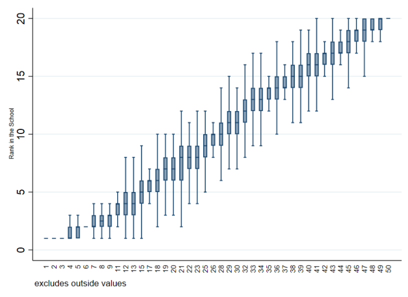
(a) Mean Q1- Variance Q1



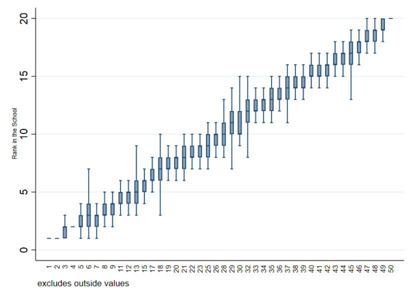
(b) Mean Q1- Variance Q4



(c) Mean Q4- Variance Q1

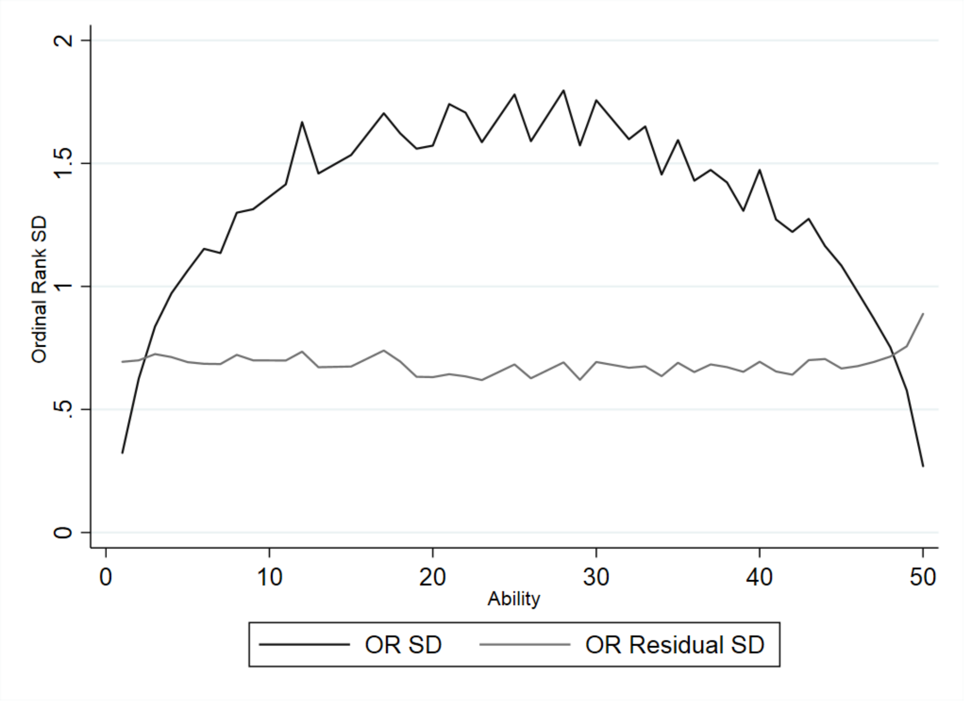


(d) Mean Q4- Variance Q4



Note: Figure 3 shows the box plot of students' ability and their ordinal rank. Each box represents the median, 25 percentile, and 75 percentile of ordinal rank for each ability point. Four types of schools are shown in panels: schools with low mean and low variance (panel a); schools with low mean and high variance (panel b); schools with high mean and low variance (panel c); and schools with high mean and high variance (panel d).

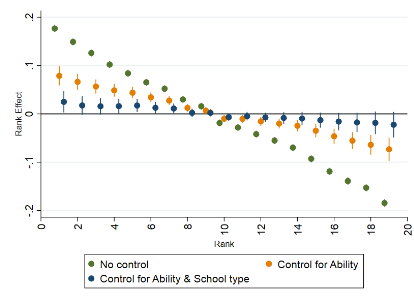
Figure (4) The standard deviation of ordinal rank by the ability



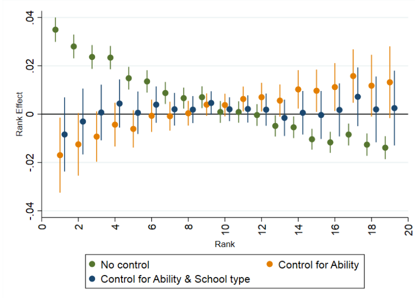
Note: This figure shows the standard deviation of the students' ordinal rank (black line) and the residuals (grey line) of ordinal rank. When calculating the residual of ordinal rank, the rank is used as a dependent variable regressed on ability, school type, pre-characteristics control, and school-cohort fixed effect.

Figure (5) Non-linear Balance Test

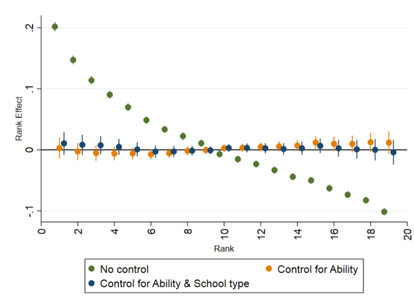
(a) Male



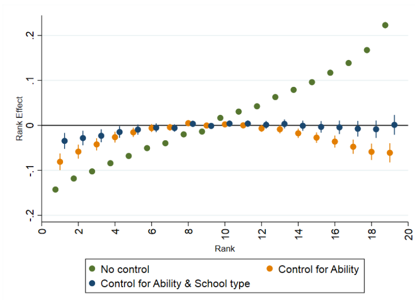
(b) Immigrant



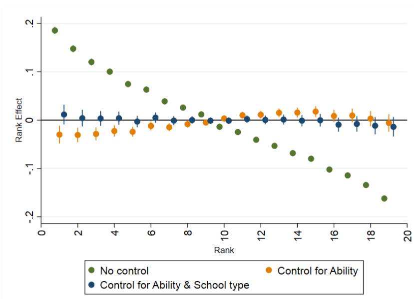
(c) Low Income Parent



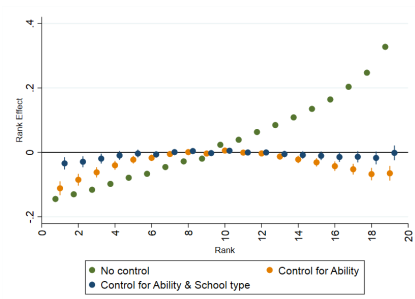
(d) High Income Parent



(e) Low Education Parent



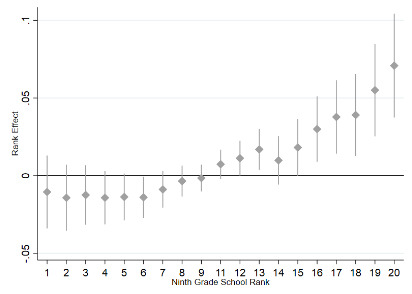
(f) High Education Parent



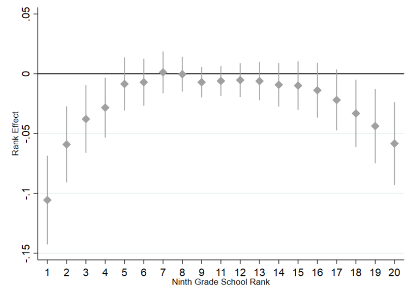
Note: This figure shows the non-linear balance test of different characteristics of the pupils across ranks. Each point is the estimate and 95% confidence interval of the corresponding rank. The green dots mark the estimates when no control is included in the model, the yellow dots mark the estimates when 50 dummies of ability are included in the model, and the blue dots mark the estimates when the interaction of school types and 50 dummies of ability were included.

Figure (6) The effect of ninth grade ordinal rank on different outcomes

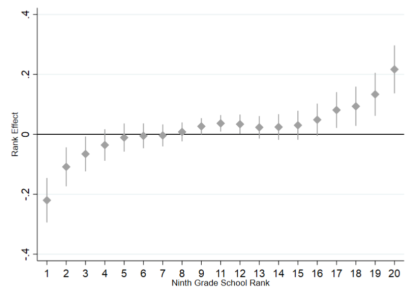
(a) Attending STEM



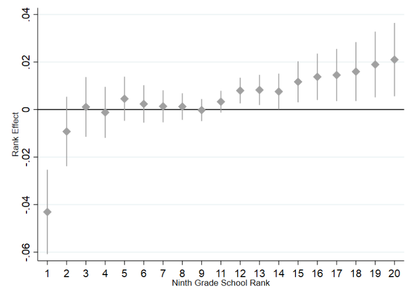
(b) Attending Vocational Track



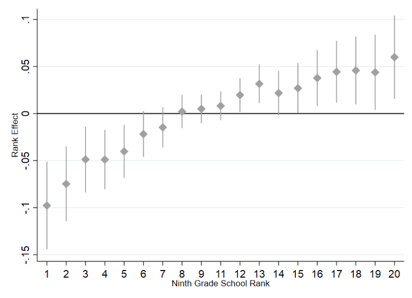
(c) Years of Schooling age 33



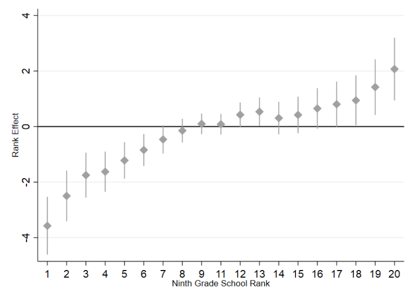
(d) Finished Upper Secondary school



(e) Income, age 30-33



(f) Income Rank age 30-33



Note: Each dot in the figures marks the estimates of 1 to 20 school rank with 95% confidence interval. The 10th rank is the reference point. In all models, dummy variables for 16 types of schools (based on mean and variances) are interacted with the pupils' national rank. The following controls are included in the model: immigration status, gender, parental education, parent income, and school-cohort fixed effect.

Table (1) Descriptive statistics

	Mean	Standard Deviation	Observations
Demographics			
Female	48.9	49.99	794076
Father income log	8.9	2.18	754112
Mother income log	8.6	2.14	727636
Father years of education	11.2	2.86	726405
Mother years of education	11.3	2.54	757699
Born outside of sweden	7.5	26.33	794076
Parent born outside of sweden	13.0	33.59	794076
Academic outcome			
GPA	3.2	0.70	794076
English (advanced)	3.0	0.90	211072
English	3.3	0.87	536233
Mathematics (advanced)	3.0	0.95	309044
Mathematics	3.3	0.90	445494
Sweidish	3.2	0.90	770575
Long term outcome			
STEM	22.4	41.67	276062
Vocational Track	37.3	48.35	276062
Income (age 30-33)	7.6	1.00	684328
Years of education (age 33)	13.1	2.17	756340
Finished USS	86.1	34.60	794076
Number of students in each school	121.8	40.30	794076
Number of students in each rank	6.1	2.01	794076

Note: This table contains descriptive statistics of the variable used in the study. It covers all students in compulsory school between 1990 and 1997.

Table (2) Balance Test

	Male	Low edu	High edu	Low income	High income	Immigrant
			A. Un-interacted			
Rank	-0.35***	-0.32***	0.44***	-0.26***	0.33***	-0.060***
	(0.0031)	(0.0029)	(0.0043)	(0.0028)	(0.0037)	(0.0028)
			B. School Mean Quartiles			
Rank	-0.056***	-0.0064	0.019	0.0068	0.012	0.0028
	(0.016)	(0.013)	(0.014)	(0.015)	(0.013)	(0.0053)
			C. School Mean Deciles			
Rank	-0.054***	-0.014	0.032**	0.0042	0.019	0.0035
	(0.016)	(0.013)	(0.014)	(0.015)	(0.014)	(0.0055)
			D. School Variance Quartiles			
Rank	-0.026*	0.020	-0.018	0.010	-0.0098	0.000071
	(0.015)	(0.012)	(0.014)	(0.014)	(0.013)	(0.0047)
			E. School Variance Deciles			
Rank	-0.018	0.016	-0.016	0.0088	-0.0061	-0.00092
	(0.015)	(0.012)	(0.013)	(0.014)	(0.013)	(0.0047)
			F. School Mean quartiles and variance quartiles			
Rank	-0.00081	-0.012	0.010	0.0076	0.020	-0.00024
	(0.016)	(0.013)	(0.013)	(0.015)	(0.013)	(0.0052)

Note: This table shows the balance test when the different characteristics of the students were regressed on rank and 50 categories of abilities in various specifications. This table resembles Table 2 in Denning et al. (2020). In panel A, the ability does not interact with school types. Panel B interacts ability with dummies for quartiles of school mean GPA. Panel C interacts ability with deciles of school mean. Panel D and E interact ability with quartiles of school mean and variance, respectively. Panel D interacts ability with 16 dummy variables for school mean and variance.

Table (3) Linear Results

	(a)	(b)	(c)	(d)	(e)	(f)
Income Rank						
School Rank	0.24***	0.27***	0.28***	0.18***	0.17***	0.24***
	(0.037)	(0.040)	(0.041)	(0.039)	(0.039)	(0.042)
Effect size (% of mean)	0.48	0.54	0.56	0.35	0.33	0.47
Income Log						
School Rank	0.0077***	0.0082***	0.0082***	0.0057***	0.0052***	0.0074***
	(0.0015)	(0.0017)	(0.0017)	(0.0016)	(0.0017)	(0.0018)
Effect size (% of mean)	0.10	0.11	0.11	0.076	0.069	0.099
Years of Schooling						
School Rank	0.020***	0.029***	0.032***	0.0062**	0.0043	0.015***
	(0.0027)	(0.0029)	(0.0030)	(0.0028)	(0.0029)	(0.0030)
Effect size (% of mean)	0.15	0.22	0.24	0.047	0.033	0.12
Finished Upper Secondary School						
School Rank	0.0014**	0.0018***	0.0022***	0.0015***	0.0016***	0.0022***
	(0.00055)	(0.00061)	(0.00062)	(0.00059)	(0.00060)	(0.00063)
Effect size (% of mean)	0.16	0.21	0.26	0.18	0.18	0.25
STEM						
School Rank	0.0058***	0.0063***	0.0063***	0.0029***	0.0028***	0.0039***
	(0.00097)	(0.0011)	(0.0011)	(0.00099)	(0.00099)	(0.0011)
Effect size (% of mean)	2.60	2.82	2.83	1.31	1.24	1.73
Vocational track						
School Rank	0.0054***	-0.0013	-0.0018	0.0057***	0.0058***	0.0014
	(0.0012)	(0.0013)	(0.0014)	(0.0013)	(0.0013)	(0.0014)
Effect size (% of mean)	1.44	-0.35	-0.49	1.52	1.55	0.38

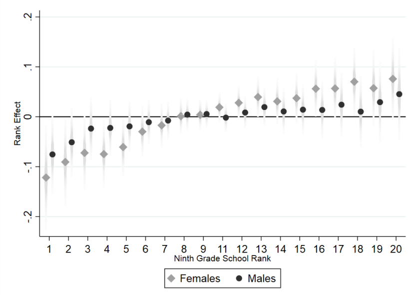
Note: This table shows the linear effect of students' ordinal rank on different short- and long-term outcomes. Each row in the table shows the estimates of each outcome on ordinal rank (1–20). All regressions included ability, school-cohort fixed effect, parent income, parent education, gender, and immigration status.

Appendices

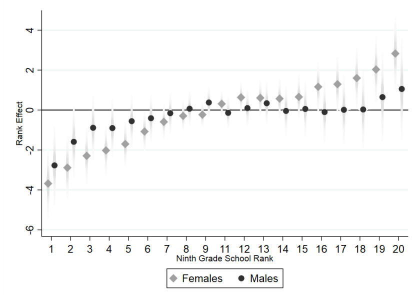
A Figures

Figure (A.1) Heterogeneity by Gender

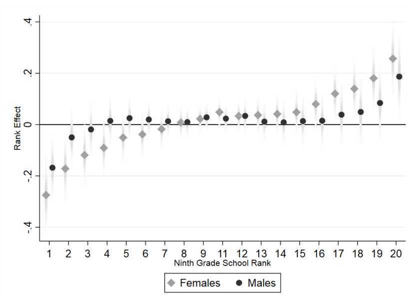
(a) Income, age 30-33



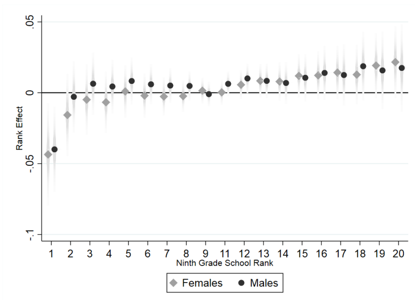
(b) Income Rank age 30-33



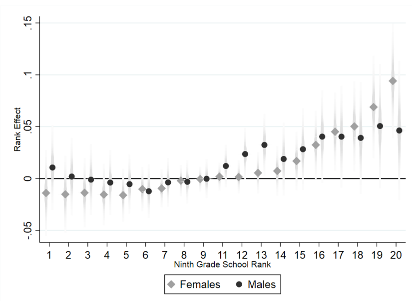
(c) Years of Schooling age 33



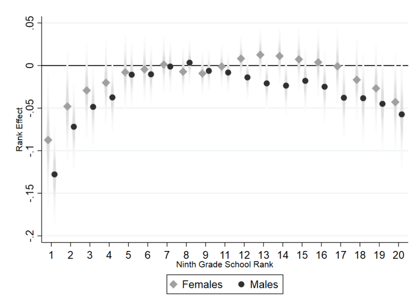
(d) Finished Upper Secondary school



(e) Attending STEM



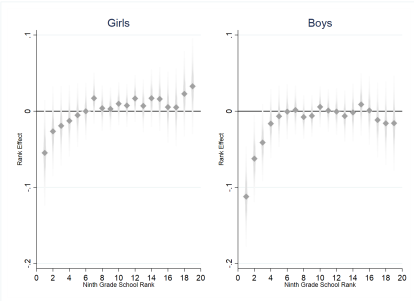
(f) Attending Vocational Track



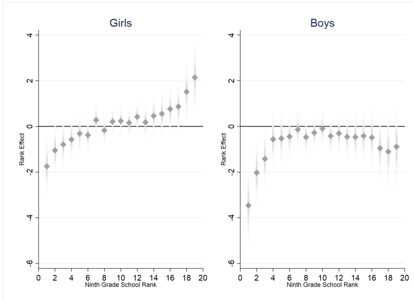
Note: Each dot in the figures shows the estimates of 1–20 school rank with 95% confidence interval for females (grey) and males (black). The 10th rank is the reference point.

Figure (A.2) Heterogeneity by Gender— Gender group

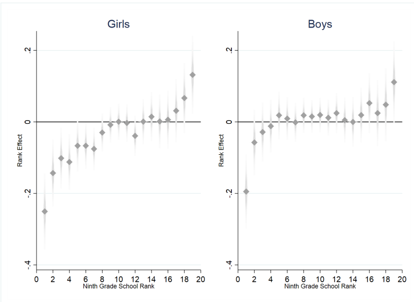
(a) Income, age 30-33



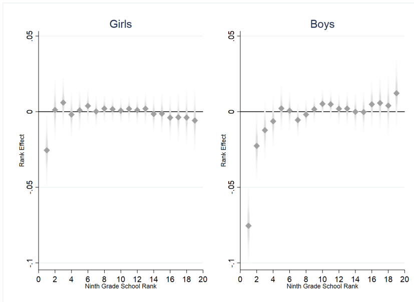
(b) Income Rank age 30-33



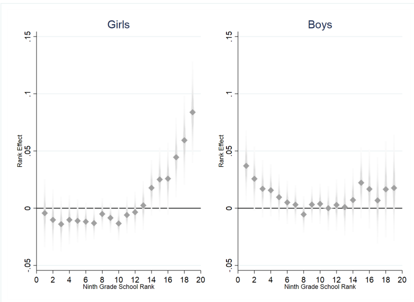
(c) Years of Schooling age 33



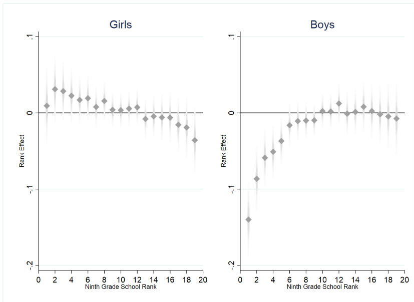
(d) Finished Upper Secondary school



(e) Attending STEM



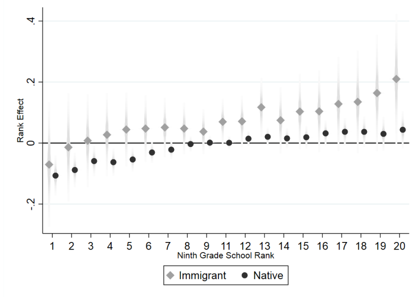
(f) Attending Vocational Track



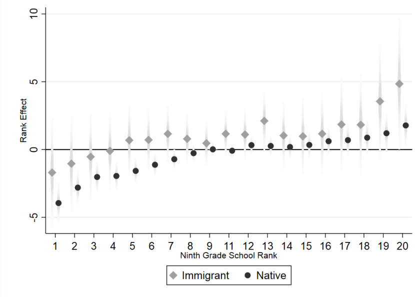
Note: This figure shows the estimates of girl's rank among girls (left panel) and estimates of boy's rank among boys (right panel) for different outcomes.

Figure (A.3) Heterogeneity by Immigration status

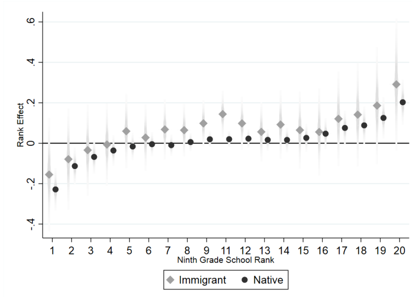
(a) Income, age 30-33



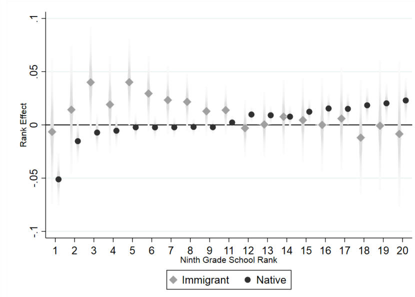
(b) Income Rank age 30-33



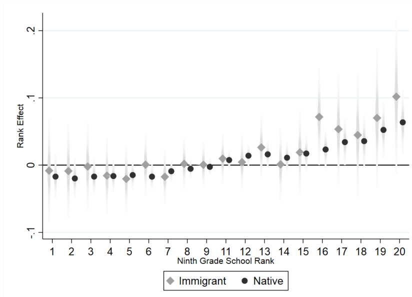
(c) Years of Schooling age 33



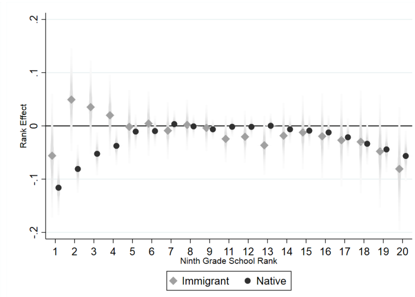
(d) Finished Upper Secondary school



(e) Attending STEM



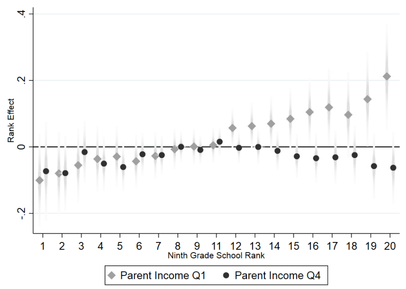
(f) Attending Vocational Track



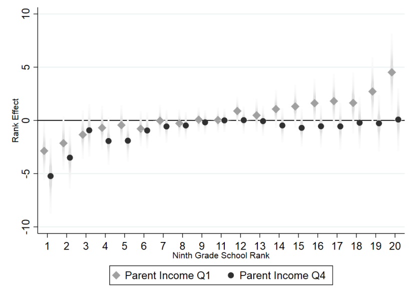
Note: Each dot in the figures shows the estimates of 1–20 school rank with 95% confidence interval for immigrant (grey) and Native (black). The 10th rank is the reference point. In all models, dummy variables for 16 types of schools (based on mean and variances) are interacted with the students' national rank. Gender of the student, parental education, parent income, and school-cohort fixed effect are included in the model.

Figure (A.4) Heterogeneity by Parent's Income

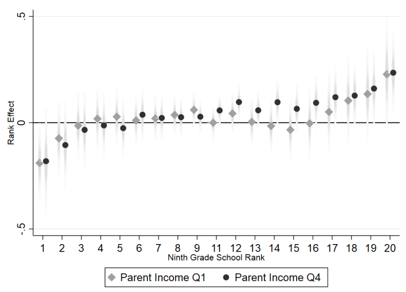
(a) Income, age 30-33



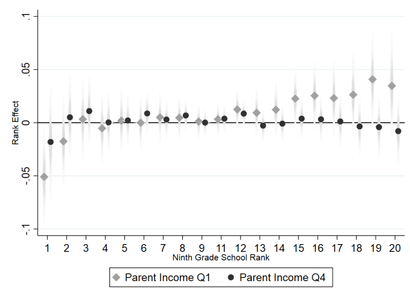
(b) Income Rank age 30-33



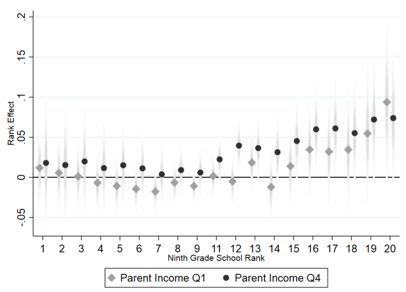
(c) Years of Schooling age 33



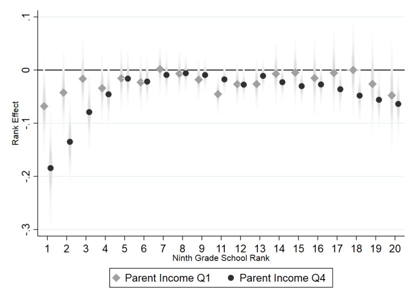
(d) Finished Upper Secondary school



(e) Attending STEM



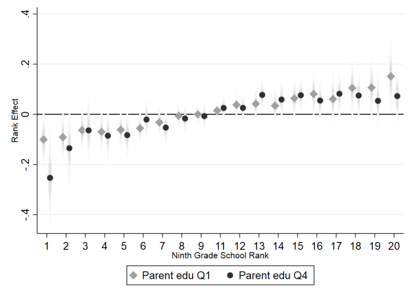
(f) Attending Vocational Track



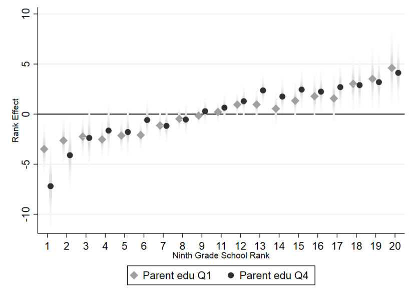
Note: Each dot in the figures shows the estimates of 1–20 school rank with 95% confidence interval for low income parent (grey) and high-income parent (black). The 10^{th} rank is the reference point. In all models, dummy variables for 16 types of schools (based on mean and variances) are interacted with the students' national rank. Gender of the students, parental education, parent income, and school-cohort fixed effect are included in the model.

Figure (A.5) Heterogeneity by Parent's Education

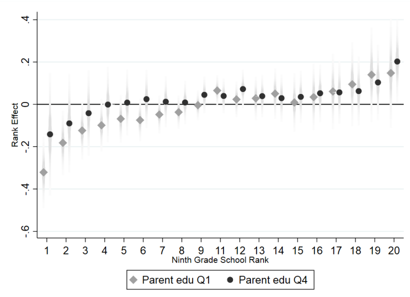
(a) Income, age 30-33



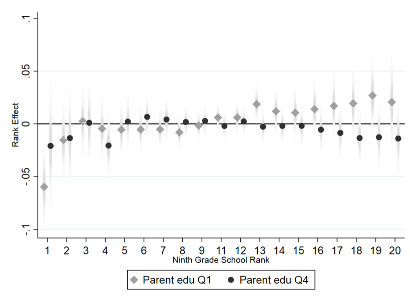
(b) Income Rank age 30-33



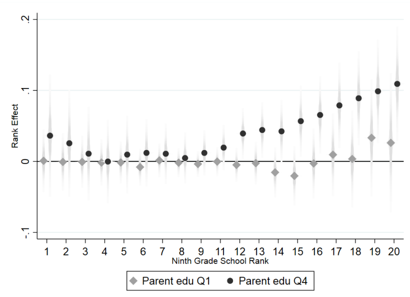
(c) Years of Schooling age 33



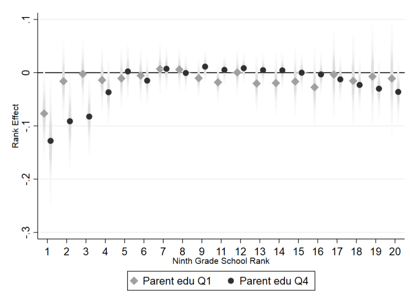
(d) Finished Upper Secondary school



(e) Attending STEM



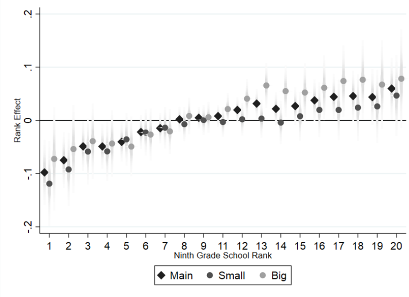
(f) Attending Vocational Track



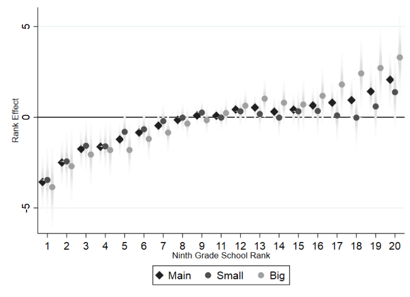
Note: Each dot in the figures shows the estimates of 1–20 school rank with 95% confidence interval for low-educated parent (grey) and high-educated parent (black). The 10th rank is the reference point. In all models, dummy variables for 16 types of schools (based on mean and variances) are interacted with the students' national rank. Gender of the students, parental education, parent income, and school-cohort fixed effect are included in the model.

Figure (A.6) Robustness check– school size

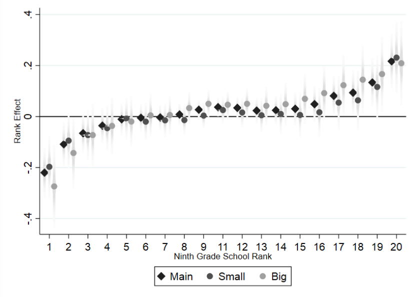
(a) Income, age 30-33



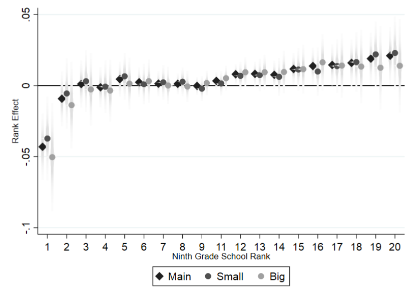
(b) Income Rank age 30-33



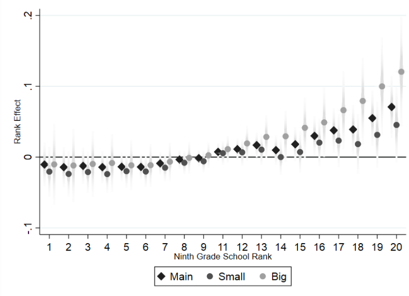
(c) Years of Schooling age 33



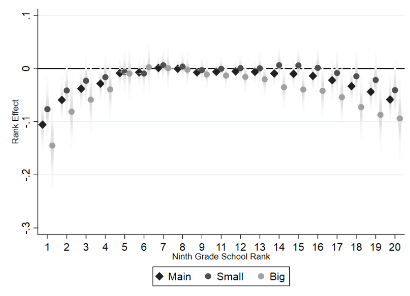
(d) Finished Upper Secondary school



(e) Attending STEM



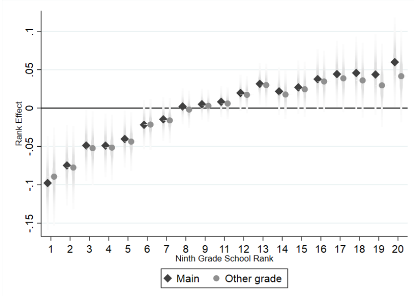
(f) Attending Vocational Track



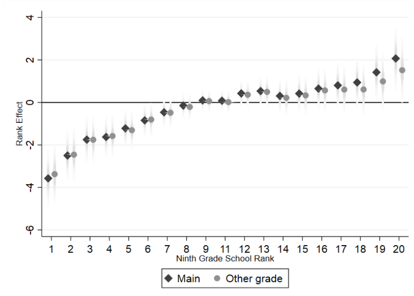
Note: Each dot in the figures shows the estimates of 1–20 school rank with 95% confidence interval for the small school (grey) and big school (black). Small (big)schools are defined if the school has less (more) than 120 students. The 10th rank is the reference point. In all models, dummy variables for 16 types of schools (based on mean and variances) are interacted with the student's national rank. Gender of the students, parental education, parent income, immigration status and school-cohort fixed effect are included in the model.

Figure (A.7) Robustness check—other grades

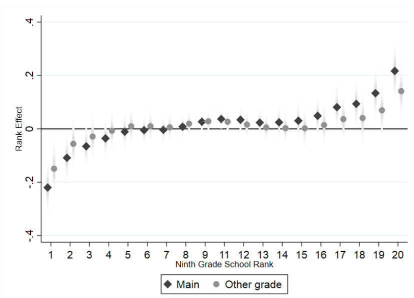
(a) Income, age 30-33



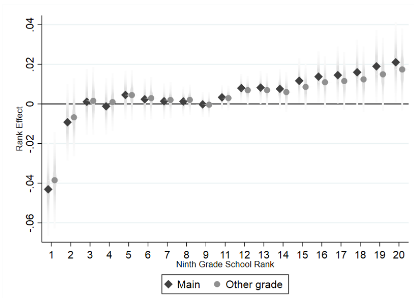
(b) Income Rank age 30-33



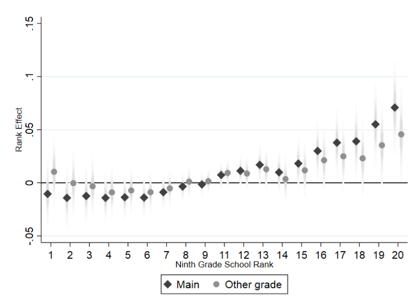
(c) Years of Schooling age 33



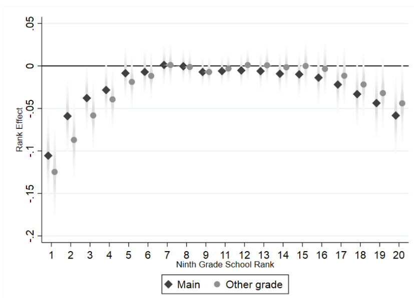
(d) Finished Upper Secondary school



(e) Attending STEM



(f) Attending Vocational Track



Note: Each dot in the figures shows the estimates of 1–20 school rank with 95% confidence interval. The main specification is similar to the result in Figure 6. The grey dots showed the estimates when Swedish, Mathematics, and English are added to the main model. The 10th rank is the reference point. In all models, dummy variables for 16 types of schools (based on mean and variances) are interacted with the students' national rank. Gender of the students, parental education, parent income, and school-cohort fixed effect are included in the model.

B Data description

This section explains the data used in this study. First, the information about how each variable was constructed is provided, followed by detailed descriptive statistics (Table B.1 and Table B.2).

B.1 Data description and variable definitions

B.1.1 GPA, ability, school rank, and school types

- Final grades from grade nine, GPA:
  Students were graded on a 5-point scale. The grade point average (GPA) is defined as the average of the subject grades, ranging from 1 to 5. The GPA is recoded with one decimal in the registry and is provided at a 50-level range, from 1–50.

- Swedish and Mathematics:
  The final grade for mathematics, Swedish and English, from 1–5.

- School code:
  Identifier of the schools, each school has one unique code.

- School rank among girls:
  Sorting each female student in her female cohort and school and grouped in 20 equal groups, each group contains 5% of female students in school-cohort.

- School Rank among boys:
  Sorting each male student in his male cohort and school and grouped in 20 equal groups, each group contains 5% of male students in school cohort.

- Country Rank (ability):
  Sorting each student on GPA among his or her cohort and grouped in 50 groups. The groups are not of equal size.

- School type based on GPA mean:

  To calculate this variable, the average of the school GPA is calculated. Then schools were sorted based on the mean from the lowest to the highest in 4 and 10 groups.

- School type based on GPA variances :

  To construct these variables, first the average of the school GPA variance was calculated. Then schools were sorted based on the variances from the lowest to the highest in 4 and 10 groups.

- School size:

  The number of the students in each school and year indicate the school size.

B.1.2 Students outcomes:

- Students' years of educations:

  To build students years of education, the students were followed until the age of 33 and the years of education were calculated based on the highest level of completed education. This is based on the latest degree (finished study) at the age of 33.

- Students' earned income:

  The average yearly earned income of the students between age 30 and 33. The analysis used the log of this variable.

- Students income rank:

  I used student income rank for students when they were between 30 and 33 years old. To find the income rank, the income for all students in the same cohort were ranked from lowest to the highest and grouped in 100 ranks. The percentile rank for the students was averaged between age 30 and 33.

- Students study track:

  After finishing grade nine, students could choose to attend different tracks, which are categorized as vocational, academic, or preparatory. The preparatory track normally consists of a year of studies aimed at preparing the student to enter a regular vocational or academic track, and it is for students who lack sufficient qualifications (have too low grades) after grade nine. Each of the vocational and academic tracks is divided into sub-tracks. The vocational track includes a sub-track related to industry and technology, health and childcare, and business and administration, and the science track consists of sub-tracks related to STEM, social sciences, and the arts.

B.1.3 Students demographic characteristics:

- Female:

  A dummy variable is equal to one if the students is female, and zero otherwise.

- Immigration:

  A dummy variable is equal to one if the student is born outside Sweden, and zero if the student is born in Sweden.

- Immigration year:

  A set of dummy variables indicate the age of immigration to Sweden.

- Immigrant studying Swedish as a second language:

  A dummy variable of one is used if an immigrant student was taking the course “Swedish as second language” and zero otherwise. This indicator is highly correlated with the “immigration year” as students who immigrated later in life more often attend this course.

- Students year and month of birth:

  Dummy variables indicate each student’s month and year of birth.

B.1.4 Family characteristics

- Parents years of education:

  The mother's and father's years of education when students were in grade nine. The variables are based on the highest level of completed education.

- Parents income:

  The mother's and father's incomes were measured when the students were in grade nine. Their incomes were summed and sorted from the lowest to the highest. Four dummies indicating the lowest (Q1) to the highest (Q4) income level were used.

B.2 Data sources

The data were obtained from registers from Statistics Sweden, for example Inkomst- och taxeringsregistret (IoT); hogskoleregistret (HR); Skolverkets elevregister (SE). Table B.3 shows the years that data were collected, the age of the students at the time of the variable used, and the data sources.

Table (B.1) Descriptive statistics

	Mean	SD	Min	Max	Obs
GPA, Rank and ability					
GPA	3.22	0.70	1.00	5.00	794076
English (advance)	2.99	0.90	1.00	5.00	211072
English	3.32	0.87	1.00	5.00	536233
Mathematics (advance)	3.00	0.95	1.00	5.00	309044
Mathematics	3.26	0.90	1.00	5.00	445494
Swedish	3.19	0.90	1.00	5.00	770575
Outcomes					
Academic track	0.57	0.50	0.00	1.00	276062
Social science	0.05	0.22	0.00	1.00	276062
STEM	0.22	0.42	0.00	1.00	276062
Vocational Track	0.37	0.48	0.00	1.00	276062
Vocational Track Technology	0.17	0.38	0.00	1.00	276062
Vocational Track Health Care	0.10	0.30	0.00	1.00	276062
Vocational Track Trade	0.10	0.30	0.00	1.00	276062
Income	7.65	1.00	0.00	12.27	684328
Years of Schooling	13.14	2.17	7.00	20.00	756340
Finished USS	86.09	34.60	0.00	100.00	794076
Controls					
Female share	48.87	49.99	0.00	100.00	794076
Father income log	8.92	2.18	0.00	14.23	754112
Mother income log	8.56	2.14	0.00	13.06	727636
Father years of education	11.23	2.86	7.00	20.00	726405
Mother years of education	11.32	2.54	7.00	20.00	757699
Parent income Q1	19940.25	28324.35	0.00	106991.40	193395
Parent income Q2	47236.09	50784.44	1739.00	157236.00	193392
Parent income Q3	62938.19	66818.10	2712.80	199728.81	193390
Parent income Q4	90613.31	100444.99	3512.20	1550109.00	193392
Parent Education Q1	17.11	1.76	14.00	20.00	181642
Parent Education Q2	20.91	0.88	19.00	22.00	174572
Parent Education Q3	23.55	1.12	22.00	26.00	175557
Parent Education Q4	28.97	2.59	26.00	40.00	176264
Number of students in each school	121.83	40.30	1.00	296.00	794076
Number of students in each rank	6.09	2.01	0.05	14.80	794076

Note: This table contains detail descriptive statistics of the variable used in the study. It covers all students in compulsory school for the period 1990–1997.

Table (B.2) Descriptive statistics—School types

	School Mean					School SD				
	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max	Obs
Quartile 1	3.08	0.07	1.65	3.14	198667	0.62	0.03	0.07	0.65	198623
Quartile 2	3.17	0.02	3.14	3.20	198726	0.67	0.01	0.65	0.69	198518
Quartile 3	3.23	0.02	3.20	3.27	198286	0.71	0.01	0.69	0.73	198739
Quartile 4	3.38	0.11	3.27	3.99	198397	0.76	0.03	0.73	1.23	198188
Decile 1	3.03	0.08	1.65	3.09	79936	0.59	0.03	0.07	0.62	80209
Decile 2	3.11	0.01	3.09	3.13	79481	0.64	0.01	0.62	0.65	79505
Decile 3	3.14	0.01	3.13	3.15	79124	0.65	0.00	0.65	0.66	78509
Decile 4	3.16	0.01	3.15	3.18	79152	0.67	0.00	0.66	0.68	79406
Decile 5	3.19	0.01	3.18	3.20	79700	0.68	0.00	0.68	0.69	79512
Decile 6	3.21	0.01	3.20	3.22	79467	0.69	0.00	0.69	0.70	80228
Decile 7	3.24	0.01	3.22	3.25	79048	0.71	0.00	0.70	0.72	78779
Decile 8	3.27	0.01	3.25	3.29	80888	0.73	0.01	0.72	0.74	79468
Decile 9	3.33	0.02	3.29	3.37	78337	0.74	0.01	0.74	0.76	79747
Decile 10	3.48	0.11	3.37	3.99	78943	0.79	0.03	0.76	1.23	78705

Note: This table reports descriptive statistics of the school-level data for the period 1990–1997. School's were categorized based on GPA mean (left panel) and GPA variances (right panel). The mean and the variance are sorted in quartile and decile.

Table (B.3) Data Sources

Variables	Year	Age of students	Register
GPA, Rank and ability			
GPA	1990-1997	15/16	SE
Swedish 9th grade	1990-1997	15/16	SE
Math 9th grade (advanced)	1990-1997	15/16	SE
Math 9th grade (not advanced)	1990-1997	15/16	SE
English 9th grade (advanced)	1990-1997	15/16	SE
English 9th grade (not advanced)	1990-1997	15/16	SE
Rank in school (all)	1990-1997	15/16	SE
Rank in school (only girls)	1990-1997	15/16	SE
Rank in school (only boys)	1990-1997	15/16	SE
Rank in country (all)	1990-1997	15/16	SE
Outcomes			
STEM	1995-1997	15/16	HR
Vocational	1995-1997	15/16	HR
Years of education	2007-2014	33	SUN
Log income	2004-2014	30-33	IoT
Income rank	2004-2014	30-33	IoT
Control Variables			
Gender	——	—	IoT
Immigration			IoT
Year of birth	1974-1981	0	IoT
Month of birth	1974-1981	0	IoT
Mother level of edu (4 level)	1990-1997	15/16	SUN
Father level of edu (4 level)	1990-1997	15/16	SUN
Mother years of edu	1990-1997	15/16	SUN
Mother years of edu	1990-1997	15/16	SUN
Parental household income 4 quantiles	1990-1997	15/16	IoT
School dummies	1990-1997	15/16	SE
School size	1990-1997	15/16	SE
School type	1990-1997	15/16	SE

Note: This table reports the years that data were collected, the age of the students at the time of the variable used, and the data sources.

II

School Autonomy and Subject-Specific Timetables*

*Acknowledgments: I thank Karin Edmark, Matthew Lindquist, Jonas Vlachos, Hans Grönqvist, Anders Stenberg, Roujman Shahbazian, and Dan-Olof Rooth. I have significantly benefited from comments from seminar participants at the Institute for Evaluation of Labor Market and Education Policy (IFAU), the Swedish Institute for Social Research (SOFI) at Stockholm University, and the Department of Economics at Stockholm University. I would also like to thank Mari Eneroth and Sofie Burman for their help in gathering information concerning the timetable reform. Financial support from Handelsbankens forskningstiftelser and from IFAU is gratefully acknowledged.

1 Introduction

Because high-quality education is perceived as vital to a country's success, children spend a considerable amount of their time in school. Therefore, how much time is allocated to specific subjects needs careful consideration. Recently, research has focused on how increasing time spent on specific subjects (often focusing on STEM-related topics) affects students' future employment and earnings (Kirkeboen et al. 2016, Dahl et al. 2020). Little attention, however, has been given to who should decide how to allocate the instruction time across subjects. That is, should subject-specific timetables be determined and regulated by the central government and therefore be homogeneous across schools or should each school decide how to allocate the teaching time between subjects?

A school-level schedule solution has the potential advantage of enabling schools to adjust their subject-specific teaching time to the local students' needs because schools have more relevant and detailed information about their students (Hanushek et al. 2011). However, a school-based timetable increases each school's workload and administrative activities as teachers and principals need to determine the timetable (Zabojnik 2002). In addition, there is a risk that inequality will increase when the central government provides fewer regulatory guidelines as some schools might not be qualified to make decisions about timetables (Lundahl 2002*b*).

Few empirical studies have addressed this question. Hanushek et al. (2011) use panel data from 42 countries observed annually for 10 years and investigate the effect of increased school autonomy on student outcomes. Fuchs & Wöbmann (2007) use the PISA database to examine the effect of school autonomy on student performance. They found that school autonomy concerning managing personnel, teachers' decisions, and textbooks positively affects student outcomes. However, school autonomy has adverse effects in areas with a strong potential for opportunistic behavior, such as forming a school budget.

This study evaluates a Swedish policy that transferred the decision-making authority over time allocation across subjects from the national to the school level. The experiment took place in 900 of Sweden's approximately 3000 primary and lower secondary schools in the early 2000s. The central government implemented the experiment to evaluate the effects of allowing schools to allocate total teaching time across subjects. Several studies on the reform implementation were carried out in connection with the trial; these findings are reviewed in section 2.3. However, no comprehensive evaluation has been performed on the long-term impact of using extensive microdata on students, a gap in knowledge this study fills.

This study evaluates the impact that timetable reforms had on the short-term outcomes for grade 9 students–specifically, the probability of entering a STEM field, the probability of graduating from high school, and years of completed education at the age of 25. Here, we study the average effect on the students in the participating schools and estimate the impact separately for students from different socioeconomic statuses (SES). This is essential for determining whether the policy decreased the inequality in the system by primarily benefiting low-achieving students or whether it was more beneficial for already high-achieving groups and therefore increased educational inequality.

The reform was implemented in two steps. First, municipalities applied to participate. Second, the schools in the selected municipalities decided whether to participate. Because the municipalities and schools were not randomly selected (i.e., they were self-selected), participation can be correlated with both school and municipality-specific characteristics. This analysis is based on a Differences-in-Differences design (DiD) that included school and year fixed effects, which were constant for all time-invariant school characteristics and controlled for yearly national shocks. Furthermore, the preferred specification includes local labor market by year fixed effects to control for regional changes across time. In addition, the specification includes additional covariates that influence factors such as gender of the students, parental education and income, student birth month, and student immigration status.

DiD estimates the causal effects of the reform under the assumption that the included covariates and fixed effects sufficiently control for school and municipality level factors that are correlated with the treatment. This assumption is supported by the fact that I found no violation of parallel trends in any of the studies' outcomes in the pre-treatment period when the preferred specification is used.

The results of my evaluation suggest that the school-based timetable on average had no impact on student outcomes. This finding holds both for subject-specific grades and other more long-term outcomes but is in contrast to the positive results reported by earlier government reports and other qualitative studies about this reform. I argue that one crucial aspect of analyzing the reform is to consider the differences between municipalities that applied the reform.

I also checked whether the effect is heterogeneous in three levels: student characteristics, school size, and geographical area. In terms of the students' background characteristics, I found the reform impacted the probability of attending STEM fields in university among students with low parental income and education. The reform-effect also varied based on school size–i.e., there is a tendency of positive effects in larger schools. Furthermore, the reform positively affected years of education, the probability of finishing upper secondary school, and finishing STEM fields at universities in large cities, whereas the effect is zero for these outcomes in urban areas. Moreover, two robustness checks–one based on dividing the control group into two alternative control groups and one based on using matching technique to define the control group–produced similar results.

This study makes several contributions to the literature. First, this is the first large-scale microdata study that looks at school autonomy regarding the use of timetables. Second, the policy was implemented as a quasi-experimental scheme where only some municipalities could participate. That is, this study controlled for many potential confounding factors, moving closer to the causal effect of interest. Third, this study delivers novel evidence on various outcomes that can be impacted by decentralization policy. In contrast, most of the existing evaluations of decentralization policies are only focused on achieved grades. Fourth, this study focuses on the average effects and investigates heterogeneity in the school-based timetable's development across students' gender, students' immigration status, parental income and parental education. Finally, this study's sample sizes are likely large enough to estimate the potential effects of the policy precisely.

The rest of the paper is organized as follows. The next section describes how the timetable reform was implemented. Section 3 presents the data, and section 4 presents the empirical strategy. In section 5, I check the validity of the empirical strategy. The results are presented in section 6. Section 7 contains the robustness check, and section 8 presents the conclusions.

2 The Swedish time schedule pilot project

2.1 The content and background of the reform

The timetable reform under study in this paper was implemented in the context of increased decentralization of the Swedish public sector. The reform was preceded by three main decentralization reforms of the education system (Åsa Ahlin & Mörk 2008). In 1991, the responsibility for compulsory schools was shifted from the national level to the municipality level. In 1993, the grant structure changed; municipalities were paid a general grant rather than a target-based grant, so the municipalities had more flexibility in how to spend the grant money. Finally, in 1996, the wage-setting decision was shifted from central negotiation to school managers and municipalities. These reforms meant that Sweden went from being regarded as having one of the most centralized education systems among the OECD countries in 1990 to having one of the most decentralized education systems in 1999.

At the end of the 1990s, the Swedish government decided to go one step further and implement a pilot project where some schools would be given the authority to shape the timetable. Under the existing system, the timetable was determined centrally by the Ministry of Education, although schools were free to decide about some dimensions. The schools participating in the pilot project were given expanded decision power over the timetable and were free to schedule the time allocated to different subjects. The sole restriction was that the total instruction time needed to reach the minimum specified hours SOU2005:101 2005.

The project was based on a government report published in 1997 (SOU 1997:121), which proposed running a pilot project where a set of schools would be allowed to depart from the national schedule. The project's primary motivation was that a strict schedule was not compatible with the goal-based grading system implemented in 1998 for the compulsory education system nor with the goal-based grading system implemented in 1994 for the upper secondary education system. It was argued that merely participating in a class for a particular duration does not necessarily mean that students learn a subject deeply or broadly. It was also argued that schools needed more freedom to decide how the total teaching time was to be allocated between subjects to encourage the use of this new, more goal-oriented education system. The hope was that the pilot project would yield valuable information and experiences that could help determine whether a decentralized timetable system should be implemented for all schools in the country.

2.2 Selection of participating municipalities and schools

The government proposition to run the policy experiment was approved by the Parliament on November 3, 1999. The policy was implemented in two steps. First, the Ministry of Education invited all municipalities to apply for participation in the experiment. Of the country's 289 municipalities, 79 applied; of these, 70 were selected to participate. Second, these selected municipalities were allowed to decide which schools would participate in the experiment. It was basically up to the municipality to choose which schools were to be involved. Some municipalities selected all their schools, and some municipalities selected only a few of their schools (Figure A.1).

Although it is not known how the school-level decisions to participate were made in the 70 participant municipalities, a study by Rönnberg 2007*b* sheds light on this process in 16 of the municipalities. The majority of school leaders surveyed in the study emphasized that schools were selected based on their willingness–i.e., schools that volunteered were chosen to participate. However, there are three exceptions to the voluntary notification procedure for schools. In these three municipalities, either all schools participated or the municipality specifically selected schools for participation without sending a general request to every school. In total, 183 schools (20% of the students) participated in the pilot experiment in these 16 municipalities.

In the municipalities that allowed the schools to decide whether to participate in the experiment, several reasons for participating were documented–e.g., to increase individualized learning, to increase interdisciplinary work, to support a more comprehensive and overall view of learning, to develop new methods of instruction and learning, and to increase students' interests and responsibility. The main aim was to increase students' opportunities to plan their own work and let the content of the work control the planning of the day. For these municipalities and schools, the timetable represented a time constraint that was inconsistent with a goal-driven and results-driven system.

2.3 How did the participating schools implement the pilot policy?

There is no comprehensive, centrally available information about how all the schools that participated in the pilot project changed the timetable and other related aspects. However, a relatively large number of studies were carried out in relation to the implementation of the policy experiment, and this section reports what we know about the implementation in some of the participating schools based on these studies (see Barrios-Fernández & Bovini 2021, Dahl et al. 2020, Dills & Hernandez-Julian 2008, Rönnberg 2007*a*, Nyroos 2007, Rönnberg 2007*a*, Rönnberg 2007*b*, Nyroos et al. 2004, Lundahl 2005 Lundahl 2002*a*, Lundahl 2002*b*) . These were primarily written between 2000 and 2007, and some of them were part of the project Schools Without National Time Schedule (Skola Utan National Timplan, SKU).

In some schools, teachers would meet with one student at a time to discuss what activities would best meet the student's needs and interests. These meetings were intended to help students work in a more goal-oriented and independent manner. Often, students had a mentor or supervisor and the schools planned and recorded the students' weekly work in logbooks or planning books that a parent was required to sign. Planning time became a natural part of the schedule, and students, with the help of their teachers, often created their own individual development plans. In addition, some schools also allocated more time to interdisciplinary sessions.

Most schools stated that they developed their work teams extensively, that the work teams gained greater authority and independence, and that the work teams engaged in further pedagogical discussions. Another clear tendency was to increase physical activity. Several schools created profile classes where some subjects or subject areas were given increased time, such as sports, culture, media, languages, and mathematics. Several schools stated that they used a flexible starting time in the morning and ending time in the afternoon and that this contributed to a calmer environment. In addition, the reports often noted age-integrated teaching, more extended coherent teaching sessions without interruptions, level groups, and help with homework in the afternoons (Skolornas arbetssätt; SOU 2005:101).

Of the participating schools in 16 municipalities, just over half (57%) of the 55 principals said that participation in the pilot program meant that the time distribution changed to a large or considerable extent at their school, but about 40% said that the changes were small in both these aspects(Rönnberg 2007*b*). Unlike the principals, the ten surveyed teachers' association representatives believed the pilot policy did not result in any significant changes: eight representatives marked the changes as marginal and noted that the changes mainly concerned redistribution of subject-specific time of student's own work sessions (stugtid). A longitudinal study of three schools complemented the description that changes mainly related to modified working methods and only slightly to rearrangements of subjects and students (Lundahl 2005).

Based on interviews with students and teachers, Elmeroth et al. (2005), looking at how participant schools reallocated the time, found that the timetable pilot may have had a different impact on different subjects. In several of the studied experimental schools, the school day was divided into subject time, individual time, and common time. Both students and teachers believed that Social Studies, Swedish, and Swedish as a Second Language had benefited from the experimental set up. In these subjects, the knowledge goals were perceived as evident–i.e., it was easy for students to continue with the tasks in these subjects during their individual study time. The teachers in science subjects, however, were concerned that the skills learned risked being superficial as the time for these subjects was perceived as scarce and insufficient for conducting demonstrations, experiments, and laboratory work. For the language subjects, the difficulties depended partly on the limited ability to perform verbal exercises during the specific time and other teachers' lack of knowledge of different languages. The latter becomes a problem when students need help with languages during sessions led by teachers who do not know these languages (Elmeroth et al. 2005). Alm (2003) points out that mentor time has established itself in the trial schools for grades 7–9, and students generally call for more trusting relationships with their teachers (Skolornas arbetssätt; SOU 2005:101)

In sum, the pilot project seems to have induced schools to implement more individualized and student-led sessions, where students work individually according to their plans and progressions, but less time was devoted to specific subjects. However, it is difficult to know if these changes will be viewed as substantive or marginal and to what extent the experiences of the schools surveyed in the above studies are generalizable to all 900 participating schools.

3 Data

The study is based on panel data covering all grade 9 students in Sweden from 1990 to 2010. These data encompass students born between 1975 and 1998 as students start grade 9 at age 15 and include almost 100,000 students yearly.[1] All observations with missing information on student background characteristics (about 4%) were dropped from the analysis. This section describes the three categories of variables used in the regression analysis: the treatment variables, the outcome variables, and the control variables.

3.1 The treatment variables

Two alternative treatment variables were used in the main regression. The first is defined as a simple dummy variable for cohorts who attended grade 9 in a treated school and in post-treatment years. This treatment variable captures the policy experiment's extensive marginal effect on the outcomes. The second treatment variable captures the treatment's linear effects and is defined as the number of years each student can be assumed to have studied in the treated schools and post-treatment. This variable is between 0–9: 0 for control groups and 9 for students exposed to the reform for 9 years. This treatment variable captures the intensive marginal effect of the treatment. A three-level indicator of the treatment variable is also defined to test the reform's dose-related non-linear effect. This indicator takes a value of one if a student is assumed to have been treated between 1 and 3 years after compulsory school (grades 1–9), the indicator takes a value of two

[1] Some students may start school early or late and therefore be a bit younger or older

if the student is assumed to have been treated between 4 and 6 years after compulsory school, and the indicator takes a value of three if the student is assumed to have been treated between 7 and 9 years after compulsory school.

Since the data used for this study only provide links between schools and students measured at the end of grade 9 (i.e., the end of compulsory education), the treatment intensity variables are based on the school that the student attended in grade 9 in combination with school-level information on what grades are offered by the school. For example, a student who finished grade 9 in school A in 2009 is assumed to have been subject to six years of treatment if school A offered grades 4–9 but not grades 1–3. In other words, for such cases, I assume that the school that the student attended in grades 1–3 was not a treated school. That is, some error in this variable is likely because the student may not have attended school A throughout grades 4–8 and because the 1–3 grade school that the student attended may have been treated. Therefore, the treatment effect will either be downward or upward biased, depending on what is more prevalent.

Although the intensive treatment variable has this flaw, I still find it valuable to use this variable to approximate the impact of treatment duration on the effects. Columns 4, 5, and 6 of Table B.9 in the Appendix show the number of students who belong to these treatment groups each year. According to the simple treatment dummy variable, almost 18% of the students were in the treated group. Of these, 50% were exposed to the treatment for fewer than 3 years, 26% for 4 to 6 years, and 23% for more than 7 years. The distribution of the affected students across municipalities is shown in Figure A.2, where municipalities are shaded according to the share of students who attended schools that participated in the pilot project. There is a high variation between municipalities in this group. For example, in Enköping and Sundbyberg, almost all students were involved in the experiments (dark blue), whereas in Västerås no more than 10% of the students were exposed to the reform.

3.2 Outcome variables

A school-specific timetable could affect students' academic outcomes in the short term since schools could allocate more time to certain subjects or help students in private sessions. This might affect student grades and graduation rates, the study track students attend in upper secondary education, and how long students continue studying. Six dependent variables were used.

First, the immediate effect on students' school results was measured using the final grades awarded for grade 9. We studied the impact on the final grades in different subjects and the grade point average (GPA). [2] Since the grading system changed in 1998, [3] which resulted in a substantial change in the GPA distribution, the outcome measure is defined as the student's percentile rank in the cohort as this generates an arguably more consistent measure over time than a standardized measure. To construct this variable, all students in the same cohort were sorted from the lowest to the highest GPA and were clustered in 100 equal-sized bins. The subject-specific grades for mathematics, Swedish, and English were used in two forms: receiving a passing grade and receiving a high grade. A passing grade means students did not fail the course. A high grade means the students passed with distinction (VG) or passed with special distinction (MVG). The sample was restricted to 1998 and onwards for these variables as this was when the new grading system was introduced.

The second outcome variable is the student's percentile rank at the last year of upper secondary school. This was constructed in the same manner as the corresponding grade 9 measure. In Sweden, students select their track after finishing compulsory school (i.e., after grade 9). The selection is based on their grade 9 GPA, given their track choices. It is important to consider that the GPA in different study tracks is not comparable since

[2] GPA is measured as the mean value of 16 subjects

[3] Before 1998, five levels in each subject from 1 to 5 (1 is the lowest and 5 is the highest). After 1998, the grading system became "goal-based"–i.e., the evaluation was based on pre-specified goals defined for each topic, and students were graded on a 4-level scale, with the lowest being a "Fail" grade.

each track in upper secondary school covers different subjects, so it is vital to control for the track when upper secondary GPA is used as an outcome.[4] We need to consider that the study track is potentially an outcome variable affected by the treatment, so it is a "bad control". Nevertheless, I chose to present results from a specification that includes dummy variables for the tracks of attendance, because not controlling the track leads to an omitted variable bias. Interpretation of the results from this specification should be made while keeping in mind that the specification includes "bad controls" and therefore may be biased.

The third outcome variable is a dummy variable, which indicates whether students attended an academic STEM track[5] in the first year of upper secondary school. This variable, however, only shows if a student was accepted, not whether a student finished the track. The students' upper secondary track information is only available after 1995.

The fourth and fifth dependent variables are the outcomes that include years of education and a dummy variable, which indicates if a student finished upper secondary school. These two outcomes are measured when students are 25 years old and include all students who attended grade 9 between 1990 and 2007. The period is restricted to 2007 because the last year of the outcome data is 2016, when most of the students who finished grade 9 in 2007 were 25 years old.

The last outcome is a dummy variable that indicates whether the student attended and finished the STEM field in the university. This outcome is measured when most of the students were 28 years old. This age is not a perfect age to measure the university field of study in Sweden as some students attend university at older ages but still indicate finishing the STEM field at university.

[4]There are several options that students can attend in upper secondary schools. They can choose among various academic tracks that prepare for university studies or they can choose from several vocational options that prepare students for occupations in for example nursing and industry. Vocational studies can also be followed by higher vocational studies or some university/college training. Students who are not qualified to enter upper secondary school based on the grade 9 grades can take a preparatory year.

[5]This is defined as attending the Science or the Technical Science tracks in the fall of grade 9

3.3 Control variables

Several control variables that are likely relevant for the outcomes are included in the model. These controls are student's gender, age, immigration status and parental income and education. All student's background characteristics were measured when the students finished grade 9.[6] Empirical and theoretical studies have shown that these variables are highly correlated with educational outcomes in Sweden. I also controlled for the age of the students using the month and year of the birth of the students.

3.4 Local Labor Markets (LLMs)

In the next section, I discuss the importance of controlling for the differences between different geographical areas since the overall environment of these areas changed significantly during the study period. Two variables were used to control these differences: the Local Labor Markets (LLMs) and the number of the private schools in the municipalities.

Statistics Sweden defines LLMs based on the observed commuting patterns across municipalities. Each LLM is supposed to be independent of the outside supply and demand for labor. Therefore, LLMs change over time based on the commuting distance of the labor market. At the beginning of the study period (1990), there were 112 LLMs. By the last year of the study period (2010), there were only 74 LLMs. To use a consistent definition, I use the number of LLMs in 1999, which was 90. The share of the private primary schools in each municipality is also used to capture municipality differences. This variable defines the share of students who attended the private schools in each year and each municipality.

[6]This means that we assume that the treatment itself did not affect these covariates. Given the nature of the covariates (measuring either students' characteristics such as age and gender or the parental education level and income), this seems like a fairly innocuous assumption.

4 Empirical Strategy

To identify the policy's causal effect separately from potential confounding factors such as the quality of schools and unobserved parental and student background variables, I use a Differences-in-Differences (DiD) estimation strategy. This model allowed me to control for fixed unobserved variables that could be correlated with the policy and educational outcomes.

The DiD equation for estimating the average treatment effect of the reform is:

$$y_{ist} = \alpha + \beta_0 Treat_{st} + \gamma \boldsymbol{X}_{ist} + T_t + S_s + LLM * T_t + P_m + \beta_b Tr_{sb} + \varepsilon_{ist} \tag{1}$$

y_{ist} are the outcomes of student i in school s at time t. $Treat_{st}$ is a treatment variable, defined as one for students who attend a treated school in a post-treatment year and zero. S_s is a school-fixed effect that controls for permanent school characteristics that can affect educational outcomes. T_t is year fixed effects handle time-changing factors that influence outcomes in the whole country at the same time, for example, nation-wide education policies. $\boldsymbol{X}_{ist}$ is set of student's characteristics such as gender, immigration background, parents' education, and parents' income. ε_{ist} is an error term.

Previous literature shows that the educational outcomes in Sweden are different across regions. The descriptive statistic in the next section for the selected and non-selected municipalities also shows that the characteristics between these municipalities are different from the beginning. Although the fixed effect of the school controls for the fixed school differences, there is a concern that there is a difference in the regional trend that may confound the estimated treatment effect. This might lead to failure of the parallel trend assumption. To capture local trends during the study period, the LLM-year fixed effects were added to the model. There were 90 LLM in Sweden in 1999[7] , so 1800 dummy variables (90 labor market area * 20 years) were used. To further control

[7]The definition of the LLMs has changed slightly over the time. To have a balanced definition of LLM, I used the definition in 1999 for all years.

for differences across municipalities, the share of private primary schools each year (P_m) in the municipalities was also added.

To get a quick indication of whether the parallel trends' assumption behind the DiD estimation holds for each outcome variable, a placebo pre-treatment dummy (Tr_{sb}) was added to the regression equation. This variable is defined as one for students graduating from a treated school (i.e., a school that will eventually be treated) in the first half of the pre-treatment period, 1990–94, and is zero otherwise. [89] Non-zero β_b estimates suggest that the outcomes of the treated and untreated groups of students develop differently before the reform and therefore suggest that the assumptions of the DID specification are not fulfilled. Non-significant estimates, on the other hand, indicate that the specification is valid. The placebo coefficient is reported in all baseline results in the result section.

The parameter of interest is β_0, which captures the average effect of the policy on the treated group. Although the β_0 estimator captures the impact of attending a school that participated in the timetable reform in a year when the reform was in place, we lack comprehensive and precise information about how the participating schools implemented the reform (section 2.3). Thus, I was not able to tell if any estimated impact is due to any particular component of the reform, for example, if some particular change to the schedule was more or less beneficial to students. I will, however, discuss the results concerning the partial information on the reform available from the studies carried out in relation to its implementation (section 2.3).

As explained in the previous sections, we use both the simple treatment dummy variable (0 and 1) and the exposure treatment variables (0–9) . In addition to the average treatment equation above, several heterogeneity regressions are estimated by running the model separately for subsets of students to test whether the effect differs across the

[8]This means that the reference period is 1995–1999, and both the placebo coefficient and the treatment effect are estimated in relation to this period.

[9]One of the studied outcomes–i.e., the likelihood to enroll in a STEM track in the first grade of upper secondary school–had no observations prior to 1995, which means that the placebo estimate is omitted from the estimation for this outcome.

groups of students.

5 Empirical strategy validity

The DiD model assumes the outcome variables in the control and treatment groups follow parallel trends in the absence of the treatment. This assumption cannot be tested formally as the counterfactual of "no treatment" for the treated group in the treatment period cannot be observed. However, an informal way to evaluate the likelihood of this assumption is to study whether the pre-treatment period trends are parallel–i.e., determining whether the differences in school outcomes between the control and treated groups are constant overtime before introducing the reform. In addition, the DiD model assumes that no other changes were implemented at the time of policy implementation that would confound the estimation of the reform effects. To test if this assumption, balance checks were performed to see if the treatment control group was balanced. In other words, both the treated municipalities and the control group needed to implement further reforms in 2000. This section provides a detailed evaluation of the assumptions underlying the DID equation.

5.1 Reform selection process

The treatment selection process for this pilot project needs to be considered when evaluating the more likely threats to the validity of the DiD model. As described in section 2, a selection was made at the municipality level before the school level. As a result, there are two types of non-participant schools: non-participant schools in participating municipalities (the left branch of the flowchart) and schools in the non-participating municipalities (right branch of the flowcharts) (Figure A.1). Judging from the studies referenced in section 2, it is likely that many of the schools in the former group could have participated in the experiment if they chose to; I call this Group A. Group B is the schools that did not have a chance to participate in the experiment due to decisions made at the municipality level.

Using either of these control groups or the combination of the two comes with advantages as well as disadvantages. Control group A helps control for the schools' municipality-related unobservable characteristics since these groups are in the same municipalities as the treated schools. This is beneficial since the municipalities in Sweden are responsible for providing compulsory school. These schools have the same local school funding system, are subject to the same municipality-level decisions, and share the same demographic and economic environment. The disadvantage is that these schools potentially voluntarily chose to opt-out of the treatment. This increases the likelihood of selection bias at the school level. Schools in Control group B, on the other hand, are not in the same municipalities, which means that they may be subject to different municipality-level trends. However, they were not allowed to participate in the experiment even if they wanted to, which means that I might have found good matches in terms of characteristics correlated with school-level selection in treatment at the school level within this group.

I use the following estimation strategy with respect to the definition of the control group. First, I evaluate the appropriateness of both of these potential control groups as well as of the combination of the two based on how well they balance with the treatment group with respect to pre-treatment characteristics and trends. This evaluation determines what control group is used for the main estimations of the paper. Second, I present some results for the groups that are not chosen as the main strategy to evaluate the robustness of the main results. Third, I further investigate whether the results are sensitive to the composition of the control group by combining the DiD with a matching strategy to increase balance in a robustness section.

5.2 Descriptive statistics and balance check

This section first presents descriptive statistics for the municipalities and schools in the treated and (potential alternative) control groups. Next, it evaluates whether the assumptions underlying the DiD specification of equation 1 are valid.

The following information is depicted in Table 1. The municipalities selected for treatment had a higher population density, more students per school, and more immigrant students. In addition, the standard deviation of most variables is much higher in the participating municipalities, which means the variation within this group is higher than within the non-participating group. Finally, based on the 1998 election, the selected municipalities have a slightly higher vote share of right-leaning political parties.

The school-level characteristics for the three alternative control groups are depicted in Table 2 The schools in Control group B are, on average, more similar to the treated schools, in particular when it comes to the share of immigrants, type of municipality (urban or rural)[10], GPA, and the number of students in each school. For a better sense of the differences, we calculated the normalized values for the differences of the average values of the characteristics for the set of alternative control groups. In other words, we scale the differences with the pooled standard deviation of the two groups. (See Imbens & Rubin (2010) for normalized differences). [11] In Table 2, Δ_i indicates the normalized differences–i.e., the smaller the value of Δ_i , the more similar the control group is to the treatment group regarding that variable. This indicator also has another advantage for our study: we could determine which group gives the smallest index to find the most comparable control groups. Group ALL (i.e., all control schools) and group B are more similar to the treatment group than control group A.

[10] 1 is more urban and 9 is more rural.

[11] The normalized difference is calculated by $\Delta_i = \frac{\bar{X}_t - \bar{X}_c}{\frac{(s_t{}^2 - s_c{}^2)}{2}^{0.5}}$, where $\bar{X}_t$ and $\bar{X}_c$ are the mean value of treated and control groups, and s_c and s_t are the standard deviation of these two groups. This indicator's advantages are that it does not depend on the sample size (in contrast to the t-test) and is more informative in groups with many observations.

More critical for the DiD model's validity is the balance in terms of changes over time in the treated and non-treated units. To understand whether the covariates of the students in the treated and control groups are balanced in terms of the type of time and cross-sectional variation used in the DiD analysis, we replaced Y_{ist} in equation 1 with the predetermined characteristics of the whole period without adding other covariates. That is, we regressed each covariant on the pre-treatment placebo dummy (1995–1999) [12], LLM-year fixed effect, and school fixed effect to see how these characteristics were affected by the treatment. Therefore, a coefficient value close to zero for the pre-treatment period (1990–1999) indicated that the treatment and control groups are more balanced in terms of the changes in the outcome variable. The coefficients after treatment (2000–2010) show whether the treatment had any effect on the students with different characteristics.

The event study type analysis is presented in Figure 1 when all schools in the control (ALL) are used. The analysis is performed for male students (panel a), immigrant (panel b), parents' incomes (panels c and d), and parents' education levels (panels e and f). Each dot shows treatment estimates for these characteristics in each year. The figure reveals that there is slightly fewer (more) students with low (high) educated parents in the treatment group before the reform; however, this difference is not significant in most of the years. The same analysis is done for the Control Group A and B. Figure A.3 in the appendixes show that the estimates are almost the same with these control groups.

5.3 Parallel trends in the outcome variables

The DiD model assumes that the outcome variables in the treatment and control groups would follow parallel trends in the absence of treatment. A straightforward way of

[12] Note that the pretreatment dummy variable here takes value 1 for 1995–99, whereas in the regression specification (1) takes value 1 for 1990–94. That is, in the main specifications, we wanted to use the years before the treatment period 1995–99 as the reference period, so the treatment effect is estimated in relation to these years.

evaluating the parallel trend assumption is to run an event study analysis and see if the estimates before the reform are statistically zero.

Equation 1 is used to conduct an event-study analysis but replaces the simple post-reform dummy with separate dummy variables for each year and estimates dummy variables for each pre-treatment year. If the assumption of parallel trends is valid, the estimated coefficients for these dummies should be zero for all pre-treatment years (1999 and before). The estimated coefficients for the post-reform years (after 1999), on the other hand, capture the development of the treatment effect over time. Figures 2 and 3 show the event study analysis for all outcomes. Figure 2 shows the effect on grade 9 percentile rank GPA, upper secondary percentile rank GPA, the probability of enrolling in a STEM field in upper secondary school,[13] the probability of graduating from upper secondary school within 12 years, the years of education at the age of 25, and the probability of finishing a STEM degree at university. Figure 3 illustrates the event study analysis for Mathematics, Swedish, and English in two levels–i.e., probability of passing grade and probability of high grades. Since the grading system changed in 1998, there are only two years before the reform.

These figures suggest no statistically significantly different trends in the treatment and control group pre-treatment outcome variables. There is some evidence of divergence post-treatment for upper secondary percentile rank GPA, years of schooling, and high school graduation, although the trend returns to the pre-treatment levels after a few years and is always statistically insignificant. We take this finding as further support of the validity of the DiD and as a tentative suggestion of a treatment effect of the timetables reform.

Figures A.4 - A.7 in the Appendix shows the event study analysis for samples A and B. The analysis shows that in both samples, the pattern for the pre-reform period

[13]Note that years of schooling are computed based on information on the highest level of completed education at age 25. This means, for example, that a student who attended the first grade of high school but did not complete high school will be recorded as having the same number of years of education as an individual who did not attend high school.

is similar to the pattern for the sample All. Therefore, we chose "all schools" to avoid dropping observations for the main results. In the robustness section, we discuss that the results are similar when we use these two samples.

6 The effect of the timetable reform on the outcome variables

The empirical analysis begins by estimating the educational effects of the timetable reform in both extensive and intensive marginal settings. The section checks whether the reform has heterogeneous effects for students with different background characteristics. The section ends with a discussion of the robustness checks.

6.1 The Main Results

The reform's extensive and intensive marginal effect is shown in Table 3 (different educational outcomes) and Table 4 (Mathematics, Swedish, English Grades). The left panel shows the extensive marginal effect of the treatment when the treatment variable captures whether students were in the untreated or treated schools (0 or 1 treatment). The right panel shows the extensive marginal effect of the reform when the 0–9 treatment variable used in equation 1 (i.e., the treatment variable) captures the number of the years that students were (plausibly[14]) exposed to the treatment. The placebo shows the dummy variable's estimates for all treated schools for 1990–1994. Column 1 shows the treatment effect when no control for municipalities was used in the model. Column 2 includes the share of private schools in the municipality, and column 3 includes LLM-year fixed effects added to the model. In all models, school fixed effect, year fixed effect, parental education and income, and student immigration status, the student's birth

[14]As explained in section 3.1, since we can only link students to schools in grade 9, these alternative treatment variables are based on assumptions of the students' school trajectories.

month, and gender were added in the model. As we discussed in the empirical strategy section, it is important to consider that the participant and control municipalities may be subject to different regional trends, so the preferred specification is the last column.

The effect size of the estimates is small. The only significant estimate (with a 10% significance level) in column 3 is the probability of attending an academic STEM field in upper secondary school. The coefficient is 0.006, which means that attending a treated school increases the probability of choosing and being accepted to an academic STEM field in upper secondary school by almost 0.6 percent. The right panel of Table 3 shows the extensive marginal effect. The coefficient for the years of education is significant and positive (with a 10% significance level). It shows that each year attending a treated school could increase the years of education by 0.0054. Considering the variable's mean value (12.51 years), the effect size is 0.34 percent of mean if students study for 9 years in treated schools (0.0054/12.51 * 9 * 100). The coefficient in column 6 for other outcomes is insignificant and small.

How students were graded in various subjects was also analyzed by looking at two outcomes: receiving a passing grade and the probability of having a high grade for several subjects.[15] Table 4 shows the reform's effect on the probability of receiving passing grades, above passing grades, and high grades (pass with distinction and pass with excellence) for three core courses– Mathematics, English, and Swedish. These subjects are important as students need to pass them to qualify for the regular upper secondary school educational tracks. The reform has no extensive and intensive marginal effect on earning a pass or high grade in any of the outcomes.

The next question is whether the treatment effect is non-linear and whether the exposure to the treatment in a non-linear format needs to be considered. To do this, the treatment effect was divided into three levels: when students were plausibly attending

[15] The period of analysis is restricted to 1998–2010 because the grading system changed in 1998 and the definition of passing and high grade changed. In the new system, students could receive one of four grades: fail (f), pass, pass with distinction, and pass with excellence.

treated schools for less than three years, 4–6 years, and 6–9 years. The results are shown in Figure 4 and Figure 5. In all outcomes, the point estimates are zero for students exposed to the treatment for fewer than 3 years (level 1). After that, the effects start increasing for years of education and percentile rank at upper secondary school. The most considerable results are seen for those exposed for 7 to 9 years (level 3). The percentile rank of GPA at upper secondary school increases by one rank if students are exposed to the treatment 7–9 years. The probability of finishing STEM at the university is increased by 0.5%. Figure 5 shows the dose-related estimates for the three core subjects. The estimates are not significant in all cases except the probability of passing Swedish, which is affected negatively with more exposure to the treatment.

6.2 Heterogeneity Effect

In this section, we examine whether some subgroups were more affected by the reform than others. According to the government, one of the experiment's motivations was to give schools the freedom to offer more help to students with a foreign background. The government also emphasized that students with low SES needed more individual instruction, which can be provided with the school-based timetable. Furthermore, large schools were more likely to be treated (Table 2). Large schools might need more flexibility since they have more diverse students, although this is only speculation. We also looked at the differences between rural and urban areas since it might be the case that schools in rural areas are more capable of managing the timetable. Therefore, the heterogeneity analysis was performed in the following areas: large cities, urban areas, small school, large schools, low and high parent education, low and high parent income, and gender and immigration background of the students. [16]

The effect of the treatment on all outcomes is shown in Figure 6 and Figure 7. The outcomes are presented in panels a–f. The intensive marginal estimates and 95

[16] Appendix C explained how these categories are defined.

confidence interval (CI) of 11 subgroups of students are shown in each panel. Except for grades in Swedish and English, the intensive marginal effect of the treatment on outcomes are significant and positive in large schools. Students with low educated and low-income parents are also more likely to finish STEM at university. The estimates for percentile rank in upper secondary schools seem to be positive in all groups; however, as the fraction of the students who go to a different study track are different, the estimates of this outcome are not reliable.

7 Robustness Check

To check whether the results are robust to the choice of the control groups, we performed two robustness checks. First, we analyzed whether the effects in Sample A and Sample B are different. Second, we constructed a matched control group to see how results changed when using a control group more similar to the treatment group.

7.1 Sample A and Sample B

For the sake of completeness, this section presents the results using two alternative control groups (A and B), which are described in section 5.1, rather than the full control sample. Figure A.4-A.7 in the appendix show the event study of sample A and sample B. Before the reform, there was no relation between outcomes and the treatment variable in almost all outcomes, so these cases support the parallel trend assumption. The estimated treatment effect is small and significant in some outcomes in sample B but in most cases estimates are similar for both groups (Table B.5-B.8)

7.2 Matching

One way to improve the likelihood that the parallel trend assumption holds is to limit the sample of analysis to more similar schools–i.e., to find the schools in the non-participating group that are more comparable with the participating schools. Matching helps minimize unobservable variable bias due to selection by matching pre-reform characteristics. I used a standard matching strategy to match school-level data from the year before the reform. The following variables are used for matching schools: immigration share, the average parental income, the average level of parents' education, the school size, and the municipality size. We matched three control groups: ALL controls (i.e., matching schools in all municipalities), Control A, and Control B. The matching improves both the balance test (Table A.3) and parallel trend assumptions (Figures A.8–A.13) slightly. The results based on matched samples also show that the estimates are almost similar to the main results.

8 Concluding

This paper analyzes the impact of a school-based timetable on several student outcomes by investigating a program that gave some schools in Sweden the opportunity to allocate time to different topics as they deemed appropriate. The reform was implemented as a pilot project, and municipalities and schools could choose to be part of the treatment. I use the whole population registry data of students who finished grade 9 from 1990 to 2010 and implemented a DiD specification. The identifying assumption is that exposure to the school-based timetable is as good as random and restricted by school-fixed and cohort-LLM fixed effects and the included covariates. A set of validation checks supports this assumption.

After estimating the effect of the reform on different educational outcomes, we found that there is no average treatment effect. We checked both intensive and extensive marginal effects as well as non-linear exposure to the treatment effect. The study found that more prolonged exposure to the school-based timetable only led to small positive impacts on years of education. Studying in a treated school increased the years of schooling by 0.4%. The results also suggest that timetable reform does not affect the students' average grades. This paper argues that it is essential to consider the area-specific trend in the model since without controlling for that the selection process of the experience generates a bias. These results are contradicted in a previous study and governmental reports about this reform that suggest a positive effect of the timetable experience.

For the external validity of the effect of this reform in another context, we should consider the differences between the Swedish education system at the time of the reform with other systems. First, at the beginning of 2000, the Swedish education system was among the most decentralized education systems in the world, so the schools and municipalities were familiar with these kinds of decisions. In another world, timetable reform was implemented in an environment with high level of flexibility in the school system. Second, the reform was implemented some years after the introduction of private schools in Sweden, so the more autonomy in public schools could be a reaction to this change.

The results also shows a heterogeneous effect of the reform on low background SES students and different school sizes. Students who studied in large schools positively reacted to the reform. The students from larger schools were twice as likely to choose a STEM track. Large schools can potentially benefit from this type of reform since they have more diversity with respect to students' abilities and backgrounds. They also have more teachers who can make flexible changes to the timetable and use these timetables to respond to pupils' individual needs. Further research is needed to develop a more complete picture.

References

Alm, F. (2003), *Skolämnen och alternativen: Schemat som indikator på vad som händer i skolor utan timplan*, Linköping University Electronic Press.

Barrios-Fernández, A. & Bovini, G. (2021), 'It's time to learn: School institutions and returns to instruction time', *Economics of Education Review* **80**, 102068.

Bimber, B. (1993), *School Decentralization: Lessons from the Study of Bureaucracy*, Rand Corporation.

Bruns, B., Filmer, D. & Patrinos, H. A. (2011), *Making Schools Work : New Evidence on Accountability Reforms*, number 2270 *in* 'World Bank Publications', The World Bank.

Dahl, G., Rooth, D.-O. & Stenberg, A. (2020), Long-run returns to field of study in secondary school, Technical report, National Bureau of Economic Research.

Dills, A. K. & Hernandez-Julian, R. (2008), 'Course scheduling and academic performance', *Economics of Education Review* **27**(6), 646–654.

Elmeroth, E., Eek-Karlsson, L., Olsson, R. & Valve, L.-O. (2005), 'Tid för målstyrning: utvärdering av kvalitetsarbete i försöksverksamheten med utbildning utan timplan i grundskolan'.

Fasih, T., Barrera-Osorio, F. & Patrinos, H. A. (2009), *Decentralized Decision-Making in Schools*, The World Bank.

Fuchs, T. & Wöbmann, L. (2007), 'What accounts for international differences in student performance? a re-examination using pisa data', *Empirical Economics* **32**(2-3), 433–464.

Hanushek, E. A., Link, S. & Woessmann, L. (2011), Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA, CESifo Working Paper Series 3648, CESifo Group Munich.

Imbens, G. W. & Rubin, D. B. (2010), *Rubin causal model*, Springer.

Kirkeboen, L. J., Leuven, E. & Mogstad, M. (2016), 'Field of study, earnings, and self-selection', *The Quarterly Journal of Economics* **131**(3), 1057–1111.

Lavy, V. (2020), 'Expanding school resources and increasing time on task: Effects on students' academic and noncognitive outcomes', *Journal of the European Economic Association* **18**(1), 232–265.

Lundahl, L. (2002*a*), 'From centralisation to decentralisation: Governance of education in sweden', *European educational research journal* **1**(4), 625–636.

Lundahl, L. (2002*b*), 'Sweden: decentralization, deregulation, quasi-markets-and then what?', *Journal of education policy* **17**(6), 687–697.

Lundahl, L. (2005), 'A matter of self-governance and control the reconstruction of swedish education policy: 1980-2003', *European education* **37**(1), 10–25.

Nyroos, M. (2007), 'Time to learn, time to develop? change processes in three schools with weak national time regulation', *Pedagogy, Culture & Society* **15**(1), 37–54.

Nyroos, M., Rönnberg, L. & Lundahl, L. (2004), 'A matter of timing: time use, freedom and influence in school from a pupil perspective', *European Educational Research Journal* **3**(4), 743–758.

Rönnberg, L. (2007*a*), 'A recent swedish attempt to weaken state control and strengthen school autonomy: The experiment with local time schedules', *European Educational Research Journal* **6**(3), 214–231.

Rönnberg, L. (2007*b*), 'The swedish experiment with localised control of time schedules: Policy problem representations', *Scandinavian Journal of Educational Research* **51**(2), 119–139.

Åsa Ahlin & Mörk, E. (2008), 'Effects of decentralization on school resources', *Economics of Education Review* **27**(3), 276 – 284.

SOU2005:101 (2005), Utan timplan – för målinriktat lärande, Technical report, statens offentliga utredningar, Stockholm.

SOU2005:102 (2005), Utan timplan – forskning och utvärdering , Technical report, statens offentliga utredningar, Stockholm.

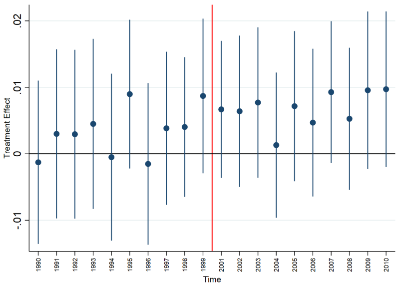
SOU2007:121 (2007), Utan timplan – forskning och utvärdering , Technical report, statens offentliga utredningar, Stockholm.

Wallberg, E., Eliasson, R., Fernstedt, L., Gullstam, I., Lundstrom, M., Molander, M. & å, M. W. (2005), Utan timplan, Technical report, Slutbetä nkande från Timplanedelegationen.

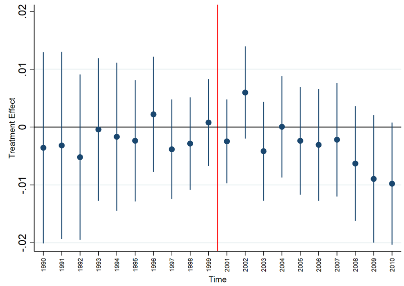
Zabojnik, J. (2002), 'Centralized and decentralized decision making in organizations', *Journal of Labor Economics* **20**(1), 1–22.

Figure (1) Balance Test

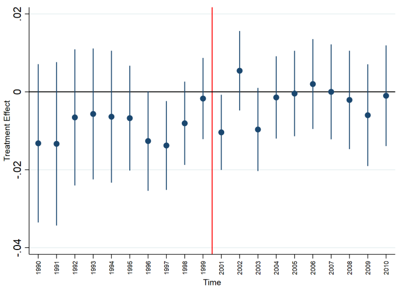
(a) Male



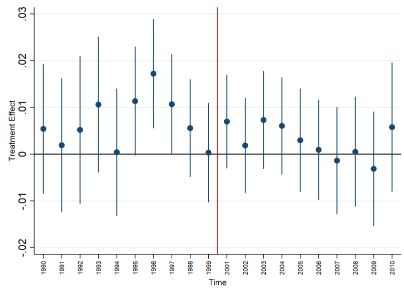
(b) Immigrant



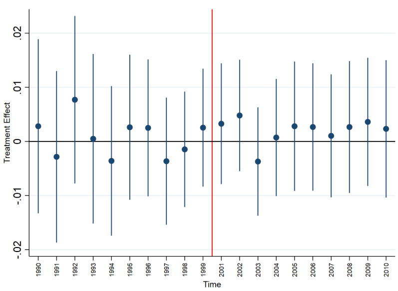
(c) Low Income Parent



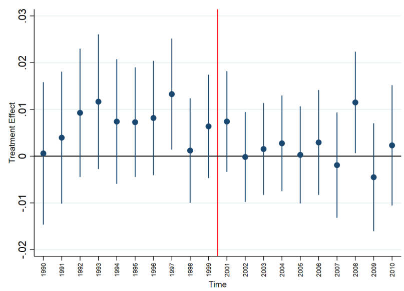
(d) High Income Parent



(e) Low Education Parent

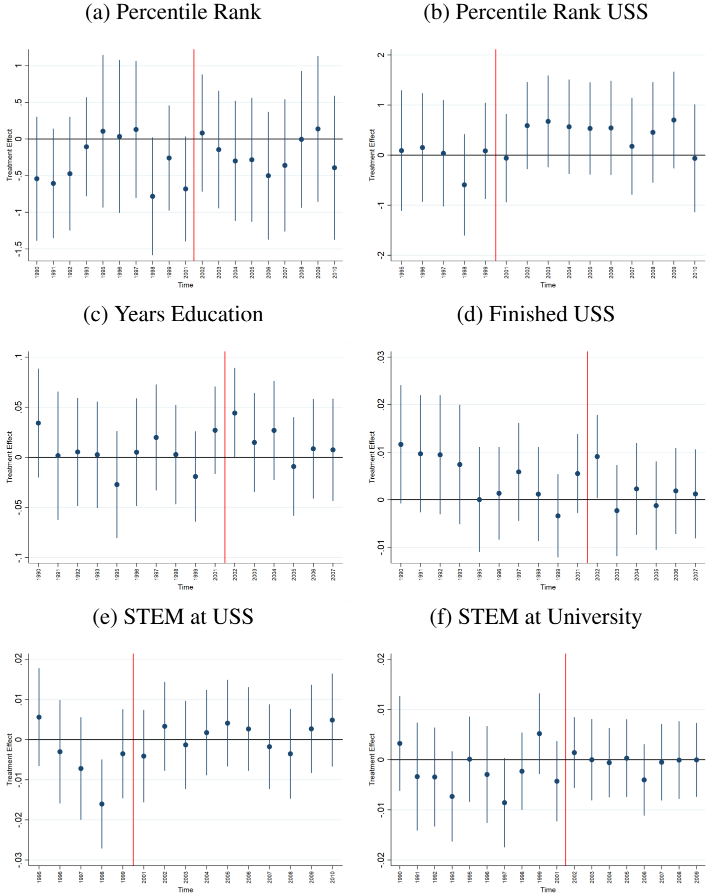


(f) High Education Parent



Note: This figure shows the event study type analysis when all schools in the control group are used. The analysis is performed for male (panel a) immigrant students (panel b), parents' incomes (panels c and d), and parents' education levels (panels e and f). Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) each year.

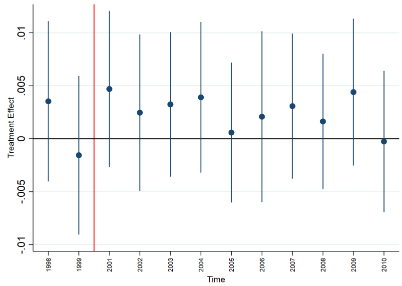
Figure (2) Event study graph of the effect of school-based timetable on different outcomes



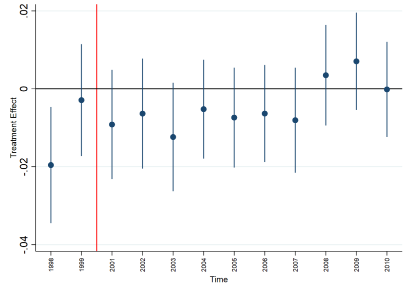
Note: This figure shows the event study of treatment on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f). Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (3) Event study graph of the effect of school-based timetable on Swedish, English, and Mathematics grades

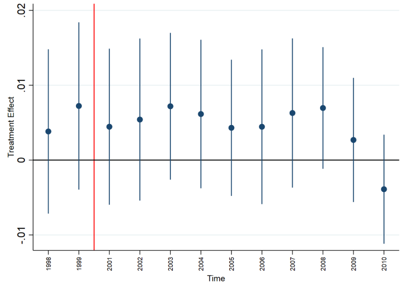
(a) Mathematics pass



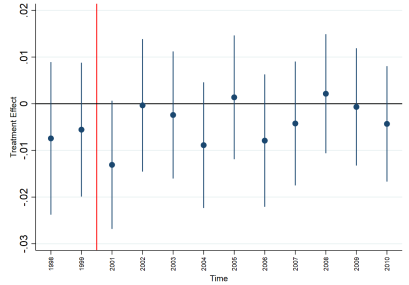
(b) Mathematics high grade



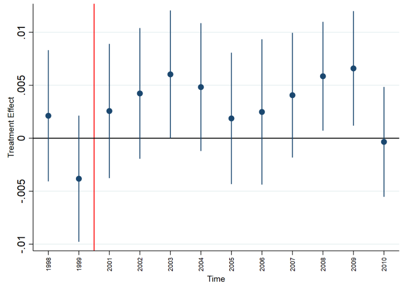
(c) Swedish pass



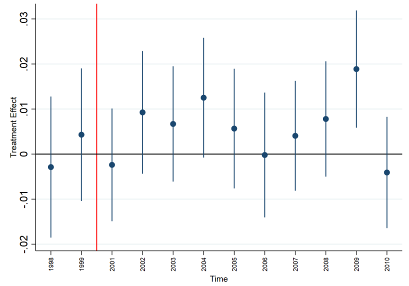
(d) Swedish high grade



(e) English pass

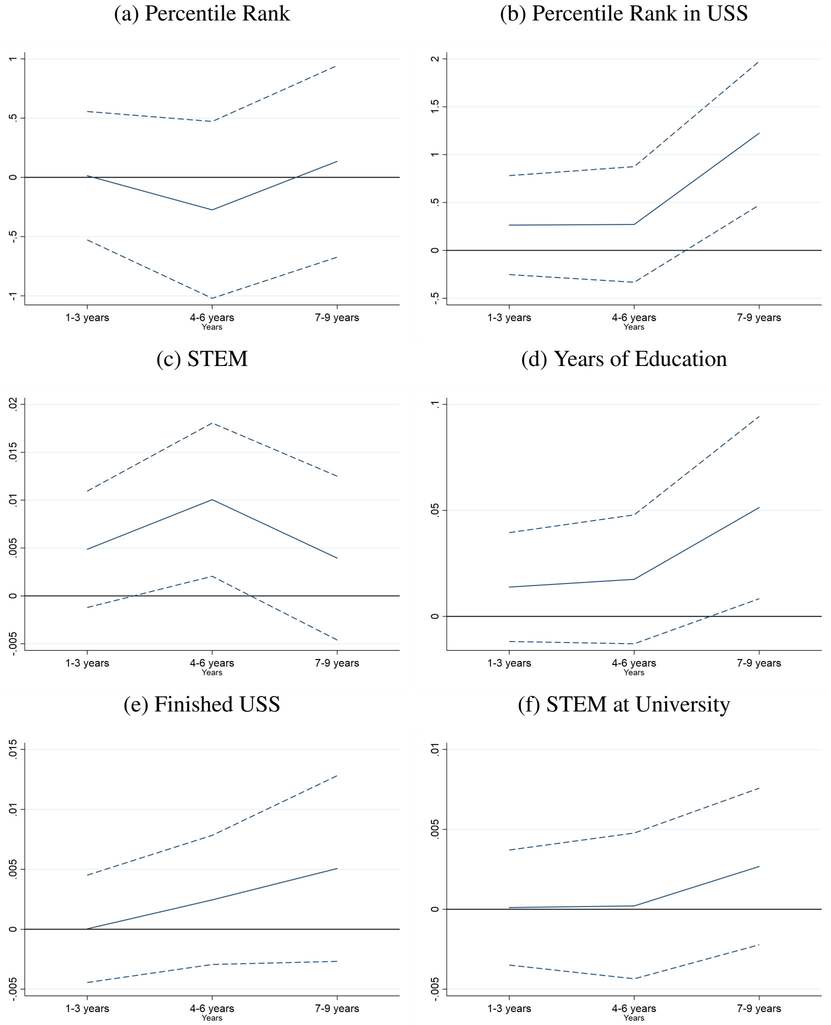


(f) English high grade



Note: This figure shows the event study of treatment on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d) English pass grade (panel e), and English high grade (panel f). Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

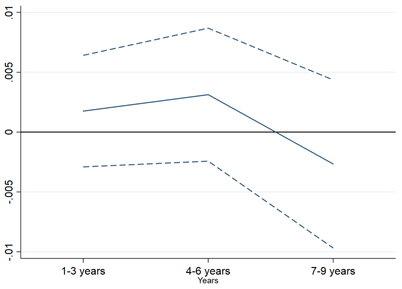
Figure (4) Effect of years of exposure to the school-based timetable on different outcomes



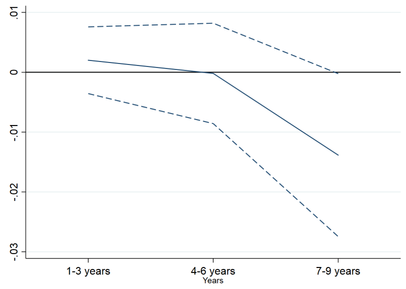
*Note:*The figure shows estimates and 95% confidence intervals (CI) from a regression on the effect of years of exposure to the school-based timetable on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f). The x-axis shows the level of exposure to the treatment: 1–3 years, 4–6 years, and 6–9 years. Control variables include LLM-year, school-fixed effect, immigration status, parent income (4 levels), and parent education (4 levels). Standard errors are clustered at the school level.

Figure (5) Effect of years of exposure to the school-based timetable on Swedish, English, and Mathematics grades.

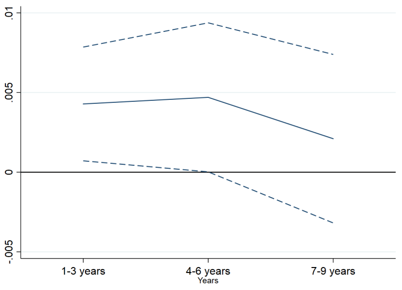
(a) Mathematics pass



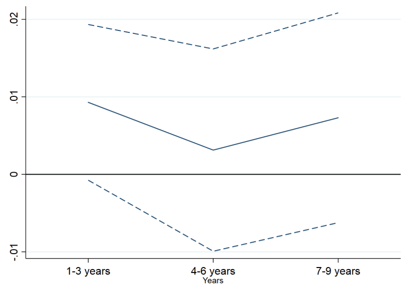
(b) Swedish pass



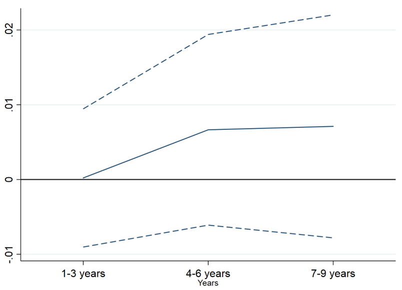
(c) English pass



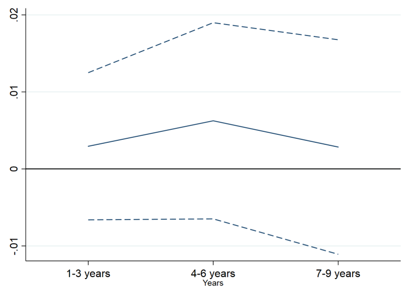
(d) Mathematics high grade



(e) Swedish high grade



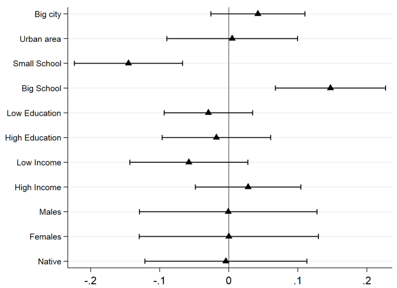
(f) English high grade



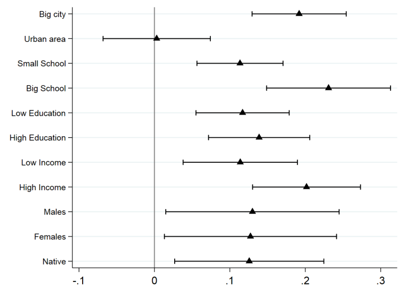
Note: The figure shows estimates and 95% confidence intervals (CI) from a regression on the effect of years of exposure to the school-based timetable on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d), English pass grade (panel e), and English high grade (panel f). The x-axis shows the level of exposure to the treatment: 1–3 years, 4–6 years, and 6–9 years. Control variables include LLM-year, school-fixed effect, immigration status, parent income (4 levels), and parent education (4 levels). Standard errors are clustered at the school level.

Figure (6) Heterogeneity analyses, linear effect of school-based timetable on different outcomes

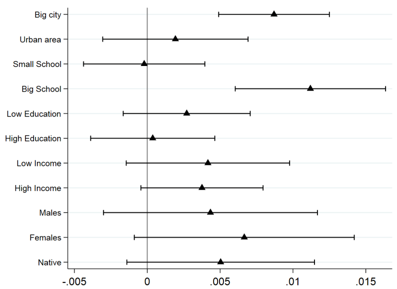
(a) Percentile Rank



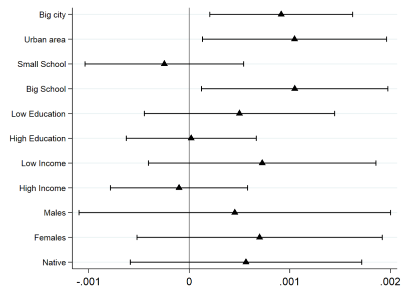
(b) Percentile Rank USS



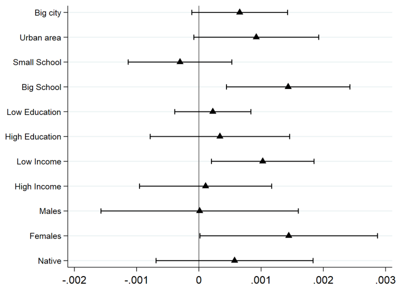
(c) Years of Education



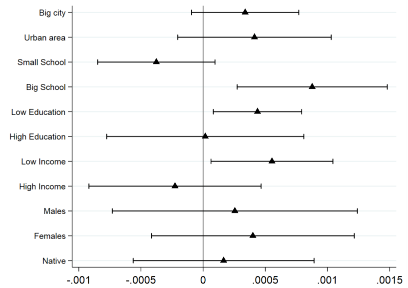
(d) Finished USS



(e) STEM at uss



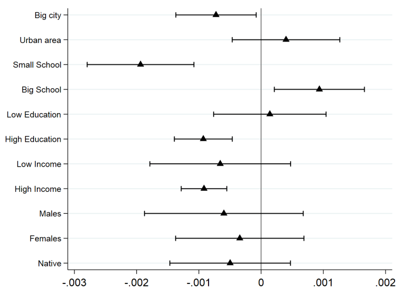
(f) STEM at University



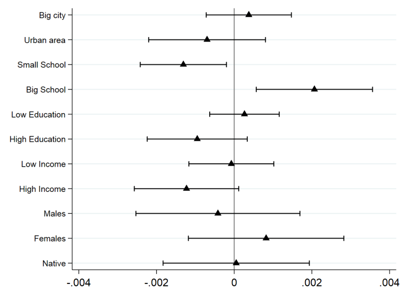
Note: This figure shows the heterogeneity analyses on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f). Each panel consists of eleven subgroups of students. The definition of these groups are presented in Appendix C. Each dot shows treatment estimates for these subgroups and 95% confidence intervals (CI). Control variables include LLM-year, school-fixed effect, and other background characteristics, excluding the main subgroups. Standard errors are clustered at the school level.

Figure (7) Heterogeneity analyses, linear effect of school-based timetable on Swedish, English, and Mathematics grades

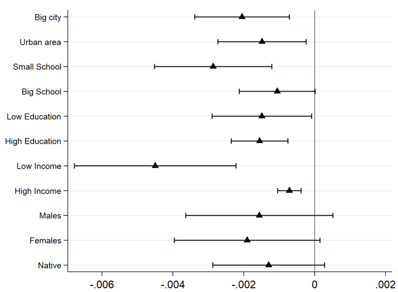
(a) Mathematics pass



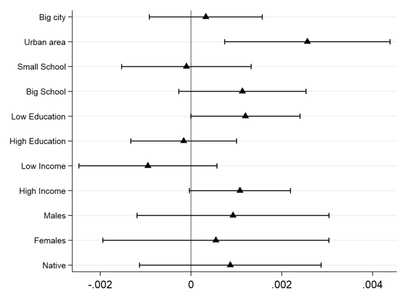
(b) Mathematics high grade



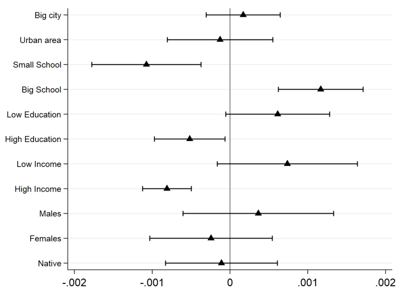
(c) Swedish pass grade



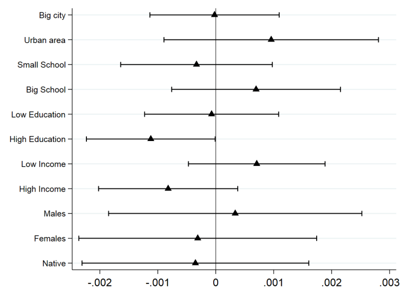
(d) Swedish high grade



(e) English pass rade



(f) English high grade



Note: Note:This figure shows the heterogeneity analyses on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f). Each panel consists of eleven subgroups of students. The definition of these groups are presented in Appendix C. Each dot shows treatment estimates for these subgroups and 95% confidence intervals (CI). Control variables include LLM-year, school-fixed effect, and other background characteristics, excluding the main subgroups. Standard errors are clustered at the school level.

Table (1) Descriptive statistics for participating and non-participating municipalities one year before the treatment (1999)

	Participant	Non-Participant
Number of Students	683.8 (841.3)	243.9 (357.5)
Female (%)	48.37 (2.461)	48.62 (3.941)
Immigrants (%)	13.43 (7.656)	9.297 (5.861)
Born Out of Sweden (%)	9.433 (4.503)	6.358 (3.632)
Average Income	4.436 (0.624)	4.181 (0.720)
Median Income	3.976 (0.261)	3.870 (0.259)
Employment	76.82 (3.986)	76.28 (4.181)
Inhabitants Per Square Kilometer	342 (788.6)	59 (144.7)
Total Cost per student	56288.4 (7688.9)	54740.6 (4848.4)
Right-wing share (%)	0.41 (0.49)	0.33 (0.47)
Number of municipalities	70	219

Note: This table shows the descriptive statistics of participating and non-participating municipalities in the reform one year before the experiment.

Table (2) Descriptive statistics for treated and control schools one year before the treatment (1999)

	Treatment	Control All	Δ All	Control A	Δ A	Control B	Δ B
Immigrant(%)	3.38	3.73	-0.03	5.89	-0.20	2.59	0.09
	(8.20)	(11.80)		(15.47)		(9.13)	
Females(%)	48.66	47.38	0.12	47.19	0.12	47.48	0.12
	(5.84)	(13.98)		(15.81)		(12.91)	
Fathers YoE	11.49	11.47	0.03	11.75	-0.24	11.31	0.19
	(0.91)	(1.11)		(1.23)		(1.01)	
Mother YoE	11.53	11.51	0.02	11.67	-0.13	11.43	0.12
	(0.81)	(1.00)		(1.24)		(0.83)	
Father income (log)	7.17	7.12	0.12	7.13	0.09	7.12	0.15
	(0.37)	(0.47)		(0.56)		(0.41)	
Mother income (log)	6.86	6.81	0.13	6.86	0.01	6.78	0.21
	(0.33)	(0.43)		(0.47)		(0.41)	
Municipality size (1–10)	3.89	4.15	-0.12	3.15	0.36	4.69	-0.36
	(2.09)	(2.39)		(1.98)		(2.42)	
School size	100.52	77.44	0.51	75.61	0.54	78.41	0.50
	(41.64)	(48.18)		(50.00)		(47.20)	
GPA (0–320)	200.67	198.45	0.10	195.62	0.17	199.93	0.04
	(13.98)	(29.89)		(38.57)		(24.01)	
Percentile Rank (1–100)	50.11	49.38	0.07	48.70	0.12	49.74	0.04
	(6.81)	(12.32)		(15.18)		(10.51)	
Choosing STEM (%)	21.49	20.46	0.09	21.70	-0.02	19.81	0.16
	(8.66)	(12.61)		(14.12)		(11.69)	
Years of education	12.51	12.39	0.19	12.32	0.26	12.43	0.15
	(0.43)	(0.84)		(1.00)		(0.73)	
Low education (%)	16.98	19.53	-0.18	21.86	-0.30	18.30	-0.11
	(8.50)	(17.85)		(21.35)		(15.55)	
Finished USS(%)	84.28	81.94	0.17	78.00	0.38	84.03	0.02
	(7.87)	(17.66)		(22.09)		(14.36)	
Observations	183	1009		350	1	659	

Note: This table shows mean value and standard deviation for different variables in the analysis for the treated group, ALL control, Control A, and Control B. The normalized difference is indicated next to each control group.

Table (3) Difference-in-Difference estimates of effect of school timetable reform on a different outcome

	Extensive Marginal Effect			Intensive Marginal Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Nine grade GPA rank						
Treatment	0.19	0.16	-0.037	0.071	0.056	-0.0013
	(0.31)	(0.30)	(0.34)	(0.059)	(0.058)	(0.059)
Placebo	0.27	0.24	0.32	0.34	0.29	0.34
	(0.26)	(0.26)	(0.28)	(0.28)	(0.27)	(0.30)
Upper Secondary GPA rank						
Treatment	0.19	0.16	-0.037	0.071	0.056	-0.0013
	(0.31)	(0.30)	(0.34)	(0.059)	(0.058)	(0.059)
Placebo	0.27	0.24	0.32	0.34	0.29	0.34
	(0.26)	(0.26)	(0.28)	(0.28)	(0.27)	(0.30)
STEM at Upper Secondary School						
Treatment	0.0047	0.0045	0.0059*	0.0013**	0.0012**	0.00066
	(0.0034)	(0.0034)	(0.0035)	(0.00059)	(0.00059)	(0.00062)
Placebo	-	-	-	-	-	-
Years of education at age of 25						
Treatment	0.032**	0.030**	0.019	0.012***	0.010***	0.0054*
	(0.015)	(0.015)	(0.015)	(0.0035)	(0.0034)	(0.0032)
Placebo	0.015	0.013	0.017	0.021	0.019	0.017
	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	(0.014)
Finished upper secondary school at age of 25						
Treatment	0.0031	0.0030	0.0013	0.0013**	0.0011**	0.00055
	(0.0025)	(0.0025)	(0.0026)	(0.00058)	(0.00057)	(0.00058)
Placebo	0.0046	0.0046	0.0035	0.0055**	0.0053*	0.0039
	(0.0028)	(0.0029)	(0.0029)	(0.0028)	(0.0028)	(0.0029)
Finished STEM at university						
Treatment	0.0024	0.0023	0.00058	0.0010***	0.00096***	0.00032
	(0.0020)	(0.0020)	(0.0021)	(0.00035)	(0.00034)	(0.00035)
Placebo	-0.00026	-0.00038	-0.00043	0.00091	0.00068	0.000032
	(0.0022)	(0.0022)	(0.0022)	(0.0021)	(0.0021)	(0.0021)
Share of private school	N	Y	Y	N	Y	Y
LLM-Year fixed effect	N	N	Y	N	N	Y

Note: This table reports the estimated effects of the timetable reform on different educational outcomes using the Difference-in-Difference (DiD) design as specified by equation 1. Columns 1-3 show the extensive marginal effect of the treatment (0, 1). Columns 4–6 show the intensive marginal (0–9 treatment). The placebo shows the dummy variable's estimates for all treated schools for 1990–1994. Columns 1 and 4 show the treatment effect when no control for municipalities was used. In columns 2 and 5, the share of private schools in the municipality was added. In columns 3 and 6, LLM-year fixed effects are also added to the model. In all models, school fixed effect, year fixed effect, parental education, parental income, immigration status, the month of birth, and the gender of the students were added to the model.

Table (4) Difference-in-Difference estimates of effect of school timetable reform on Swedish, English, and mathematics grades

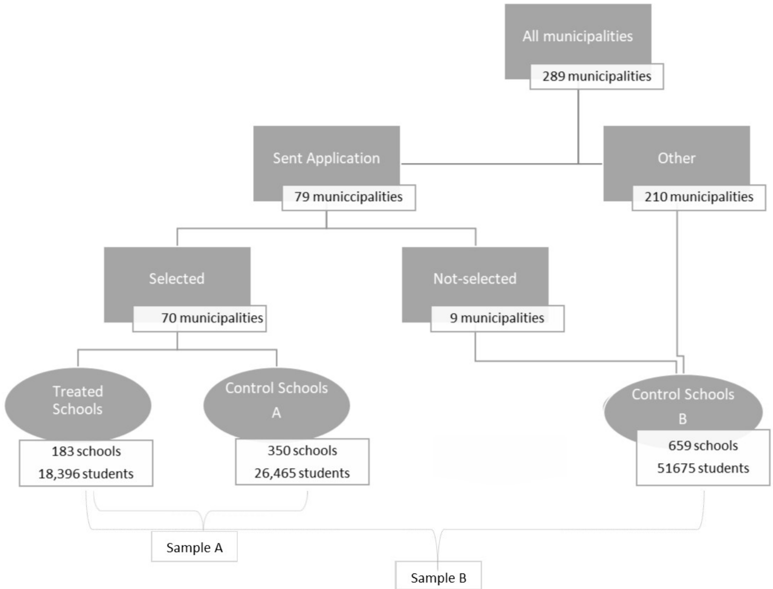
	Extensive Marginal Effect			Intensive Marginal Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Swedish Pass						
Treatment	-0.0033	-0.0032	-0.0020	-0.0022**	-0.0022**	-0.0018*
	(0.0038)	(0.0038)	(0.0041)	(0.00095)	(0.00095)	(0.0010)
Swedish High Grade						
Treatment	0.0059	0.0054	0.0033	0.0010	0.00091	0.00075
	(0.0055)	(0.0056)	(0.0060)	(0.00097)	(0.00097)	(0.0010)
English Pass						
Treatment	0.0060***	0.0059***	0.0039*	0.00075**	0.00069*	0.000051
	(0.0021)	(0.0021)	(0.0021)	(0.00036)	(0.00036)	(0.00038)
English High Grade						
Treatment	0.0072	0.0067	0.0037	0.0014	0.0013	7.2e-06
	(0.0055)	(0.0055)	(0.0060)	(0.00092)	(0.00092)	(0.00095)
Mathematics Pass Grade						
Treatment	0.0032	0.0031	0.0011	0.00025	0.00020	-0.00050
	(0.0026)	(0.0026)	(0.0028)	(0.00051)	(0.00051)	(0.00052)
Mathematics High Grade						
Treatment	0.014**	0.014**	0.0073	0.0018**	0.0017*	0.00017
	(0.0057)	(0.0057)	(0.0061)	(0.00086)	(0.00086)	(0.00091)
Share of private school	N	Y	Y	N	Y	Y
LLM-Year fixed effect	N	N	Y	N	N	Y

Note: This table reports the estimated effects of the timetable reform on grades in Swedish, English, and Mathematics using the Difference-in-Difference (DiD) design as specified by equation 1. Columns 1–3 show the extensive marginal effect of the treatment (0, 1). Columns 4–6 show the intensive marginal (0–9 treatment). Columns 1 and 4 show the treatment effect when no control for municipalities was used. In columns 2 and 5, the share of private schools in the municipality was added. In columns 3 and 6, LLM-year fixed effects are also added to the model. In all models, school fixed effect, year fixed effect, parental education, parental income, immigration status, the month of birth, and the gender of the students were added to the model.

Appendices

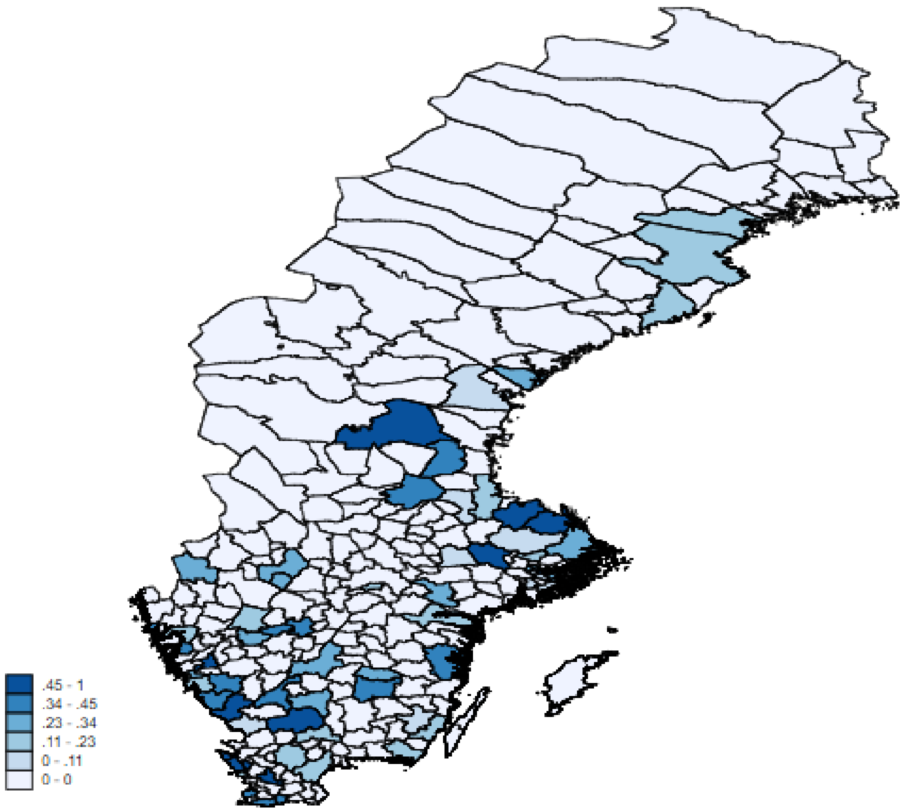
A Figures

Figure (A.1) The experiment implementation process



Note: This figure shows how the reform was implemented in 2000. First, the Ministry of Education invited all municipalities to apply for participating in the experiment. Next, selected municipalities were allowed to decide which of their schools would participate in the experiment.

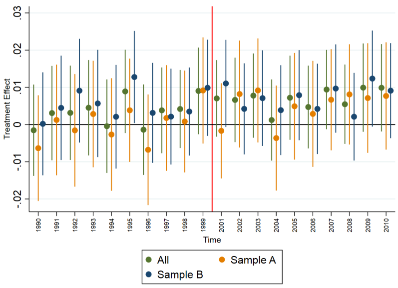
Figure (A.2)   The distribution of the affected students across municipalities in Sweden



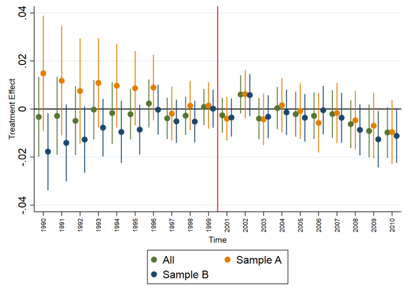
Note: This figure shows the share of students affected by the reform in 2000 in each municipalities in Sweden. The dark blue municipalities had a higher share of the treated students, and white municipalities were the control group.

Figure (A.3) Balance Test– Sample A and B

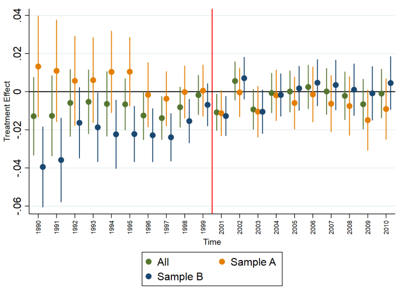
(a) Male



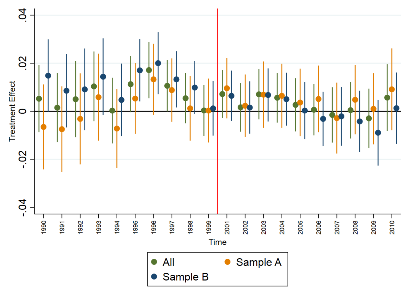
(b) Immigrant



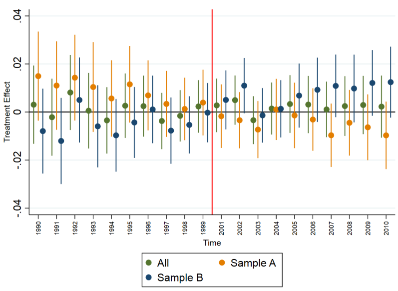
(c) Low Income Parent



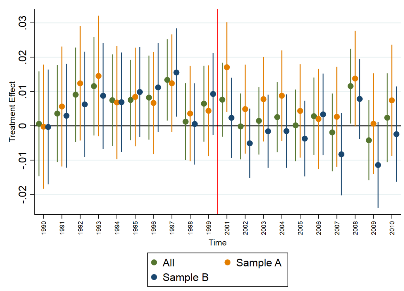
(d) High Income Parent



(e) Low Education Parent



(f) High Education Parent



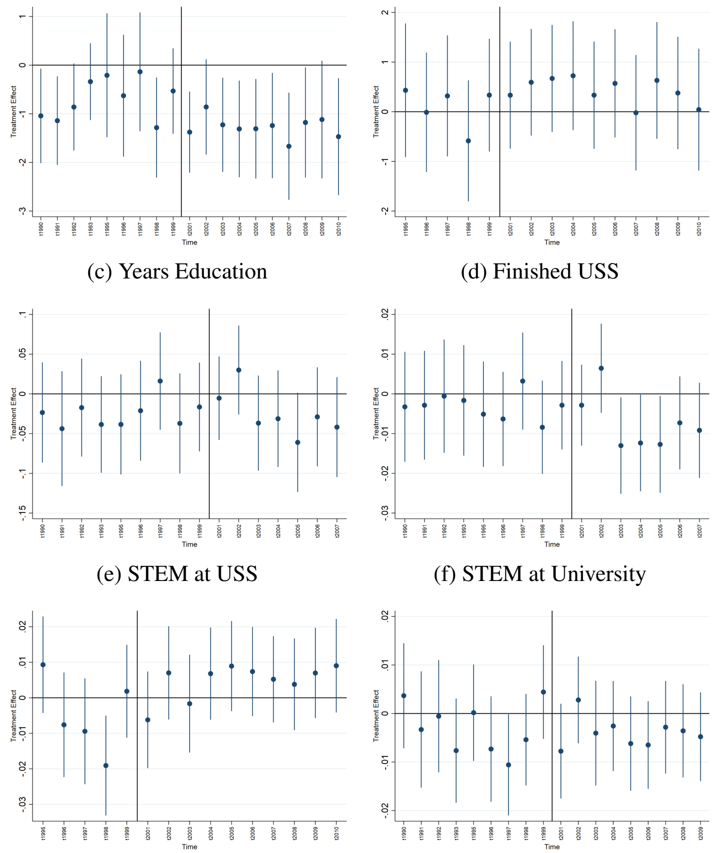
Note: This figure shows the event study type analysis for sample A, sample B and sample All. The analysis is performed for male (panel a) immigrant students (panel b), parents' incomes (panel c and d), and parents' education levels (panel e and f). Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) each year.

Figure (A.4) Event study graph of the effect of school timetable on different outcomes– Sample A

(a) Percentile Rank (b) Percentile Rank USS

(c) Years Education (d) Finished USS

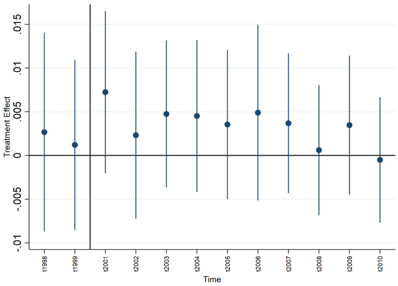
(e) STEM at USS (f) STEM at University



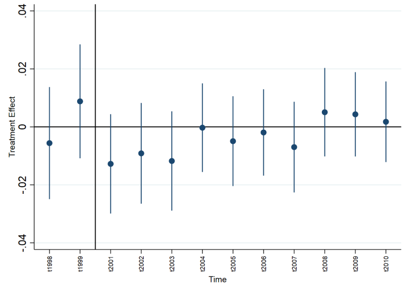
Note: This figure shows the event study of treatment on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f) in sample A. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.5) Event study graph of the effect of school timetable on Swedish, English, and Mathematics grades– Sample A

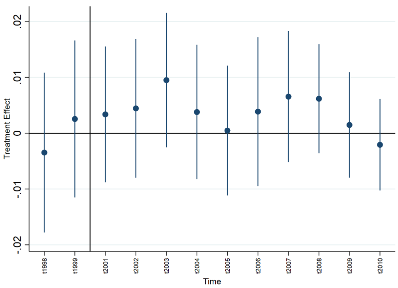
(a) Mathematics pass



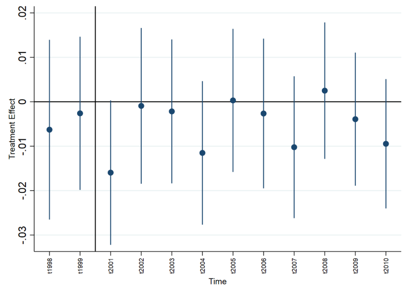
(b) Mathematics high grade



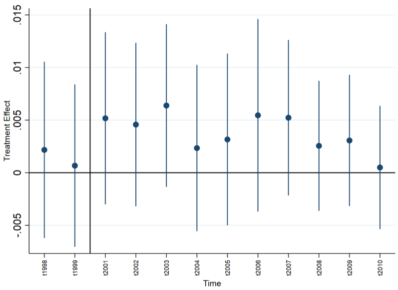
(c) Swedish pass



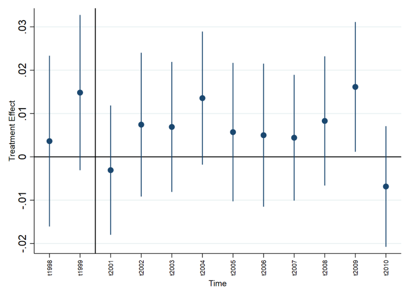
(d) Swedish high grade



(e) English pass

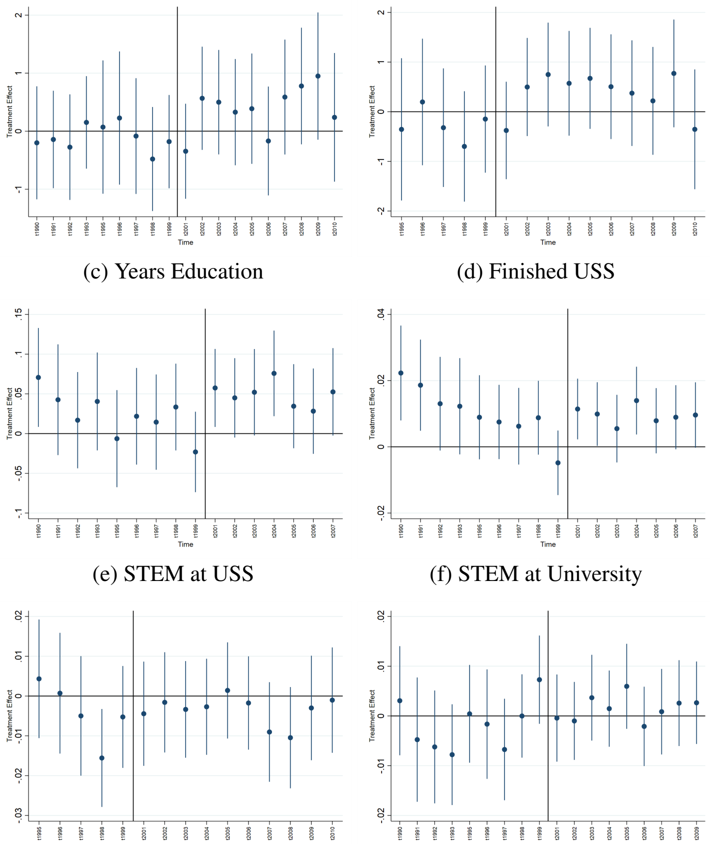


(f) English high grade



Note: This figure shows the event study of treatment on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d) English pass grade (panel e), and English high grade (panel f) in sample A. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.6) Event study graph of the effect of school timetable on different outcomes– Sample B

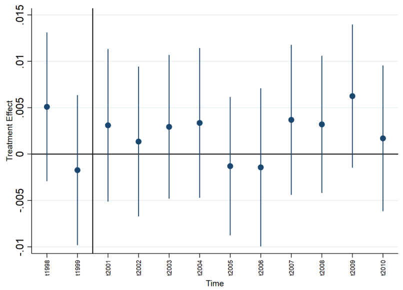


(a) Percentile Rank
(b) Percentile Rank USS
(c) Years Education
(d) Finished USS
(e) STEM at USS
(f) STEM at University

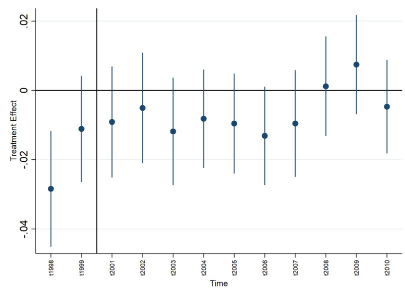
Note: This figure shows the event study of treatment on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f) in sample B. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.7) Event study graph of the effect of school timetable on Swedish, English, and Mathematics grades– Sample B

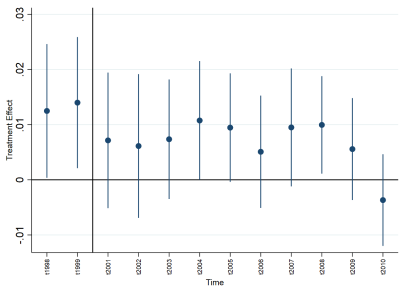
(a) Mathematics pass



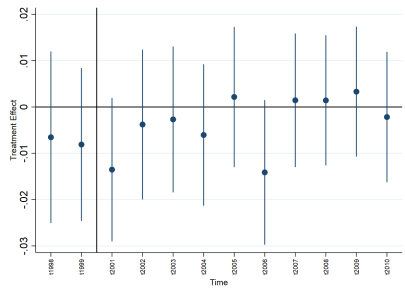
(b) Mathematics high grade



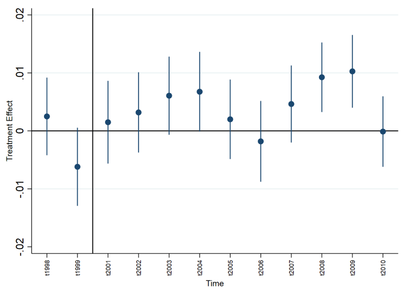
(c) Swedish pass



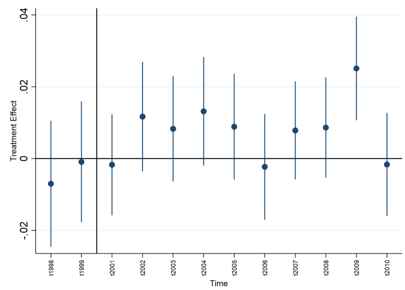
(d) Swedish high grade



(e) English pass

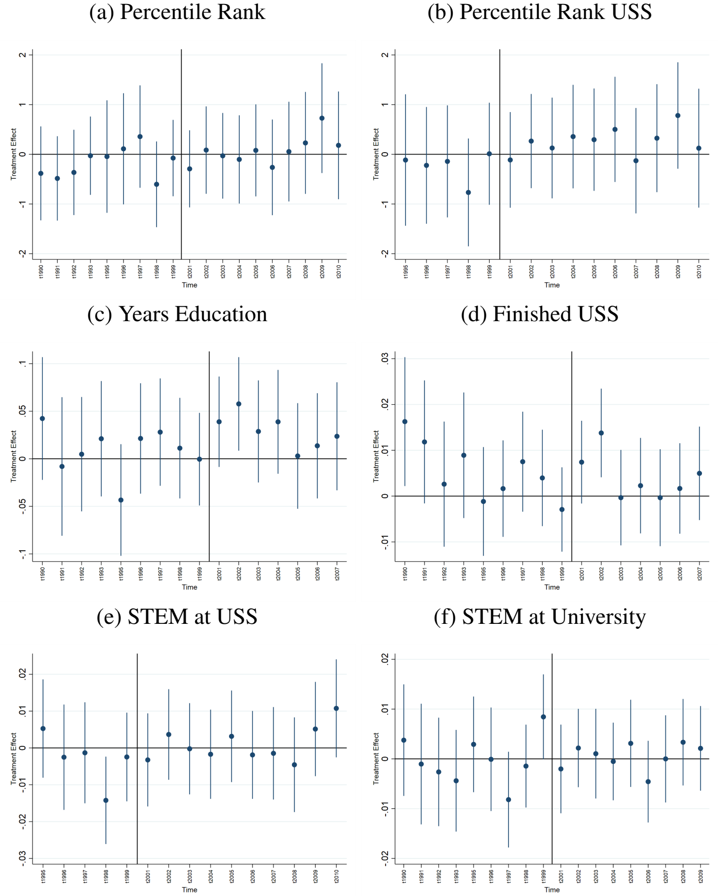


(f) English high grade



Note: This figure shows the event study of treatment on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d) English pass grade (panel e), and English high grade (panel f) in sample B. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

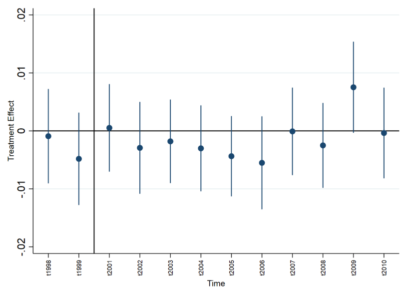
Figure (A.8) Event study graph of the effect of school timetable on different outcomes– Matched All



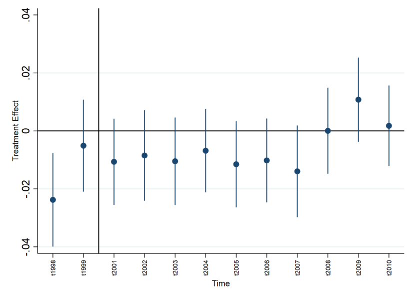
Note: This figure shows the event study of treatment on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f) in sample matched all. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.9) Event study graph of the effect of school timetable on Swedish, English, and Mathematics grades– Matched All

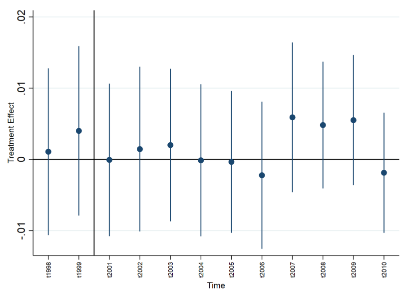
(a) Mathematics pass



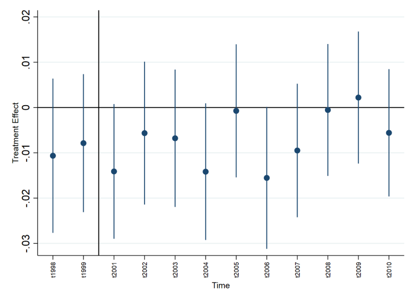
(b) Mathematics high grade



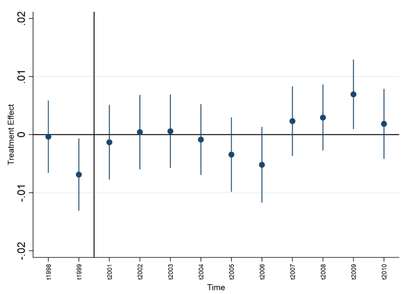
(c) Swedish pass



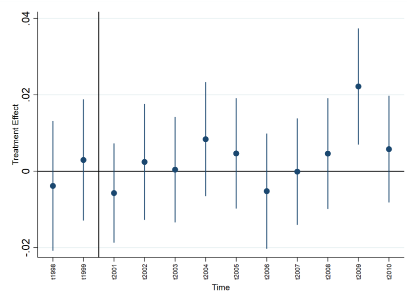
(d) Swedish high grade



(e) English pass



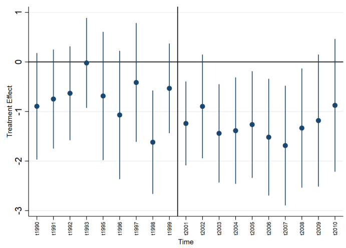
(f) English high grade



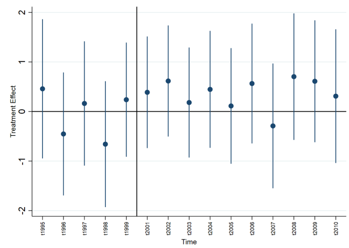
Note: This figure shows the event study of treatment on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d) English pass grade (panel e), and English high grade (panel f) in the sample matched all. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.10) Event study graph of the effect of school timetable on different outcomes– Matched A

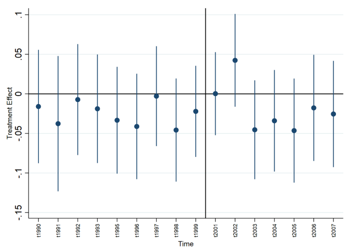
(a) Percentile Rank



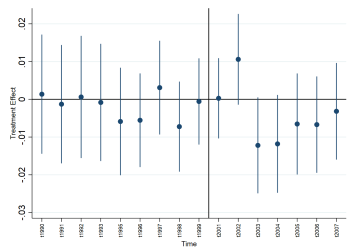
(b) Percentile Rank USS



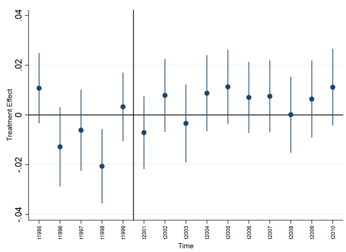
(c) Years Education



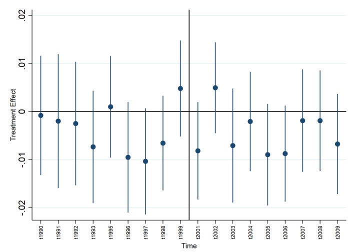
(d) Finished USS



(e) STEM at USS



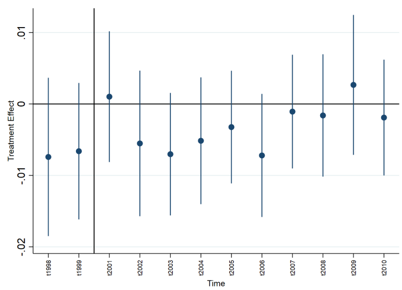
(f) STEM at University



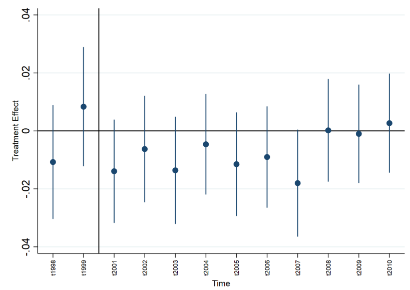
Note: This figure shows the event study of treatment on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f) in sample matched A. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.11) Event study graph of the effect of school timetable on Swedish, English, and Mathematics grades– Matched A

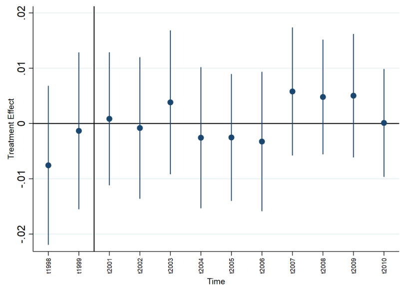
(a) Mathematics pass



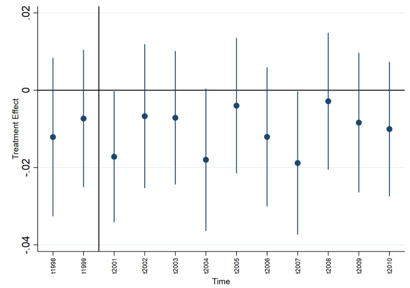
(b) Mathematics high grade



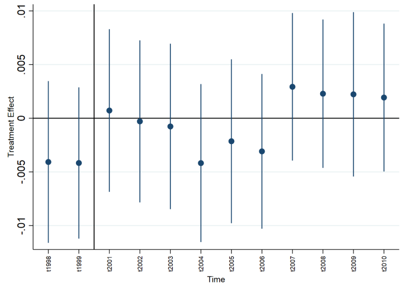
(c) Swedish pass



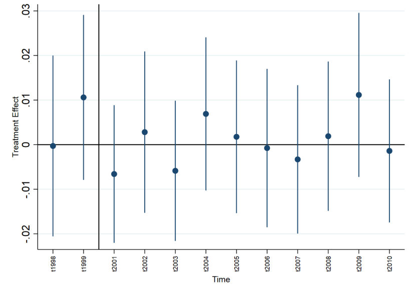
(d) Swedish high grade



(e) English pass



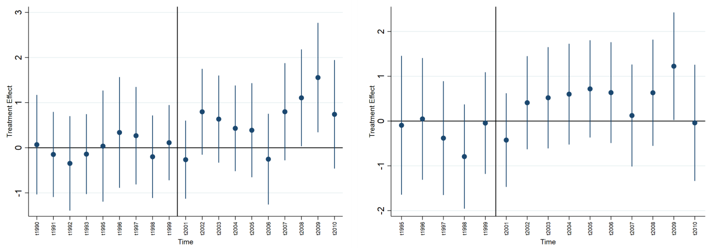
(f) English high grade



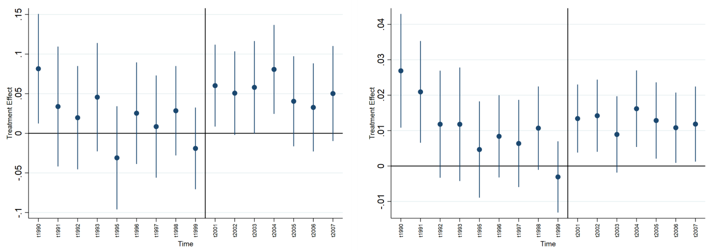
Note: This figure shows the event study of treatment on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d) English pass grade (panel e), and English high grade (panel f) in the sample matched A. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.12) Event study graph of the effect of school timetable on different outcomes– Matched B

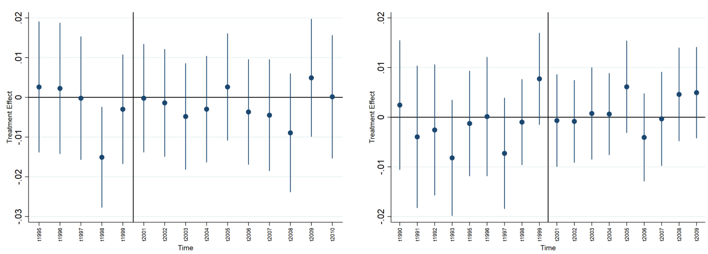
(a) Percentile Rank (b) Percentile Rank USS



(c) Years Education (d) Finished USS



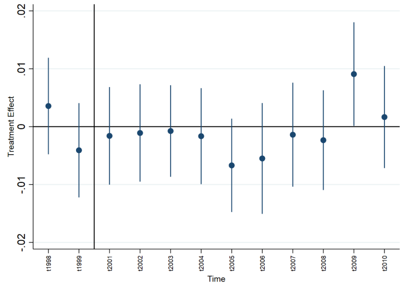
(e) STEM at USS (f) STEM at University



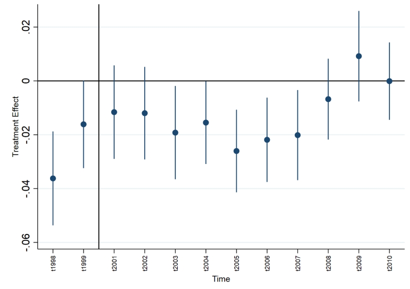
Note: This figure shows the event study of treatment on grade 9 percentile rank GPA (panel a), upper secondary percentile rank GPA (panel b), the years of education at the age of 25 (panel c), the probability of graduating from upper secondary school within 12 years (panel d), the probability of enrolling in a STEM field in upper secondary school (panel e), and the probability to finish STEM field at university (panel f) in sample matched B. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

Figure (A.13) Event study graph of the effect of school timetable on Swedish, English, and Mathematics grades– Matched B

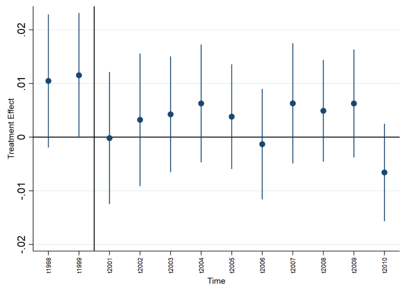
(a) Mathematics pass



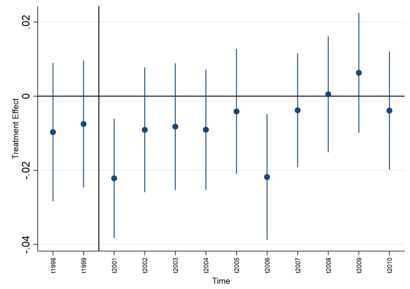
(b) Mathematics high grade



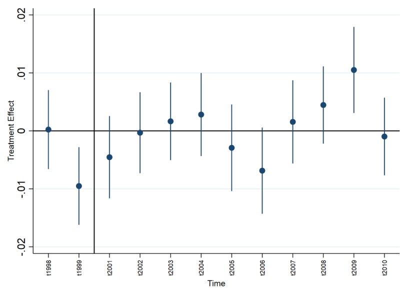
(c) Swedish pass



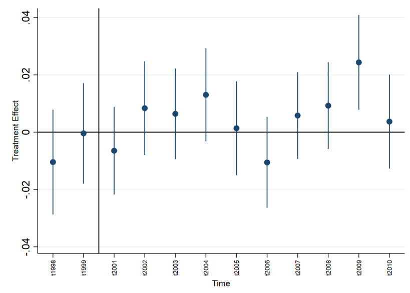
(d) Swedish high grade



(e) English pass



(f) English high grade



Note: This figure shows the event study of treatment on mathematic pass grade (panel a), mathematics high grade (panel b), Swedish pass grade (panel c), Swedish high grade (panel d) English pass grade (panel e), and English high grade (panel f) in the sample matched B. Each dot shows treatment estimates for these characteristics and 95% confidence intervals (CI) in each year. The vertical line is the reform year.

B Tables

Table (B.5) Difference-in-Difference estimates of effect of school timetable reform on a different outcome –Sample A

	Extensive Marginal Effect			Intensive Marginal Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Nine grade GPA rank						
Treatment	-0.80**	-0.61	-0.57	-0.088	-0.067	-0.098
	(0.41)	(0.39)	(0.46)	(0.071)	(0.070)	(0.073)
Placebo	0.15	0.032	0.0045	0.42	0.23	0.094
	(0.29)	(0.29)	(0.32)	(0.32)	(0.31)	(0.35)
Upper Secondary GPA rank						
Treatment	0.50	0.42	0.53	0.15**	0.14**	0.13**
	(0.34)	(0.36)	(0.36)	(0.058)	(0.060)	(0.055)
Placebo	-0.33	-0.28	-0.23	-0.22	-0.16	-0.19
	(0.39)	(0.39)	(0.38)	(0.37)	(0.36)	(0.37)
Years of education at age of 25						
Treatment	-0.013	-0.0050	-0.0016	0.0024	0.0035	0.00074
	(0.020)	(0.019)	(0.018)	(0.0043)	(0.0041)	(0.0038)
Placebo	0.0049	-0.0044	0.014	0.017	0.0055	0.016
	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)
Finished upper secondary school at age of 25						
Treatment	-0.0022	-0.0013	-0.0022	-0.00014	-0.000022	-0.00021
	(0.0031)	(0.0031)	(0.0033)	(0.00071)	(0.00070)	(0.00071)
Placebo	0.0016	0.00044	0.0022	0.0026	0.0011	0.0030
	(0.0034)	(0.0035)	(0.0036)	(0.0033)	(0.0034)	(0.0036)
STEM at Upper Secondary School						
Treatment	0.0062	0.0073*	0.0083*	0.0011	0.0012*	0.00060
	(0.0043)	(0.0043)	(0.0043)	(0.00071)	(0.00072)	(0.00071)
Placebo	-	-	-	-	-	-
Observations	717,408	716,219	716,219	717,408	716,219	716,219
Finished STEM at university						
Treatment	0.00039	0.0011	-2.4e-06	0.00064	0.00072*	0.00011
	(0.0025)	(0.0025)	(0.0026)	(0.00041)	(0.00041)	(0.00042)
Placebo	0.00037	-0.000043	-0.0018	0.0017	0.0011	-0.0016
	(0.0025)	(0.0025)	(0.0026)	(0.0024)	(0.0024)	(0.0025)
Share of private school	N	Y	Y	N	Y	Y
LLM-Year fixed effect	N	N	Y	N	N	Y

Note: This table reports the estimated effects of the timetable reform on different educational outcomes using the Difference-in-Difference (DiD) design as specified by equation 1 in sample A. Columns 1-3 show the extensive marginal effect of the treatment (0, 1). Columns 4–6 show the extensive marginal (0–9 treatment). The placebo shows the dummy variable's estimates for all treated schools for 1990–1994. Columns 1 and 4 show the treatment effect when no control for municipalities was used. In columns 2 and 5, the share of private schools in the municipality was added. In columns 3 and 6, LLM-year fixed effects are also added to the model. In all models, school fixed effect, year fixed effect, parental education, parental income, immigration status, the month of birth, and the gender of the students were added to the model

Table (B.6) Difference-in-Difference estimates of effect of school timetable reform on Swedish, English, and mathematics grades– Sample A

	Extensive Marginal Effect			Intensive Marginal Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Swedish Pass						
Treatment	0.0019	0.0016	0.0033	-0.0020*	-0.0020*	-0.00098
	(0.0049)	(0.0048)	(0.0055)	(0.0011)	(0.0011)	(0.0011)
Swedish High Grade						
Treatment	-0.0029	-0.0021	-0.000032	-0.00035	-0.00023	0.00014
	(0.0066)	(0.0066)	(0.0075)	(0.0011)	(0.0011)	(0.0012)
English Pass						
Treatment	0.00087	0.0014	0.0017	-0.00031	-0.00024	-0.00053
	(0.0027)	(0.0026)	(0.0028)	(0.00047)	(0.00047)	(0.00049)
English High Grade						
Treatment	-0.010	-0.0089	-0.0051	-0.00081	-0.00065	-0.0011
	(0.0068)	(0.0067)	(0.0076)	(0.0011)	(0.0011)	(0.0011)
Mathematics Pass Grade						
Treatment	-0.00067	-0.00029	0.00092	-0.00079	-0.00074	-0.00073
	(0.0034)	(0.0033)	(0.0041)	(0.00062)	(0.00062)	(0.00064)
Mathematics High Grade						
Treatment	0.00062	0.00089	-0.0052	-0.00013	-0.000097	-0.0013
	(0.0071)	(0.0071)	(0.0081)	(0.0010)	(0.0010)	(0.0011)
Share of private school	N	Y	Y	N	Y	Y
LLM-Year fixed effect	N	N	Y	N	N	Y

Note: This table reports the estimated effects of the timetable reform on different educational outcomes using the Difference-in-Difference (DiD) design as specified by equation 1 in sample A. Columns 1-3 show the extensive marginal effect of the treatment (0, 1). Columns 4–6 show the extensive marginal (0–9 treatment). The placebo shows the dummy variable's estimates for all treated schools for 1990–1994. Columns 1 and 4 show the treatment effect when no control for municipalities was used. In columns 2 and 5, the share of private schools in the municipality was added. In columns 3 and 6, LLM-year fixed effects are also added to the model. In all models, school fixed effect, year fixed effect, parental education, parental income, immigration status, the month of birth, and the gender of the students were added to the model.

Table (B.7) Difference-in-Difference estimates of effect of school timetable reform on a different outcome – Sample B

	Extensive Marginal Effect			Intensive Marginal Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Nine grade GPA rank						
Treatment	0.68**	0.60**	0.45	0.16***	0.14**	0.097
	(0.31)	(0.30)	(0.34)	(0.059)	(0.058)	(0.060)
Placebo	0.32	0.33	0.42	0.32	0.32	0.40
	(0.27)	(0.27)	(0.32)	(0.29)	(0.29)	(0.35)
Upper Secondary GPA rank						
Treatment	0.27	0.29	0.44	0.057	0.060	0.11**
	(0.30)	(0.30)	(0.35)	(0.054)	(0.054)	(0.056)
Placebo	0.65*	0.62*	0.21	0.64*	0.61*	0.28
	(0.36)	(0.36)	(0.39)	(0.34)	(0.34)	(0.39)
Years of education at age of 25						
Treatment	0.054***	0.050***	0.035**	0.018***	0.016***	0.011***
	(0.016)	(0.015)	(0.016)	(0.0035)	(0.0034)	(0.0034)
Placebo	0.020	0.021	0.019	0.026	0.026*	0.022
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)
Observations	1,191,881	1,188,129	1,188,129	1,191,881	1,188,129	1,188,129
Finished upper secondary school at age of 25						
Treatment	0.0056**	0.0055**	0.0035	0.0021***	0.0020***	0.0016***
	(0.0026)	(0.0026)	(0.0028)	(0.00058)	(0.00057)	(0.00061)
Placebo	0.0060**	0.0063**	0.0029	0.0074**	0.0075***	0.0042
	(0.0029)	(0.0029)	(0.0033)	(0.0029)	(0.0029)	(0.0031)
STEM at Upper Secondary School						
Treatment	0.0040	0.0032	0.0024	0.0013**	0.0011*	0.000019
	(0.0036)	(0.0036)	(0.0043)	(0.00060)	(0.00061)	(0.00072)
Placebo	-	-	-	-	-	-
Finished STEM at university						
Treatment	0.0035*	0.0031	0.00099	0.0013***	0.0012***	0.00060
	(0.0021)	(0.0021)	(0.0023)	(0.00035)	(0.00035)	(0.00038)
Placebo	-0.00048	-0.00047	-0.000061	0.00080	0.00074	0.00087
	(0.0023)	(0.0023)	(0.0026)	(0.0022)	(0.0022)	(0.0025)
Observations	1,378,648	1,374,718	1,374,718	1,378,648	1,374,718	1,374,718
Share of private school	N	Y	Y	N	Y	Y
LLM-Year fixed effect	N	N	Y	N	N	Y

Note: This table reports the estimated effects of the timetable reform on different educational outcomes using the Difference-in-Difference (DiD) design as specified by equation 1 in sample B. Columns 1-3 show the extensive marginal effect of the treatment (0, 1). Columns 4–6 show the extensive marginal (0–9 treatment). The placebo shows the dummy variable's estimates for all treated schools for 1990–1994. Columns 1 and 4 show the treatment effect when no control for municipalities was used. In columns 2 and 5, the share of private schools in the municipality was added. In columns 3 and 6, LLM-year fixed effects are also added to the model. In all models, school fixed effect, year fixed effect, parental education, parental income, immigration status, the month of birth, and the gender of the students were added to the model

Table (B.8) Difference-in-Difference estimates of effect of school timetable reform on Swedish, English, and mathematics grades– Sample B

	Extensive Marginal Effect			Intensive Marginal Effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Swedish Pass						
Treatment	-0.0059	-0.0057	-0.0078*	-0.0025***	-0.0024**	-0.0026**
	(0.0038)	(0.0039)	(0.0044)	(0.00096)	(0.00097)	(0.0011)
Swedish High Grade						
Treatment	0.010*	0.0092	0.0044	0.0018*	0.0016	0.0015
	(0.0058)	(0.0058)	(0.0068)	(0.00100)	(0.00100)	(0.0011)
English Pass						
Treatment	0.0085***	0.0081***	0.0052**	0.0012***	0.0011***	0.00054
	(0.0022)	(0.0022)	(0.0024)	(0.00036)	(0.00036)	(0.00040)
English High Grade						
Treatment	0.015***	0.014**	0.0100	0.0026***	0.0023**	0.00095
	(0.0058)	(0.0057)	(0.0068)	(0.00095)	(0.00094)	(0.0010)
Mathematics Pass Grade						
Treatment	0.0051*	0.0049*	0.00014	0.00073	0.00065	-0.00027
	(0.0027)	(0.0027)	(0.0029)	(0.00052)	(0.00051)	(0.00053)
Mathematics High Grade						
Treatment	0.021***	0.020***	0.014**	0.0026***	0.0023***	0.00071
	(0.0059)	(0.0059)	(0.0064)	(0.00088)	(0.00089)	(0.00098)
Share of private school	N	Y	Y	N	Y	Y
LLM-Year fixed effect	N	N	Y	N	N	Y

Note: This table reports the estimated effects of the timetable reform on different educational outcomes using the Difference-in-Difference (DiD) design as specified by equation 1 in sample B. Columns 1-3 show the extensive marginal effect of the treatment (0, 1). Columns 4–6 show the extensive marginal (0–9 treatment). The placebo shows the dummy variable's estimates for all treated schools for 1990–1994. Columns 1 and 4 show the treatment effect when no control for municipalities was used. In columns 2 and 5, the share of private schools in the municipality was added. In columns 3 and 6, LLM-year fixed effects are also added to the model. In all models, school fixed effect, year fixed effect, parental education, parental income, immigration status, the month of birth, and the gender of the students were added to the model.

Table (B.9) Number of students with different level of treatment

Year	Control Group			Level of Treatment			Total
	All	A	B	1–3 years	4–6 years	7–9 years	
2000	79,967	26,876	53,091	19,551	0	0	99,518
2001	83,660	28,126	55,534	20,355	0	0	104,015
2002	85,560	28,912	56,648	21,050	0	0	106,610
2003	87,817	29,517	58,300	7,305	13,684	0	108,806
2004	93,811	31,915	61,896	7,943	13,972	0	115,726
2005	96,822	33,605	63,217	7,782	14,650	0	119,254
2006	102,843	36,098	66,745	8,199	4,626	10,305	125,973
2007	102,430	36,622	65,808	8,122	4,614	9,931	125,097
2008	100,996	36,697	64,299	7,590	4,538	9,701	122,825
2009	98,185	36,464	61,721	7,063	4,214	9,046	118,508
2010	94,753	35,703	59,050	6,795	3,914	8,565	114,027
2011	88,667	34,015	54,652	5,913	3,670	8,225	106,475
2012	83,456	32,574	50,882	5,539	3,360	7,339	99,694
Total	1,198,967	427,124	771,843	133,207	71,242	63,112	1,466,528

Note: This table shows the number of students who belong to treatment and control groups.

Table (B.10) Descriptive statistics for treated and control group one year before the treatment (1999)–Matched group

	Treatment	Control All	Δ All	Control A	Δ A	Control B	Δ B
Immigrant(%)	3.38	3.51	-0.01	4.15	-0.08	2.68	0.09
	(8.20)	(10.40)		(10.08)		(8.25)	
Females(%)	48.66	47.84	0.11	47.72	0.09	48.08	0.07
	(5.84)	(9.14)		(12.97)		(9.17)	
Fathers Years of education	11.49	11.47	0.02	11.63	-0.14	11.40	0.10
	(0.91)	(1.00)		(1.03)		(0.97)	
Mother Years of education	11.53	11.50	0.04	11.60	-0.07	11.49	0.05
	(0.81)	(0.89)		(1.04)		(0.78)	
Father income (log)	7.17	7.18	-0.01	7.16	0.04	7.18	-0.00
	(0.37)	(0.36)		(0.49)		(0.33)	
Mother income (log)	6.86	6.85	0.04	6.84	0.04	6.85	0.03
	(0.33)	(0.37)		(0.45)		(0.29)	
Municipality size (1–10)	3.89	3.95	-0.03	3.32	0.28	4.45	-0.25
	(2.09)	(2.35)		(1.99)		(2.36)	
School size	100.52	92.60	0.19	86.68	0.31	90.83	0.23
	(41.64)	(42.96)		(47.74)		(43.60)	
GPA (0–320)	200.67	200.69	-0.00	196.51	0.15	201.50	-0.05
	(13.98)	(23.98)		(35.69)		(20.51)	
Percentile Rank (1–100)	50.11	50.29	-0.02	49.02	0.10	50.46	-0.04
	(6.81)	(10.21)		(13.63)		(8.82)	
Choosing STEM (%)	21.49	21.33	0.02	21.52	-0.00	20.71	0.08
	(8.66)	(10.10)		(12.76)		(9.67)	
Years of education	12.51	12.49	0.04	12.41	0.14	12.54	-0.05
	(0.43)	(0.64)		(0.88)		(0.52)	
Low education (%)	16.98	17.43	-0.04	19.75	-0.20	16.13	0.09
	(8.50)	(12.84)		(17.80)		(10.32)	
Finished Upper Secondarly (%)	84.28	83.95	0.03	80.69	0.25	85.91	-0.19
	(7.87)	(13.01)		(18.51)		(9.18)	
Observations	183	539		273		435	

Note: This table shows descriptive statistics for the treated group, All control, control A, and control B after matching. Besides each control group, the normalized difference is calculated.

C Defining Variable in Heterogeneity Section

- Large city and Urban area: This variable followed Swedish municipalities' 2017 classification. The municipalities are categorized into nine groups, from very large municipalities (category 1) to rural municipalities with a visitor industry (category 9). Large cities are categorized 1–4 and rural area are categorized 5-9.

- Small and large schools: To define the school size, the number of students in each school in each year was sorted from the smallest to the largest in 100 equal groups. Small schools were ranked 1–50, and large schools were ranked 51–100.

- Low education parents: A dummy variable were used to indicate whether both parents had a low level of education (more than 11 years)

- High education parents: High education parents meant that at least one of the parents had education of more than 15 years.

- Low and high parent income: To construct this variable, the income of both parents was added when students were in grade 9. The parent income was sorted from lowest to the highest. The first quartile of the parents' income rank in each year is categorized as low income, and the fourth quartile is classified as high income.

III

Short-term and long-term effects of GDP on traffic deaths in 18 OECD countries, 1960–2011

Iman Dadgar, Thor Norström

▸ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jech-2015-207138).

Swedish Institute for Social Research, Stockholm University, Stockholm, Sweden

Correspondence to
Professor Thor Norström, Swedish Institute for Social Research, Stockholm University, S-106 91 Stockholm, Sweden; totto@sofi.su.se

Received 23 December 2015
Revised 20 June 2016
Accepted 19 July 2016
Published Online First
16 August 2016

To cite: Dadgar I, Norström T. *J Epidemiol Community Health* 2017;**71**:146–153.

ABSTRACT

Background Research suggests that increases in gross domestic product (GDP) lead to increases in traffic deaths plausibly due to the increased road traffic induced by an expanding economy. However, there also seems to exist a long-term effect of economic growth that is manifested in improved traffic safety and reduced rates of traffic deaths. Previous studies focus on either the short-term, procyclical effect, or the long-term, protective effect. The aim of the present study is to estimate the short-term and long-term effects jointly in order to assess the net impact of GDP on traffic mortality.

Methods We extracted traffic death rates for the period 1960–2011 from the WHO Mortality Database for 18 OECD countries. Data on GDP/capita were obtained from the Maddison Project. We performed error correction modelling to estimate the short-term and long-term effects of GDP on the traffic death rates.

Results The estimates from the error correction modelling for the entire study period suggested that a one-unit increase (US\$1000) in GDP/capita yields an instantaneous short-term increase in the traffic death rate by 0.58 ($p<0.001$), and a long-term decrease equal to −1.59 ($p<0.001$). However, period-specific analyses revealed a structural break implying that the procyclical effect outweighs the protective effect in the period prior to 1976, whereas the reverse is true for the period 1976–2011.

Conclusions An increase in GDP leads to an immediate increase in traffic deaths. However, after the mid-1970s this short-term effect is more than outweighed by a markedly stronger protective long-term effect, whereas the reverse is true for the period before the mid-1970s.

INTRODUCTION

In 2013, 1.25 million lives worldwide were lost on the roads, which makes traffic crashes the ninth leading cause of death.[1] It is thus of great importance to get a better understanding of the driving forces behind changes in traffic deaths. The present paper will focus on the role of economic development as indicated by per-capita gross domestic product (GDP).

Previous research shows that increases in GDP are associated with increases in traffic deaths; this is a short-term effect mainly due to the increased road traffic induced by an expanding economy. However, at least in high-income countries, there seems to exist a long-term effect of economic growth that is manifested in improved traffic safety and reduced rates of traffic deaths. Extant research in the field tends to focus on either the short-term, procyclical effect, or the long-term, protective effect. However, both of these effects need to be considered jointly in order to assess the net impact of GDP on traffic mortality. In the present paper, we achieve this by analysing cross-sectional time series data for 18 affluent countries spanning the time period 1960–2011.

BACKGROUND

The relation between economic fluctuations and population health is complex and seemingly contradictory. This may explain why the received wisdom concerning this relationship has undergone some quite substantial shifts. It is clear that economic downturns in past historical centuries led to severe malnutrition and starvation and thus worsened population health. Economic growth, on the other hand, was conducive to education, improved sanitation and living conditions and, in the end, lowered mortality.[2 3] However, as demonstrated by Preston,[4] there is a diminishing health return to economic growth, and there are even indications that economic downturns in highly industrialised societies may improve population health. The explanation to this counterintuitive finding is that although a downturn in all probability has a detrimental health effect on those who are severely hit, for example, by losing their jobs, this negative effect may be more or less offset by a beneficial health effect on the remaining, and much larger, part of the population. Several plausible mechanisms underlying the latter effect have been suggested and substantiated. A slowdown in the economy is thus associated with reduced overtime and work-related stress, less driving and car crashes, less air pollutions and reduced intake of unhealthy products such as alcohol and tobacco.[5–7] Already in the early 20th century, there were reports[8] suggesting that economic booms were associated with above average mortality, whereas the opposite was true for economic downturns. However, these results were ignored for a long time, probably because they appeared to run counter to intuition.[9]

The investigation by Ruhm[7] was one of the first well-designed studies in the field; on the basis of fixed-effects modelling of US state data for the period 1972–1991, he found that recessions are associated with lowered all-cause mortality. More detailed, cause-specific, analyses revealed that traffic deaths especially decreased during bad times. The procyclical relation between macroeconomic conditions and traffic deaths is echoed in other single-country studies, including Neumayer,[10] who analysed German state-panel data, Farmer[11] using US monthly time-series data and studies relying on annual US state-panel data,[12 13] as well as in large-scale studies based on cross-sectional time-series data covering a large number of countries[14–17] (see Hakim *et al*[18] for a review of older studies pointing in the same direction).

Various mechanisms underlying the procyclical effect on traffic deaths have been suggested in the literature. The most self-evident is that increased income tends to increase exposure, that is, driving, including commuting and freight transportation.[18] Macroeconomic fluctuations also tend to affect the composition of drivers in a way that impacts on traffic risks. Thus, young people, who have an elevated accident risk, are often more likely to become unemployed and thus drive less compared to others in bad times. Further, the number of inexperienced drivers may decrease in recessions due to a decreased number of new driving license holders.

However, considering the steady growth in GDP and the marked downward trend in traffic death rates in affluent countries during the last half-century,[19] there must reasonably exist some mechanism countervailing the procyclical effect. In fact, although the procyclical short-term effect of GDP on traffic deaths seems plausible and well substantiated empirically, a long-term protective effect seems equally likely. Thus, safer roads, safer vehicles[20 21] and improved medical treatment[22 23] are three factors that have been found important for improving traffic safety and reducing traffic deaths, and these three factors are in all probability correlated with GDP. Additional efficient preventive measures that are likely to be linked to GDP include speed limits,[24 25] seat-belt usage,[26] and legal maximum alcohol limits for driving.[27] In regard to empirical evidence of a protective effect, the studies by Kopits and Cropper[28] and van Beeck *et al*,[19] based on data for a large number of countries, suggest that increasing prosperity is protective against traffic deaths in developed countries.

In conclusion, the hypothesis of a procyclical short-term and the hypothesis of a protective long-term effect of GDP on traffic deaths seem well corroborated and empirically supported. However, extant research has tended to focus on one or the other of these effects, but to get insights about the net effect it is necessary to consider them jointly by applying a more comprehensive approach. Such an approach is indeed a logical sequel of two of the more recent studies in the field.[15 16] Although both of them focus the procyclical short-term effect, Chen[16] hints at possible beneficial effects of economic prosperity on road safety from a long-term perspective, whereas Yannis *et al*[15] emphasise that future research should also consider the long-term relationship between GDP and traffic deaths by applying the type of statistical techniques that we will actually make use of. The main aim of our study is thus to apply a modelling technique that estimates the short-term as well as long-term impact of GDP on traffic deaths.

However, there are two additional topics that we will address; the possibility of a structural shift and the potential impact of seat belt legislation. On the basis of data for 21 OECD countries, van Beeck *et al*[19] report a reversal in the cross-sectional relation between GDP and traffic death rates; the correlation was positive prior to the mid-1970s, thereafter it became negative. A plausible explanation of this shift, offered by the authors, is that in the early, less prosperous period, there was a stronger link between GDP and exposure (driving) than in the later period when mobility had levelled off. In this later period, GDP instead became protective by facilitating, for example, improvements in traffic infrastructure. To investigate whether a corresponding shift is present in the temporal association between GDP and traffic deaths, we analysed two subperiods, 1960–1975 and 1976–2011.

Although it would be of interest to include additional factors potentially impacting traffic death rates, lack of comparable data makes us confine ourselves to one additional factor, namely, the implementation of seat belt legislation. Seat belt use is considered to be the single most effective means of reducing injuries in the event of a motor vehicle crash.[26] Mandatory seat belt laws should thus have a considerable potential in affecting traffic mortality rates. This is also borne out in a review of evaluations of such laws.[26] Such evaluations are typically before- and after-trials without control areas, although there are certainly more sophisticated studies as well, for example[29] relying on US state-panel data.

DATA AND METHOD

The study comprises 18 OECD countries, and the longest observation period is 1960–2011, although it is appreciably shorter for some countries (see table 1). Age-specific road traffic mortality data for women and men were obtained from the WHO Mortality Database (Geneva). (Table 2 shows which ICD codes were included.) Age-standardised mortality rates (number of deaths per 100 000 population) were constructed following WHO World Standard.[30] Different ICD classifications have been used during the study period, from ICD-7 to ICD-10. Possible influences of revisions of ICD classification were captured by dummy variables. Missing mortality data (table 1) were imputed through linear interpolation; dummy variables were created for these years. Data on per-capita GDP, expressed in Purchasing Power Parity (PPP), converted into US dollars of 1990 years value, were obtained from the Maddison Project.[31] We performed age-specific analyses in addition to analyses for the adult population (20+), which we regard as the main outcome. Data on mandatory seat belt legislation were obtained from ref. [32] and various national sources. A dummy variable was created that took the value 1 at the year of legislation and onwards, and 0 otherwise. An alternative coding assumed a

Table 1 Descriptive statistics (period average) for GDP/capita (US$1000) and traffic deaths per 100 000 in the age group 20 years and above

Country	Observation period	GDP	Mortality
Australia	1960–2011	13.05	20.55
Austria	1960–2009	11.26	21.59
Belgium	1960–2010	11.92	21.94
Canada	1960–2009	13.46	18.93
Denmark	1960–2009	13.35	13.94
Finland	1960–2009	10.9	14.63
France	1960–2010	11.79	20.85
Germany	1960–2011	11.27	17.48
Ireland	1960–2009	9.17	14.82
Italy	1960–2010	9.98	18.97
Japan	1960–2011	10.53	13.07
New Zealand	1960–2009	11.49	13.16
Norway	1960–2010	13.07	19.89
Sweden	1960–2009	12.8	8.98
Switzerland	1960–2009	15.71	10.69
The Netherlands	1960–2010	12.54	16.97
UK	1960–2010	12.56	10.24
USA	1960–2011	16.92	21.78
Total		12.32	16.58

GDP, gross domestic product.

Table 2 ICD codes for traffic mortality data

	Cause of death	ICD-10	ICD-9	ICD-8	ICD-7
1	Pedestrian injured in collision with two-wheeled or three-wheeled motor vehicle	V02			
2	Pedestrian injured in collision with car, pick-up truck or van	V03			
3	Pedestrian injured in collision with heavy transport vehicle or bus	V04			
4	Pedal cyclist injured in collision with two-wheeled or three-wheeled motor vehicle	V12			
5	Pedal cyclist injured in collision with car, pick-up truck or van	V13			
6	Pedal cyclist injured in collision with heavy transport vehicle or bus	V14			
7	Motorcycle rider injured in transport accident	V20–V29			
8	Occupant of three-wheeled motor vehicle injured in transport accident	V30–V39			
9	Car occupant injured in transport accident	V40–V49			
10	Occupant of pick-up truck or van injured in transport accident	V50–V59			
11	Occupant of heavy transport vehicle injured in transport accident	V60–V69			
12	Bus occupant injured in transport accident	V70–V79			
13	Person injured in unspecified motor vehicle accident, traffic	V89.2			
	Total of 1–13		B471	A138	A138

gradual impact where the legislation year was coded 0.5, next year 0.75 and then 1.

We included an interaction term to capture the possible excess effect of GDP during the years of the financial crisis. The interaction term was constructed as follows:

$$\mathrm{GDPcrisis}_{it} = \mathrm{GDP}_{it} \times \mathrm{Crisis}_{it} \tag{1}$$

where Crisis is a country-specific variable that takes the value 0 in years with no recession, 0.25 in years with a 1-quarter recession and so forth, and 1 in years with 4 quarters of recession. The common recession definition was used, that is, that a recession occurred when GDP has contracted at least two consecutive quarters. Data were obtained from Eurostat and OECD.

We used two different methodological techniques to investigate the relation between GDP and traffic deaths. The rationale for this is that triangulating findings from different methods should reduce the risk of obtaining method-bound results. Both methods explore within-country variation only (fixed-effects models). The first method relies on the error correction model (ECM), whereas the second method is based on a model including contemporaneous GDP (to gauge the short-term effect), and a weighted sum of past GDP (to assess the long-term effect). A brief description of the two methods is given below.

Although error correction modelling is a standard tool in economics, it is, as pointed out by De Boef and Keele,[33] underused in other branches of social science. Error correction modelling is useful when short-term and long-term dynamics are focused;[34] its feasibility in the present context is highlighted by Yannis *et al*,[15] although they described it as demanding and did not apply it. We chose the single-equation approach for estimating our ECMs. The simulation results presented by Durr[35] suggest that this approach performs at least as well as the more complex two-step procedure developed by Engle and Granger.[36] Following standard specification,[35 37 38] our ECM looks as follows in its most basic form:

$$\Delta \mathrm{Mortality}_{it} = \alpha + \beta_0 \Delta \mathrm{GDP}_{it} + \beta_1 \mathrm{Mortality}_{it-1} + \beta_2 \mathrm{GDP}_{it-1} + \varepsilon_{it} \tag{2}$$

In this equation, β_0 indicates the instantaneous, short-term effect of a change in GDP on mortality, whereas β_1 estimates the speed at which the long-term effect operates. If such an effect does exist, the estimate of β_1 should be negative and statistically significant. The model assumes that the long-term effect decays geometrically; thus $1-(-1\times\beta_1)$ corresponds to the lag parameter in a lag scheme with geometrically declining lag weights (which we will make use of in our second modelling approach). The total long-term effect is calculated as $\beta_2/(-1\times\beta_1)$.

Prior to estimating an ECM, it is necessary to carry out some key tests. These analyses comprised two steps; first, we tested for unit root using the Fisher-type ADF panel unit root test.[39] If the independent and dependent variables prove to be integrated of the order $I(1)$, the next step is to test whether they are cointegrated. We used the panel cointegration tests developed by Westerlund,[40] denoted P_t and P_a. Simulation results[40] indicate that the tests have better small-sample properties and power than other commonly used panel cointegration tests, eg, the Pedroni tests.[41] The simulations further indicate that each of the two tests has its own merits and limitations and should thus be considered jointly. The tests accommodate various forms of heterogeneity and also generate p values that are robust against cross-sectional dependencies via bootstrapping.[40] Provided the tests indicate cointegration, it is appropriate to proceed to error correction modelling.

Our second methodological technique is a modified version of an approach that is commonly applied in alcohol epidemiology to assess a relation that involves a marked lag-structure, for example, the relation between per-capita alcohol consumption and liver cirrhosis mortality.[42] We will refer to the model as weighted lag model (WLM), and it specified as follows:

$$\Delta \mathrm{Mortality}_{it} = \alpha + \beta_0 \Delta \mathrm{GDP}_{it} + \beta_1 \Delta \mathrm{GDPW}_{it} + \varepsilon_{it} \tag{3}$$

In this model, β_0 indicates the instantaneous, short-term effect of a change in GDP on mortality, whereas the long-term effect is assessed by the estimated effect of a weighted sum of lagged values of GDP, computed as follows:

$$\mathrm{GDPW}_{it} = (\lambda \mathrm{GDP}_{it-1} + \lambda^2 \mathrm{GDP}_{it-2} + \lambda^3 \mathrm{GDP}_{it-3} + \cdots + \lambda^n \mathrm{GDP}_{it-n}) / \sum_{k=1}^{n} \lambda^k \tag{4}$$

The lag scheme was truncated at lag 15, and the lag parameter (λ) was fixed a priori to the value estimated by the ECM, as described above. In the age-specific estimations of model (4), we used the estimated lag parameter (λ) from the corresponding ECM. As noted above, the observation period for the mortality data starts 1960. However, to not lose observations in the analyses that include the weighted GDP-indicator, the series for GDP begin 1945.

All estimated models included the crises variable (as specified above), dummy variables for the interpolations and various ICD classifications. We also included country-specific dummies to account for the possible heterogeneity due to unobserved characteristics that may remain after differencing.

A complication with time-series cross-sectional data is the likely presence of serial and spatial (cross-country) dependence of the errors, which yields a downward bias of the OLS estimates of the SEs. We thus chose a modelling technique that addresses this complication in two ways. First, it accounts for spatial dependence of the errors by applying the more conservative panel-corrected SEs suggested by Beck and Katz.[43] Simulation results indicated that the panel-corrected SEs performed excellently; the procedure also yields a correction for any panel heteroscedasticity.[43] Secondly, our modelling technique accounts for serial dependence by including panel-specific autoregressive parameters for estimation of residual autocorrelation.

On the basis of the panel-corrected SEs, we used the Bewley transformation regression[44] (also described in De Boef and Keele[33]) to estimate SEs and significance levels of the long-term effect in the ECMs.

As a robustness test, we estimated equation (2) by using a heterogeneous method, that is, Pesaran and Smith's[45] mean group estimator (MG), which accommodates heterogeneous effects (slope coefficients) across panels.

All statistical analyses were performed with Stata V.14 (StataCorp, College Station, Texas, USA).

RESULTS

Descriptive statistics are found in table 1. As can be seen in figure 1, all countries experienced a steady growth in GDP during the study period. Another trait common to most countries is the decreasing trend in the death rate following an initial increase, although the length of the initial increase varies across countries.

Mortality data are missing 1998–2001 for Belgium, 2004–2005 for Italy, 2000 for UK, 2005 for Australia and 2006 for Canada.

The outcome of the panel unit root tests of GDP and various traffic death rates (table 3) suggests that for most of the eight variables the null hypothesis of unit root cannot be rejected by any of the four statistics, and the null cannot be rejected for any of the variables by the Pm-test, which is a recommended test in large panels.[39] We thus regard all our variables as having a unit root and proceed to test whether the relation between GDP and traffic deaths is cointegrated. Table 4 shows that the null hypothesis of no cointegration was rejected by at least one of the two panel tests in all age groups, except for the age group 0–19 years. We thus proceed to the estimation of the ECMs for the age groups above 19 years.

Table 5 displays the estimates of the ECMs. According to the outcome, the short-term effect implies that a one-unit increase (US\$1000) in GDP/capita yields an instantaneous increase in the total death rate by 0.58 in the adult population (20+). As expected, the long-term effect has a negative sign and is estimated at −1.59. Both of these estimates were strongly statistically significant. The estimates from the alternative model (WLM), displayed in table 6, were fairly consistent with those from the ECM. All age-specific estimates but one were statistically significant; the variation in effects across age groups does not show any systematic pattern.

The interaction term (GDPcrisis) capturing the possible excess effect of GDP during the years of the financial crisis was clearly insignificant in all model estimations (estimates not shown). The dummy variables for changes in ICD classifications were also statistically insignificant, except for the ICD-10 dummy variable that was significant in some of the age-specific analyses (estimates not shown). The estimated effects of seat belt legislation had the expected negative sign, but did not reach statistical significance in any of the age groups (estimates not shown). The outcome from the robustness test (reported in online supplementary appendix) where we used a method[45] that allows for heterogeneous effects across panels is consistent with the estimates reported above.

The period-specific model estimates (table 5, last rows) suggest a structural shift in the relation between GDP and traffic deaths. The protective long-term effect is about equally strong in both periods, whereas the procyclical short-term effect is markedly stronger in the early period than in the late period (t-value for difference=3.15, $p<0.002$). Further, in the late period the protective effect outweighs the procyclical effect, whereas the reverse is true for the early period.

The diagnostics of the residuals are satisfactory with regard to stationarity, whereas the autocorrelation is significant in the models for the three oldest age groups, but not in the model for our main outcome (20+). The cross-unit correlations are not very strong, but still statistically significant. However, this should not be a concern as the SEs we use are corrected for this kind of spatial correlation as described above. (The uncorrected SEs in the model for our main outcome (20+) are about 35% smaller.)

DISCUSSION

Previous research suggests that an increase in GDP is associated with an increase in traffic deaths; the most important mechanism underlying this relation is likely increased private and commercial road transport spurred by an expanding economy. On the other hand, there is also empirical support for the obvious assumption that economic growth creates resources that can be invested in safer traffic infrastructure leading to a long-term decrease in death rates. In the present study, we have strived to integrate these two strands of the literature and to apply a more comprehensive modelling approach in which the short-term and long-term effects were estimated jointly. Our results are indeed in line with these previously reported findings that road mortality is procyclical in the short run, but protective in the long run. However, the novelty of our findings is that they indicate the net of these opposing effects. In the analysis of the entire period, the long-term effect was markedly stronger than the short-term effect. However, period-specific analyses revealed a structural break implying that the protective effect outweighs the procyclical effect only in the period after 1975, whereas the reverse is true for the period 1960–1975. This outcome accords with the common pattern of a positive trend in GDP accompanied by an initial increase in the death rate, which was followed by a decreasing trend.

Our findings should also be regarded in a wider context. As noted in the introduction, there is a large number of studies suggesting that overall mortality, a common proxy for population health, is procyclical. It is worth pointing out that this is to a substantial extent driven by the procyclical character of traffic deaths and that the long-term protective dynamics are typically not considered in this literature.

Our finding that there is no excess effect of the economic crisis that bursted in the fall of 2007 (the Great Recession) accords with the outcome reported in a study[49] with a similar design as the present study. That investigation found a statistically significant effect of the unemployment rate on suicide, but this effect was thus not reinforced by the Great Recession. One possible reason for the absence of any significant impact of seat belt legislation is that the implementation of this regulation was fairly synchronised across countries, occurring typically in 1975

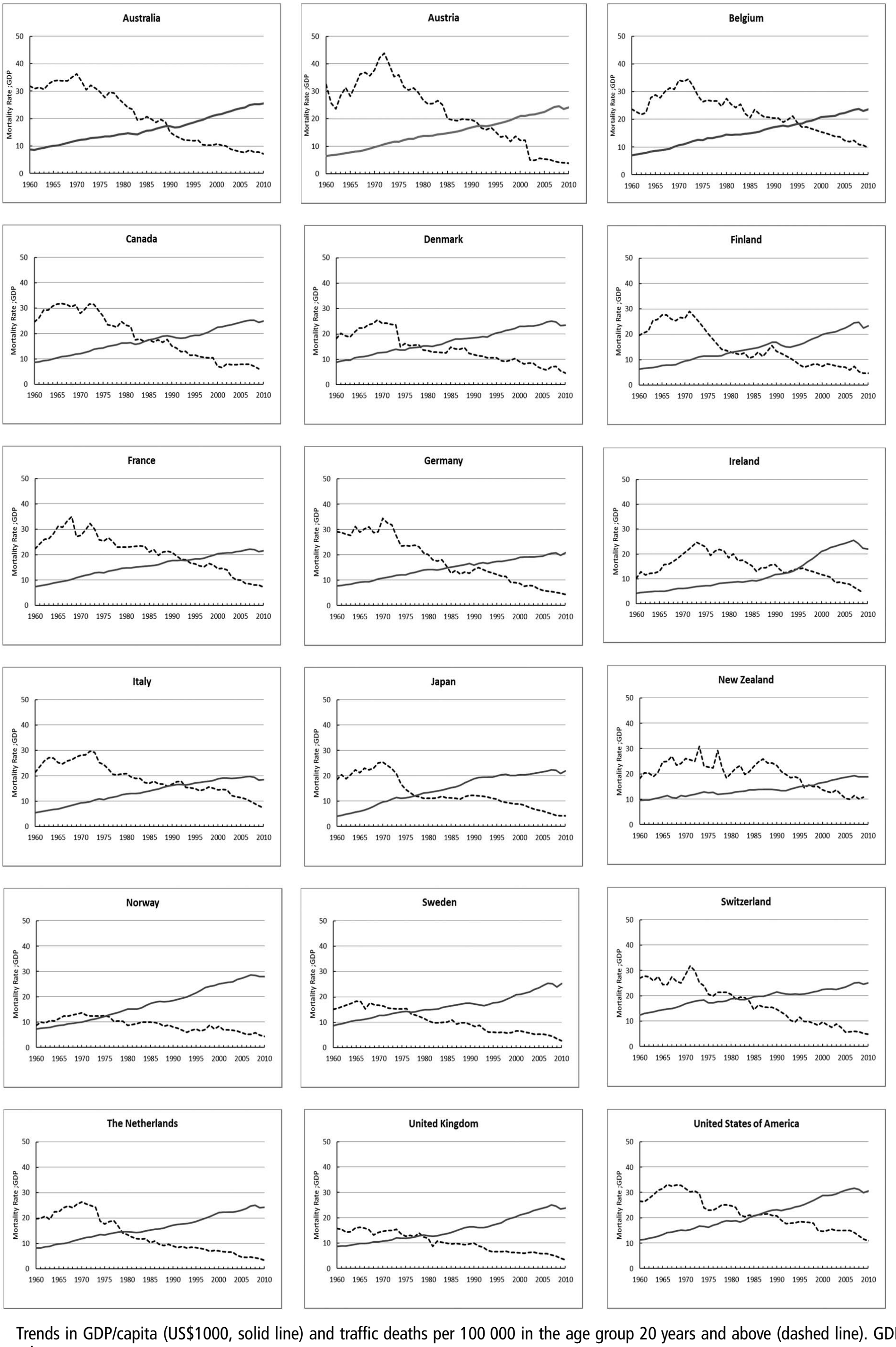


Figure 1 Trends in GDP/capita (US$1000, solid line) and traffic deaths per 100 000 in the age group 20 years and above (dashed line). GDP, gross domestic product.

Table 3 Fisher-type ADF panel unit root tests of H0: all panels contain unit roots against H1: at least one panel is stationary

Test		GDP/capita		0–19		20–34		35–49		50–64		65+		0+		20+	
		Statistic	p Value	Statistic	p Value	Statistic	p Value	Statistic	p Value	Statistic	p Value	Statistic	p Value	Statistic	p Value	Statistic	p Value
Inverse χ^2	P	14.12	0.999	31.42	0.686	30.32	0.735	40.96	0.262	45.08	0.143	42.12	0.223	30.80	0.71	32.10	0.66
Inverse normal	Z	2.84	0.998	−0.71	0.238	−0.43	0.334	−1.78	0.038	−2.25	0.012	−1.86	0.032	−0.55	0.29	−0.74	0.23
Inverse logit	L	2.81	0.997	−0.64	0.262	−0.39	0.348	−1.62	0.054	−2.05	0.022	−1.70	0.046	−0.50	0.31	−0.67	0.25
Modified inv. χ^2	Pm	−2.58	0.995	−0.54	0.706	−0.67	0.748	0.59	0.279	1.07	0.142	0.72	0.235	−0.61	0.73	−0.46	0.68

Table 4 Westerlund panel cointegration tests of H0: no cointegration for panels against H1: cointegration for all panels

Test	0–19			20–34			35–49			50–64			65+			0+			20+		
	Statistic	p Value	Robust P	Statistic	p Value	Robust P	Statistic	p Value	Robust P	Statistic	p Value	Robust P	Statistic	p Value	Robust P	Statistic	p Value	Robust P	Statistic	p Value	Robust P
Pa	−7.27	0.875	0.781	−9.99	0.233	0.255	−12.39	0.008	0.035	−10.71	0.109	0.135	−10.87	0.090	0.131	−9.69	0.300	0.255	−11.48	0.039	0.057
Pt	−9.36	0.326	0.369	−10.53	0.037	0.100	−12.51	<0.001	0.004	−11.46	0.002	0.025	−11.08	0.008	0.049	−11.21	0.005	0.030	−12.08	<0.001	0.016

Robust P are p values which are robust against cross-sectional dependencies obtained via bootstrapping (number of bootstraps was set to 800).

Table 5 Estimates of GDP/capita (US$1000) on traffic death rates (per 100 000) based on ECMs

		ΔGDP_t			$Mortality_{t-1}$			GDP_{t-1}			Long-term effect			Residual diagnostics						
														1		2		3		4
Age group	N	Est	SE	p Value	Est	SE	p Value	Est	SE	p Value	Est	SE	p Value	Statistics	p Value	Statistics	p Value	Statistics	p Value	Correlation
20+	906	0.58	0.18	0.001	−0.10	0.02	<0.001	−0.16	0.02	<0.001	−1.59	0.02	<0.001	−18.89	<0.001	1.40	0.25	12.03	<0.001	0.183
20–34	906	0.90	0.20	<0.001	−0.09	0.02	<0.001	−0.14	0.02	<0.001	−1.50	0.02	<0.001	−19.48	<0.001	0.44	0.51	6.52	<0.001	0.151
35–49	906	0.32	0.17	0.06	−0.12	0.02	<0.001	−0.13	0.02	<0.001	−1.05	0.02	<0.001	−19.64	<0.001	8.62	0.01	9.02	<0.001	0.190
50–64	906	0.48	0.21	0.02	−0.11	0.02	<0.001	−0.17	0.03	<0.001	−1.63	0.02	<0.001	−19.51	<0.001	8.77	0.01	10.53	<0.001	0.195
65+	906	0.78	0.33	0.02	−0.09	0.02	<0.001	−0.25	0.04	<0.001	−2.72	0.03	<0.001	−19.29	<0.001	4.70	0.04	10.14	<0.001	0.167
20+;<=1975	270	2.49	0.66	<0.001	−0.27	0.07	<0.001	−0.26	0.12	0.030	−0.98	0.14	<0.001	−17.25	<0.001	2.65	0.12	7.98	<0.001	0.263
20+; >1975	636	0.43	0.10	<0.001	−0.08	0.02	<0.001	−0.07	0.02	<0.001	−0.87	0.01	<0.001	−19.14	<0.001	1.64	0.22	1.89	0.06	0.168

Residual diagnostics:
1. Pesaran's panel data unit root test for stationarity (CIPS, robust against cross-sectional dependencies).[46] H0: panels contain unit roots; H1: panels are stationary.
2. Wooldridge test for autocorrelation in panel data.[47] H0: no first-order autocorrelation; H1: first-order autocorrelation.
3. Pesaran's test of cross-sectional independence.[48] H0: cross-units are independent; H1: cross-units are dependent.
4. Averaged absolute cross-panel correlation coefficient.
ECMs, error correction models; GDP, gross domestic product.

Table 6 Estimates of GDP/capita (US$1.000) on traffic death rates (per 100 000) based on WLMs

		ΔGDP_t			$\Delta GDPW_t$			Residual diagnostics						
								1		2		3		4
Age group	N	Est	SE	p Value	Est	SE	p Value	Statistics	p Value	Statistics	p Value	Statistics	p Value	Correlation
20+	906	0.67	0.19	<0.001	−1.31	0.47	0.005	−19.17	<0.001	1.22	0.285	14.27	<0.001	0.193
20–34	906	1.02	0.22	<0.001	−1.52	0.56	0.006	−19.61	<0.001	0.44	0.516	8.78	<0.001	0.154
35–49	906	0.39	0.18	0.027	−1.04	0.42	0.013	−19.71	<0.001	8.92	0.008	9.54	<0.001	0.160
50–64	906	0.53	0.22	0.015	−1.31	0.55	0.017	−19.83	<0.001	9.45	0.007	10.83	<0.001	0.195
65+	906	0.93	0.35	0.007	−1.70	0.78	0.029	−19.50	<0.001	4.97	0.040	11.51	<0.001	0.179

Residual diagnostics:
1. Pesaran's panel data unit root test for stationarity (CIPS, robust against cross-sectional dependencies).[46] H0: panels contain unit roots; H1: panels are stationary.
2. Wooldridge test for autocorrelation in panel data.[47] H0: no first-order autocorrelation; H1: first-order autocorrelation.
3. Pesaran's test of cross-sectional independence.[48] H0: cross-units are independent; H1: cross-units are dependent.
4. Averaged absolute cross-panel correlation coefficient.
GDP, gross domestic product; WLMs, weighted lag models.

or 1976. Further, the degree and pace by which legislation affected actual seat belt usage probably varies between countries.

Before concluding, we wish to highlight the major strengths and limitations of our study. Our findings are based on two conservative methods each of which has its own, but not identical weaknesses. Although the consistency of the outcomes thus seems reassuring, the risk of omitted variable bias can never be dismissed in the present kind of research. However, it should be noted that although there are numerous factors that affect the traffic death rate, only omitted factors that also are synchronised with changes in traffic mortality as well as GDP would bias our outcomes. Our data comprise a large number of countries spanning quite a long time period. However, these data represent affluent countries during a fairly prosperous historical epoch, which limits the generalisability of our findings.

What is already known on this subject

Increases in gross domestic product (GDP) lead to increases in traffic deaths, plausibly due to the increased road traffic induced by an expanding economy. However, there also seems to exist a long-term effect of economic growth that is manifested in improved traffic safety and reduced rates of traffic deaths. It is not known which of these opposing effects that is dominating.

What this study adds

Our study, based on cross-sectional time-series data for 18 OECD countries spanning the period 1960–2011, suggests that an increase in GDP leads to an immediate increase in traffic deaths. However, after the mid-1970s this short-term effect is more than outweighed by a markedly stronger protective long-term effect, whereas the reverse is true for the period before the mid-1970s.

In conclusion, an increase in GDP leads to an immediate increase in traffic deaths. However, after the mid-1970s this short-term effect is more than outweighed by a markedly stronger protective long-term effect, whereas the reverse is true for the period before the mid-1970s.

Acknowledgements The authors thank Professor Matthew Lindquist, Professor Gerard Pfann, Professor Joakim Westerlund and two anonymous reviewers for valuable advice and comments on previous drafts of this paper.

Contributors ID and TN conceived the design and drafted the paper. ID compiled the data and conducted the analyses. Both authors approved the final draft.

Funding This research was funded by FORTE: Swedish Research Council for Health, Working Life and Welfare (2013-0376), and The Swedish Research Council (421-2012-5503).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 WHO. *Global status report on road safety 2015*. World Health Organization, 2015.
2 Bengtsson T, Ohlsson R. Age-specific mortality and short-term changes in the standard of living: Sweden, 1751–1859. *Eur J Popul* 1985;1:309–26.
3 Norström T. Real wages, alcohol consumption and mortality in Sweden, 1861–1913. *Eur J Popul* 1988;4:183–96.
4 Preston SH. The changing relation between mortality and level of economic development. *Popul Stud* 1975;29:231–48.
5 Catalano R, *et al*. The health effects of economic decline. *Ann Rev Public Health* 2011;32:431–50.
6 Gerdtham U-G, Ruhm CJ. Deaths rise in good economic times: Evidence from the OECD. *Econ Hum Biol* 2006;4:298–16.
7 Ruhm CJ. Are recessions good for your health? *Quart J Econ* 2000;115:617–50.
8 Ogburn WF, Thomas DS. The influence of the business cycle on certain social conditions. *J Am Stat Assoc* 1922;18:324–40.
9 Tapia Granados JA. Economic growth and health progress in England and Wales: 160 years of a changing relation. *Soci Sci Med* 2012;74:688–95.
10 Neumayer E. Recessions lower (some) mortality rates: evidence from Germany. *Soc Sci Med* 2004;58:1037–47.
11 Farmer CM. Trends in motor vehicle fatalities. *J Saf Res* 1997;28:37–48.
12 Ruhm CJ. Recessions, healthy no more? *J Health Econ* 2015;42:17–28.
13 He MM. Driving through the great recession: why does motor vehicle fatality decrease when the economy slows down? *Soc Sci Med* 2016;155:1–11.
14 Stuckler D, Basu S, Suhrcke M, *et al*. The public health effect of economic crises and alternative policy responses in Europe: an empirical analysis. *Lancet* 2009;374:315–23.
15 Yannis GE, Papadimitriou E, Folla K. Effect of GDP changes on road traffic fatalities. *Saf Sci* 2014;63:42–9.
16 Chen G. Association between economic fluctuations and road mortality in OECD countries. *Eur J Public Health* 2014;24:612–14.
17 Burke PJ, Nishitateno S. Gasoline prices and road fatalities: international evidence. *Econ Inq* 2015;53:1437–50.

18 Hakim S, Shefer D, Hakkert AS, *et al*. A critical review of macro models for road accidents. *Accid Anal Prev* 1991;23:379–400.
19 van Beeck EF, Borsboom GJ, Mackenbach JP. Economic development and traffic accident mortality in the industrialized world, 1962–1990. *Int J Epidemiol* 2000;29:503–9.
20 Ameratunga S, Hijar M, Norton R. Road-traffic injuries: confronting disparities to address a global-health problem. *Lancet* 2006;367:1533–40.
21 Beenstock M, Gafni D. Globalization in road safety: explaining the downward trend in road accident rates in a single country (Israel). *Accid Anal Prev* 2000;32:71–84.
22 Noland RB. Medical treatment and traffic fatality reductions in industrialized countries. *Accid Anal Prev* 2003;35:877–83.
23 Noland RB, Quddus MA. Improvements in medical care and technology and reductions in traffic-related fatalities in Great Britain. *Accid Anal Prev* 2004;36:103–13.
24 McCarthy P. Effect of speed limits on speed distributions and highway safety: a survey of recent literature. *Transport Rev* 2001;21:31–50.
25 Elvik R. Speed limits, enforcement, and health consequences. *Annu Rev Public Health* 2012;33:225–38.
26 Dinh-Zarr TB, Sleet DA, Shults RA, *et al*. Reviews of evidence regarding interventions to increase the use of safety belts. *Am J Prev Med* 2001;21:48–65.
27 Mann RE, Macdonald S, Stoduto LG, *et al*. The effects of introducing or lowering legal per se blood alcohol limits for driving: an international review. *Accid Anal Prev* 2001;33:569–83.
28 Kopits E, Cropper M. Traffic fatalities and economic growth. *Accid Anal Prev* 2005;37:169–78.
29 Cohen A, Einav L. The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *Rev Econ Stat* 2003;85:828–43.
30 Ahmad OB, Boschi-Pinto C, Lopez AD, *et al*. *Age standardization of rates: a new WHO standard*. Geneva: World Health Organization, 2001.
31 The Maddison Project. http://www.ggdc.net/maddison/maddison-project/home.htm, 2013 version. 2013.
32 Resolution no. 38 Concerning Seat Belts. 2009, OECD/ITF.
33 De Boef S, Keele L. Taking time seriously. *Am J Political Sci* 2008;52:184–200.
34 Banerjee A, Dolado J, Galbraith JW, *et al*. *Co-integration, error correction, and the econometric analysis of non-stationary data. OUP Catalogue*, OUP, 1993.
35 Durr RH. An essay on cointegration and error correction models. *Political Anal* 1992;4:185–228.
36 Engle RF, Granger CWJ. Co-integration and error correction: representation, estimation, and testing. *Econometrica* 1987;55:251–76.
37 Beck N. Comparing dynamic specifications: the case of presidential approval. *Political Anal* 1991;3:51–87.
38 Grant T, Lebo MJ. Error correction methods with political time series. *Political Anal* 2016;24:3–30.
39 Choi I. Unit root tests for panel data. *J Int Money Finance* 2001;20:249–72.
40 Westerlund J. Testing for error correction in panel data. *Oxf Bull Econ Stat* 2007;69:709–48.
41 Pedroni P. Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Economet Theor* 2004;20:597–625.
42 Norström T, Skog OJ. Alcohol and mortality: methodological and analytical issues in aggregate analyses. *Addiction* 2001;96:5–17.
43 Beck N, Katz JN. What to do (and not to do) with time-series cross-section data. *Am Political Sci Rev* 1995;89:634–47.
44 Bewley RA. The direct estimation of the equilibrium response in a linear dynamic model. *Econ Lett* 1979;3:357–61.
45 Pesaran MH, Smith R. Estimating long-run relationships from dynamic heterogeneous panels. *J Econom* 1995;68:79–113.
46 Pesaran MH. A simple panel unit root test in the presence of cross-section dependence. *J Appl Econom* 2007;22:265–312.
47 Wooldridge JM. *Econometric analysis of cross section and panel data*. MIT Press, 2010.
48 Pesaran MH. *General diagnostic tests for cross section dependence in panels*. CESifo Working Paper Series No. 1229; IZA Discussion Paper No. 1240. 2004. SSRN: http://ssrn.com/abstract=572504
49 Norström T, Grönqvist H. The Great Recession, unemployment and suicide. *J Epidemiol Community Health* 2015;69:110–16.

IV

Scandinavian Journal of Public Health, 1–10

ORIGINAL ARTICLE

Is there a link between all-cause mortality and economic fluctuations?

IMAN DADGAR & THOR NORSTRÖM

Swedish Institute for Social Research, Stockholm University, Stockholm, Sweden

Abstract

Background: All-cause mortality is a global indicator of the overall health of the population, and its relation to the macro economy is thus of vital interest. The main aim was to estimate the short-term and the long-term impact of macroeconomic change on all-cause mortality. Variations in the unemployment rate were used as indicator of temporary fluctuations in the economy. *Methods:* We used time-series data for 21 OECD countries spanning the period 1960–2018. We used four outcomes: total mortality (0+), infant mortality (<1), mortality in the age-group 20–64, and old-age mortality (65+). Data on GDP/capita were obtained from the Maddison Project. Unemployment data (% unemployed in the work force) were sourced from Eurostat. We applied error correction modelling to estimate the short-term and the long-term impact of macroeconomic change on all-cause mortality. *Results:* We found that increases in unemployment were statistically significantly associated with decreases in all mortality outcomes except old-age mortality. Increases in GDP were associated with significant lowering long-term effects on mortality. ***Conclusions:*** **Our findings, based on data from predominantly affluent countries, suggest that an increase in unemployment leads to a decrease in all-cause mortality. However, economic growth, as indicated by increased GDP, has a long-term protective health impact as indexed by lowered mortality.**

Keywords: *all-cause mortality, GDP, unemployment. Great Recession, error correction model*

Introduction

All-cause mortality is a classic indicator of the overall health of the population [1]. It is therefore of great concern to get a better understanding of the driving forces behind changes in mortality. Using cross-sectional time-series data for 21 countries, this work studied the potential role of macroeconomic fluctuations as indicated by changes in unemployment and per capita gross domestic product (GDP).

Intuition could easily make one believe that recessions can only be for the worse, but as described below the relation between economic fluctuations and population health is complex and seemingly contradictory. This may explain why the received wisdom concerning this relationship has undergone some quite substantial shifts. It is clear that economic downturns in past historical centuries led to severe malnutrition and starvation and thus worsened population health. Economic growth, on the other hand, was conducive to education, improved sanitation and living conditions, and, in the end, lowered mortality [2]. However, as demonstrated by Preston [3] there is a diminishing health return to economic growth, and there are even indications that economic downturns in highly industrialized societies may improve population health. The explanation to this counterintuitive finding is that although a downturn in all probability has a detrimental effect on some outcomes, such as mental health as indexed by suicide [4], this negative effect may be more or less offset by a beneficial impact on other dimensions of health. Several examples of such beneficial effects have been suggested and substantiated. A slow-down in the economy is thus associated with reduced overtime and work-related stress, less driving and fewer car crashes, less air pollution, and reduced intake of

Correspondence: Thor Norström, Swedish Institute for Social Research, Stockholm University, Stockholm, S-106 91, Sweden.
E-mail: totto@sofi.su.se

Date received 5 March 2021; reviewed 12 August 2021; accepted 9 September 2021

Article reuse guidelines: sagepub.com/journals-permissions
DOI: 10.1177/14034948211049979
journals.sagepub.com/home/sjp
SAGE

unhealthy products such as alcohol and tobacco [4-6]. Already in the early 20th century there were reports (e.g. Ogburn and Thomas [7]) suggesting that economic booms were associated with increased mortality, while the opposite was true for economic downturns. However, these results were ignored for a long time, probably because they appeared to run counter to intuition. In the 1970s and 1980s, Harvey Brenner published a series of papers [8] suggesting marked negative influences of recessions on population health, as indexed by mortality rates. These findings attracted much interest; however, closer examinations of Brenner's work [9] revealed serious methodological flaws, such as correlating trending time-series, and arbitrary specifications of lagged effects. The investigation by Ruhm [4] was one of the first well-designed studies in the field. On the basis of fixed-effects modelling of US state data for the period 1972–1991, his findings suggested that recessions are associated with improved health. More specifically, mortality from eight out of ten causes of death under study decreased during bad times, especially traffic fatalities. An important exception was suicide, which increased in downturns. Alluding to the title of his article, Ruhm ends with: 'Are recessions good for your health? Surprisingly, the answer appears to be yes'. This finding was replicated by Tapia Granados' study [10], based on US data for the period 1900–1996. Findings from some additional country studies [11, 12] point in the same direction. Some studies use cross-sectional time-series data including a large set of countries. Thus Gerdtham and Ruhm [6] analysed time-series data for 23 Organisation for Economic Co-operation and Development (OECD) countries for the period 1960–1997, and concluded that economic expansion was associated with increased all-cause mortality; these results were replicated in a similar study by Bilal et al. [13]. There are some studies that report non-significant associations between all-cause mortality and economic fluctuations, including one investigation based on Danish data [14], and another study relying on time-series data for 26 European Union countries [15]. Further, one study[16], based on Danish and US data, found that upturns in the economy were associated with decreases in mortality. In sum, although most of the well-designed studies on this field suggest a procyclical effect, the overall pattern of the findings is far from conclusive (see the review by Catalano et al. [5] for a similar conclusion). Moreover, previous research has generally focused on the immediate, contemporaneous health effect of the economy, typically gauged by oscillations of the unemployment rate. A large body of research suggests that unemployment is associated with a range of adverse health outcomes (see elsewhere for reviews [5, 17]). Perhaps the most succinct unemployment effect is noted on indicators of mental health [18]. For instance, studies at individual as well as aggregate level suggest a link between unemployment and suicide risk [18]. Other outcomes associated with unemployment status include heart disease mortality and all-cause mortality [17]. However, two studies based on Finish data [19, 20] found that the unemployment effect on mortality was weaker the higher the unemployment rate, which indicates the presence of health selection; that people with poor health run an elevated risk of becoming unemployed. To minimize this source of bias, Böckerman and Ilmakunnas [21] applied fixed-effects modelling of panel data pertaining to Finland. Their findings showed no association between unemployment and self-assessed health. As noted above, several studies have even reported negative relations between unemployment rates and various fatality rates.

However, the presence of health-protective long-term effects of economic growth seems quite plausible. A case in point is traffic fatalities. Research suggests that economic upturns are associated with increases in traffic deaths [4]; this is likely a short-term effect mainly due to the increased road traffic induced by an expanding economy. However, at least in high-income countries, it seems reasonable to expect a long-term effect of economic growth that is manifested in safer vehicles and roads that leads to improved traffic safety and reduced rates of traffic deaths. The hypothesis of such a protective long-term effect of GDP on traffic deaths was supported by a study based on panel data for 18 OECD countries [22]. A similar line of reasoning should be applicable to all-cause mortality. During the last half-century, the period we focus upon, there has been a marked and steady decrease in all-cause mortality in affluent countries (see below). The driving forces behind this development, as suggested in the literature [1, 3, 23, 24], include improvements in nutrition, housing, educational level and medical treatment. Because all of these factors are to varying degrees linked to economic growth, it seems reasonable to hypothesize a long-term beneficial impact of GDP on population health.

Another issue concerns the possible heterogeneity in the association between economic change and mortality, that is, that certain country characteristics may modify the association at issue. For instance, previous research suggests that the pernicious unemployment effect on suicide is weaker the more generous the unemployment protection of the country [15, 25]. In the present context it may be hypothesized that a possible beneficial association between economic growth and population health would be stronger the larger share of GDP that is spent on

Table I. Descriptive statistics (period average) for all-cause mortality (number of deaths per 100,000), unemployment (% of the work force), and GDP/capita ($1,000). Country-group signifies degree of public spending on social insurance systems as % of GDP where 1=low, 2=medium and 3=high public spending.

Country	Mortality				Unemployment	GDP	Country-group	Observation period
	Infant	20-64	65+	Total				
Australia	978.6	303.7	4823.1	601.0	5.5	32.1	1	1960–2018
Austria	1283.1	328.6	5371.6	666.4	3.3	28.3	3	1960–2018
Belgium	1154.8	332.9	5398.2	667.5	7.4	27.1	2	1960–2016
Canada	1232.3	306.8	4574.9	633.1	7.4	31.0	1	1960–2005
Denmark	854.5	320.5	5209.0	637.5	5.5	32.8	3	1960–2018
Finland	744.3	364.9	5526.5	688.2	6.5	25.5	3	1960–2018
France	956.8	337.8	4578.7	599.4	6.8	27.3	3	1960–2014
Germany	1162.9	322.3	5361.3	658.6	5.2	31.6	3	1960–2018
Greece	1470.6	254.2	4872.0	584.7	10.3	18.5	1	1974–2017
Ireland	1166.2	337.2	6056.1	723.9	9.1	29.1	1	1960–2015
Italy	1467.3	282.7	4935.3	606.3	8.7	27.0	2	1960–2017
Japan	791.5	265.5	4516.2	549.9	2.8	25.4	1	1960–2018
New Zealand	1105.3	329.2	5165.0	649.0	3.9	24.3	2	1960–2016
Norway	795.9	266.4	4841.7	576.5	2.9	56.2	3	1960–2016
Portugal	2505.2	352.3	5783.8	748.4	7.3	17.3	1	1974–2018
Spain	1193.2	282.8	4790.5	590.1	15.1	21.3	2	1972–2017
Sweden	672.3	250.9	4730.4	553.8	4.6	29.9	3	1960–2018
Switzerland	887.2	265.0	4638.9	561.3	2.0	47.0	2	1960–2017
The Netherlands	801.5	264.6	4886.9	578.3	4.9	31.5	3	1960–2018
United Kingdom	1077.8	319.7	5406.7	657.8	5.8	25.5	2	1960–2016
United States of America	1364.5	395.9	5065.1	694.4	6.0	36.4	1	1960–2007

welfare provisions. To test this notion, the countries were sorted into three groups (low, medium and high; see Table I) based on their ranking on spending on public insurance systems (sources: OECD Social Expenditure Database and Gerdtham and Ruhm [6]). The main areas for social public spending include policies related to pensions, health, family and unemployment. In the low-country group the average spending on public insurance systems was 16.6 % of GDP during the period 1980–2018; the corresponding figures for the medium- and high-spending groups were 20.3 and 24.5 %, respectively.

Study aims

The main aim of this paper is thus to assess the short-term as well as the long-term impact of macroeconomic change on all-cause mortality, using data for 21 OECD countries spanning the period 1960–2018. In keeping with most previous studies, we will use changes in the unemployment rate as indicator of temporary fluctuations in the economy. The possible long-term impact of economic growth on mortality will be assessed by error correction modelling of the effect of GDP.

In sum, the main potential contributions of our paper are (i) that we span a long time period and include a large set of countries, making the findings more generalizable than those from previous studies; and (ii) that we assess the short-term as well as the long-term effect of economic change on mortality – this is an important issue because the short-term and the long-term effect may have opposite signs, a phenomenon that has not been addressed in previous research.

Data

The study comprises 21 OECD countries, and the longest observation period is 1960–2018, though it is somewhat shorter for some countries (see Table I). Age-specific mortality data were obtained from the World Health Organization (WHO) Mortality Data Base (Geneva). We used four outcomes: total mortality (0+), infant mortality (<1), working-age mortality (20–64), and old-age mortality (65+). The outcomes were expressed as number of deaths per 100,000 population, and were age-standardized following WHO World Standard. Unemployment data (% unemployed in the work force) were sourced from Eurostat. Data on gross domestic product/capita (GDP), expressed in Purchasing Power Parity (PPP), converted into US dollars of 1990 years value, were obtained from the Maddison Project [26].

Statistical analysis

Our analytical strategy for estimating the relation between mortality and the two economic indicators is to apply error correction modelling (ECM), which is a feasible approach when short- and long-term dynamics are addressed [27]. Although ECM is a standard modelling tool in economics, it is, as pointed out by De Boef and Keele [27], under-utilized in other branches of social science.

However, prior to performing ECM it is necessary to carry out some initial analyses with respect to the variables where a long-term effect may be expected, that is, GDP and mortality. These analyses comprised two steps; first, we tested for unit root using the Fisher-Type ADF panel unit root test [28]. If the independent and dependent variables prove to be integrated of the order $I(1)$, the next step is to test whether they are cointegrated. Two variables, X and Y, are cointegrated if there exists a linear combination of X and Y that is stationary around which the two series fluctuate. This implies that if X drifts off, Y is bound to follow suit, and in the long run the series will not diverge far apart. The theory of cointegration stems from Engle and Granger [29], and empirical examples include the relation between GDP and traffic fatalities [22]. We used the panel cointegration tests developed by Westerlund [30], denoted P_t and P_a. Simulation results [30] indicate that the tests have better small-sample properties and power than other commonly used panel cointegration tests. The simulations further indicate that each of the two tests has its own merits and limitations, and should thus be considered jointly. The tests accommodate various forms of heterogeneity, and also generate p-values that are robust against cross-sectional dependencies via bootstrapping [30]. As detailed below, the outcome of these initial analyses suggested that both GDP and mortality were integrated of the order $I(1)$, and that they were cointegrated according to at least of one of the two tests; the conditions for performing ECM were thus fulfilled.

Following standard specifications [27], our error correction model was specified as follows:

$$\begin{aligned}\Delta LnMortality_{it} &= \alpha + \beta_0 \Delta GDP_{it} \\ &+ \beta_1 LnMortality_{it-1} + \beta_2 GDP_{it-1} \\ &+ \beta_3 \Delta Unemployment_{it} + \beta_i CD_i + \varepsilon_{it}\end{aligned} \tag{2}$$

Following common practice (e.g. Ruhm [4]), we used the natural log of mortality as outcome. In this equation, β_0 indicates the instantaneous, short-term effect of a change in GDP on mortality, while β_1 estimates the speed at which the long-term effect operates. If such an effect actually exists, the estimate of β_1 should be negative and statistically significant. The model assumes that the long-term effect decays geometrically. The total long-term effect is calculated as $-\beta_2/\beta_1$. CD is a vector of country dummies.

A complication with cross-sectional time-series data is the likely presence of serial and spatial (cross-country) dependence of the errors, which yields a downward bias of the estimated standard errors (SEs). As a remedy, we applied a modelling technique that addresses this complication as follows. First, it accounts for spatial dependence of the errors by applying the more conservative panel-corrected SEs suggested by Beck and Katz [31]. Simulation results indicated that the panel-corrected SEs performed excellently; the procedure also yields a correction for any panel heteroscedasticity [31]. Secondly, our modelling technique accounts for temporal dependence by including panel-specific autoregressive parameters for estimation of residual autocorrelation. In addition, we included country-specific dummies to account for possible country-specific heterogeneity. It should be emphasized that our analytical design implies that only temporal within-country variation is exploited.

The analyses reported in the main tables were carried out on unweighted data. However, as a sensitivity test, the main analyses were also performed on data where the observations were weighted by the square root of the country population. To test for possible gender-specific effects, we estimated separate models for females and males. We also estimated separate models for the three country-groups with different levels of spending on public insurance systems.

We used the Bewley transformation regression [32] to estimate SEs and significance levels of the long-term effect in the ECMs. All statistical analyses were performed with Stata, V.15 (StataCorp).

Results

Table I displays descriptive statistics. As appears in Figure 1, there was a steady growth in GDP in all countries during the study period. Another trait common to all countries is the decreasing trend in mortality. In contrast, the trajectories in unemployment do not display any common pattern; for most countries there is no marked trend, but rather irregularly occurring peaks and troughs.

The results of the panel unit root tests of GDP and various mortality rates (Table II, panel A) suggest that for all variables the null hypothesis of unit root cannot be rejected by any of the four statistics. Given this outcome, we proceed to test whether the relation between GDP and mortality is cointegrated. Table II (panel B) shows that the null hypothesis of no cointegration was rejected by at least one of the two panel tests in all age groups. We thus proceeded to estimate the error correction models. The outcome is shown in Table III. The percentage change in mortality from a one-unit increase

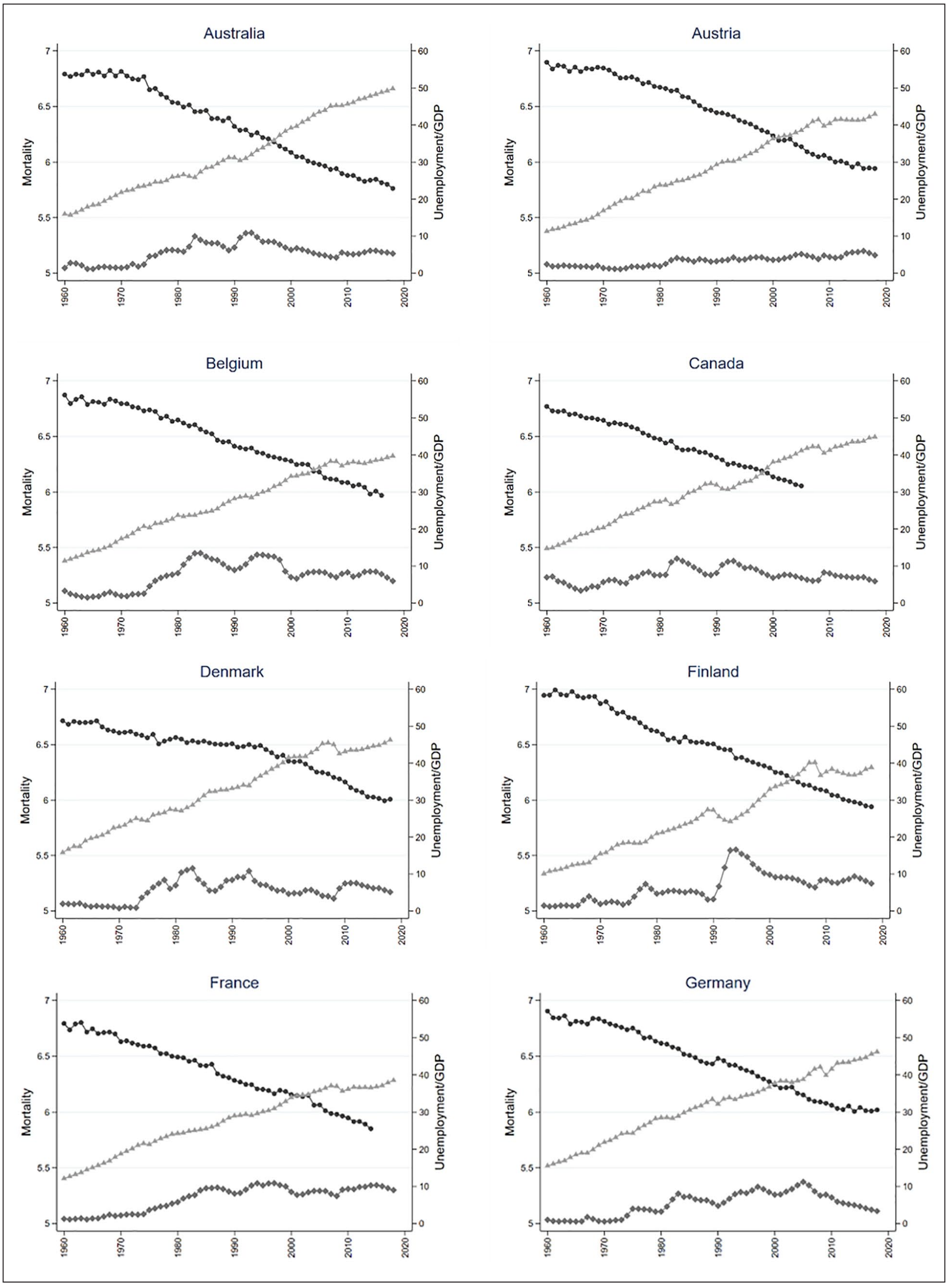


Figure 1. *(Continued)*

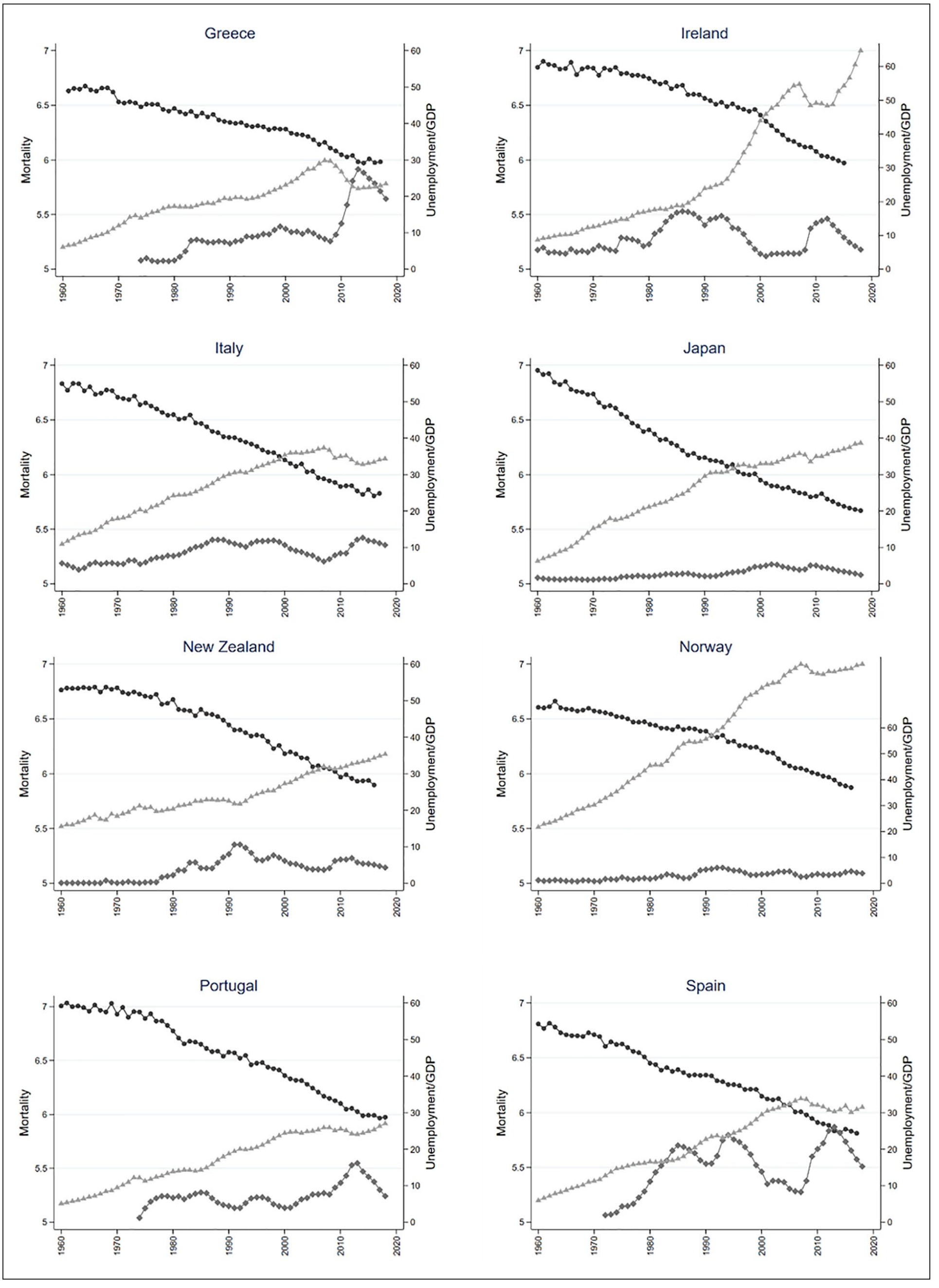


Figure 1. *(Continued)*

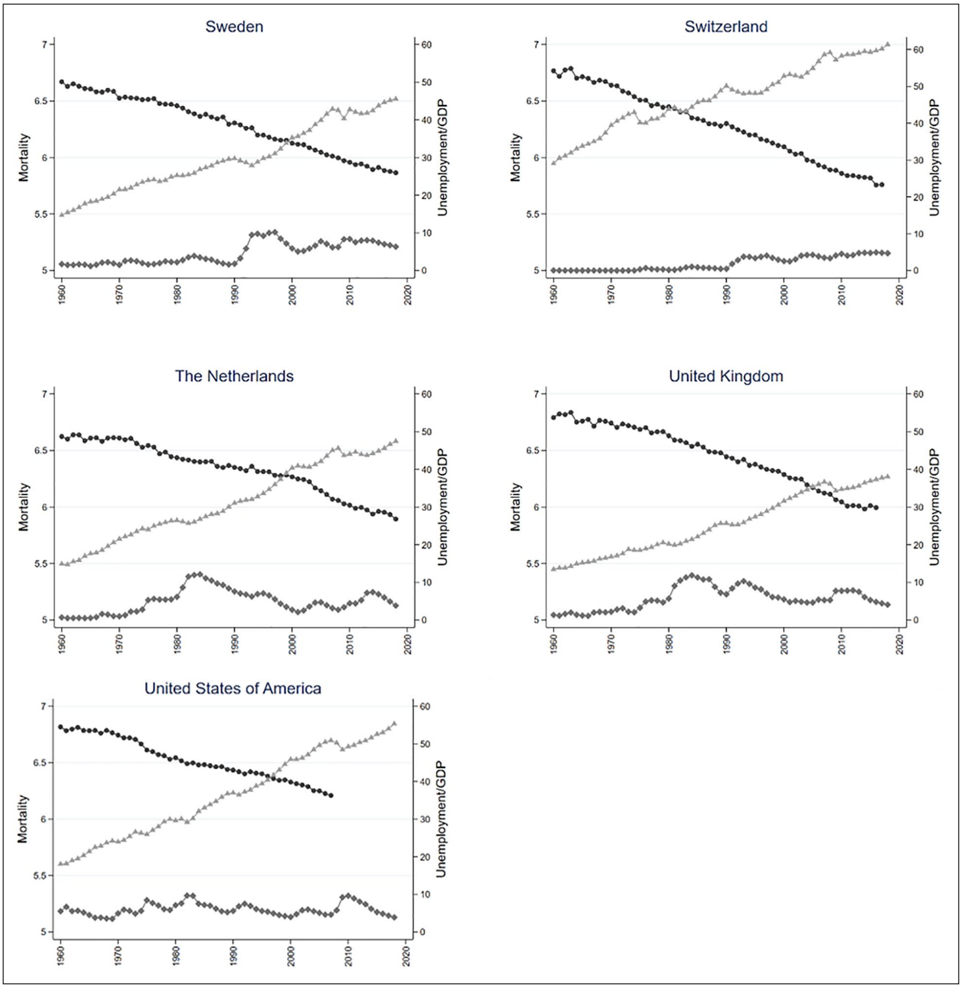


Figure 1. Trends in all-cause mortality per 100,000 in log (black circle), unemployment rate (diamond) and GDP per capita (US$1000, triangle).

in an explanatory variable, X, is obtained by the expression, $\left(exp\left[\hat{\beta}\right]-1\right)*100$, where $\hat{\beta}$ is the estimated effect of X. An increase in the unemployment rate by 1 percentage point was thus estimated to give a decrease in total mortality by 0.3%. The corresponding figure for infant mortality and mortality in the age-group 20–64 years were 0.8 and 0.3%, respectively, while the estimate for old-age mortality was not statistically significant. A one-unit increase in GDP (in $1000, which on average corresponds to a relative increase in GDP by 3.3%) had no statistically significant instantaneous effect on infant mortality or old-age mortality, while the significant estimates for 20–64 years mortality and total mortality imply a reduction in mortality by 0.3 and 0.4%, respectively. Now turning to the long-term effect of a one-unit increase in GDP (i.e. corresponding to an increase by 3.3%), the results suggest a reduction in total mortality by 3.8%; the corresponding figures for infant mortality, 20–64 years mortality and old-age mortality were 4.6, 7.0 and 3.1%, respectively (all these estimates were statistically significant).

To put the key estimates into perspective, and to facilitate comparisons among them, we converted

Table II. Unit root tests (Panel A), and cointegration tests (Panel B).

Panel A. Fisher-Type ADF panel unit root tests of H0: All panels contain unit roots against H1: At least one panel is stationary.

Test		GDP/capita		Unemployment		Infant mortality		20–64		Old-age mortality		Total	
		Statistic	*p*	Statistic	*p*	Statistic	*p*	Statistic	*p*	Statistic	*p*	Statistic	P
Inverse chi–squared(36)	P	34.26	0.80	27.99	0.95	54.14	0.10	5.96	>0.99	2.23	>0.99	2.60	>0.99
Inverse normal	Z	2.33	0.99	0.65	0.74	−0.44	0.33	10.51	>0.99	8.23	>0.99	9.19	>0.99
Inverse logit t(94)	L	2.49	>0.99	0.60	0.73	−0.86	0.20	12.35	>0.99	8.47	>0.99	9.89	>0.99
Modified inv. chi–squared	pm	−0.84	0.80	−1.53	0.94	1.32	0.09	−3.93	>0.99	−4.34	>0.99	−4.30	>0.99

Panel B. Westerlund panel cointegration tests of H0: no cointegration for panels against H1: cointegration for all panels.

	Infant mortality			20–64			Old-age mortality			Total mortality		
	Statistic	*p*	Robust P	Statistic	*p*	Robust P	Statistic	*p*	Robust P	Statistic	*p*	Robust P
Pa	−13.04	0.001	<0.001	−0.58	0.759	0.620	−13.51	<0.001	<0.001	−14.40	<0.001	<0.001
Pt	−11.55	0.016	<0.001	−6.88	<0.001	<0.001	−14.73	<0.001	<0.001	−13.44	<0.001	<0.001

them into elasticities, confining ourselves to the effects on mortality in the whole population (Total). The outcome suggests that the elasticity for the short-term effect in GDP is −0.1259; i.e. an increase in GDP by 1% would yield an instantaneous decrease in total mortality of 0.1259%. The corresponding figure for the long-term effect of GDP is −1.1429, and for unemployment −0.0181.

The estimates from the analyses based on weighted data (Table SI) differ little from these based on unweighted data. With regard to gender differences, the only more systematic pattern is that the long-term effect of GDP tends to be somewhat stronger in males than in females (Table SII). Finally, public spending on social insurance systems does not seem to modify the response of mortality to macroeconomic changes; the effect estimates display no systematic differences across the three country-groups (Table SIII).

Discussion

Economic upturns seem to have beneficial effects on some causes of death, and detrimental on others; to assess the net effect of macroeconomic change it is thus feasible to focus upon a global outcome, such as all-cause mortality. In this study we used panel data for 21 OECD countries, spanning the period 1960–2018 to estimate the association between all-cause mortality and two key macroeconomic indicators, GDP and unemployment. The aim was to assess not only the short-term health effect of temporary fluctuations in the economy, as indicated by fluctuations in unemployment, but in addition to estimate the long-term effect of economic growth in GDP. We found that an increase in unemployment is associated with an improvement in population health (as indexed by total mortality). The size of the estimated effect, 0.3% decrease in mortality following a 1 percentage point increase in unemployment, is somewhat lower than most of the previously reported estimates, ranging from 0.4% [4] over 1.1% [33] to 2.2% [10]. It may seem surprising that also infant mortality was found to be negatively related to unemployment. However, a couple of previous studies that have reported similar findings substantiate some plausible underlying mechanisms, viz. that recessions tend to lower levels of air pollution [34], and to generate improved health behaviour in mothers (e.g. less smoking and drinking) [35].

In contrast, our findings suggested a short-term as well as a long-term protective effect of growth in GDP, where the long-term effect was markedly stronger than the short-term impact. To our knowledge, there is no other study focusing on all-cause mortality that has elucidated this issue, implying that we lack a basis for comparisons. However, one study that applied the same analytical strategy (ECM), but a more narrow outcome (traffic fatalities), also reported a long-term beneficial impact of GDP [22].

An important issue is, of course, what implications our findings have with regard to future research and policy measures. The health-economic research that has emerged in the vein of Ruhm's influential work (e.g. [4]) makes it tempting to conclude that 'the procyclical character of mortality fluctuations is beginning to be a proven fact' [36]. Our finding of a marked protective long-term effect of economic growth makes such a conclusion disputable. However, notwithstanding any beneficial long-term effect of economic growth, our findings do suggest that temporary economic upturns, as indicated by decreased unemployment, tend to have deleterious effects on population

Table III. Estimates of GDP/capita ($1,000) and unemployment on all-cause mortality rates (per 100,000) based on error correction models (ECM). Models include country dummies. Panel-corrected standard errors accounting for spatial dependence and panel heteroscedasticity.

Age-group	N	ΔGDP_t			$\Delta Unemployment_t$			$Mortality_{t-1}$			GDP_{t-1}			Long-term effect of GDP		
		Est	SE	*p*	Est	SE	*p*	Est	SE	*p*	Est	SE	*p*	Est	SE	*p*
Infant	1135	−0.0018	0.0040	0.6629	−0.0079	0.0027	0.0034	−0.0404	0.0070	<0.001	−0.0019	0.0005	<0.001	−0.0466	0.0003	<0.001
20–64	1155	−0.0035	0.0011	0.0018	−0.0031	0.0008	<0.001	−0.0103	0.0049	0.0342	−0.0007	0.0001	<0.001	−0.0724	0.0003	<0.001
65+	1146	−0.0043	0.0023	0.0641	−0.0025	0.0014	0.0659	−0.0396	0.0098	<0.001	−0.0012	0.0002	<0.001	−0.0312	0.0002	<0.001
Total	1135	−0.0042	0.0018	0.0219	−0.0030	0.0011	0.0053	−0.0244	0.0073	<0.001	−0.0009	0.0002	<0.001	−0.0388	0.0002	<0.001

health. An obvious task is to identify the mechanisms underlying these effects. As noted above in the Introduction, several mechanisms have been suggested and at least partly corroborated, including increased work-related stress, more road traffic and car crashes, higher levels of air pollution, and increased consumption of unhealthy products such as alcohol and tobacco. To make some progress in this area, it seems urgent to investigate the possible presence of socio-cultural contingencies; that is, is the relation between economic fluctuations and these mechanisms modified by social and cultural characteristics of the country? A better understanding of these relations and their socio-cultural contingencies would potentially enable the tailoring of policy measures to mitigate these adverse health effects of economic upturns.

Before concluding, we will note the major strengths and limitations of the study. Our data comprise a large number of countries, and cover a fairly long time period. However, these data are confined to affluent countries during a prosperous historical epoch, which of course limits the generalizability of our findings. Our estimates rely on within-country variation only, thus avoiding the potential bias that likely arises from cross-country co-variation. However, the risk of omitted variable bias cannot be dismissed in the present kind of research; i.e. that the findings have been distorted by the omission of some factor that is related to mortality as well as to the macroeconomic indicators. We applied a modelling approach (ECM) that is novel to the field, and which yielded new insights into the dynamics of the relation between mortality and macroeconomic change. However, the uniqueness of the findings regarding the long-term effect of GDP implies that we have little external evidence to validate them against, so these findings should be probed further in future research.

Bearing the above caveats in mind, we conclude that our findings suggest that an increase in unemployment yields an instantaneous decrease in all-cause mortality among infants and in the working-age population. Further, we found a protective short-term as well as long-term effect of GDP.

Key points

- All-cause mortality is a global indicator of the overall health of the population, and its relation to the macro economy is thus of vital interest.
- On the basis of time-series data for 21 OECD countries spanning the period 1960–2016, we found that increases in unemployment had a statistically significant association with decreases in mortality.

- Economic growth, as indicated by increased GDP, had a long-term protective health impact as indexed by lowered mortality.

Contributors

ID and TN conceived the design and drafted the paper. ID compiled the data and conducted the analyses. Both authors approved the final draft.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by FORTE: Swedish Research Council for Health, Working Life and Welfare (2013-0376), and The Swedish Research Council (421-2012-5503).

ORCID iD

Thor Norström https://orcid.org/0000-0002-5746-7723

Supplemental material

Supplemental material for this article is available online.

References

[1] Sen A. Mortality as an indicator of economic success and failure. *Econ J* 1998;108:1–25.
[2] Tapia Granados JA and Ionides EL. The reversal of the relation between economic growth and health progress: Sweden in the 19th and 20th centuries. *J Health Econ* 2008;27: 544–563.
[3] Preston SH. The changing relation between mortality and level of economic development. *Popul Stud* 1975;29: 231–248.
[4] Ruhm CJ. Are recessions good for your health? *Quart J Econ* 2000;115: 617–650.
[5] Catalano R, et al. The health effects of economic decline. *Annu Rev Public Health* 2011;32: 431–450.
[6] Gerdtham U-G and Ruhm CJ. Deaths rise in good economic times: Evidence from the OECD. *Econ Human Biol* 2006;4:298–316.
[7] Ogburn WF and Thomas DS. The influence of the business cycle on certain social conditions. *J Am Stat Assoc* 1922;18:324–340.
[8] Brenner MH. Mortality and the national economy: A review, and the experience of England and Wales 1936-1976. *Lancet* 1979;314:568–573.
[9] Gravelle HSE, Hutchinson G and Stern J. Mortality and unemployment: A critique of Brenner's time-series analysis. *Lancet* 1981;2:675–679.
[10] Tapia Granados JA. Increasing mortality during the expansions of the US economy, 1900–1996. *Int J Epidemiol* 2005;34:1194–1202.
[11] Tapia Granados JA. Recessions and mortality in Spain, 1980-1997. *Eur J Popul* 2005;21:393–422.
[12] Tapia Granados JA and Ionides EL. Mortality and macroeconomic fluctuations in contemporary Sweden. *Eur J Popul* 2011;27:157–184.
[13] Bilal U R, Cooper F, Abreu C, et al. Economic growth and mortality: Do social protection policies matter? *Int J Epidemiol* 2017;46:1147–1156.
[14] Søgaard J. Econometric critique of the economic change model of mortality. *Soc Sci Med* 1992;34:947–957.
[15] Stuckler D, Basu S, Suhrcke M, et al. The public health effect of economic crises and alternative policy responses in Europe: an empirical analysis. *Lancet* 2009;374:315–323.
[16] Catalano R. The effect of deviations from trends in national income on mortality: The Danish and USA data revisited. *Eur J Epidemiol* 1997;13:737–743.
[17] Jin RL, Shah CP and Svoboda TJ. The impact of unemployment on health: A review of the evidence. *J Public Health Policy* 1997;18:275–301.
[18] Paul KI and Moser K. Unemployment impairs mental health: Meta-analyses. *J Vocat Behav* 2009;74:264–282.
[19] Martikainen PT and Valkonen T. Excess mortality of unemployed men and women during a period of rapidly increasing unemployment. *Lancet* 1996;348:909–912.
[20] Martikainen P, Mäki N and Jäntti M. The effects of unemployment on mortality following workplace downsizing and workplace closure: A register-based follow-up study of Finnish men and women during economic boom and recession. *Am J Epidemiol* 2007;165:1070–1075.
[21] Böckerman P and Ilmakunnas P. Unemployment and self-assessed health: Evidence from panel data. *Health Econ* 2009;18:161–179.
[22] Dadgar I and Norström T. Short-term and long-term effects of GDP on traffic deaths in 18 OECD countries, 1960–2011. *J Epidemiol Commun Health* 2016;71:146–153.
[23] Cutler D, Deaton A and Lleras-Muney A. The determinants of mortality. *J Econ Perspect* 2006;20:97–120.
[24] Pritchett L and Summers LH. Wealthier is healthier. *J Human Resource* 1996;31:841–868.
[25] Norström T and Grönqvist H. The Great Recession, unemployment and suicide. *J Epidemiol Commun Health* 2015;69:110–116.
[26] The Maddison-Project. http://www.ggdc.net/maddison/maddison-project/home.htm. 2013 version. 2013. (accessed 1 June 2021).
[27] De Boef S and Keele L. Taking time seriously. *Am J Polit Sci* 2008;52:184–200.
[28] Choi I. Unit root tests for panel data. *J Int Money Finance* 2001;20:249–272.
[29] Engle RF and Granger CWJ. Co-integration and error correction: representation, estimation, and testing. *Econometrica* 1987;55:251–276.
[30] Westerlund J. Testing for error correction in panel data. *Ox Bull Econ Stat* 2007;69:709–748.
[31] Beck N and Katz JN. What to do (and not to do) with time-series cross-section data. *Am Polit Sci Rev* 1995;89: 634–647.
[32] Bewley RA. The direct estimation of the equilibrium response in a linear dynamic model. *Econ Lett* 1979;3:357–361.
[33] Neumayer E. Recessions lower (some) mortality rates: Evidence from Germany. *Soc Sci Med* 2004;58:1037–1047.
[34] Chay KY and Greenstone M. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Quart J Econ* 2003;118:1121–1167.
[35] Dehejia R and Lleras-Muney A. Booms, busts, and babies' health. *Quart J Econ* 2004;119:1091–1130.
[36] Granados JAT. *Mortality and economic fluctuations: Theories and empirical results from Spain and Sweden*. New School University, 2003.

Swedish Institute for Social Research

Dissertation Series

If not otherwise stated, the dissertation has been submitted at Stockholm University. The dissertations in the Swedish language contain an English summary.

1. Ante Farm (1986): *A Model of the Price Mechanism*

2. Michael Tåhlin (1987): *Arbetets värde och kostnader. En studie av lönearbetets konsekvenser för individen* (The Value and Costs of Work. A Study of the Consequences of Wage Labour for the Individual)

3. Lucienne Portocarero (1987): *Social Mobility in Industrial Societies: Women in France and Sweden*

4. Lennart Erixon (1987): *Profitability in Swedish Manufacturing - Trends and Explanations*

5. Peter Hedström (1988): *Structures of Inequality: A Study of Stratification within Work Organizations* (Harvard University)

6. Jan O. Jonsson (1988): *Utbildning, social reproduktion och social skiktning* (Education, Social Reproduction, and Social Stratification)

7. Jaime Behar (1989): *Trade and Employment in Mexico*

8. Carl le Grand (1989): *Interna arbetsmarknader, ekonomisk segmentering och social skiktning. En studie av arbetslivsstrukturer, anställningsstabilitet och löneskillnader* (Internal Labour Markets, Economic Segmentation and Social Stratification)

9. Ryszard Szulkin (1989): *Privat eller offentligt? Organisationsstruktur och arbetsförhållanden under olika ägandeformer* (Private or Public? Organizational Structure and Working Conditions under Different Forms of Ownership)

10. Sten-Åke Stenberg (1990): *Vräkt ur folkhemmet. En studie av vräkningarna i Sverige under 1900-talet* (Evictions in the Welfare State)

11. Olle Lundberg (1990): *Den ojämlika ohälsan. Om klass- och könsskillnader i sjuklighet* (Inequality in Ill Health. On Class and Sex Differences in Illness)

12. Susanne Oxenstierna (1990): *From Labour Shortage to Unemployment? The Soviet Labour Market in the 1980s*

13. Sven E. Olsson (1990): *Social Policy and Welfare State in Sweden*

14. Joakim Palme (1990): *Pension Rights in Welfare Capitalism. The Development of Old-Age Pensions in 18 OECD Countries 1930 to 1985*

15. Mahmood Arai (1990): *Essays on Non-Competitive Wage Differentials*

16. Johan Fritzell (1991): *Icke av marknaden allena. Inkomstfördelningen i Sverige* (Not Solely by the Market: Income Distribution in Sweden)

17. Eugenia Kazamaki (1991): *Firm Search, Sectoral Shifts, and Unemployment*

18. Lena Schröder (1991): *Springpojkar och språngbrädor. Om orsaker till och åtgärder mot ungdomars arbetslöshet* (Dead-end Jobs and Upgrading Plans. On Reasons Behind and Programmes Against Youth Unemployment) (Uppsala universitet)

19. Olli Kangas (1991): *The Politics of Social Rights. Studies on the Dimensions of Sickness Insurance in OECD Countries* (Helsingfors universitet)

20. Göran Sidebäck (1992): *Kampen om barnets själ. Barn- och ungdomsorganisationer för fostran och normbildning 1850-1980* (The Struggle for the Soul of the Child. Child- and Youth Organizations for Rearing and Normbuilding 1850-1980)

21. Christina Axelsson (1992): *Hemmafrun som försvann. Övergången till lönearbete bland gifta kvinnor i Sverige 1968-1981* (The Housewife that Disappeared. Married Women's Transition to Paid Employment in Sweden 1968-1981)

22. Hjördis D'Agostino (1992): *Why Do Workers Join Unions? A Comparison of Sweden and OECD Countries*

23. Maria Nyström Peck (1994): *Childhood Class, Body Height and Adult Health. Studies on the Relationship between Childhood Social Class, Adult Height and Illness and Mortality in Adulthood*

24. Tomas Korpi (1994): *Escaping Unemployment. Studies in the Individual Consequences of Unemployment and Labour Market Policy*

25. Irene Wennemo (1994): *Sharing the Costs of Children. Studies on the Development of Family Support in the OECD Countries*

26. Viveca Östberg (1996): *Social Structure and Children's Life Chances. An Analysis of Child Mortality in Sweden*

27. Stig Blomskog (1997): *Essays on the Functioning of the Swedish Labour Market*

28. Katarina Richardson (1997): *Essays on Family and Labor Economics*

29. Håkan Regnér (1997): *Training at the Job and Training for a New Job: Two Swedish Studies*

30. Kristiina Manderbacka (1998): *Questions on Survey Questions on Health* (Helsingfors universitet)

31. Helen Dryler (1998): *Educational Choice in Sweden: Studies on the Importance of Gender and Social Contexts*

32. Michael Gähler (1998): *Life After Divorce. Economic, Social and Psychological Well-being Among Swedish Adults and Children Following Family Dissolution*

33. Lena Granqvist (1998): *A Study of Fringe Benefits. Analysis Based on Finnish Micro Data* (Åbo Akademi)

34. Olof Bäckman (1998): *Longitudinal Studies on Sickness Absence in Sweden*

35. Anna Thoursie (1998): *Studies on Unemployment Duration and on the Gender Wage Gap*

36. Christian Kjellström (1999): *Essays on Investment in Human Capital*

37. Gunnar Isacsson (1999): *Essays on the Twins Approach in Empirical Labor Economics*

38. Eero Carroll (1999): *Emergence and Structuring of Social Insurance Institutions: Comparative Studies on Social Policy and Unemployment Insurance*

39. Peter Skogman Thoursie (1999): *Disability and Work in Sweden*

40. Helena Persson (1999): *Essays on Labour Demand and Career Mobility*

41. Magnus Nermo (1999): *Structured by Gender. Patterns of Sex Segregation in the Swedish Labour Market. Historical and Cross-national Comparisons*

42. Ola Sjöberg (2000): *Duties in the Welfare State. Working and Paying for Social Rights*

43. Mikael Lindahl (2000): *Studies of Causal Effects in Empirical Labor Economics*

44. Ingemar Kåreholt (2000): *Social Class and Mortality Risk*

45. Ingalill Montanari (2000): *Social Citizenship and Work in Welfare States: Comparative Studies on Convergence and on Gender*

46. Ann-Zofie E. Duvander (2000): *Couples in Sweden. Studies on Family and Work*

47. Mia Hultin (2001): *Consider Her Adversity. Four Essays on Gender Inequality in the Labor Market*

48. Carin Lennartsson (2001): *Still in Touch. Family Contact, Activities and Health Among Elderly in Sweden*

49. Per Båvner (2001): *Half Full or Half Empty? Part-time Work and Well-being Among Swedish Women*

50. Per Gillström (2001): *Fair Care: Four Essays on the Allocation and Utilization of Health Care*

51. Magnus Bygren (2001): *Career Outcomes in the Swedish Labor Market: Three Contextual Studies*

52. Björn Öckert (2001): *Effects of Higher Education and the Role of Admission Selection*

53. Susanne Alm (2001): *The Resurgence of Mass Unemployment. Studies in Social Consequences of Joblessness in Sweden in the 1990s*

54. Ann-Christin Jans (2002): *Notifications and Job Losses on the Swedish Labour Market*

55. Sara Ström (2002): *A Shared Experience. Studies on Families and Unemployment*

56. Roger Vilhelmsson (2002): *Wages and Unemployment of Immigrants and Natives in Sweden*

57. Charlotte Samuelsson (2002): *Att göra eller inte göra... Arbetslösas fritidsdeltagande, sökaktivitet, anställningsmöjligheter och tidsstruktur (*To Do or Not to Do…Unemployed's Leisure Participation, Search Activity, Job Opportunities and Time Structure*)*

58. Tommy Ferrarini (2003): *Parental Leave Institutions in Eighteen Post-war Welfare States*

59. Jenny Säve-Söderbergh (2003): *Essays on Gender Differences in Economic Decision-making*

60. Kenneth Nelson (2003): *Fighting Poverty: Comparative Studies on Social Insurance, Means-tested Benefits and Income Redistribution*

61. Marie Evertsson (2004): *Facets of Gender: Analyses of the Family and the Labour Market*

62. Gabriella Sjögren (2004): *Essays on Personnel Economics and Gender Issues*

63. Kent Friberg (2004): *Essays on Wage and Price Formation in Sweden*

64. Ingrid Esser (2005): *Why Work? Comparative Studies on Welfare Regimes and Individuals' Work Orientations*

65. Åsa Olli Segendorf (2005): *Job Search Strategies and Wage Effects for Immigrants*

66. Pathric Hägglund (2006): *Natural and Classical Experiments in Swedish Labour Market Policy*

67. Lars Brännström (2006): *Phantom of the Neighbourhood. Longitudinal Studies on Area-based Conditions and Individual Outcomes*

68. Helena Holmlund (2006): *Education and the Family. Essays in Empirical Labour Economics*

69. Pernilla Andersson (2006): *Four Essays on Self-Employment*

70. Johanna Kumlin (2007): *Disentangling Sex Segregation. Studies on the Roots and Routes of Labour Market Sex Segregation*

71. Anders Böhlmark (2007): *School Reform, Educational Achievement and Lifetime Income. Essays in Empirical Labor Economics*

72. Krister Sund (2007): *Teachers, Family and Friends. Essays in Economics of Education*

73. Christer Gerdes (2008): *Studying the Interplay of Immigration and Welfare States*

74. Katarina Boye (2008): *Happy Hour? Studies on Well-Being and Time Spent on Paid and Unpaid Work*

75. Lena Lindahl (2008): *Family Background and Individual Achievement – Essays in Empirical Labour Economics*

76. Richard Baltander (2009): *Education, Labour Market and Incomes for the Deaf/Hearing Impaired and the Blind/Visually Impaired*

77. Sara Brolin Låftman (2010): *Children's Living Conditions. Studies on Health, Family and School*

78. Charlotta Magnusson (2010): *Mind the Gap. Essays on Explanations of Gender Wage Inequality*

79. Lalaina Hirvonen (2010): *Essays in Empirical Labour Economics: Family Background, Gender and Earnings*

80. Martin Hällsten (2010): *Essays on Social Reproduction and Lifelong Learning*

81. Marta Lachowska (2010): *Essays in Labor Economics and Consumer Behavior*

82. Marieke Bos (2010): *Essays in Household Finance*

83. Patrik Gränsmark (2010): *Essays on Economic Behavior, Gender and Strategic Learning*

84. Elin Olsson (2011): *Social Relations in Youth: Determinants and Consequences of Relations to Parents, Teachers, and Peers*

85. Karin Halldén (2011): *What's Sex Got to Do with It? Women and Men in European Labour Markets*

86. Frida Rudolphi (2011): *Inequality in Educational Outcomes. How Aspirations, Performance, and Choice Shape School Careers in Sweden*

87. Susan Niknami (2012): *Essays on Inequality and Social Policy. Education, Crime and Health.*

88. Martin Nybom (2013): *Essays on Educational Choice and Intergenerational Mobility.*

89. Jenny Torssander (2013): *Equality in Death? How the Social Positions of Individuals and Families are Linked to Mortality.*

90. Yerko Rojas (2014): *Childhood Social Exclusion and Suicidal Behavior in Adolescence and Young Adulthood*

91. Cecilia von Otter (2014): *Educational and Occupational Careers in a Swedish Cohort*

92. Per Olof Robling (2015): *Essays on the Origins of Human capital, Crime and Income Inequality*

93. Hrvoje Kap (2015): *Comparative Studies of Vocational Education and Training*

94. André Richter (2016): *Essays on the Intergenerational Transmission of Disadvantage: The Role of Prenatal Health and Fertility*

95. Per Engzell (2016): *Intergenerational Persistence and Ethnic Disparities in Education*

96. Niklas Kaunitz (2017): *Workers, Firms and Welfare*

97. Martin Berlin (2017): *Essays on the Determinants and Measurement of Subjective Well-Being*

98. Julia Boguslaw (2017): *When The Kids Are Not Alright: Essays on Childhood Disadvantage and Its Consequences*

99. Sara Kjellsson (2017): *Sick of Work? Questions of Class, Gender and Self-Rated Health*

100. Roujman Shahbazian (2018): *Sibling Configuration and Adulthood Outcomes: The Case of Two-Child Families*

101. Daniel Fredriksson (2019): *Enabling employment? Drivers and outcomes of active labour market policies in comparative perspective*

102. Linus Andersson (2019): *Essays on Family Dynamics: Parenting, Fertility and Divorce in Sweden*

103. Charlotta Boström (2019): *Education, skills and gender: The impact of a grading reform and the business cycle on labor market outcomes*

104. Johan Westerman (2020): *Motives Matter: Intrinsic motivation in work learning and labor market performance*

105. Simon Hjalmarsson (2021): *Taking part on equal terms? Associations between Economic Resources and Social Participation among Swedish Adolescents"*

106. Anni Erlandsson (2022): *Gender, Parenthood, Ethnicity and Discrimination in the Labor Market. Experimental Studies on Discrimination in Recruitment in Sweden.*

This thesis consists of four self-contained essays on the economics of education and health.

1. The effect of ordinal rank in school on educational achievement and income in Sweden

2. School autonomy and subject-specific timetables

3. Short-term and long-term effects of GDP on traffic deaths

4. Is there a link between cardiovascular mortality and economic fluctuations?

Iman Dadgar
is an economist interested in the economics of education and health. He holds M.Sc. in Economics from Stockholm University.

ISBN 978-91-7911-830-3
ISSN 0283-8222

Department of Economics

Stockholm University