Mobility in Science

John Källström

Lund University

DOCTORAL DISSERTATION

by due permission of the School of Economics and Management,

Lund University, Sweden.

To be defended at Holger Crafoord EC3:211, Lund

on October 11, 2019 at 10:15.

Faculty opponent

John McHale, National University of Ireland Galway

Organization
Lund University
Department of Economics
P.O. Box 7082
S-220 07 Lund, Sweden

Document name
Doctoral dissertation

Date of issue
October 11, 2019

Author
John Källström

Sponsoring organization

Title and subtitle
Mobility in Science

Abstract

The thesis consists of three independent chapters. The chapters are all empirical studies of scientists and the role that different types of mobility play in shaping researchers' scientific output and careers.

Chapter 1, *The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform*, uses the liberalization of Swedish work migration, in 2008, to study the effect of immigration on the publishing productivity of incumbent academics in Sweden. The reform led to a sharp increase in the number of Asian academic researchers and PhD-students coming to Sweden. Identification relies on both the suddenness of the supply shock and that departments with past exposure to Asian migration saw relatively larger inflows of Asian migrants. Results show that the supply shock increased the publication output of incumbent researchers. Positive effects are found to be mainly explained by increased publishing productivity of already prolific incumbent researchers. For less productive incumbents, evidence instead suggests crowding-out effects and reduced productivity.

Chapter 2, *Does mobility across universities raise scientific productivity?*, studies the effects of inter-university mobility on researcher productivity. The study suggests substantial gains from mobility on scientific output. Mobility effects are not explained by promotions taking place jointly with a move. Positive effects are found among individuals who move between universities and not for those who move to or from university colleges. Moreover, we find that the positive effect of moving only applies to researchers in medicine, natural sciences and engineering, and technology, with no effect of mobility found in the social sciences and the humanities.

Chapter 3, *On the social origins of scientists: How intergenerational (im-)mobility shape science*, investigates the social background of PhD-graduates. Results suggest that parents' characteristics are essential determinants for obtaining a PhD-level education. Of particular importance is if the parent also holds a PhD. This association is gender- and field-specific and is large in comparison to other sources of exposure to researcher careers in the childhood environment. Taken together, the results suggest that the family environment is crucial for obtaining a PhD-level education. Moreover, the study reveals that the existence of intergenerational spillovers also affects patenting and publishing behavior in a later research career.

Keywords
Economics of science, mobility, scientists, scientific productivity, migration

Classification system and/or index terms (if any)
JEL Classification: J24, J61, O31, I23

Supplementary bibliographical information

Language
English

ISSN and key title
0460-0029 Lund Economic Studies no. 216

ISBN
978-91-7895-111-6 (print)
978-91-7895-112-3 (pdf)

Recipient's notes

Number of pages
197

Price

Security classification

Distributed by Department of Economics, P.O. Box 7082, S-220 07 Lund, Sweden

Signature ______________________ Date September 9, 2019

Mobility in Science

John Källström

LUND UNIVERSITY

Lund Economic Studies Number 216

Distributed by the Department of Economics
Lund University
P.O. Box 7082
S-220 07 Lund
Sweden

ISBN 978-91-7895-111-6 (print)
ISBN 978-91-7895-112-3 (pdf)
ISSN 0460-0029

Printed in Sweden by Media-Tryck, Lund University
Lund 2019

To Carin

Contents

Abstract i

Acknowledgment iii

Introduction 1

1 Background 1
- 1.1 The nature of knowledge and scientific knowledge production . . . 3
- 1.2 How knowledge is produced . . . 5
- 1.3 Mobility and spillovers in science . . . 7

2 Contribution of the thesis 10
- 2.1 The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform . . . 10
- 2.2 Does mobility across universities raise scientific productivity? . . . 12
- 2.3 On the social origins of scientists: How intergenerational (im-)mobility shapes science . . . 13

References 15

Chapter I: The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform 23

1 Introduction 24

2 Institutional setting: work migration in Sweden 27
- 2.1 Pre-reform period . . . 27
- 2.2 The 2008-reform . . . 28

3 Data and summary statistics 29
- 3.1 Descriptive statistics on foreign-born researchers in Sweden 31
- 3.2 Defining treatment . . . 36

4 Empirical strategy 42
- 4.1 Finding the appropriate control group . . . 42
- 4.2 Econometric specification . . . 46

5 The productivity of incumbent Swedish researchers **48**
5.1 Main results 48
5.2 Robustness of main results 51
5.3 Heterogeneity 52

6 Conclusion **58**

References **61**

A Supplementary tables and figures **65**

B Robustness analysis **77**

C Matched sample results **82**

Chapter II: Does mobility across universities raise scientific productivity? **89**

1 Introduction **90**

2 Job mobility and researcher productivity **92**

3 Data **94**

4 Empirical strategy **98**
4.1 Econometric specification 98
4.2 Threats to identification 99
4.3 Inverse probability of treatment censoring weights 101

5 Results **103**
5.1 Main effect of mobility 103
5.2 Heterogeneous effects of mobility 104

6 Concluding discussion **111**

References **115**

A Tables and graphs **119**

B Construction of our matched sample **124**

C Matched sample results **127**

D Estimations of IPTC weights **134**

Chapter III: On the social origins of scientists: How intergenerational (im-)mobility shapes science 139

1 Introduction 140

2 Postgraduate education in Sweden 144

3 Data 146

4 Obtaining a PhD 149

4.1 Descriptive statistics 150
4.2 Regression specification 155
4.3 Baseline results 155
4.4 Differences across genders 159
4.5 The choice of educational field for PhD 161
4.6 Alternative sources of exposure to PhD education and careers 165

5 Scientific and inventive productivity and social background 167

6 Conclusion 172

References 174

A Supplementary Materials for Section 4 178

B Robustness analysis for Section 4 189

C Supplementary Materials for Section 5 192

Abstract

The thesis consists of three independent chapters. The chapters are all empirical studies of scientists and the role that different types of mobility play in shaping researchers' scientific output and careers.

Chapter 1, *The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform*, uses the liberalization of Swedish work migration in 2008, to study the effect of immigration on the publishing productivity of incumbent academics in Sweden. The reform led to a sharp increase in the number of Asian academic researchers and PhD-students coming to Sweden. Identification relies on both the suddenness of the supply shock and that departments with past exposure to Asian migration saw relatively larger inflows of Asian migrants. Results show that the supply shock increased the publication output of incumbent researchers. Positive effects are found to be mainly explained by increased publishing productivity of already prolific incumbent researchers. For less productive incumbents, evidence instead suggests crowding-out effects and reduced productivity.

Chapter 2, *Does mobility across universities raise scientific productivity?*, studies the effects of inter-university mobility on researcher productivity. The study suggests substantial gains from mobility on scientific output. Mobility effects are not explained by promotions taking place jointly with a move. Positive effects are found among individuals who move between universities and not for those who move to or from university colleges. Moreover, we find that the positive effect of moving only applies to researchers in medicine, natural sciences and engineering, and technology, with no effect of mobility found in the social sciences and the humanities.

Chapter 3, *On the social origins of scientists: How intergenerational (im-)mobility shapes science*, investigates the social background of PhD-graduates. Results suggest that parents' characteristics are essential determinants for obtaining a PhD-level education. Of particular importance is if the parent also holds a PhD. This association is gender- and field-specific and is large in comparison to other sources of exposure to researcher careers in the childhood environment. Taken together, the results suggest that the family environment is crucial for obtaining a PhD-level education. Moreover, the study reveals that the existence of intergenerational spillovers also affects patenting and publishing behavior in a later research career.

Keywords: Economics of science, mobility, scientists, scientific productivity, migration
JEL Classification: J24, J61, O31, I23

Acknowledgment

When telling people what my thesis is about, I usually say that I do research about researchers. This prompts two responses depending on the background of the person. When speaking to other academics, they tend to become very interested. I guess that people like it when someone shows an interest in what they do, also (especially?) academics. However, when speaking to people outside academia, the response tends to be something like: *'How very meta.'* I tend to agree. Researching researchers allows oneself to reflect on the research process and academic life in general. As an example, in the introductory chapter of this thesis, I discuss how time, effort, and resources are the three primary inputs in scientific knowledge production. Indeed, this thesis would not have been possible without many late nights, struggles, and a personal computer. Although, also mentioned in the introduction is the growing importance of collaboration for research. In the research on researchers, this usually implies some formal scientific cooperation, such as co-authoring a paper, grant applications, or some research project. Having now produced a bit of knowledge, I would add to that list that knowledge production would not be possible without also the *informal* research collaborations one forms during the process. To be specific, colleagues, friends, and family without whom writing this thesis would not have been possible. Here I would like to thank some of them.

First of all, I wish to thank my supervisor, Olof Ejermo. Olof, I believe you did an excellent job. You always made time for meetings and to comment and suggest improvements for my work. Your eye for details in everything from inconsistencies in arguments to (not in the least!) formatting and language, really helped me improve my thinking and writing. I also thank you for providing an open and forgiving atmosphere where we could unconditionally discuss my — as well as your — research (often over a BMC-, LTH-, or (lately) Medicon-lunch). In academia, people tend to be quick to critique, but you always took care to listen to and made an effort to understand my sometimes loosely formed ideas. I never hesitated to ask you a question or for advice.

Likewise, my co-supervisor, Kaveh Majlesi, deserves a big thanks. Kaveh, your expertise in everything empirical, really helped me improve my research designs. Especially the first chapter of this thesis benefitted from your comments. I do not know if you realize it, but you helped tremendously in developing the empirical strategy in that paper. I also learned a lot from you during seminars, meetings, and courses. I am grateful to all recurring participants at the brown bag seminars, who at various stages commented and provided input on my work. None mentioned, none forgotten. I am thankful to Joakim Gullstrand, who was my supervisor for my bachelor's and master's theses, and who encouraged me to apply to the PhD-program.

Moreover, all the CIRCLE-people deserve some recognition. Before starting my PhD, I was a research assistant at CIRCLE, and I kept my affiliation throughout my PhD. It has been instructive and intellectually enriching to be around people from various disciplines who all share an interest in topics related to innovation. I think my affiliation with CIRCLE helped me broaden my perspective on research, methods, and theory. Among the people at CIRCLE, I would especially like to thank Martin Andersson and Torben Schubert, who at various stages delivered insightful and always constructive criticism.

Additionally, I am thankful to Claudio Fassio, who together with Olof Ejermo co-authored the second chapter of this thesis. Your understanding of academic life both helped shape that chapter and my thinking. A special thanks also go to the discussant on my final seminar, Patrick Gaule, from the University of Bath. Who read and provided helpful feedback on the chapters in my thesis.

I am also grateful to the two consecutive coordinators of the PhD-program at the department, Jerker Holm and Tommy Andersson, for always looking out for the best interest of your PhD-students. I want to thank Pontus Hansson for entrusting me with teaching duties, including the responsibilities of being the main teacher for a course. A big thanks also to the admin at the department who were always helpful in all matters big and small: Nathalie, Mariana, Jenny, Rikke, Marie, Peter, and Azra.

All my friends in the PhD-student collegiate at the department, both past and present, deserve a big thanks. Karl, things were never the same after you left. Anna, it was great to have you as a roommate during the last year. I believe I am the sixth study buddy from the old study group to get a PhD. Fairly good turnout, I would say. Lina thank you for putting in a good word for me. Without that, I do not think I would have started this journey in the first place. Also, thanks to Danial, Demid, Devon, Dinh-Vihn, Dominice, Dominika, Erik, Fredrik, Hampus, Hjördis, Jim, Jonas, Josefin, Jörgen, Kristoffer, Marco, Margaret, Pol, Sandra, Sanna, Sara Mi, Sara Mo, Thomas, Yana Pe, Yana Pr, and Zahra who all contributed in different ways to make my PhD so much more enjoyable.

I want to thank my parents, Ulf and Gerd. Your continual support and encouragement have been essential. You taught me the value of hard work, endurance, and of education. I am forever grateful. Also, thanks to my sister, Selma, although I do not know if you will understand for what. Finally, all my gratitude goes out to Carin. I would not have been able to do this without you at my side. You are simply the best.

Lund, September 2019
John

Introduction

Introduction

1 Background

According to the definitions laid out by Audretsch et al. (2002), science is the search for new knowledge based on observation of facts and truths. It starts with something known and searches for something new and unknown. In turn, technology is defined as the application of new knowledge learned through science to solve a problem. Technological change is the rate at which new knowledge is diffused and put into use in the economy. Closely related are the concepts of invention and innovation. In a sense, invention is to innovation what science is to technology. An invention is the creation of something new. An invention becomes an innovation once it is put to use. Innovations may be new products, new processes, or new organizational methods that are novel and add value to economic activity. Broadly speaking, underlying an innovation there is a new technology.

The above definitions are useful since they explain link between science, technology, innovation and the creation of economic value. In economics, Hicks (1932) were perhaps the first to make the link between technology and production explicit. Later, Solow (1956) and Swan (1956)[1] would independently develop theoretical models in which technological change drives productivity growth in the economy, although these early theories considered technological change as an exogenous factor, something given, outside the scope of the model. It was not until the 1980s and the work by Romer (1986) and Lucas (1988) and the new endogenous growth theory, that the formal connection between economic growth and the accumulation of knowledge and human capital was made explicit. This connection was then further developed by Romer (1990), Grossman & Helpman (1991) and Aghion & Howitt (1992), in which innovation, and thereby technological change, are the result of idea creation and research activities in the economy. In line with this, economic historians have put the formation of the institution of science — as a medium for the creation and diffusion of knowledge — as a cornerstone of the modern industrial sector and the subsequent centuries of unprecedented technological change (Landes, 2010; Mokyr, 2002; Rosenberg, 1974).

Thus, there seems to be a consensus among scholars that science is at the heart of the creation of new technologies and innovations. However, how do new scientific ideas come about? At its core, the generation of

[1]Building on the earlier work by Harrod (1939) and Domar (1946).

new ideas is a human activity. On the one hand, apocryphal anecdotes, such as Archimedes' jumping out of his bath yelling '*Eureka!*,' or Newton being hit on the head by a falling apple, make it seem as if luck or coincidence — serendipity — is the essential feature of scientific progress. If, on the other hand, scientific knowledge production depends upon individuals, institutions, and incentives, then economic research should play an important role in increasing our understanding in this area.

Following the pioneering work of Zvi Grilishes on the diffusion of hybrid corn (see e.g., Griliches, 1957, 1960), there has been ample empirical work on the diffusion of new technologies and its effect on the economy (see e.g., Bloom et al., 2013; Ellison et al., 2010; Greenstone et al., 2010; Jaffe et al., 1993; Thompson & Fox-Kean, 2005). Yet, comparatively little is still known about the origin of ideas and the creation of the technologies in the first place. This is likely not due to a lack of interest, but rather to past data limitations. However, in recent years we have seen a surge of studies using new data and methods that allows economists to link mentors and students, collaborators, and intellectual peers to characterize the scientific production team.[2] In particular, the data on scientists, papers, and citations together with methodological advances in the analysis of quasi-experimental data have allowed scholars to begin to credibly characterize the knowledge production process, the conditions under which scientists create new knowledge and the benefits that follow.

One important theme that has emerged from the recent literature is the notion that an increasing burden of knowledge, because of an ever-expanding scientific frontier, has led to greater scientific specialization, longer training periods, and to an increased propensity to collaborate (Wuchty et al., 2007). This realization has cast a shadow over the potential for ideas-based growth because it implies that discoveries and innovation are becoming more difficult over time (Gordon, 2017; Jones, 2009). However, whether this concern is warranted or not is an empirical question and will depend on the institutions and incentives knowledge producers face.

In this thesis, I heed this concern and will in a series of empirical studies explore scientific knowledge production and the knowledge producers. The purpose of this chapter is to provide a brief background to the topics related to these issues, to provide a context for the chapters, and to specify the contribution of each of the chapters in the thesis. I start in the two following sections to discuss the properties of knowledge and knowledge production. I then discuss the role of mobility in relation

[2] See e.g., Azoulay et al. (2018) and the references therein for an overview of some of the recent advances.

to knowledge production. Finally, I summarize the contributions of the individual chapters contained in the thesis.

1.1 The nature of knowledge and scientific knowledge production

Knowledge production implies a good is being produced. However, knowledge is not like most other goods. Instead, knowledge is what economists refer to as a public good (Dasgupta & David, 1994).[3] Knowledge — as the product of research — is not depleted once it is shared; when it is made public you cannot easily exclude others from using it; and, the incremental cost of adding additional users is close to zero. However, unlike a typical public good, the stock of knowledge is not diminished by extensive use. Instead, it is usually enlarged because fundamental advances can serve as inputs in other research, leading to further discoveries, as well new products and processes. Although this implies a high social return for the dissemination of new knowledge, it creates a problem for the individual producers of knowledge because others can take advantage of their discoveries and inventions without paying the full price for developing them (Arrow, 1962). Moreover, even though there is a considerable consensus for science being central to technological change and economic growth, due to the uncertain nature of scientific work, the economic value of basic research is difficult to forecast and sometimes even to gauge in retrospect (Dasgupta & David, 1994). From this, it follows that investment in research is highly uncertain.

Taken together, the public good-like nature of knowledge and the uncertainty of the value of the results from research endeavors, explain why economists usually consider the social returns of knowledge creation to be higher than the private returns. This means that markets will invest less in research than what is socially optimal. The recognition of the social benefits of research, in combination with this fundamental *market failure*, is the main argument for public investment in science and for institutional arrangements such as intellectual property rights

[3] To be precise, knowledge becomes a public good once it is codified and disseminated to others. Usually, there is a distinction made between knowledge and information, information being the codification of knowledge (Dasgupta & David, 1994, p. 493). Here I make no such distinction. Moreover, there are several forms of friction associated with knowledge diffusion. For example, cognitive limits determine what and how much knowledge an individual can acquire, or whether certain skills or prior knowledge might be needed to decode it. For this reason, the correct definition of knowledge/information might rather be a *quasi*-public good.

(e.g., patents), which allow the researcher to earn temporary monopoly rents on his/her invention (Arrow, 1962).

How these institutional arrangements are structured differs across sectors of the economy. There is an important difference between public and private knowledge producers. To fix ideas, we can follow David and Dasgupta's (1994) discussion where the authors make a distinction between the science and technology domains of basic research. According to David and Dasgupta's definitions, *science* is associated with (public) academic science, whereas *technology* refers to the domain of (private) industrial research and development. What distinguishes the two communities is not the methods, the nature of knowledge, nor (even) the source of funding. Instead, the distinction has to do with the social organization of research in the public and private sectors and with the reward structure in the respective sector.

It is important to recognize that the line between the science and technology domains is often blurred. The same individual can move between domains or be in both at the same time. What matters is what the scientists do with the findings from their research efforts. Research that is undertaken in secret in order to sell the results — by, for example, patents or copyrights — belong to the industry domain of science according to Dasgupta and David. What distinguishes academic science is the institution of codifying and disseminating research findings to others. In effect, creating a public good, which is usually done by publishing research results. That is not to say that academic scientists do not patent, nor that private scientists do not publish research results. Instead, it has to do with the motives of research: is the motive to create a public good from research that is useable to others, or is it to exclude others from using the knowledge so as to accrue economic rent, in practice creating a private good? Thus, there are good reasons to suspect that scientific knowledge will *leak out*, or *spill over*, to other sectors of the economy where it may contribute to growth and innovation (Adams, 1990; Adams et al., 2006; Zucker et al., 1998).

Why do academic scientists publish research results? The question has occupied scholars of science for decades. The answer has to do with the institution of *priority of discovery*. The term, which was first coined by sociologist Robert K. Merton (1957; 1973), indicates that the rewards in academia are related to collegiate recognition, in the sense that the goal of academic scientists is to be first to communicate an advance in knowledge, and that their reward is the recognition of the scientific community. In this way *priority of discovery* awards intellectual property rights to the scientist who is first to make an important discovery. Examples of such rewards range from the use of eponyms for important discoveries (e.g., Haley's comet, Planck's constant, or Alzheimer's disease) to awarding prizes (the Nobel Prize probably the most famous). However, the most common, and necessary form, to earn recognition is to codify research results into a publication (Stephan, 2012). That attribution rights to discoveries are of economic importance in academia shows not in the least when considering that junior academic researchers may be willing to trade intellectual property rights in order to secure authorship rights on publications with senior colleagues, even in cases where they have rights to both (Lissoni et al., 2013).

Thus, what distinguishes academic from industrial scientists is the non-market reward system in academia, which has evolved to provide incentives for scientists to produce and share their knowledge, thus behaving in socially desirable ways. In this thesis, the focus is directed towards investigating the publishing activity of academic scientists. That is to say, the science domain in Dasgupta and David's definition.

1.2 How knowledge is produced

In the previous section, we discussed knowledge as the output of research and how its public-good nature, in combination with high (but uncertain) social benefits, has evolved various institutions and incentive structures to make knowledge producers productive. We now turn to how knowledge is produced — that is to say, the inputs that go into knowledge production. In his Nobel lecture, Stigler (1983, p. 536) made the following remark:

> *Any new idea — a new conceptualization of an existing problem, a new methodology, or the investigation of a new area — can not be fully mastered, developed into the stage of a tentatively acceptable hypothesis, and possibly exposed to some empirical tests without a large expenditure of time, intelligence, and research resources.*

Stigler's observation describes the inputs in the scientific knowledge production function, namely: time, cognitive effort, and research resources (Stephan, 2012).

The first inputs in research are time and cognitive effort. Research builds on prior research, so the knowledge base of the individual researcher plays a crucial role. As Isaac Newton famously put it: "*If I have been able to see further, it was only because I stood on the shoulders of giants*." Thus, individual researchers need to accumulate knowledge to be able to contribute to knowledge production. However, as the knowledge stock in society grows, the process of accumulating knowledge may take a longer time, which in turn will imply a longer time for making discoveries (Jones, 2010). Making society's ability to stand on the shoulders of giants dependent not only on generating knowledge, but also on the quality of the mechanisms for storing, certifying, and accessing that knowledge (Furman & Stern, 2011).

Related to knowledge accumulation is the ability of the researchers themselves. Scientists are often characterized as having above-average intelligence. Recent work confirms that this is many times true and is related to research output (Aghion et al., 2017). Though, stamina and motivation also play an important part in the research process (Fox, 1983; Hagstrom, 1965). Moreover, researchers are often portrayed as deriving pleasure or utility from solving problems and making discoveries (Sauermann & Roach, 2010; Stern, 2004). However, little is known about how abilities and preferences shape knowledge production, or how and when they are formed. For instance, much literature on human-capital accumulations is concerned with the importance of nature versus nurture in the development of skills and adult outcomes (see e.g., Black & Devereux, 2011). However, when and how scientific aptitude and ability are formed is still something of a black box. Although, recent studies have begun to disentangle the role of childhood environment from inborn characteristics in shaping the career of scientists (e.g., Bell et al., 2018).

The final major input in research is resources. Stephan (2012) describes how the need for equipment and costly infrastructure differs across scientific disciplines. For instance, in the social sciences, the necessary equipment for cutting-edge research is generally a personal computer, access to a database, and maybe a research assistant or two. This is in stark contrast to the natural sciences where the resource requirements are usually considerably more extensive, often involving access to substantial equipment. For example, for a researcher in particle physics, time on an accelerator might be a necessary condition to do research, while an astronomer requires time on a telescope. That access to equipment and resources are important for researchers also show in their behavior. For instance, studies have shown how curtailed access to certain important research instruments may jeopardize entire research streams (see e.g., Murray & Stern, 2007; Murray, 2010). Moreover, the importance of resources also shows in the willingness of researchers to trade recognition in the form of co-authorship, or citations to get access to certain equipment (Hagstrom, 1965; Walsh et al., 2005, 2007).

One way that scientists respond to a lack of time, knowledge, cognitive efforts, ability, or resources is through teamwork. That the scientific research team is crucial for the knowledge-production process is becoming increasingly clear (Azoulay et al., 2010; Jaravel et al., 2018; Jones et al., 2008; Waldinger, 2011). An indication of the increasing importance of teamwork in science is the growing number of co-authors on scientific publications (Wuchty et al., 2007). One explanation relates to the ever-expanding accumulation of knowledge in society, to which scientists respond by acquiring narrower expertise, which in turn requires larger team sizes (Jones, 2009). An additional explanation, which does not rule out the former, is that with the advent of information technology, communication costs are reduced, making it cheaper to collaborate and form teams across institutions (Agrawal & Goldfarb, 2008). In that way researchers may easier form collaborations to access distant inputs not available in their current institution. Stephan (2012) concludes that it is not enough to decide to do research; one must also have access to the necessary research inputs.

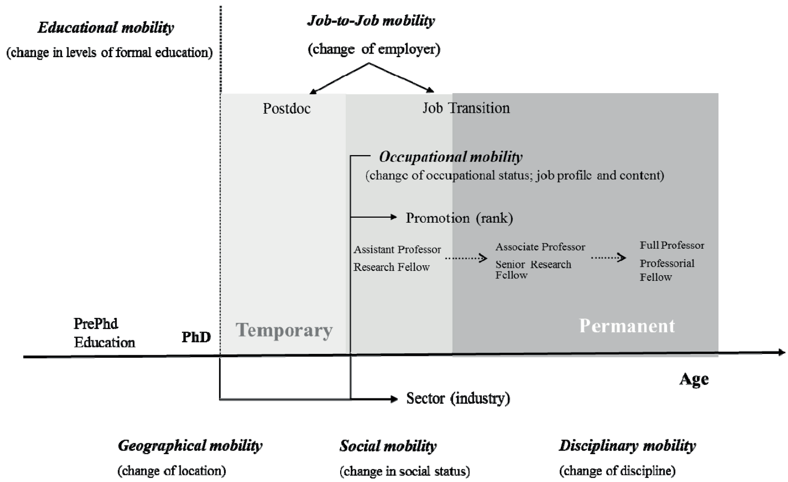
1.3 Mobility and spillovers in science

In the previous section, we discussed the inputs to knowledge production. Mobility is an important aspect in relation to this because it allows the researchers to change the inputs for their scientific knowledge production. In particular, mobility allows researchers to access resources, collaborators, or other intellectual inputs that otherwise would be unaccessible, which presumably will shift their scientific productivity. Moreover, mobility may increase the visibility of the individual scientist extending his/her network of citing peers (Azoulay et al., 2011). In addition, mobility may affect not only the productivity of the mover but also that of the new peers since it also allows them to use the inputs that the mobile scientists bring with them.

To clarify the concept of mobility, Figure 1 shows a typology of different types of mobility at various stages in the career of an academic scientist according to Fernández-Zubieta et al. (2015). The figure is useful since it shows how mobility can be broadly defined and classified based on the motives and nature of changes in the environment of a researcher. In particular, the authors distinguishes between educational (change across levels of formal education), geographical (change of location), job-to-job (change of employer), occupational (change of occupational status, e.g. job profile and content), sectoral (change in the sector of employment), social (change in social position) and disciplinary mobility (change of disciplinary focus).

When considering mobility, it is important to note that one type of mobility does not rule out another. One can easily see how geographical, job, and occupational mobility often occur simultaneously when researchers move to another university to obtain promotion to a higher academic rank. Likewise, it seems likely that sectoral mobility often happens together with educational mobility once a PhD has been completed and the recent graduate chooses whether to work in academia or industry.

Figure 1: Researcher mobility in a life course perspective



Source: Fernández-Zubieta et al. (2015)

The various types of mobility has the potential change the inputs into a researcher's scientific production in different ways. Perhaps the most obvious examples in this regard are geographical and job mobility, which will move the researcher to a new location and/or employer with different colleagues and resources. Moreover, changes in job profile (occupational mobility) may change the time and effort that the researcher has at his/her disposal for research. Similarly, moving to a different sector or discipline may completely change the inputs and direction of scientific work. In addition, social mobility is in this context related to access to research inputs. For example, an academic researcher who joins a high-quality department could be considered upwardly mobile in the social hierarchy of academia whereas a researcher who joins a lower quality department could be considered downwardly mobile. Since movements up and down university hierarchies are associated with different access to resources and peers, social mobility aspects are also relevant to an analysis of researcher mobility and its effects.

However, there are ample reasons to suspect that the effects of mobility may also *spill over* to peers. To understand why, consider that there exists friction in the diffusion of knowledge. Publication is not synonymous with replicability (Stephan, 2010). Some knowledge, or techniques, are not — or cannot be — codified and disseminated. When such knowledge is embedded in an individual, it is referred to as *tacit knowledge* (cf., Polanyi, 1966). Tacit knowledge is not public, and the obstacles to diffusion are usually considered great. For example, colocation and face-to-face interaction are sometimes required for diffusion to take place (von Hippel, 1994).[4]

In the context of academia, we can think about much knowledge that is *sticky* in the sense that it is not easily codified and may instead be embedded in individuals or groups. Examples include skills concerning how to run an experiment, the know-how necessary to extract a particular protein strain, or how to operate specific equipment in the lab. If such knowledge complement the existing knowledge in a university department or lab it may enhance the productivity of co-workers. Thus, labor mobility has the potential to act as a channel for knowledge-flows between academic institutions, firms, or sectors of the economy (Azoulay et al., 2011; Moen, 2005; Zucker et al., 2002). Moreover, studies have shown how knowledge is many times bounded by geographical distance and social networks (Agrawal et al., 2006; Breschi & Lissoni, 2009; Jaffe et al., 1993). Thus, researcher mobility can influence the diffusion of knowledge by bridging these boundaries.

In light of this, it is unsurprising that the mobility of researchers and the establishment of research networks garners interest from policymakers (e.g., OECD, 2008). However, recent studies on geographical mobility between academic institutions suggest that spillovers and productivity effects from researcher mobility depend on complementarities in scientific trajectories existing between the incoming mover and the incumbent peers (Agrawal et al., 2017). In cases where such complementarities are lacking, or when researchers compete for local resources and intellectual space, mobility may instead diminish the knowledge production of the individual scientist (Borjas & Doran, 2012, 2015a,b).

[4] Although it has long been assumed that interaction and colocation facilitate knowledge diffusion, we have only recently seen convincing empirical studies using exogenous variation to identify the effects of colocation. For example, Catalini (2018) provides evidence of the importance of colocation for collaboration among researchers using forced relocation of research labs due to asbestos removal. His findings corroborate that research labs that were randomly allocated close to each other were more likely to author joint work. Moreover, colocation was found to be especially important for the creation of new research ventures, implying that there was considerable friction in the diffusion of knowledge. Furthermore, using data from a large-scale field experiment, Boudreau et al. (2017) confirm that the formation of new research collaborations is subject to considerable friction.

Thus, whether mobility enhances or diminishes productivity in largely an empirical question.

2 Contribution of the thesis

This thesis consists of three independent chapters. The chapters are all empirical studies that fall under the auspices of the economics of science and innovation. All three chapters deal with scientists and the role that different types of mobility play in shaping researchers' scientific output and careers.

Chapter I deals with geographical mobility. The chapter examines how the immigration of foreign-born researchers affects the publication production of incumbent researchers. That is, it is a study on the existence of spillovers between immigrant and incumbent researchers. On a similar note, Chapter II is about the effects of job-to-job and geographical mobility on productivity, but from the point of view of the mover. The chapter emphasizes the relationship between geographical, job, and social mobility across the university hierarchy. Finally, in Chapter III, the focus is shifted to the socio-economic determinants of entry into scientific careers and PhD-level education, as well as the impact that such selection has on inventive and publishing outcomes. The chapter concerns itself with the educational and social aspect of mobility, trying to answer the question of who becomes a scientist in the first place.

All three chapters use a recently collected database that follows the publication activity of about 30,000 academic researchers in Sweden over time. The database is unique in that it links publication data to Swedish administrative registers, allowing researchers to access a host of detailed information about the researchers. The database is further described in Ejermo et al. (2016) and `http://paris.circle.lu.se`.

2.1 The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform

The first chapter, *The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform*, contributes to the literature on the effect of immigration on scientists' productivity. It estimates the effect that increased immigration of researchers has on the publishing productivity of incumbent researchers. The following observation motivates the paper: on the one hand, migrants are new colleagues, collaborators, and team members that bring with them new knowledge and skills that could complement and enhance incumbents' research output. On the other hand, migrants and incumbents also compete for the same scarce resources, which means that an increase in the supply of researchers through immigration might instead crowd-out incumbents, reducing their productivity.

To identify a causal effect of immigration on incumbents' productivity, I use a recent migration policy reform in Sweden. In 2008, Sweden liberalized work migration from non-European countries. In effect, this meant giving the same comprehensive rights to non-European work migrants as citizens from other EU member states. I find that this led to a sharp and sudden increase in the number of Asian academic researchers and PhD-students coming to Sweden. To estimate the effect of this Asian-supply shock on incumbent Swedish scientists, I follow a recent paper by Borjas et al. (2018) and compare the publishing productivity between incumbent researchers employed at departments with larger and smaller inflows in a difference-in-differences setup. Identification relies on both the suddenness of the supply shock and that university departments with a larger past exposure to Asian migration saw relatively larger inflows of Asian migrates after the reform.

In the chapter, I find that the average publication productivity increased for incumbents, but only in departments with the largest exposure to Asian immigration. On average, incumbent researchers in these departments published about 0.4 additional publications each year after the reform relative to controls. Compared to the average publication rate per person and year in this group, the estimates correspond to a 45 percent increase. In total, this translates to about 1,080 additional scientific papers in the post-reform period 2008—2011. As outcomes, I also consider citations per year and if collaboration with Asian immigrants increased after the reform, the latter quantified as the proportion of publications with at least one Asian migrant co-author to all publications each year. The results for the additional outcomes reveal that highly exposed incumbents neither published lower-quality papers in response to increased competition, nor by them collaborating relatively more with Asian migrants after the reform. Instead, the results suggest that the inflow of Asians enhanced the productivity of incumbents.

To gauge possible mechanisms to explain the results, I leverage the richness of the data to investigate heterogeneity in the overall response to the supply shock across scientific fields, academic positions, ethnicity and relative position in the productivity distribution. As one might expect, the effect is heterogeneous across all these dimensions. Moreover, the results reveal that the increased productivity of already above-median productive researchers explain most of the positive effects, both across and within fields and positions. At the same time, I also find evidence of less-prolific researchers decreasing publishing productivity, indicating some instances of crowding-out effects for this group. Furthermore, the supply shock increased the productivity of incumbent researchers that have an Asian ethnic background, which also increased collaboration with the Asian migrants.

In sum, the paper shows how of benefits and cost of an increase in the supply of scientists unevenly distribute across incumbents. I find that productivity spillovers differed across field, position, and productivity of incumbents, suggesting that the potential of migration to influence the rate of knowledge production depends crucially on how the newcomers interact with the pre-existing workforce.

2.2 Does mobility across universities raise scientific productivity?

In the second chapter, *Does mobility across universities raise scientific productivity?*[5], we provide the first country-level analysis of the effect of researchers' mobility on productivity. Thus, in contrast to the first chapter, this paper deals with the productivity of the mover. The lack of mobility among researchers has attracted substantial interest from policymakers and scholars because it is often claimed that low mobility across academic institutions hampers the diffusion of ideas and may lead to intellectual inbreeding. This observation has inspired literature concerned with the determinants of mobility among academic researchers (e.g., Azoulay et al., 2017). However, less is known about the effect of inter-university mobility for the moving researcher.

In the chapter, we examine the effects of mobility on scientific productivity in terms of both publication output and the quality of scientific output, gauged through citation-weighted publication output. We also distinguish the pure mobility effect from various factors that may have an impact on the overall effect, such as the initial level of productivity of researchers, the interaction between mobility and promotion, the importance of the status of the university of origin and destination, as well as differences across disciplinary fields. The empirical analysis addresses selection using inverse probability treatment censoring weights (Azoulay et al., 2009).

Our results suggest that mobility induces a long-lasting increase in a researcher's publications by 32 percent and citations by 63 percent. Moreover, such mobility effects are not explained by promotions taking

[5] Co-authored with Olof Ejermo and Claudio Fassio.

place at the same time as the move, suggesting that it is the mobility event itself that explains a large part of the productivity increases. Furthermore, positive effects are found among individuals who move between universities and not for those who move to or from lower-ranked university colleges. Finally, we find that the positive effect of moving only applies to researchers in the *hard sciences*, with no effect found in the social sciences and the humanities.

Overall, we argue that our results suggest considerable sorting effects in the university system where more research-intensive departments will both hire better researchers and get a bigger effect in terms of publications from their hires. In addition, the role of mobility for productivity differs by discipline. We find that the positive effect of mobility on productivity only applies to the capital- and team-intensive hard sciences, suggesting that access to equipment, labs, or collaborators are especially important for explaining the effects of mobility.

2.3 On the social origins of scientists: How intergenerational (im-)mobility shapes science

The third chapter, *On the social origins of scientists: How intergenerational (im-)mobility shape science*, contributes to the emerging literature on the social origins of knowledge workers (see, Aghion et al., 2017; Bell et al., 2018). In particular, the paper investigates the role of parents for obtaining a PhD-level education. Who becomes a researcher is not random, and parents may play a crucial role in shaping the preferences, skills, and abilities necessary to enter a career as a scientist. Although it is well-established that children tend to obtain the same level of education as their parents (e.g., Björklund & Salvanes, 2011), this chapter takes a first step towards establishing the link for PhD-level education and to investigate the implications of such intergenerational correlation on patenting and publishing behavior of PhD holders.

Using intergenerational-linked register data, the chapter first establishes the importance of socio-economic background for obtaining a PhD-level education. Parents of PhD students tend to earn higher wages and are more educated than the average parent. In particular, they are also more likely to have a PhD-level education themselves. Specifically, having a PhD-educated parent increases the probability of children also having a PhD-level education by 2—3 percentage points, which is equivalent to moving from the 5th to the 9th decile of grades for a child's chances of obtaining a PhD. Moreover, I find that this correlation has a strong gender-specific component, where the boy-father and mother-daughter relationships are of importance. Also, it is more likely for children with PhD-educated parents to obtain a PhD-level education in the same narrowly defined field as their parent compared to PhD students whose parents hold a lesser university degree. The results taken together provide suggestive evidence that the observed intergenerational correlations are not mainly driven by inherited differences in inborn ability across children. Instead, the results are indicative of the importance of the childhood environment.

In the second part of the analysis, I investigate the implications of intergenerational spillovers for inventive and scientific performance in terms of patent and publication output of PhD graduates. The results suggest a complex relationship: while PhD-educated men with a PhD-educated parent do not perform differently from other PhD holders, I find that PhD women are less likely to apply for a patent if the mother is PhD-educated and are less likely to publish if the father is PhD-educated. These relationships also extend to the case were I instead consider the number of patents and publications.

Taken together, the results from the first and second parts of the analysis suggest that parents' have an effect not only on their children's decision to enter a career in science but also on what type of career the children pursue after entering the field. That the effect is gender- and field-specific further suggests that the answer can be found in the childhood environment. In particular, parents provide different role models, access to networks, or other specific human capital that help children pursue careers in certain fields. Thus, overall parents have an important influence on the allocation of human capital and for scientific and inventive activity in society.

References

Adams, J., Clemmons, J. R., & Stephan, P. (2006). How rapidly does science leak out? NBER Working Paper 11997, National Bureau of Economic Research.

Adams, J. D. (1990). Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy*, *98*(4), 673–702.

Aghion, P., Akcigit, U., Hyytinen, A., & Toivanen, O. (2017). The social origins of inventors. NBER Working Paper 24110, National Bureau of Economic Research.

Aghion, P. & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2).

Agrawal, A., Cockburn, I., & Mchale, J. (2006). Gone but not forgotten: Knowledge flows, labor mobility, and enduring social relationships. *Journal Of Economic Geography*, *6*(5), 571–591.

Agrawal, A. & Goldfarb, A. (2008). Restructuring research: Communication costs and the democratization of university innovation. *American Economic Review*, *98*(4), 1578–1590.

Agrawal, A., McHale, J., & Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*, *46*(4), 853–867.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The Rate And Direction Of Inventive Activity: Economic And Social Factors* (pp. 609–626). National Bureau of Economic Research.

Audretsch, D. B., Bozeman, B., Combs, K. L., Feldman, M., Link, A. N., Siegel, D. S., Stephan, P., Tassey, G., & Wessner, C. (2002). The economics of science and technology. *The Journal Of Technology Transfer*, *27*(2), 155–203.

Azoulay, P., Ding, W., & Stuart, T. (2009). The impact of academic patenting on the rate, quality and direction of (public) research output. *The Journal Of Industrial Economics*, *57*(4), 637–676.

Azoulay, P., Ganguli, I., & Zivin, J. G. (2017). The mobility of elite life scientists: Professional and personal determinants. *Research Policy*, *46*(3), 573–590.

Azoulay, P., Graff-Zivin, J., Uzzi, B., Wang, D., Williams, H., Evans, J. A., Jin, G. Z., Lu, S. F., Jones, B. F., Börner, K., Lakhani, K. R., Boudreau, K. J., & Guinan, E. C. (2018). Toward a more scientific science. *Science*, *361*(6408), 1194–1197.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal Of Economics*, *125*(2), 549–589.

Azoulay, P., Zivin, J. S. G., & Sampat, B. N. (2011). The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. In J. Lerner & S. Stern (Eds.), *The Rate and Direction of Inventive Activity Revisited* (pp. 107–155). University of Chicago Press.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2018). Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*, *134*(2), 647–713.

Björklund, A. & Salvanes, K. G. (2011). Education and family background: Mechanisms and policies. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education*, volume 3 (pp. 201–247). Elsevier.

Black, S. & Devereux, P. (2011). Recent developments in intergenerational mobility. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (1 ed.)., volume 4 chapter 16, (pp. 1487–1541). Elsevier.

Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, *81*(4), 1347–1393.

Borjas, G. J. & Doran, K. B. (2012). The collapse of the Soviet Union and the productivity of American mathematicians. *The Quarterly Journal Of Economics*, *127*(3), 1143–1203.

Borjas, G. J. & Doran, K. B. (2015a). Cognitive mobility: Labor market responses to supply shocks in the space of ideas. *Journal Of Labor Economics*, *33*(S1), S109–S145.

Borjas, G. J. & Doran, K. B. (2015b). Which peers matter? The relative impacts of collaborators, colleagues, and competitors. *Review Of Economics And Statistics*, *97*(5), 1104–1117.

Borjas, G. J., Doran, K. B., & Shen, Y. (2018). Ethnic complementarities after the opening of China: How Chinese graduate students affected the productivity of their advisors. *Journal of Human Resources*, *53*(1), 1–31.

Boudreau, K. J., Brady, T., Ganguli, I., Gaule, P., Guinan, E., Hollenberg, A., & Lakhani, K. R. (2017). A field experiment on search costs and the formation of scientific collaborations. *The Review of Economics and Statistics*, *99*(4), 565–576.

Breschi, S. & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439–468.

Catalini, C. (2018). Microgeography and the direction of inventive activity. *Management Science*, *64*(9), 4348–4364.

Dasgupta, P. & David, P. A. (1994). Toward a new economics of science. *Research Policy*, *23*(5), 487–521.

Domar, E. D. (1946). Capital expansion, rate of growth, and employment. *Econometrica*, *14*(2), 137.

Ejermo, O., Alder, C., Fassio, C., & Källström, J. (2016). Publications of Academic Researchers In Sweden (PARIS). Technical report, CIRCLE, Lund University. `http://paris.circle.lu.se`.

Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, *100*(3), 1195–1213.

Fernández-Zubieta, A., Geuna, A., & Lawson, C. (2015). What do we know of the mobility of research scientists and impact on scientific production. In A. Geuna (Ed.), *Global Mobility of Research Scientists* (pp. 1–33). Elsevier.

Fox, M. F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*, *13*(2), 285–305.

Furman, J. L. & Stern, S. (2011). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review*, *101*(5), 1933–1963.

Gordon, R. J. (2017). *Rise and Fall of American Growth*. Princeton Univers. Press.

Greenstone, M., Hornbeck, R., & Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, *118*(3), 536–598.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, *25*(4), 501.

Griliches, Z. (1960). Hybrid corn and the economics of innovation. *Science*, *132*(3422), 275–280.

Grossman, G. M. & Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, *58*(1), 43.

Hagstrom, R. G. (1965). *Scientific Community*. Basic Books.

Harrod, R. F. (1939). An essay in dynamic theory. *The Economic Journal*, *49*(193), 14.

Hicks, J. R. (1932). *The Theory of Wages*. Macmillan.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577–598.

Jaravel, X., Petkova, N., & Bell, A. (2018). Team-specific capital and innovation. *American Economic Review*, *108*(4-5), 1034–73.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review Of Economic Studies*, *76*(1), 283–317.

Jones, B. F. (2010). Age and great invention. *Review Of Economics And Statistics*, *92*(1), 1–14.

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, *322*(5905), 1259–1262.

Landes, D. S. (2010). *The Unbound Prometheus*. Cambridge University Press.

Lissoni, F., Montobbio, F., & Zirulia, L. (2013). Inventorship and authorship as attribution rights: An enquiry into the economics of scientific credit. *Journal of Economic Behavior & Organization*, *95*, 49–69.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3–42.

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, *22*(6), 635–659.

Merton, R. K. (1973). *The Sociology Of Science: Theoretical And Empirical Investigations*. University Of Chicago Press.

Moen, J. (2005). Is mobility of technical personnel a source of R&D spillovers? *Journal Of Labor Economics*, *23*(1), 81–114.

Mokyr, J. (2002). *The Gifts Of Athena: Historical Origins Of The Knowledge Economy*. Princeton University Press.

Murray, F. (2010). The oncomouse that roared: Hybrid exchange strategies as a source of distinction at the boundary of overlapping institutions. *American Journal of Sociology*, *116*(2), 341–388.

Murray, F. & Stern, S. (2007). Do formal intellectual property rights hinder the free flow of scientific knowledge? *Journal of Economic Behavior & Organization*, *63*(4), 648–687.

OECD (2008). *The Global Competition For Talent: Mobility Of The Highly Skilled*. Paris: Organization for Economic Cooperation and Development.

Polanyi, M. (1966). *The Tacit Dimension*. The University of Chicago Press.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*(5), 1002–1037.

Romer, P. M. (1990). Endogenous technological change. *Journal Of Political Economy*, *98*(5 Pt 2).

Rosenberg, N. (1974). Science, invention and economic growth. *The Economic Journal*, *84*(333), 90–108.

Sauermann, H. & Roach, M. (2010). A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. *Research Policy*, *39*(3), 422–434.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65.

Stephan, P. E. (2010). The economics of science. In B. Hall & N. Rosenberg (Eds.), *Handbook of The Economics of Innovation, Vol. 1*, volume 1 (pp. 217–273). Elsevier.

Stephan, P. E. (2012). *How Economics Shapes Science*, volume 1. Harvard University Press Cambridge, Ma.

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, *50*(6), 835–853.

Stigler, G. J. (1983). Nobel lecture: The process and progress of economics. *Journal of Political Economy*, *91*(4), 529–545.

Swan, T. W. (1956). Economic Growth and Capital Accumulation. *Economic Record*, *32*(2), 334–361.

Thompson, P. & Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, *95*(1), 450–460.

von Hippel, E. (1994). "sticky information" and the locus of problem solving: Implications for innovation. *Management Science*, *40*(4), 429–439.

Waldinger, F. (2011). Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge. *Review Of Economics And Statistics*, *98*(5), 811–831.

Walsh, J. P., Cho, C., & Cohen, W. M. (2005). View from the bench: Patents and material transfers. *Science*, *309*(5743), 2002–2003.

Walsh, J. P., Cohen, W. M., & Cho, C. (2007). Where excludability matters: Material versus intellectual property in academic biomedical research. *Research Policy*, *36*(8), 1184–1203.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

Zucker, L., Darby, M. R., & Torero, M. (2002). Labor mobility from academe to commerce. *Journal Of Labor Economics*, *20*(3), 629–660.

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1998). Intellectual human capital and the birth of U.S. biotechnology enterprises. *American Economic Review*, *88*(1), 290–306.

Chapter I

The impact of immigration on scientists' productivity: Evidence from a Swedish policy reform

Abstract

In 2008, Sweden liberalized its work migration policy. The reform led to a sharp and sudden increase in the number of Asian academic researchers and PhD-students coming to Sweden. I use this immigration shock to study the effect of an increase in the supply of foreign researchers on the publication productivity of incumbent academic researchers. The identification relies on both the suddenness of the supply shock and that university departments with a larger past exposure to Asian migration saw relatively larger inflows of Asian migrates after the reform. I find that the supply shock increased the publication output of incumbent researchers in departments with the largest degree of past exposure by 45 percent. Effects are heterogeneous across fields, faculty positions, ethnicity and past productivity of incumbents. Positive effects are mainly explained by increased publishing productivity of already prolific incumbent researchers. For less productive incumbents, I instead find evidence of crowding-out effects. Thus, the supply shock widened the gap between highly productive and less-producitive incumbents.

Keywords: Skilled migration, scientific productity, spillovers, academics
JEL Classification: J24, J61, O31, I23

The author is grateful for valuable comments and suggestions from conference and seminar attendants at Utrecht University, Bordeaux University and Lund University. Moreover, I wish to especially thank Patrick Gaule for providing helpful comments on an earlier version of this paper.

1 Introduction

The academic labor market is increasingly international in scope with more scientists pursuing a career or training outside their home country (see e.g., Hunter et al., 2009; Weinberg, 2011). According to the GlobSci survey of academic scientists across 16 countries in Europe and North America, in 2011 the average share of foreign-born faculty ranged from around 20 to almost 60 percent (Franzoni et al., 2012). In Sweden, the setting of this paper, the share of foreign-born researchers has almost doubled in recent years — from around 15 percent in the mid-1990s to just under 30 percent by 2012.

As countries institute policies such as tax incentives or fast-tracked visa applications to attract foreign high-skilled workers and students (see OECD, 2008), understanding how skilled migration shapes knowledge production is important and has wide-ranging implications for the design of migration, innovation, and science policy. Moreover, the existence of spillovers and peer effects in knowledge production has further implications for the optimal allocation of public research and development (R&D) funds, the distribution of credit among scientists, scientific reputation building, and ultimately the design of research incentives that foster innovation and continued economic growth (Azoulay et al., 2010).

It is often noted that internationally mobile scientists are positively selected in terms of both past and future performance (Franzoni et al., 2014; Levin & Stephan, 1999; Gaulé & Piacentini, 2013). However, a large body of literature has found that migrant researchers also influence incumbent researchers' knowledge production (e.g., Borjas & Doran, 2012; Ganguli, 2015; Kerr & Lincoln, 2010; Moser et al., 2014), although the direction of this effect remains unclear. On the one hand, immigrant researchers may possess specialized human capital, or ideas and skills, that "spill over" to incumbent researchers, increasing their productivity. A growing body of literature on the importance of teamwork and spillovers in scientific work speaks to this view (Jones, 2009; Jaravel et al., 2018; Azoulay et al., 2010; Oettl, 2012). On the other hand, newcomers may also compete with incumbents for limited location- and field-specific resources such as research funding, lab space, department slots, or journal space (Borjas & Doran, 2012; Borjas et al., 2018). Hence, when demand cannot shift in response to changed supply, increased inflows may instead crowd out native scientists, diminishing their opportunities for research and reducing their productivity. However, it is worth noting that successful immigrant researchers may also increase resources in the long run, which could equally increase the productivity of incumbents down the line.

In order to study spillovers between migrant and incumbent scientists, I use an exogenous change to the inflow of work or study migrants due to recent immigration policy reform in Sweden. The reform implied a wide-ranging deregulation of migration from countries outside the European Economic Area (EEA) to Sweden in 2008. For Swedish academia, I find that the reform acted as a powerful pull-factor for researchers and PhD-students of predominantly Asian descent, leading to a sharp and sudden increased inflow of academics from Asian countries. To examine how this Asian supply shock affected incumbent researchers' publication productivity, I employ unique data matching on the Scopus publication output of over 57,510 academic researchers in Sweden from 2001 through 2011 to university-employee-linked administrative registers.

To tease out a causal effect, I follow a similar approach as Borjas et al. (2018), who showed how Chinese doctoral students in the US tended to gravitate to university departments with an already relatively large ethnic Chinese faculty following the opening of China in 1989. In a similar vein to their paper, I use the fact that the Swedish university departments that attracted most Asian researchers after the reform were also the departments with a history of Asian-born faculty to create distinct treatment and control groups of incumbent Swedish researchers. Treatment is assigned based on incumbent researchers' pre-reform affiliation to departments with more or less past exposure to Asian migration. Employing a difference-in-differences strategy, I contrast the publication performance of the treatment and control groups before and after the 2008-reform. To account for pre-reform differences across researchers in exposed and less exposed university departments in a flexible way, I rely on a non-parametric matched sample approach to construct a control group of researchers with similar characteristics as treated researchers.[1]

In the main analysis I find that the average publication productivity, as measured by number of publications per year, increased for incumbents working in departments with the largest exposure to Asian migration. On average, incumbent researchers in these departments published about 0.4 additional publications each year after 2008 relative to controls. Compared to the average publication rate per person and year in this group, the estimates correspond to a 45 percent increase. In total, this translates to about 1,080 additional scientific papers in the post-reform period 2008–2011. As outcomes, I also consider citations

[1]For recent studies using a similar design, see: Azoulay et al. (2010, 2011) and Ganguli (2015).

per year and if collaboration with Asian immigrants increased after the reform, the latter quantified as the proportion of publications with at least one Asian migrant co-author to all publications each year. The results for the additional outcomes reveal that exposed incumbents did not publish lower-quality publications in response to increased competition nor collaborate relatively more with Asian migrants after the reform. Instead, these results suggest that the inflow of Asians generated positive productivity spillovers to incumbents.

Although a few recent studies have found evidence for positive productivity spillovers existing between scientists (e.g., Moser et al., 2014), the overall evidence is mixed. Borjas et al. (2018) suggest that there seem to be powerful productivity spillovers among scientists who are directly collaborating with each other (as in, e.g., Azoulay et al., 2010; Waldinger, 2010). Whereas, competition effects seem to dominate researchers on rivalling scientific trajectories (see e.g., Waldinger, 2012; Borjas & Doran, 2012). Instead, recent evidence points to spillovers being mediated through complementarities existing between the newcomer and the incumbent scientists (Agrawal et al., 2017). Moreover, the strength of spillovers or peer effects is tightly linked to the nature of the academic discipline itself, some fields such as science and engineering being characterized by a faster-paced rise in the number of coauthors over time (see e.g., Wuchty et al., 2007; Agrawal et al., 2014; Jaravel et al., 2018), while in other disciplines research remains more of a solitary endeavor (e.g., mathematics in Borjas & Doran, 2012).

To explore under what conditions inflows of immigrants give rise to positive productivity spillovers, I leverage the richness of the data to investigate heterogeneity in the overall response to the supply shock across scientific fields, academic positions, ethnicity, and relative position in the productivity distribution. As expected, the effects are heterogeneous across all these dimensions. Moreover, the results reveal that researchers with an above-median number of cumulative publications before 2008 drive most of the positive effects, both across and within fields and positions. At the same time, I also find evidence of less prolific researchers instead decreasing publishing productivity after 2008, indicating that crowding-out effects occurred for this group. Effects are especially strong in the field of engineering, which is also the field most exposed to the supply shock. Furthermore, in line with earlier literature, the supply shock increased the productivity of incumbent researchers with an Asian ethnic background, who also disproportionally increased collaboration with the Asian migrants.

To summarize, productivity spillovers were generated by a supply shock of Asian migrants to Sweden. However, specific disciplinary or positional effects, as well as the ethnicity and past productivity of incumbents, attenuate the overall effect. This reconciles some of the varying findings in the earlier literature on high-skilled immigrants and knowledge production. The importance of complementarity between scientists will differ depending on the scientific field and the relative position of incumbents. Hence, both the characteristics of the newcomers and those of the incumbents moderate what we can expect from increased inflows of foreign scientists. I find that established and prominent researchers are mainly able to better leverage the supply shock to increase productivity, whereas others may instead find themselves facing increased competition. This suggest that the migrants possessed skills and characteristics that complemented the knowledge production of incumbent star scientists, enhancing their productivity.

The main contribution of this paper is to provide credible causal estimates on the impact of high-skilled immigration on domestic knowledge production and to highlight how effects differ across the characteristics of incumbents.

2 Institutional setting: work migration in Sweden

2.1 Pre-reform period

Swedish policy vis-á-vis non-Nordic labor migrations has undergone two major shifts in the post-war era: first, in the early 1970s, with the introduction of restrictions; then, in 2008, with substantial liberalization (OECD, 2011).[2] In the period following the end of the Second World War, the growing Swedish industry needed workers — especially low-skilled workers — and policy was set up to meet labor shortages in industry. As in several other European countries, recruitment from abroad was centralized and managed by the government, usually in the form of bilateral agreements (the first such agreements were signed with Italy and Hungary in 1947).

The centralized system persisted until 1972, when large inflows of immigrants from Yugoslavia, Greece and Turkey, in combination with a general economic downturn, made trade unions push for further restrictions. In the new system, Sweden allowed two types of non-Nordic labor migration: i) employment to meet shortages that could not imme-

[2]Citizens of other Nordic countries where exempt from the labor migration system and have been able to move freely to, and work in, Sweden since 1954.

diately be filled by local workers; ii) permanent status offered to those in highly specialized occupations. For the first type, public authorities, in consultation with trade unions, examined the labor market situation for a particular occupation to establish if a shortage existed. Twice a year, the Public Employment Agency would publish a ”shortage list” and the Swedish Migration Board consulted it before granting work permits to applicants. In the years leading up to 2008, university teachers were not considered to be in short supply (Swedish Public Employment Service, 2008). This meant that it was unlikely that work permits were given to foreign academics before 2008. However, shortages were established for many occupations related to scientific fields where there was a larger surge of foreign academics after 2008, such as engineers, IT-specialists, and mathematicians.

For the second type of migration, permanent permits were issued to the most qualified workers, generally those with skills not readily available in Sweden, holding open-ended contracts with relatively high salaries. However, such a permit was issued in small numbers and seldom to universities. Between 2005 and 2008, only about 15 percent of these permits were issued to employees recruited by universities and research institutes. Instead, most permits were issued to a wide range of ICT, industrial and professional services employers (OECD, 2011).

In 1994, in a step to join the European Community, Sweden became a member of the European Economic Association (EEA)[3] Membership entailed accepting free mobility of workers from other member states. Thus, from the mid-1990s onwards, work migration policy was shifted completely towards third-country nationals outside the EEA/EU.

2.2 The 2008-reform

By the late 1990s, in response to worries about skill shortages and changing demographics, policymakeers pressed for reforms of Swedish labor migration policy. Thus, in 2008, after almost a decade of debate, the ruling center-right coalition in cooperation with the Green party, one of the opposition parties, pushed for a reform of Sweden’s work migration policy. The decision on the new law was taken by Parliament in November 2008 and became valid from December 15, 2008.

Under the new law, if a non-EEA/EU citizen got any kind of job offer with compensation according to a collective agreement or at the same level as the collective agreement in the sector, a work permit should be granted. The only requirement was that the vacancy should be publicly

[3] See Ejermo & Zheng (2018), for an analysis on the implications that the EEA membership had on the skill composition of immigration to Sweden.

announced. There were no quotas or skill requirements. The veto rights of unions were abolished. Note that to be hired as a researcher for longer than three month a work permit was needed.[4] Hence, the reform simplified hiring of non-EEA/EU citizens, also to universities. The law also gave foreign students (including doctoral students) some options after completing their education in Sweden. In this case, a job offer for at least six months after graduating was a necessary condition for getting a work permit.

The reform implied a major liberalization in Sweden's approach towards non-EEA/EU migrant workers. According to one study, Sweden went from being the most restrictive country in the OECD to the most open (Cerna, 2016). It is noteworthy that Sweden's work migration system marks a departure from the European and global trend towards more selective labor migration, where countries shape policy to attract highly qualified immigrants and avoid low-skilled immigrants (Emilsson & Magnusson, 2015). There is nothing in the Swedish rules that makes it easier for highly skilled persons to obtain a work visa. The same legislation applies to them as, for example, seasonal workers or lower-skilled labor migrants. Despite worries that the reform would lead to large inflows of mainly low-skilled workers, it is mainly the share of permits issued to workers with high qualifications that has increased, from 15 percent in 2009 to 26 percent in 2012 (Bellini, 2016). In the next section, we will look at the impact the reform had on the academic labor market.

3 Data and summary statistics

To proceed with the analysis, I constructed a data set linking publication records to official register data available from Statistics Sweden.

The publication data were collected in a project to link Swedish university researchers to their publication output.[5] In the project, we first contacted each Swedish research university requesting staff directories going back as far as possible. Out of 28 contacted universities, 25 responded and sent the requested information.[6] The universities were

[4] For non-EU/EEA citizens, new legislation came into force on July 1, 2008, based on the EUs Researchers Visa Directive. According to these rules no work permit is needed for third country citizens, if the purpose is to teach or lecture for a period shorter than three months.

[5] For more information, see Ejermo et al. (2016) and `https://paris.circle.lu.se/`

[6] Non-responding universities were Gävle University, Swedish University of Agricultural Sciences (SLU) and Stockholm School of Economics (SSE). Gävle University is a small

asked to supply personal identity number (PIN), first and last name, affiliation information and the work e-mail address of all their teaching and research staff. This encompassed all full professors, associate professors, postdocs, and PhD students, as well as individuals in the staff categories 'other research and teaching staff' and 'guest researchers.' The two latter categories are not academic positions per se but constitutes about 25 percent of all university employees in Sweden. According to a recent report from the Swedish Agency of Higher Education, just over 60 percent of "other research and teaching staff" consists of researchers, research engineers and research assistants. In addition, the other category also includes individuals without a PhD but working in a researcher capacity or cases where information about a PhD-degree is missing (p. 101 Swedish Higher Education Authority, 2017). Even though the other staff category is somewhat diffuse, I chose to include them in my main analysis since they commonly publish.

Using the information on names, affiliations and e-mail addresses, we matched the staff directories to publication records found in the Scopus database[7]. An advantage of Scopus is that the database assigns each registered author a unique identifier, an author-id. Thus, when a researcher is matched to a publication all other publications belonging to that researcher can be found through the associated author-id. In some cases, were Scopus confused individuals by assigning more than one researcher the same author-id, usually due to homonyms, we manually distinguished between matched individuals. In the end, using this strategy we identified 27,123 unique individuals from the staff directories who were subsequently matched to their publication records in Scopus. Although this sample constitutes only about 35 percent of all author-ids with a Swedish affiliation in Scopus, it covers about 85 percent of all publications with Swedish affiliated authors. This makes me confident to include non-matched researchers as non-publishing researchers in my analysis.[8] For further details on the construction of the database and an analysis showing that it indeed constitutes a representative sample of Swedish publishing researchers, see Ejermo et al. (2016). From each matched Swedish author-id in Scopus, we extracted yearly publication counts, and the number of citations accrued to each publication (counted in a three-year window from publication), as well

university college, SLU is a large research university that brings together much of Swedish research in agricultural and veterinary sciences, and SSE is privately-funded university that specializes in economics and business management.

[7] See urlhttps://www.scopus.com/

[8] In further robustness analysis, I also removed non-matched researchers. This does not impact the results. These results are available from the author upon request.

as links to co-authors identified at other Swedish universities.

In the last step, the researchers' PIN is used to link them to longitudinal employer-employee linked records of all university personnel in Sweden (*Registret för personal vid universitet och högskolor*) housed at Statistics Sweden. The data are register-based and contains detailed information on all Swedish university personnel employed at a Swedish university in November each year. For our purposes, I extract information on university and scientific field based on the classification of national research subjects (*Nationellt forskningsämne*) and staff category (i.e., as described above: full professors, associate professors, postdocs, PhD students, other research and teaching staff, and guest researchers). Since there is no reliable information on department affiliation in the register data, I use the information on the university of employment and national research subjects to create university-subject combinations, which are closely related to university departments — I will refer to these as "departments" in the subsequent analysis. The data on scientific field starts in 2001, however, due to breaks in the data where the information on fields is reclassified in 2012 it is impossible to use the data on the university-subject level for the last year.

In the final sample, I have 57,510 individuals in an unbalanced panel running from 2001 to 2011 working at 937 "departments" at 24 universities and 92 scientific fields.[9] From other registers, I link information on year of birth and for the foreign born region-of-birth based on broad country groups[10] and reason for settlement. To define non-EU academic researchers, I use two conditions: i) the first year the individual shows up in the register he or she should be employed at a university in a staff category considered, according to our definition, a researcher; and, ii) the reason for settlement should be work or study as affected by the 2008-reform.

3.1 Descriptive statistics on foreign-born researchers in Sweden

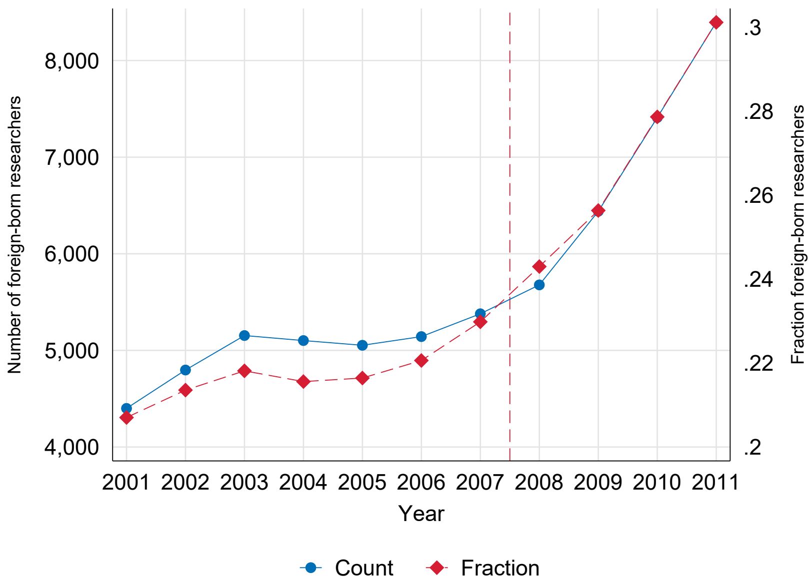
We now turn to the importance of foreign-born researchers in Swedish academia. I use the university-employee data linked to information on

[9] The reason that only 24 universities are included in the analysis is that the University College of Boras only supplied staff directories covering 2012 and is thus excluded. The school is a small college with 11,000 students and 730 employees, focused on social sciences and textile sciences.

[10] This information is only available for broader regions of origin: Nordic (excl. EU), EU (excl. Nordic), Europe (excl. Nordic and EU), Africa, Asia, Oceania, North America, South America and Former USSR (excl. Europe and EU).

the region of birth described in the preceding section, to compute the stock of foreign-born researchers at Swedish universities. Figure 1, below, plots the number and share of researchers born outside of Sweden who immigrated as adults. Looking at the figure, the importance of this group has grown over the studied period. In 2001, just over 4,000 researchers in Swedish academia were born outside Sweden, corresponding to about 21 percent of all academic researchers at that time. In 2011, the number of foreign-born researchers had doubled to more than 8,000, or almost 30 percent of Swedish academia.

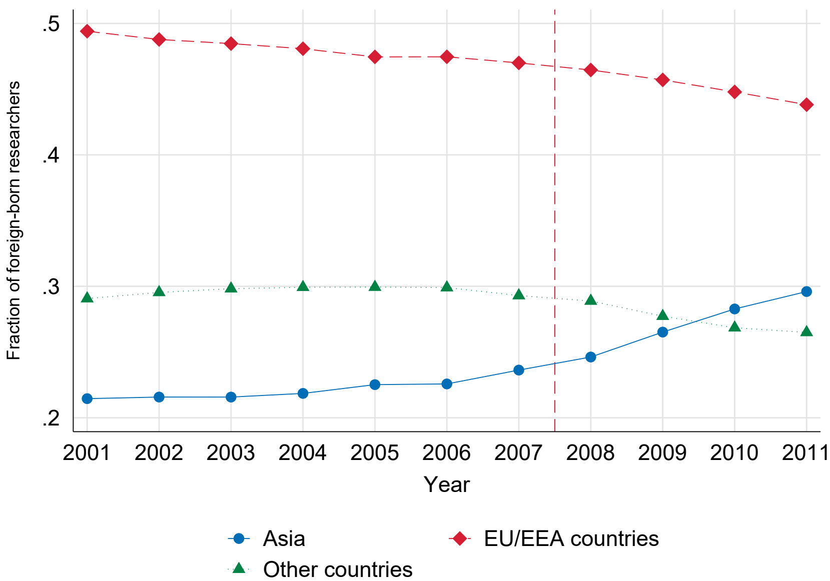
Figure 1: Stock of foreign-born academic scientists in Sweden, count (left axis) and as share of all university researchers (right axis) 2001–2011.



Note: Includes all immigrant who immigrated as adults.

Figure 2 instead plots the composition of the stock of foreign-born across regions-of-birth, distinguishing between researchers coming from EU/EEA countries, Asia and other countries. The figure reveals that the proportion of researchers of Asian origin increased after 2008 from around 23 percent of all foreign-born researchers in Sweden in 2007, to about 29 percent in 2011. At the same time the share of both EU/EEA and non-Asian groups decreased. To study non-EU/EEA and Asian immigration more carefully, Figure 3 illustrates the flow non-EU/EEA researchers arriving in Swedish universities between 2001 and 2011 with the stated reason for settlement either work or study as decomposed by region of origin. From this figure, we clearly see the size and suddenness of the supply shock after 2008. Compared to the number of arrivals from other non-EU/EEA country groups, the impact of the reform for these groups is very small in relation to the increased inflow from Asia. Thus, the deregulation of work migration to Sweden in 2008 acted as a pull-factor for Asian PhD students and researchers.

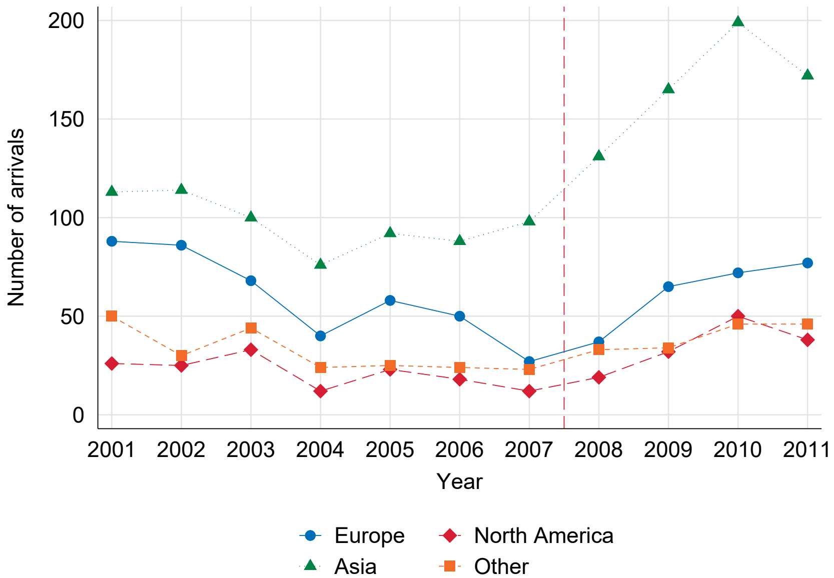
Figure 2: Fraction of foreign-born academic scientists in Sweden by region of origin, 2001–2011.



Note: Graph based on the stock of foreign-born academic researchers. Includes all immigrant who immigrated as adults.

To characterize the Asian migrants arriving after the reform in 2008, Table 1 shows descriptive statistics for the Asian-born migrants in the year of arrival. The total number arriving after 2008 is 784 by 2011. On average, the migrants published almost 1.68 publications per year; these publications have, on average, 6.67 citations after three years, and about 0.6 publications per year (35 percent) are co-authored with other

Figure 3: Flow of Non-EEA academic scientists to Sweden by region of birth, 2001–2011.



Note: Figure plots number of arrivals of migrant researchers from broad region-of-birth groups each year. The year of arrival is the first year a migrant appears in the data. Only includes migrants with work or study as their stated reason for settlement in this year.

researchers working in Sweden. About 50 percent of Asian immigrants are PhD students. But the group 'Other' is also large, mostly consisting of the staff category 'Other teaching and research staff' described in section 3. Next, we can note that the hard sciences make up the lion's share — 41 percent of Asian immigrants work at departments classified as Engineering/Technology, and another 50 percent are equally divided between the Life Sciences and Natural Sciences and Mathematics. Only 9 percent of Asian-born migrants post-2008 work in Social Sciences and Humanities. Lastly, a majority of Asian migrants work at Swedish universities that appeared on the top-100 of the 'Shanghai list' at some point between 2003 (the first year the list was published) and 2011.[11]

[11] The Shanghai list or Academic Ranking of World Universities is an annual ranking of all universities in the world. Appearing on the list is seen as an indicator of institutional quality. The list is especially important in many Asian countries for selecting a uni-

Table 1: Descriptive statistics of Asian immigrant researchers, 2008–2011

	Sum	Mean	SD	Min	Max
Scientific productivity:					
No. publications 2008–11	1,317	1.68	4.58	0	50
No. citations 2008–11	5,228	6.67	27.08	0	428
No. of publications w./ incumbents 2008–11	464	0.59	3.61	0	81
Age:					
19–31	145	0.18	0.39	0	1
32–39	138	0.18	0.38	0	1
40–47	139	0.18	0.38	0	1
48–56	179	0.23	0.42	0	1
57–	183	0.23	0.42	0	1
Male	535	0.68	0.47	0	1
Faculty position at arrival:					
PhD-student	396	0.51	0.50	0	1
Postdoc	31	0.04	0.20	0	1
Tenured	24	0.03	0.17	0	1
Other	333	0.42	0.49	0	1
Scientific field:					
Engineering & Technology	325	0.41	0.49	0	1
Life sciences	186	0.24	0.43	0	1
Natural sciences & Mathematics	201	0.26	0.44	0	1
Social sciences & Humanities	72	0.09	0.29	0	1
Shanghai ranking, top-100 university	329	0.42	0.49	0	1
N	784				

Note: Table shows summary statistics for Asian migrants to Swedish universities in the year of arrival. The year of arrival is the first year a migrant appears in the data. Only includes migrants with work or study as their stated reason for settlement in this year.

In ancillary regressions, I confirm that the Asian migrants are more productive compared to the average Swedish-born researchers. They are about 5 percentage points more likely to publish and publish about 0.2 additional papers per year compared to other academics in Sweden, holding department, faculty position, age and gender fixed. However, Asian migrants are not more cited — although, due to the imprecision of the estimate, there is large heterogeneity in the citation rates. The regressions are available in Appendix A Table A1.

As mentioned, the available data from Statistics Sweden only distinguish between broader region-of-birth groups, so we cannot be certain exactly which countries are driving the Asian supply shock. According to a report from the Swedish National Agency for Higher Education from 2009, the most common source countries among first-year PhD students from Asia are China, Iran, India, and Pakistan in decreasing order of relevance (Swedish National Agency for Higher Education, 2009). Likewise, looking at data from the GlobSci project,[12] which surveyed 47,304 researchers across 16 countries in 2011, the most common Asian countries of birth for foreign-born scientists in Sweden are China and India (about 14 percent of surveyed foreign-born scientists in Sweden stated these as their country of origin). It is likely that migrants from the same countries are driving the increasing number of Asian-born researchers after 2008.

3.2 Defining treatment

The preceding section details how the number of Asian-born university researchers coming to Sweden increased after the policy reform in 2008. To be able to estimate the causal impact that this had has on Swedish academics, we must define appropriate treatment and control groups. In this paper, I follow a similar design as Borjas et al. (2018) and compare the changes in outcomes between incumbent researchers most affected by the reform to those least affected.

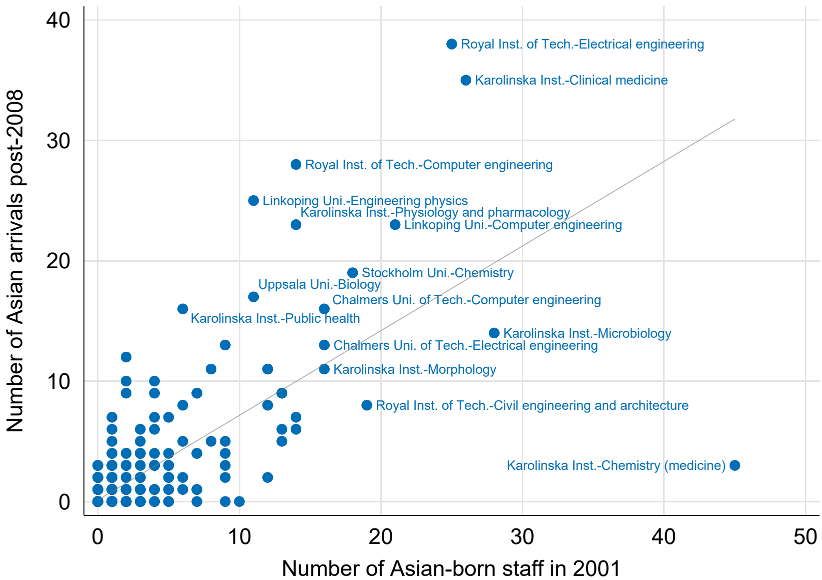
As a first step in partitioning incumbent researchers into groups affected by the supply shock (and not), I note that the Asian migrants gravitate in departments with an already large presence of Asian-born staff. Figure 4 shows the number of Asian arrivals from 2008 through 2011, plotted against the number of Asian faculty staff in 2001. It is

versity. For more information, see `http://www.shanghairanking.com/`. Between 2003 and 2011, the following five Swedish universities appeared on the top-100 of the list: Karolinska Institute, Lund University, Stockholm University, and Uppsala University.

[12]Described in Franzoni et al. (2012) and made available from `http://www.nber.org/globsci/`.

clear from the figure that departments with an already large presence of Asian-born researchers also received a relatively larger inflow after the 2008 reform. For instance, we see that the department that received the largest number of Asian immigrants after the 2008 reform, Electrical Engineering at the Royal Institute of Technology, which received almost 40 Asian immigrant researchers after the reform, was also the department with the fourth-highest number of Asian-born staff in 2001. Similarly, Clinical Medicine at Karolinska Institute, with the third-largest exposure in 2001, has the third-largest number of arrivals after 2008, with about 35 Asian immigrants.

Figure 4: Number of Asian arrivals to departments post 2008-reform, by number of Asian-born faculty in 2001



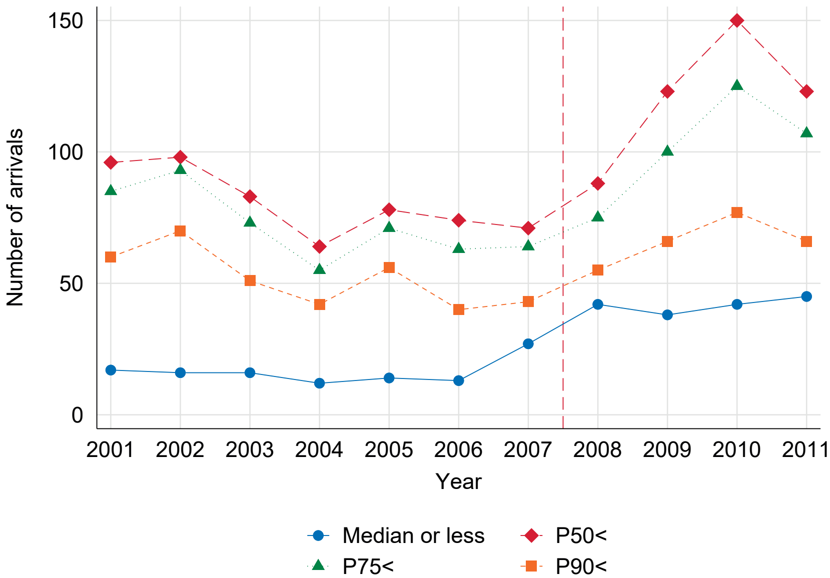
There are several possible reasons Asian migrants gravitate to departments with a large past presence of Asian staff. Earlier research has highlighted ethnic complementarities in knowledge production as a possible explanation (e.g., Borjas et al., 2018; Freeman & Huang, 2015b). This literature argues that common ethnicity, language, etc. makes collaboration between scientists more likely due to facilitating communication. Thus, an interpretation is that Asian migrants seek out departments where they will find other Asian-born researchers that share their language, nationality, and/or culture, explaining why certain universities and departments attracted disproportional numbers of Asian immigrants. Another explanation is that Asian researchers working in Sweden will keep some contact with the home country through social or work networks. Diaspora groups have been shown to be an important for foreign direct investment, technology transfer and international trade (Kerr, 2008; Leblang, 2010), as well as for transferring scientific knowledge back to the home country (Agrawal et al., 2011; Docquier & Rapoport, 2012). It is likely that information, such as job openings in Sweden or simply knowledge about the existence of a university and the possibility of working in Sweden, is also transmitted back to the home country. Some universities will also be better known in Asia, based on, for example, appearing on the Shanghai ranking lists, as discussed above.

Out of 817 departments in 2001, 238 had an Asian presence. Suppose we classify departments into exposed or non-exposed departments based on the number of Asian faculty staff and PhD students employed at the department in 2001. The discussion above suggests that the size of the supply shock might differ depending on the degree of past exposure. That is, the more Asian staff in 2001 the larger the probability that other Asian academics will know about the department and decide to migrate to Sweden. Following this, I consider departments with a presence over the median, over the 75th percentile and over the 90th percentile of the number of Asian research staff (faculty staff or PhD students) among the departments with an Asian presence in 2001 as exposed departments. This corresponds to above 2, 4, or 11 Asian staff present in 2001 for each of the percentiles of exposure, respectively. I then compared these to those departments with Asian-born staff corresponding to the median number or less (i.e. one or zero).

Figure 5 illustrates the supply shock by percentile of past exposure. From the figure we clearly see that the increase in the number of arrivals after 2008 is mainly driven by immigration to the most exposed departments (over the 75th and 90th percentile). The group of departments over the 75th percentile of exposure especially seems to drive much of the supply shock. By contrast, the group of departments with exposure at the median or less hardly seem affected at all.

Following Borjas et al. (2018), the division of departments into groups more or less exposed to the supply shock allows us to classify incumbent researchers into treatment and control groups based on their department affiliation, pre-shock. Table 2 shows descriptive statistics for researchers active in 2007, just before the reform, for all and by degree

Figure 5: Annual number of Asian migrants by departmental exposure, 2001–2012.



Note: Plots number of Asian migrant arrivals by the percentile of exposure in 2001 for all exposed departments. Departments are classified into degree of exposure based on: Median or less exposure; P50< is above median exposure; P75< is above the third quartile of exposure; and, P90< is above the ninth decile of exposure. The year of arrival is the first year a migrant appears in the data. Only includes migrants with work or study as their stated reason for settlement in this year.

of departmental exposure (median and less, above the 50th, the 75th, or the 90th percentile). The table also show the differences in means comparing characteristics of incumbents in departments with high past exposure and incumbents in low-exposure departments, as well as the result of t-test of these differences.

Table 2: Descriptive statistics for incumbent researchers in 2007 (t-test of differences in means in brackets), by degree of department exposure

		Department exposure in 2001			
	All	≤P50	P50< [Diff. (2) - (3)]	P75< [Diff. (2) - (4)]	P90< [Diff. (2) - (5)]
	(1)	(2)	(3)	(4)	(5)
Scientific output:					
No. publications 2001–07	4.01	2.22	5.87 [-3.653***]	6.11 [-3.895***]	5.92 [-3.700***]
No. citations 2001–07	24.20	9.80	39.11 [29.317***]	42.17 [-32.374***]	38.67 [-28.868***]
Age:					
19–31	0.26	0.19	0.34 [-0.147***]	0.35 [-0.165***]	0.35 [-0.166***]
32–39	0.20	0.19	0.21 [-0.015**]	0.21 [-0.015*]	0.22 [-0.021**]
40–47	0.20	0.22	0.18 [0.041***]	0.17 [0.046***]	0.18 [0.038***]
48–56	0.16	0.18	0.14 [0.040***]	0.14 [0.040***]	0.13 [0.046***]
57–	0.18	0.22	0.14 [0.081***]	0.13 [0.093***]	0.12 [0.102***]
Male	0.61	0.59	0.64 [-0.048***]	0.63 [-0.043***]	0.65 [-0.056***]
Faculty position:					
PhD-student	0.35	0.29	0.41 [-0.112***]	0.42 [-0.127***]	0.43 [-0.141***]
Postdoc	0.04	0.03	0.04 [-0.011***]	0.04 [-0.012***]	0.04 [-0.009**]
Tenured	0.42	0.53	0.30 [0.227***]	0.28 [0.252***]	0.25 [0.282***]
Other	0.19	0.15	0.25 [-0.104***]	0.26 [-0.113***]	0.28 [-0.133***]
Field of science:					
Engineering & Technology	0.23	0.13	0.33 [-0.209***]	0.35 [-0.228***]	0.43 [-0.300***]
Life sciences	0.21	0.08	0.35 [-0.263***]	0.38 [-0.299***]	0.37 [-0.287***]
Natural sciences & Mathematics	0.22	0.17	0.26 [-0.089***]	0.26 [-0.091***]	0.20 [-0.031***]
Social sciences & Humanities	0.34	0.62	0.06 [0.561***]	0.00 [0.618***]	0.00 [0.618***]

– *Table continued on next page* –

– Table continued from previous page –

Shanghai ranking, top-100 university	0.45	0.33	0.57 [-0.240***]	0.61 [-0.275***]	0.60 [-0.271***]
N	23,003	11,700	11,303	8,556	4,531

Note: Table shows mean values for all incumbents as well as by degree of department exposure. Incumbent researchers are classified into type of departments based on 2007 department affiliation. Exposure is based on the presence of Asian-born researchers in 2001 among departments with any exposure: ≤P50 is less or equal to the median exposure; P50< is above median of exposure; P75< is above the third quartile of exposure; and, P90< is above the ninth decile of exposure. Numbers in brackets shows the results of t-tests of equality of means between columns (2) and (3), (4) and (5) with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We can draw some conclusions from Table 2. As revealed by the t-tests, researchers at more exposed departments are more productive, both compared to the average researcher and to researchers working in less-exposed departments. On average, they produce more publications and higher-impact publications measured by citations rates. To some extent, this difference can likely be explained by the significantly larger proportion working in engineering, natural sciences, and medicine, fields that have a much higher publication pace compared to the social sciences or humanities. The significant larger share of male researchers likely reflects such field differences. In addition, researchers at highly exposed departments are also more likely to work in elite institutions, as indicated by the significant share working at universities on the Shanghai list. Taken together with the findings in Table 1 and A1, this indicates at least a positive selection of Asian migrants. Asian migrants are more productive and will seek out departments with more productive peers. The existence of positive selection will bias estimations of the causal effect of the supply shock. We will return to this issue in the next section, but for now we can conclude that researchers working at highly exposed departments are very different from researchers in less-exposed departments.

4 Empirical strategy

Having in the previous section classified incumbent researchers into groups more and less exposed to supply shock, we are now able to estimate the impact of the shock on the productivity of Swedish researchers. I will in the remaining analysis refer to incumbents affiliated with exposed departments (above 50th, 75th, or 90th percentile of exposure) as *treated* and incumbents working in low-exposure departments (less than or equal to 50th percentile of exposure) as *controls*.

4.1 Finding the appropriate control group

A natural starting point for identifying the effect of the supply shock on treated incumbents' productivity is to examine the changes in productivity compared to control incumbents in less exposed departments, before and after the policy reform in 2008. However, as is evident from Table 2 above, there are large differences between treated and control researchers across multiple dimensions. Since Asian migrants gravitate toward certain fields and universities, the treatment group naturally reflects this, meaning that the control group is very different compared to the treated group. An ideal experiment would assign treatment (i.e. inflows of Asian migrants) randomly across university departments. However, lacking such an experiment, I will instead employ a nonparametric matching method, coarsened exact matching (CEM; Blackwell et al., 2009; Iacus et al., 2012, 2011), which in a flexible way allows me to select appropriate control groups that are balanced on observable pre-treatment differences.

The basic approach of CEM is to choose a small set of matching covariates on which we would like to guarantee balance between treatment and control group. Subsequently, the procedure creates a number of unique strata to cover the joint distribution of the matching covariates. Next, each observation is assigned to a unique stratum and any observations in a stratum in which there is not a control observation for each treatment observation is dropped. Note that this allows matching not only on the mean of a variable but also the higher moments of the distribution, guaranteeing covariate balance and common support *ex ante*. This distinguishes the method from other matching methods based on, for example, propensity score matching. However, the method is susceptible to the *curse of dimensionality*, in the sense that the more fine-grained the partition of the support for the joint distribution (i.e., the higher the number of strata incorporated into the analysis), the larger the number of unmatched, treated observations. In general, the researcher must trade off the quality of the matches with external validity (see, e.g., discussion in Azoulay et al., 2013).

Using CEM, control researchers are selected so that they in 2007 share the following characteristics with treated researchers: i) work in the same broad scientific field (engineering, natural sciences and mathematics, medicine, or social science and humanities); ii) have the same type of position (tenured, postdoc, PhD student, or other research staff); and iii) work at the same university. Note that in this case the matching procedure is not coarse since I require a exact match on all matching covariates. In cases of ties, I randomly select one control researcher so that I end up with one control for each treated researcher that is balanced on the covariates selected to match on.[13]

Table 3 shows the pre-treatment characteristics of the matched sam-

[13] I have also included gender and age in the matching algorithm to capture gender-specific and life-cycle effects. This reduced the matching rate but did not change the overall results of my estimations. Since the estimated results did not differ to a large extent, I decided to not include these covariates in the final matching to increase the sample size. It is likely that gender and age effects are captured by the matching on position and field.

Table 3: Descriptive statistics for matched sample, above-median treatment

	Control (1)	Treated (2)	Difference (1) - (2) (3)
Scientific output:			
No. publications 2001–07	4.55	5.09	-0.533
No. citations 2001–07	24.90	30.85	-5.948*
Age:			
19–31	0.26	0.29	-0.029*
32–39	0.20	0.19	0.009
40–47	0.20	0.19	0.011
48–56	0.15	0.14	0.003
57–	0.20	0.19	0.005
Male	0.60	0.60	-0.009
Faculty position:			
PhD-sudent	0.37	0.37	0.000
Postdoc	0.05	0.05	0.000
Tenured	0.41	0.41	0.000
Other	0.17	0.17	0.000
Field of science:			
Engineering & Technology	0.25	0.25	0.000
Life sciences	0.17	0.17	0.000
Natural sciences & Mathematics	0.35	0.35	0.000
Social sciences & Humanities	0.23	0.23	0.000
Shanghai ranking, top-100 university	0.45	0.45	0.000
Observations	2,443	2,443	

Note: Table shows descriptive statistics for treated and control researchers. A treated researcher is an incumbent researcher at a high-exposure department and a control researcher is an incumbent at a low-exposure department. An exposed department is a university-field combination with above the 50th-percentile presence of Asian-born researchers in 2001. Incumbent researchers are classified into high- or low-exposure departments based on 2007 employment. Column (3) "Difference" shows the results of t-tests of equality of means between columns (1) and (2). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

ple using the above-median definition of treatment.[14] The matched sample consists of 4,886 matched incumbents (or 2,443 treated-control pairs). This implies a match rate of about 20 percent for treated incumbents. A low match rate is to be expected considering that about 60 percent of potential control matches are found within the social sciences or humanities, whereas only 6 percent of treated incumbents are found in these fields. Some of the covariates presented in Table 3 were used in the matching process while others were not. Looking at the table, the matched sample is well-balanced across covariates, also for those not used in matching, with only small significant differences for any variables across the groups. Importantly, we see that the pre-treatment differences in publication and citation rates are reduced compared to what was reported in Column (3) of Table 2. If there is selection based on these variables or if they are proxies for the existence of other omitted variables, e.g. research funding, such differences are reduced by the matching procedure. Considering that the matching process only takes into account differences across fields, universities and position, it appears to have worked quite well in creating a control group of researchers similar in most respects to treated researchers.

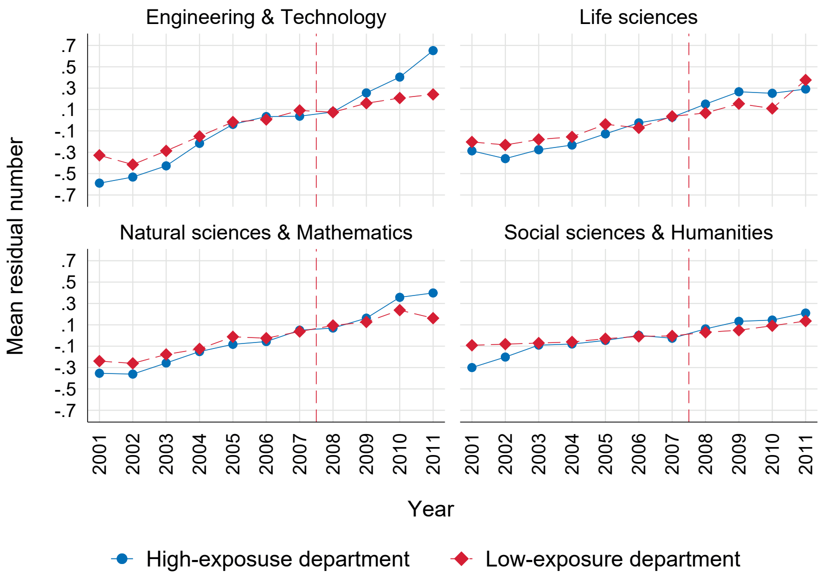
However, a worry might be that even if the matching procedure required the levels of publication to be the same across treatment and control group before the policy reform in 2008, the trends in productivity might be different between the two groups. To investigate differential publishing trajectories before the reform, I plot the mean number of annual publications for the matched above-median treatment and control groups in Figure 6.[15] Since the average publication pace is likely to differ across scientific fields, the figure plots publications separately by the field of incumbents. Furthermore, seeing as we are interested in the pre-treatment trend, I subtract the average number of publications before the 2008 reform for the respective group of departments, to more clearly see the changes in publication trends over time. The figure reveals that the pre-treatment trends in publications are very similar for high- and low-exposure departments. The figure also reveals a first — albeit a raw, but completely non-parametric — indication of the effect of the supply shock as the publication rate increases for researchers in engineering and natural sciences and mathematics at a faster pace after 2008 compared to the control group.

---

[14] For expositional purposes, I only show the matching results for the above median exposure group in the main text. For the results of the matching for the 75th and 90th percentile levels of exposure, see Tables C1 and C2 in Appendix C.

[15] For the corresponding figures for the 75th- and 90th-percentile treatment group, see Figures C1 and C2 in Appendix C.

Figure 6: Number of papers published annually by type of department and field, matched sample with above-median treatment



Note: Residual mean number of papers published annually by incumbents in high- and low-exposure departments, where the residual is defined as the difference between the actual number of papers published and the average number of papers published annually before 2008. A high-exposure department is a university-field combination with above the 50th-percentile presence of Asian-born researchers in 2001. Swedish researchers are classified into high- or low-exposure departments based on 2007 employment.

4.2 Econometric specification

The creation of a matched sample allows me to estimate the impact of the supply shock in a flexible difference-in-differences setting without having to make very strong assumptions on the functional form of the relationship (Imbens, 2004; Moffitt, 2004). Specifically, it allows me to compare the changes in productivity for incumbents by specifying the following OLS regression:

$$y_{i,t} = \phi_i + \phi_t + \beta Treated(P)_i \times Post_t + Age_{i,t} + \epsilon_{i,t} \tag{1}$$

where $y_{i,t}$ is our measures of productivity of researcher i; ϕ_i is a vector of individual-fixed effects; ϕ_t is a vector of time-fixed effects; and, $Age_{i,t}$

are dummies of incumbents' age included to capture life-cycle effects not captured by the matching procedure. The variable $\mathit{Treated(P)}_i$ is a dummy indicator equal to one if the Swedish researcher is employed at an exposed department in 2007, where P indicates the percentile of exposure: above the 50th, 75th, or 90th. The indicator is zero for the control group of Swedish researchers employed in departments with levels of past exposure at the 50th percentile or below. The indicator $Post_t$ equals one starting with the year of the policy reform (2008 or beyond). The dependent variable, $y_{i,t}$ is raw or citations-weighed publication counts, measuring productivity and the quality of output, or collaboration with Asian researchers, measured by the share of publications that are co-authored with at least one Asian-born researcher. I use robust standard errors and cluster on the department level.

The coefficient β, in Equation 1, gives the difference-in-differences estimate of the change in outcomes after the reform for incumbents in highly exposed departments relative to incumbents in low exposed departments within the same university, field and position. Also note that the regression in Equation 1 adjusts for individual-specific and time-invariant differences in productivity and collaboration, as well as calendar time fixed effects affecting trends common to all researchers.

A threat to identification of causal effects in this setting is the existence of time-variant omitted variables determining both outcome (i.e., productivity) and assignment to treatment (i.e., being exposed to Asian migration). Such variables are not captured by the fixed effects or by the matching procedure and will bias estimations. A possible omitted variable, excluded from Equation 1, could be changes over time to research funding. Changes to funding are likely to affect both productivity and recruitment, as well as vary over time so as not to be captured by fixed effects. However, seeing as research funding is correlated with both productivity and treatment (hiring of Asian researchers) it is not clear that including the variable as a control is the correct approach since this might also bias estimates.[16] Furthermore, recall that the matched sample removed differences in pre-treatment levels of publication productivity between treated and control incumbents. Since it is likely that research output is correlated to both past and future research funding (without funding there is usually less time to do research), publication levels may be considered as a proxy for the existence of funding. Thus, some differences in funding should have been accounted for in the matching process. However, in further robustness analysis, I will try to account for differences in R&D funding at the field level to see how that

[16] This problem is sometimes referred to as the 'bad control' problem, see for example Angrist & Pischke (2008).

affects the results.

5 The productivity of incumbent Swedish researchers

5.1 Main results

In Table 4, I estimate regressions corresponding with Equation 1 to analyze the effect of the Asian supply shock on the productivity of Swedish researchers. The table shows results across treatment percentile of past exposure. Specifically, Column (1) shows the results of incumbents in departments with past exposure above the median level; Column (2) uses the above 75th percentile definition of treatment; and, Column (3) shows results using the above 90th percentile definition of treatment. Note that in all specifications the control group is the same — namely, incumbents in departments with median or below levels of exposure to Asian migration in 2001.

Panel A of Table 4 show the results using annual publication counts as dependent variable. Looking at Columns (1) and (2), the estimates are positive but insignificant, indicating that, on average for these percentiles of exposure, incumbent researchers did not increase publishing productivity compared to control researchers. However, turning to Column (3) and the above 90th percentile of exposure, the estimate instead suggests productivity increased for incumbents in the most-exposed departments. The point estimates imply about 0.4 additional publications per year and person after 2008 for this group. This effect is significant at the 5% level.

To get an idea about the magnitudes involved we can compare the estimates to the pre-reform publication rate for the most exposed treatment group (see Table C2 in Appendix C). In this case, the size of the estimate corresponds to an increase of about 45 percent for the average yearly output of treated incumbents after the shock.[17] In total, the estimates translate into, on average, 1.6 additional publications produced by each treated incumbent from 2008 through 2011. Thus, in aggregate ($1.6 \times 675 =$) 1,080 additional publications in this period can be attributed to the reform and the subsequent inflow of Asian migrants.

[17] Looking at Table C2, the treated P90< group produced, on average, 6.33 papers over the 2001–2007 period, which over a seven year period corresponds to, on average, about ($6.33/7 \approx$) 0.9 papers per year before the reform.

Table 4: The 2008-reform and the productivity of incumbent researchers, by treatment intensity

	Treatment: Percentile of past exposure		
	P50<	P75<	P90<
	(1)	(2)	(3)
	A. Dependent variable: Publications per year		
Treated × Post	0.0888	0.0189	0.410**
	(0.0750)	(0.0893)	(0.193)
No. of scientists	4,886	3,246	1,350
Observations	36,567	23,827	9,830
	B. Dependent variable: Citations per year		
Treated × Post	1.686	0.0417	0.970
	(1.523)	(0.882)	(1.152)
No. of scientists	4,886	3,246	1,350
Observations	36,567	23,827	9,830
	C. Dependent variable: Collaboration per year		
Treated × Post	0.0130*	0.0170**	0.0126
	(0.00669)	(0.00759)	(0.0123)
No. of scientists	2,476	1,793	821
Observations	12,150	9,329	3,971

Notes: Table shows OLS estimates of β in Equation 1. Treatment is based on exposure to Asian migration in 2001 among departments with any exposure: P50< is above the median level of exposure; P75< is over third quartile of exposure; and, P90< is above the ninth decile of exposure. Control group is incumbents in departments with median or less level of exposure in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

It is possible that knowledge spillovers generated from the Asian migrants also affected the quality of incumbents' publication output or that migrants were able to connect incumbents to networks of citing peers in their home country (cf., Breschi et al., 2017). In these cases, the number of citations to treated incumbents' publications after 2008 should increase compared to controls. The direction that migrant researchers affect the social impact of incumbents' publication output is, however, unclear since the supply shock might also imply increased competition for incumbents. Thus, an additional response might be to increase publication pace but reduce quality of output.

To investigate, Panel B of Table 4 shows the effect of the supply shock on the quality of publication output, measured as citations-weighed annual publication output. The estimates imply that the increasing supply of Asian migrants did not affect the average quality of publications at any percentile of exposure, although the standard errors are large and cannot rule out negative effects for some treated. However, for both the above-50th and -90th percentile of exposure the confidence intervals are more clearly on the positive side. Thus, while publications increased for the most exposed incumbents, this increase did not translate into a clear decrease (or increase) in the quality of papers published compared to the control group. Thus, researchers did not publish lower (or higher) quality publications in response to increased competition.

Finally, let us turn to Panel C of Table 4, which uses the proportion of papers co-authored with Asian migrants to all publications each year as dependent variable. From Columns (1) and (2), we see that the collaboration patterns of Swedish researchers changed after 2008. For the above-median treatment group, the estimate implies that, on average, the proportion of papers co-authored with Asian migrants to all publications each year increased by about 1.3 percent after 2008 compared to the control group. The effect is significant at the 10% level. Likewise, the corresponding estimate for the above 75th-percentile treatment group implies an increase of 1.7 percent of publications co-authored with Asian migrants after 2008, on average. The effect is significant at the 5% level. However, for the treatment group in departments with above 90th percentile of past exposure, there is no significant increase in the proportion of papers co-authored with Asians.

We can draw some conclusions from the results presented above. First, only the most-exposed group of incumbents increased the number of papers published after the 2008 reform. However, they did this without co-authoring more papers with Asian migrants compared to before. Second, less-exposed treatment groups increased their collaboration with Asian immigrants, although this did not translate into an increase in productivity. The supply shock did not affect the average quality of publications for any group.

5.2 Robustness of main results

The results are robust to a broad range of alternative specifications, including alternative definitions of the treatment period, count data models, field-level R&D expenditure, the inclusion of university-specific linear time trends, and limiting the sample to a balanced panel.

Placebo treatment period.— To make sure that it is in fact the 2008-reform that drives the results, I confirm that the main results are robust by comparing them to an alternative post-treatment period. In Appendix B Table B1, I replicate the main results from Table 4, using instead the year 2004 onwards as a placebo treatment period, excluding the period after 2007. The results are clear. Moving the post-treatment period removes any significant treatment effect.

Adding field-level R&D-expenditure per faculty.— One concern might be that omitted variables, affecting both productivity and the inflow of Asian immigrants to a department, are biasing the estimates. A possible such variable discussed above is the funding environment of the incumbent researchers. More funding will mean more research time and possibly more publications, but also more money for recruitment (including of Asian researchers). To check if increased funding levels is driving the results, I collect data on total R&D expenditures (including both government base funding and external funding) at the university-field level. These data capture the general funding environment at Swedish universities at broader scientific fields (2-digit level) than used when constructing the department indicators (3-digit level). I re-estimate the main results from Table 4 including the R&D-expenditure per faculty member at the broader scientific field levels. The results are available in Appendix B, Table B2. The inclusion of R&D expenditures does not change the estimated coefficients in any qualitative way.

Including university-specific time trends.— In addition to the similar pre-trends reported in the figures above, I explicitly control for differential trends across treated and control institutions in my specifications. In particular, I re-estimate the regressions corresponding to Table 4 by including a set of university-specific linear trends. The results are reported in Appendix B, Table B3. The estimates are similar in both magnitude and precision.

Balanced panel.— To alleviate concerns about endogenous sample attrition, I confirm that the results are robust by restricting attention to a balanced panel, focusing on incumbent researchers who remain in the sample from 2001 until 2011. The results are presented in Appendix B, Table B4, and are similar to the results using the unbalanced panel.

Count data model.— Finally, to address the count data nature of publications and citations, I re-estimate the main results from Table 4 Panels A and B using a conditional fixed-effect Poisson estimator with robust standard errors. This yields comparable estimates to the OLS estimates in Table 4. I present the results using a Poisson estimator in Appendix B Table B5.

5.3 Heterogeneity

The results presented in Table 4 are the average effects of the supply shock for all treated incumbents across treatment groups. However, the prior discussion highlighted how the effect of the supply shock may differs across characteristics of incumbents. For instance, it is likely that the importance of peers and teams differ across different fields (Wuchty et al., 2007), or is based on the position and ethnicity of incumbents and newcomer (Borjas et al., 2018). Thus, it is likely that the overall effect reported in Table 4 hides much heterogeneity across the pre-treatment characteristics of incumbents.

To investigate further, Table 5 uses the same matched sample as above but introduces interaction variables between pre-treatment characteristics of incumbents and the treatment indicator. In the main text, I focus on the results using publications per year as the dependent variable and on the effect for the above-median treatment group. For results using the other dependent variables, citations per year and collaboration, see Tables A2 and A3 in Appendix A.[18] As a baseline, Column (1) of Table 5 shows an identical regression to Column (1) in Table 4.

Differences by scientific field.— Most of the early literature on productivity spillover in academia usually focuses on one scientific field. An exception is Ganguli (2015), who studies the citations of American scientists before and after the inflow of Russian scientists to the US following the fall of the USSR. She reports evidence supporting that field-specific

[18]For the results using the other treatment groups (above 75th and 90th percentiles), these results are very similar to the ones reported in the main text and are available from the author upon request.

effects mediate knowledge spillovers. Moreover, studies on the importance of teams in science show that the trends in average team sizes differ markedly across fields (Jones, 2009; Wuchty et al., 2007). This further suggests that peers will have different importance for personal productivity, depending on field.

To investigate this matter more closely, Table 5 Column (2) introduces dummies for the field of incumbents in 2007 interacted with the treatment indicator. The omitted category is researchers in engineering and technology. The results in Column (2) are markedly different compared to the baseline in Column (1). The estimate of $Treated \times Post$ is now sizeable positive and significant at the 1% level. Netting out field-specific responses, the estimate indicates that treated researchers published an additional 0.365 publications each year after the reform. We can further note that the estimate for the interaction effect of Natural Sciences and Mathematics is significantly negative. The same goes for the interaction effect for Social Science and Humanities. Taken together with the overall effect, the net effect for these fields is close to zero. Thus, the previous reported zero effect hides heterogeneous responses across scientific fields. Moreover, the responses point in different directions suggesting that some fields were better equipped for handling increased growth after the supply shock. Specifically, the results suggest that the only field that seems to have benefited was engineering, which was also the field that saw the largest inflow of Asian migrants after the reform in 2008.

Differences across faculty position.— Next, Column (3) adds interactions between the treatment indicator and dummies indicating the faculty position of incumbents in 2007. Similar to when taking field-specific effects into account, the results show how the overall effect hides much heterogeneity also across the faculty position of incumbents. The overall estimate when accounting for position is now positive and significant at the 10%-level. Moreover, the interactions show that Postdocs were able to leverage the supply shock to increase productivity the most compared to other positions, publishing an additional 0.7 papers per year. The estimate is significant at the 10%-level. Table A3 in Appendix A reveals a possible mechanism explaining this. Namely, that treated Postdocs increased the share of papers co-authored with Asian migrants from 2008. In fact, the positive effect on collaboration reported in Table 4 above is explained entirely by Posdocs publishing a larger proportion of papers together with Asian researchers. For tenured and other researchers, the point estimates are negative, but indistinguishable from zero.

Table 5: Heterogeneous impact of the 2008-supply shock on number of publications per year of incumbents, above-median treatment

	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	0.0888	0.365***	0.0634	0.0661	-0.185***	0.101
	(0.0750)	(0.105)	(0.0717)	(0.0754)	(0.0448)	(0.0994)
Scientific field (engineering and technology omitted):						
Treated × Post × Life sciences		-0.231				-0.263
		(0.221)				(0.221)
Treated × Post × Natural sciences and Mathematics		-0.389***				-0.365***
		(0.146)				(0.136)
Treated × Post × Social sciences and Humanities		-0.428***				-0.293***
		(0.103)				(0.0856)
Faculty position (PhD-students omitted):						
Treated × Post × Postdoc			0.750*			0.593
			(0.405)			(0.380)
Treated × Post × Tenured			-0.0134			-0.110
			(0.0835)			(0.0740)
Treated × Post × Other			-0.0542			-0.107
			(0.0948)			(0.0932)
Ethnicity:						
Treated × Post × Asian				0.559**		0.471*
				(0.268)		(0.245)
Productivity:						
Treated × Post × Above-median productivity					0.563***	0.527***
					(0.112)	(0.0944)
Constant	0.0103	0.0369	0.0455	0.0129	-0.0109	0.0717
	(0.136)	(0.134)	(0.128)	(0.135)	(0.133)	(0.124)

– Table continued on next page –

– Table continued from previous page –

No. of scientists	4,886	4,886	4,886	4,886	4,886	4,886
Observations	36,567	36,567	36,567	36,567	36,567	36,567

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating field, faculty position, Asian region-of-birth, and pre-2008 level productivity of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Differences across ethnicity.— Following Borjas et al. (2018), Column (4) investigates the existence of co-ethnic complementarities in knowledge production by introducing an interaction between the treatment indicator and an indicator equal to one if the incumbent was born in Asia. Although the data on region of birth are noisy, the interaction is positive and significant at the 5% level. Thus, incumbents with an Asian background benefit more from the supply shock. Moreover, Table A3 in Appendix A shows that treated incumbents of an Asian background are also more likely to co-author papers with the Asian newcomers after the reform. The findings are in line with those of Borjas et al. (2018) in the US and suggest co-ethnic complementarities as a mechanism facilitating spillovers.

Differences across the productivity distribution.— In the prior literature, positive peer effects and spillovers are frequently found when the sample consists of superstar scientists (Agrawal et al., 2017; Azoulay et al., 2010; Moser et al., 2014), whereas studies finding no or negative effects sample more representative scientists (Borjas & Doran, 2012; Waldinger, 2012). The focus on superstars is usually motivated by the disproportional importance of this group for overall knowledge production. These differing results suggest that responses to supply shocks may differ across the productivity distribution. Moreover, as discussed earlier, researcher productivity is closely linked to resources (e.g., Ejermo & Källström, 2016), making earlier productivity a proxy for the existence of funding at the individual level.

Column (5) of Table 5 introduces, as an additional dimension of heterogeneity, the pre-treatment productivity of incumbents. This is operationalized by counting the cumulative number of publications from 2001 to 2007 for each incumbent and then introducing an indicator equal to one for incumbents with above median number of publications in their respective field. This indicator is interacted with the treatment indicator to allow the response to the supply shock differ to the above-median productive group.

The estimates in Column (5) reveal that responses to the Asian supply shock differ for the low and high productivity groups of incumbents. When allowing the effect to differ across the productivity distribution, the estimates of $Treated \times Post$ become negative and significant at the 1% level, indicating that the overall effect is negative when netting out the response of high-productivity individuals. Looking at the estimate for the above-median productivity group, the estimate is sizeable and positive. The point estimate suggests an increase corresponding to about 0.55 additional papers published each year and is significant at the 1% level. Thus, the supply shock further increased the differential in productivity between high- and low-productivity researchers. Table A2 in Appendix A shows a similar pattern for citations, even though the effect is less precisely measured and only significant at the 10% level.

Finally, in Column (6) of Table 5 all interactions are included. We now see how the overall effects reported above are driven by disparate responses across fields, faculty position, ethnicity and productivity of incumbents. The overall effect remains insignificant, but there is much heterogeneity. First, the interactions remain significant and negative for both the Natural Sciences and Mathematics and Social Sciences and Humanities, indicating that exposed researchers in these fields had reduced productivity relative to incumbents in Engineering and Technology after the reform. In addition, the interaction effect with Asian ethnic background remain positive and significant, although only at the 10% level. However, the positive interaction of Above-median productivity reveals that already prolific academics, with a large past publication stock, are able to reap the most benefits from the supply shock by increasing their productivity.

To investigate further, I also estimate Table 5 separately by field and faculty position including the interactions with position, field, ethnicity, and productivity. These estimations are available in Appendix A, Tables A4–A10. The results are clear: high-productivity individuals who further increase their productivity explain the positive effects within fields and positions. In Engineering and Technology, as well as in the Life Sciences, there is some evidence of crowding-out since the overall effect turns negative when accounting for dispersed responses across faculty position, ethnicity, and productivity. Also, in Natural Sciences and Mathematics there is some evidence of crowding-out effects, but here it is driven by tenured and other staff publishing less after the reform. Moreover, the positive effect of Asian-born incumbents is explained by the response of Asian researchers within Engineering and Technology and by incumbents who were PhD students in 2007.

To sum up, productivity spillovers were generated by a supply shock of immigrant researchers to Sweden after a deregulation of work and study migration. However, responses differ across the characteristics of the incumbents. It is mainly researchers in engineering, and especially already prominent researchers in terms of past publication stock, who benefit in terms of further increased productivity. For less prominent researchers, I instead find some instances where crowding-out effects dominate, and publication productivity declined after the reform. Thus, the supply shock further increased the divide between already highly productive and less-productive incumbents.

6 Conclusion

This paper investigates the role of spillovers in knowledge production. The analysis exploits a supply shock to Swedish academia — a major deregulation of Swedish migration policy in 2008 — to tease out the causal impact of immigration on the productivity of incumbents. I show a sharp and sudden increase in the number of Asian researchers coming to Swedish universities after the reform.

To estimate the impact of this supply shock, I use novel comprehensive data on the publication output of Swedish academic researchers over the years 2001-–2011 in a difference-in-differences set-up. Asian migrants tended to gravitate towards departments with past dependence on Asian immigration. I use this fact to partition the sample into treatment and control groups based on Swedish researchers' affiliation to departments with a history of Asian faculty. To alleviate concerns about pre-treatment differences, I employ a non-parametric matching technique to find a control group of researchers less exposed to the supply shock but sharing field, type of position, and university with treated researchers.

In my main analysis, I find that only Swedish researchers at the most-exposed departments improved their publication productivity, measured by publication count, after 2008. They did this without authoring a larger share of papers jointly with Asian migrants. Less-exposed treatment groups, however, increased their collaboration with Asian immigrants, although this did not translate into an increase in productivity. The supply shock did not affect the average quality of citation-weighed publications. These results are robust to a wide range of alternative specifications, including alternative definitions of the treatment period, count data models, field-level R&D expenditure, the inclusion of university-specific linear time trends, and limiting the sample to a balanced panel.

To investigate further, I decomposed the effect of the supply shock across characteristics of the incumbent researchers. First, I note that the effect varies across scientific fields and those researchers in engineering drive the positive effects, a field where collaboration and teamwork, as well as complementarity between human and physical capital, is essential for your own productivity. Next, I investigated whether spillovers were mediated by the faculty position of incumbents. The results reveal that Postdocs gained the most in terms of improved publication productivity. Moreover, I find that the inflow of Asian migrants disproportionally benefited the productivity of Asian-born incumbents. Finally, I broke down the effect across the productivity distribution. I find that the positive effects of the supply shock are solely driven by the response of researchers in the upper distribution of the pre-treatment productivity distribution. This is also the case within fields and positions. At the same time, I find that low-productivity researchers decreased their productivity in some instances. As a result, the increased supply of Asian immigrants widened the gap between the already highly productive and less-productive researchers. This suggests an additional channel through which immigration may affect knowledge production to what has been discussed in the literature.

It is possible that other factors in the home country might also influence the migration decision of scientists and students to Sweden after 2008. For example, in 2007 the Chinese government set aside funds to subsidize Chinese academics to go abroad to work or study (Freeman & Huang, 2015a). Thus, it is possible that the supply shock is at least partially influenced by this home country reform acting as a *push factor* for Chinese researchers. However, without access to more detailed data on the country of origin this question cannot be addressed in the current paper. Moreover, this does not impact the identification strategy since the focus is on incumbent researchers.

I close with some concluding remarks on the implications of these findings. First, the main contribution of this paper rests on the detailed and comprehensive data, which allowed me to reconcile some of the mixed findings in the prior literature on high-skilled migration and knowledge production. Due to data limitations, earlier papers usually focused on one field, or specific subsets of scientists. Although these studies can be very in-depth, increasing our understanding of the specific mechanism at play, they cannot inform us about the overall impact of supply shocks on domestic knowledge production. In this paper, using data that cover a large portion of a national academic system, I show how responses to a supply shock differed across fields, positions, ethnicity, and productivity of incumbents. Thus, when evaluating the overall impact of supply shocks to science it is important to keep these heterogeneous effects in mind.

Second, the potential of migration to influence the rate of knowledge production depends crucially on how the newcomers interact with the pre-existing workforce. The evidence presented here suggests that the complementarity between migrants and incumbents will vary across characteristics of migrants and incumbents. Findings are consistent with that the reform increased the size of the academic labor market, which allowed recruits' skills to better match the skills of some incumbents. However, further studies should be devoted to investigating the exact mechanisms and conditions for when and how immigrant inflows will enhance incumbents' productivity.

Finally, it is clear that immigration policy can be a very powerful tool influencing domestic knowledge production. However, as with supply shocks in general, some incumbents will gain, some will lose, and it is difficult to foresee how gains will distribute beforehand. This modifies the extent of the effectiveness of selective supply-driven immigration policies as a way of improving scientific productivity. For example, targeting only star scientists can have a detrimental impact on incumbents' productivity if there is a lack of complementarity between them and the incumbents.

References

Agrawal, A., Kapur, D., McHale, J., & Oettl, A. (2011). Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. *Journal of Urban Economics*, *69*(1), 43–55.

Agrawal, A., McHale, J., & Oettl, A. (2014). Collaboration, stars, and the changing organization of science: Evidence from evolutionary biology. In A. B. Jaffe & B. F. Jones (Eds.), *The changing frontier: Rethinking science and innovation policy* (pp. 75–102). University of Chicago Press.

Agrawal, A., McHale, J., & Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*, *46*(4), 853–867.

Angrist, J. D. & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal Of Economics*, *125*(2), 549–589.

Azoulay, P., Stuart, T., & Wang, Y. (2013). Matthew: Effect or fable? *Management Science*, *60*(1), 92–109.

Azoulay, P., Zivin, J. S. G., & Sampat, B. N. (2011). The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. In J. Lerner & S. Stern (Eds.), *The Rate and Direction of Inventive Activity Revisited* (pp. 107–155). University of Chicago Press.

Bellini, S. (2016). EU Blue Card: a promising tool among labour migration policies? A comparative analysis of selected countries. IPE Working Papers Series 76/2016, Institute for International Political Economy Berlin.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). cem: Coarsened Exact Matching in Stata. *Stata Journal*, *9*(4), 524.

Borjas, G. J. & Doran, K. B. (2012). The collapse of the Soviet Union and the productivity of American mathematicians. *The Quarterly Journal Of Economics*, *127*(3), 1143–1203.

Borjas, G. J., Doran, K. B., & Shen, Y. (2018). Ethnic complementarities after the opening of China: How Chinese graduate students affected the productivity of their advisors. *Journal of Human Resources*, *53*(1), 1–31.

Breschi, S., Lissoni, F., & Miguelez, E. (2017). Foreign-origin inventors in the USA: Testing for diaspora and brain gain effects. *Journal of Economic Geography*, *17*(5), 1009–1038.

Cerna, L. (2016). The crisis as an opportunity for change? High-skilled immigration policies across Europe. *Journal of Ethnic and Migration Studies*, *42*(10), 1610–1630.

Docquier, F. & Rapoport, H. (2012). Globalization, brain drain, and development. *Journal of Economic Literature*, *50*(3), 681–730.

Ejermo, O., Alder, C., Fassio, C., & Källström, J. (2016). Publications of Academic Researchers In Sweden (PARIS). Technical report, CIRCLE, Lund University. `http://paris.circle.lu.se`.

Ejermo, O. & Källström, J. (2016). What is the causal effect of R&D on patenting activity in a "professor's privilege" country? Evidence from Sweden. *Small Business Economics*, 1–18.

Ejermo, O. & Zheng, Y. (2018). Liberalization of European migration and the immigration of skilled people to Sweden. *IZA Journal of Development and Migration*, *8*(1), 5.

Emilsson, H. & Magnusson, K. (2015). Högkvalificerad arbetskraftsinvandring till Sverige ["Highly qualified labor immigration to Sweden"]. In C. Calleman & P. Herzfeld Olsson (Eds.), *Arbetskraft från hela världen: Hur blev det med 2008 års reform?* chapter 3, (pp. 72–113). Delmo report.

Franzoni, C., Scellato, G., & Stephan, P. (2012). Foreign-born scientists: Mobility patterns for 16 countries. *Nat Biotech*, *30*(12), 1250–1253.

Franzoni, C., Scellato, G., & Stephan, P. (2014). The mover's advantage: The superior performance of migrant scientists. *Economics Letters*, *122*(1), 89–93.

Freeman, R. B. & Huang, W. (2015a). China's "great leap forward" in Science and Engineering. In A. Geuna (Ed.), *Global Mobility of Research Scientists* (pp. 155–175). Elsevier.

Freeman, R. B. & Huang, W. (2015b). Collaborating with people like me: Ethnic coauthorship within the United States. *Journal of Labor Economics*, *33*(S1), S289–S318.

Ganguli, I. (2015). Immigration and ideas: What did Russian scientists "bring" to the United States? *Journal Of Labor Economics*, *33*(S1), S257–S288.

Gaulé, P. & Piacentini, M. (2013). Chinese graduate students and US scientific productivity. *Review Of Economics And Statistics*, *95*(2), 698–701.

Hunter, R. S., Oswald, A. J., & Charlton, B. G. (2009). The elite brain drain. *The Economic Journal*, *119*(538), F231–F251.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, *106*(493), 345–361.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political analysis*, *20*(1), 1–24.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review Of Economics And Statistics*, *86*(1), 4–29.

Jaravel, X., Petkova, N., & Bell, A. (2018). Team-specific capital and innovation. *American Economic Review*, *108*(4-5), 1034–73.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review Of Economic Studies*, *76*(1), 283–317.

Kerr, W. R. (2008). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics*, *90*(3), 518–537.

Kerr, W. R. & Lincoln, W. F. (2010). The supply side of innovation: H-1B visa reforms and U.S. ethnic invention. *Journal Of Labor Economics*, *28*(3), 473–508.

Leblang, D. (2010). Familiarity breeds investment: Diaspora networks and international investment. *American Political Science Review*, *104*(3), 584–600.

Levin, S. G. & Stephan, P. E. (1999). Are the foreign born a source of strength for US science? *Science*, *285*(5431), 1213–1214.

Moffitt, R. A. (2004). Introduction to the symposium on the econometrics of matching. *Review Of Economics And Statistics*, *86*(1), 1–3.

Moser, P., Voena, A., & Waldinger, F. (2014). German Jewish Émigrés and US invention. *American Economic Review*, *104*(10), 3222–3255.

OECD (2008). *The Global Competition For Talent: Mobility Of The Highly Skilled*. Paris: Organization for Economic Cooperation and Development.

OECD (2011). *Recruiting Immigrant Workers: Sweden 2011*.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science*, *58*(6), 1122–1140.

Swedish Higher Education Authority (2017). *Higher Education in Sweden – 2017 status report*. Universitetskanslersämbetet.

Swedish National Agency for Higher Education (2009). *Utländska doktorander i svensk forskarutbildning ["Foreign doctoral students in Swedish doctoral education"]*. Högskoleverket.

Swedish Public Employment Service (2008). Var finns jobben år 2008? Samt bedömning på 5–10 års sikt ["Where are there jobs in 2008? And assessment in 5–10 years time"]. Technical report, Arbetsförmedlingen.

Waldinger, F. (2010). Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. *Journal Of Political Economy*, *118*(4), 787–831.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *The Review Of Economic Studies*, *79*(2), 838–861.

Weinberg, B. A. (2011). Developing science: Scientific performance and brain drains in the developing world. *Journal of Development Economics*, *95*(1), 95–104.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

A Supplementary tables and figures

Table A1: Productivity of Asian migrant researchers, 2007–2011

	(1) Ever published (0/1)	(2) Publications per year	(3) Citations per year
Asian	0.0529***	0.190**	-0.325
	(0.00987)	(0.0807)	(0.720)
R^2	0.17	0.16	0.05
Observations	102,931	102,931	102,931

Notes: Robust standard errors clustered at the department-level in parentheses. All regressions are OLS and include department- and year-fixed effects, as well as age, gender, and position dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Heterogeneous impact of the 2008-supply shock on citations per year of incumbents, above-median treatment

	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	1.686	0.470	0.111	1.716	-0.856**	-2.291
	(1.523)	(0.759)	(0.597)	(1.585)	(0.427)	(1.910)
Scientific field (engineering and technology omitted):						
Treated × Post × Life sciences		6.239				5.948
		(6.156)				(6.221)
Treated × Post × Natural sciences and Mathematics		0.936				0.955
		(2.652)				(2.563)
Treated × Post × Social sciences and Humanities		-0.647				0.510
		(0.755)				(1.000)
Faculty position (PhD-students omitted):						
Treated × Post × Postdoc			14.19			12.02
			(11.26)			(10.74)
Treated × Post × Tenured			1.598			0.229
			(1.842)			(0.983)
Treated × Post × Other			-0.761			-2.127
			(0.989)			(1.472)
Ethnicity:						
Treated × Post × Asian				-0.728		-1.194
				(1.782)		(2.004)
Productivity:						
Treated × Post × Above-median productivity					5.231*	4.319*
					(2.925)	(2.344)
Constant	1.357	1.277	1.381	1.354	1.160	1.392
	(1.468)	(1.476)	(1.316)	(1.471)	(1.509)	(1.315)

– Table continued on next page –

– Table continued from previous page –

No. of scientists	4,886	4,886	4,886	4,886	4,886	4,886
Observations	36,567	36,567	36,567	36,567	36,567	36,567

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating field, faculty position, Asian region-of-birth, and pre-2008 productivity level of incumbents. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Heterogeneous impact of the 2008-supply shock on collaboration, above-median treatment

	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	0.0130*	0.0150	0.0139	0.00947	0.0130*	0.00630
	(0.00669)	(0.0101)	(0.0100)	(0.00663)	(0.00669)	(0.0132)
Scientific field (engineering and technology omitted):						
Treated × Post × Life sciences		0.00275				0.00579
		(0.0135)				(0.0143)
Treated × Post × Natural sciences and Mathematics		0.00244				0.00429
		(0.0126)				(0.0125)
Treated × Post × Social sciences and Humanities		-0.0265**				-0.0223*
		(0.0110)				(0.0120)
Faculty position (PhD-students omitted):						
Treated × Post × Postdoc			0.0512**			0.0564**
			(0.0242)			(0.0250)
Treated × Post × Tenured			-0.00918			-0.00447
			(0.0130)			(0.0144)
Treated × Post × Other			0.00588			0.00867
			(0.0184)			(0.0199)
Ethnicity:						
Treated × Post × Asian				0.0664***		0.0669***
				(0.0245)		(0.0243)
Productivity:						
Treated × Post × Above-median productivity					.	.
					(.)	(.)
Constant	-0.00393	-0.00188	0.00660	-0.00390	-0.00393	0.00682
	(0.0128)	(0.0130)	(0.0135)	(0.0127)	(0.0128)	(0.0137)
No. of scientists	2,476	2,476	2,476	2,476	2,476	2,476

– Table continued on next page –

– *Table continued from previous page* –

Observations	12,150	12,150	12,150	12,150	12,150	12,150

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating field, faculty position, Asian region-of-birth, and pre-2008 productivity level of incumbents. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Heterogeneous effects within Engineering and Technology

	(1)	(2)	(3)	(4)	(5)
Treated × Post	0.359***	0.166	0.282**	-0.154*	-0.231*
	(0.132)	(0.127)	(0.131)	(0.0896)	(0.127)
Faculty position (PhD-students omitted):					
Treated × Post × Postdoc		-0.00669			-0.168
		(0.644)			(0.604)
Treated × Post × Tenured		0.303*			0.121
		(0.170)			(0.172)
Treated × Post × Other		0.285			0.0992
		(0.235)			(0.215)
Ethnicity:					
Treated × Post × Asian			1.510***		1.213**
			(0.565)		(0.574)
Productivity:					
Treated × Post × Above-median productivity				0.968***	0.873***
				(0.156)	(0.148)
Constant	0.123	0.0539	0.105	0.114	0.0688
	(0.185)	(0.194)	(0.185)	(0.181)	(0.197)
No. of scientists	1,214	1,214	1,214	1,214	1,214
Observations	9,065	9,065	9,065	9,065	9,065

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating faculty position, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Heterogeneous effects within Life Sciences

	(1)	(2)	(3)	(4)	(5)
Treated × Post	-0.125	-0.132	-0.112	-0.447***	-0.294**
	(0.237)	(0.161)	(0.248)	(0.107)	(0.113)
Faculty position (PhD-students omitted):					
Treated × Post × Postdoc		0.664**			0.472
		(0.324)			(0.307)
Treated × Post × Tenured		-0.0133			-0.224
		(0.210)			(0.150)
Treated × Post × Other		-0.143			-0.296
		(0.132)			(0.176)
Ethnicity:					
Treated × Post × Asian			-0.281		-0.303
			(0.279)		(0.257)
Productivity:					
Treated × Post × Above-median productivity				0.538*	0.561*
				(0.299)	(0.295)
Constant	0.831**	0.853***	0.829**	0.769**	0.848***
	(0.346)	(0.311)	(0.346)	(0.330)	(0.311)
No. of scientists	826	826	826	826	826
Observations	6,056	6,056	6,056	6,056	6,056

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating faculty position, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Heterogeneous effects within Natural sciences and Mathematics

	(1)	(2)	(3)	(4)	(5)
Treated × Post	-0.00174	0.0929	-0.0135	-0.171**	-0.0132
	(0.128)	(0.148)	(0.131)	(0.0823)	(0.131)
Faculty position (PhD-students omitted):					
Treated × Post × Postdoc		1.221			1.108
		(0.872)			(0.808)
Treated × Post × Tenured		-0.229			-0.303***
		(0.139)			(0.110)
Treated × Post × Other		-0.335**			-0.349**
		(0.158)			(0.142)
Ethnicity:					
Treated × Post × Asian			0.284		0.238
			(0.334)		(0.339)
Productivity:					
Treated × Post × Above-median productivity				0.343*	0.302**
				(0.192)	(0.148)
Constant	-0.125	0.0337	-0.121	-0.141	0.0442
	(0.286)	(0.247)	(0.285)	(0.285)	(0.242)
No. of scientists	1,718	1,718	1,718	1,718	1,718
Observations	12,796	12,796	12,796	12,796	12,796

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating faculty position, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Heterogenous effects within Social sciences and Humanities

	(1)	(2)	(3)	(4)	(5)
Treated × Post	0.110**	0.0854	0.109**	-0.0293	0.0192
	(0.0529)	(0.0620)	(0.0545)	(0.0417)	(0.0600)
Faculty position (PhD-students omitted):					
Treated × Post × Postdoc		0.591**			0.398
		(0.291)			(0.275)
Treated × Post × Tenured		-0.01000			-0.102
		(0.0627)			(0.0641)
Treated × Post × Other		0.0845			0.00560
		(0.0841)			(0.0809)
Ethnicity:					
Treated × Post × Asian			0.0401		0.0989
			(0.133)		(0.103)
Productivity:					
Treated × Post × Above-median productivity				0.396***	0.392***
				(0.0498)	(0.0544)
Constant	-0.182	-0.156	-0.181	-0.208*	-0.158
	(0.117)	(0.110)	(0.117)	(0.117)	(0.109)
No. of scientists	1,128	1,128	1,128	1,128	1,128
Observations	8,650	8,650	8,650	8,650	8,650

Notes: Table shows OLS estimates of β in Equation 1 where the treatment indicator is interacted with dummies indicating faculty position, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Heterogeneous effects for Tenured faculty

	(1)	(2)	(3)	(4)	(5)
Treated × Post	0.0723	0.200**	0.0532	-0.0367	0.0625
	(0.0737)	(0.0963)	(0.0719)	(0.0658)	(0.0708)
Scientific field (engineering and technology omitted):					
Treated × Post × Life sciences		-0.251**			-0.245**
		(0.124)			(0.114)
Treated × Post × Natural sciences and Mathematics		-0.110			-0.0900
		(0.181)			(0.160)
Treated × Post × Social sciences and Humanities		-0.277**			-0.200**
		(0.112)			(0.0933)
Ethnicity:					
Treated × Post × Asian			0.312		0.240
			(0.268)		(0.255)
Productivity:					
Treated × Post × Above-median productivity				0.370***	0.337***
				(0.111)	(0.101)
Constant	0.0241	0.0278	0.0207	0.0283	0.0276
	(0.0715)	(0.0717)	(0.0717)	(0.0702)	(0.0705)
No. of scientists	1,790	1,790	1,790	1,790	1,790
Observations	9,594	9,594	9,594	9,594	9,594

Notes: Table shows OLS estimates of β in Equation 1 for the sub-sample of tenured incumbents where the treatment indicator is interacted with dummies indicating field, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments and position based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Heterogeneous effects for Postdocs

	(1)	(2)	(3)	(4)	(5)
Treated × Post	0.639*	0.0284	0.675*	-0.477***	-0.921
	(0.372)	(0.547)	(0.400)	(0.164)	(0.562)
Scientific field (engineering and technology omitted):					
Treated × Post × Life sciences		0.566			0.504
		(0.617)			(0.646)
Treated × Post × Natural sciences and Mathematics		1.136			0.996
		(1.054)			(1.048)
Treated × Post × Social sciences and Humanities		0.244			0.0797
		(0.584)			(0.541)
Ethnicity:					
Treated × Post × Asian			-1.032		-1.338
			(0.727)		(0.993)
Productivity:					
Treated × Post × Above-median productivity				1.687***	1.643***
				(0.560)	(0.548)
Constant	0.109	0.200	0.121	0.167	0.266
	(0.752)	(0.733)	(0.749)	(0.738)	(0.719)
No. of scientists	236	236	236	236	236
Observations	1,811	1,811	1,811	1,811	1,811

Notes: Table shows OLS estimates of β in Equation 1 for the sub-sample of postdoc incumbents where the treatment indicator is interacted with dummies indicating field, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments and position based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A10: Heterogeneous effects for PhD-students

	(1)	(2)	(3)	(4)	(5)
Treated × Post	0.0490	0.456***	0.0111	-0.247***	0.110
	(0.0991)	(0.137)	(0.1000)	(0.0626)	(0.119)
Scientific field (engineering and technology omitted):					
Treated × Post × Life sciences		-0.259			-0.286
		(0.342)			(0.329)
Treated × Post × Natural sciences and Mathematics		-0.641***			-0.610***
		(0.144)			(0.135)
Treated × Post × Social sciences and Humanities		-0.528***			-0.406***
		(0.129)			(0.113)
Ethnicity:					
Treated × Post × Asian			1.175**		1.077**
			(0.535)		(0.502)
Productivity:					
Treated × Post × Above-median productivity				0.537***	0.497***
				(0.130)	(0.110)
Constant	-0.541	-0.372	-0.500	-0.522	-0.326
	(0.735)	(0.719)	(0.677)	(0.717)	(0.653)
No. of scientists	2,016	2,016	2,016	2,016	2,016
Observations	19,547	19,547	19,547	19,547	19,547

Notes: Table shows OLS estimates of β in Equation 1 for the sub-sample of PhD-student incumbents where the treatment indicator is interacted with dummies indicating field, Asian region-of-birth, and pre-2008 productivity level of incumbents. Dependent variable is number of publications per year. Treatment is based on above-median exposure to Asian migration in 2001. Incumbent researchers are classified into type of departments and position based on 2007 employment. Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Robustness analysis

Table B1: Main effect using placebo treatment period

	(1) P50<	(2) P75<	(3) P90<
	A. Publications per year		
Treated × Placebo post	0.0469	-0.0320	0.237
	(0.0673)	(0.0877)	(0.206)
R^2	0.02	0.03	0.04
No. of scientists	4,886	3,246	1,350
Observations	22,993	14,892	6,165
	B. Citations per year		
Treated × Placebo post	0.182	-0.612	-0.306
	(0.620)	(0.627)	(1.378)
R^2	0.00	0.00	0.01
No. of scientists	4,886	3,246	1,350
Observations	22,993	14,892	6,165
	C. Collaboration		
Treated × Placebo post	0.00947	0.00310	-0.00263
	(0.00794)	(0.00885)	(0.0125)
R^2	0.00	0.01	0.02
No. of scientists	1,984	1,461	654
Observations	7,059	5,443	2,263

Notes: Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Main effect, adding field-level R&D expenditure per faculty

	(1) P50<	(2) P75<	(3) P90<
		A. Publications per year	
Treated × post	0.0937	0.0224	0.363*
	(0.0759)	(0.0896)	(0.191)
R&D per faculty	0.000298	-0.000258	-0.000414
	(0.000734)	(0.000729)	(0.000462)
R^2	0.03	0.03	0.05
No. of scientists	4,883	3,246	1,350
Observations	35,488	23,299	9,530
		B. Citations per year	
Treated × post	1.672	0.0607	0.602
	(1.551)	(0.909)	(1.169)
R&D per faculty	-0.00708	-0.00452	-0.00164
	(0.00564)	(0.00487)	(0.00589)
R^2	0.00	0.00	0.01
No. of scientists	4,883	3,246	1,350
Observations	35,488	23,299	9,530
		C. Collaboration	
Treated × post	0.0136**	0.0172**	0.0158
	(0.00669)	(0.00758)	(0.0123)
R&D per faculty	0.000300	-0.0000771	-0.0000247
	(0.000423)	(0.000115)	(0.000127)
R^2	0.01	0.02	0.03
No. of scientists	2,460	1,783	816
Observations	11,821	9,081	3,823

Notes: Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B3: Main effect, adding university-time trends

	(2) P50<	(3) P75<	(4) P90<
	A. Publications per year		
Treated × post	0.0895	0.0192	0.411**
	(0.0721)	(0.0845)	(0.189)
R^2	0.03	0.04	0.06
No. of scientists	4,886	3,246	1,350
Observations	36,567	23,827	9,830
	B. Citations per year		
Treated × post	1.681	0.0624	0.938
	(1.481)	(0.842)	(1.062)
R^2	0.01	0.01	0.01
No. of scientists	4,886	3,246	1,350
Observations	36,567	23,827	9,830
	C. Collaboration		
Treated × post	0.0131**	0.0166**	0.0112
	(0.00659)	(0.00735)	(0.0121)
R^2	0.05	0.03	0.05
No. of scientists	2,476	1,793	821
Observations	12,150	9,329	3,971

Notes: Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B4: Main effect using balanced panel

	(1) P50<	(2) P75<	(3) P90<
	A. Publications per year		
Treated × post	0.0619	0.0117	0.791**
	(0.114)	(0.152)	(0.323)
R^2	0.02	0.03	0.04
No. of scientists	1,224	808	312
Observations	13,464	8,888	3,432
	B. Citations per year		
Treated × post	0.754	-1.202	2.636
	(1.944)	(1.133)	(2.161)
R^2	0.00	0.00	0.01
No. of scientists	1,224	808	312
Observations	13,464	8,888	3,432
	C. Collaboration		
Treated × post	0.0116	0.0118	0.00290
	(0.00906)	(0.0102)	(0.0150)
R^2	0.02	0.02	0.04
No. of scientists	752	529	216
Observations	5,746	4,497	1,785

Notes: Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B5: Main effect using conditional fixed-effects Poisson estimator

	(1) P50<	(2) P75<	(3) P90<
	A. Publications per year		
Treated × post	0.0205	-0.0127	0.149**
	(0.0369)	(0.0409)	(0.0675)
No. of scientists	2,469	1,789	820
Observations	19,745	13,881	6,206
	B. Citations per year		
Treated × post	0.120	-0.0283	0.0434
	(0.0973)	(0.0879)	(0.127)
No. of scientists	2,266	1,694	759
Observations	18,279	13,290	5,813

Notes: Robust standard errors clustered at the department-level in parentheses. All regressions include individual- and time-fixed effects, as well as age dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Matched sample results

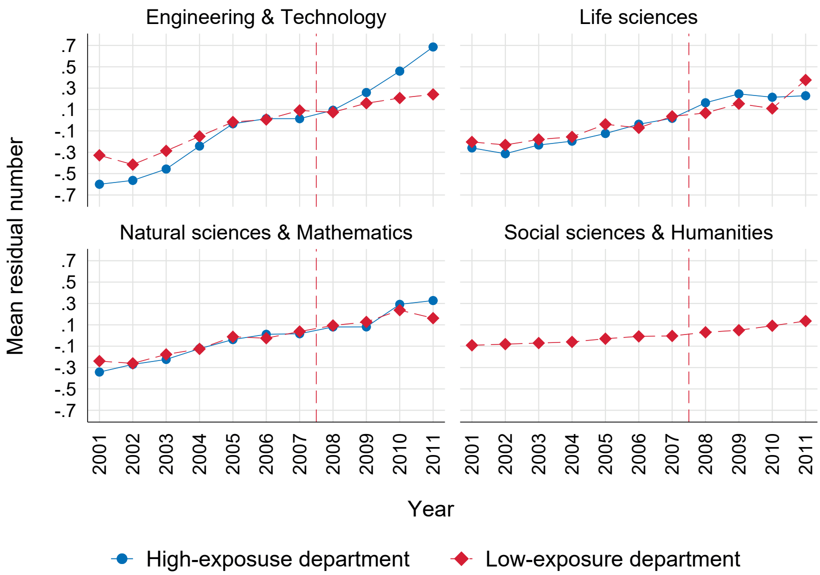
Table C1: Summary Statistics, matched sample in 2007 with 75th percentile of exposure

	Control	Treated	Difference
Scientific output:			
No. publications 2001–07	5.79	5.93	-0.137
No. citations 2001–07	32.45	36.07	-3.624
Age:			
19–31	0.31	0.35	-0.038*
32–39	0.20	0.19	0.010
40–47	0.19	0.17	0.020
48–56	0.14	0.13	0.005
57–	0.16	0.16	0.003
Male	0.63	0.64	-0.004
Faculty position:			
PhD-sudent	0.40	0.40	0.000
Postdoc	0.06	0.06	0.000
Tenured	0.36	0.36	0.000
Other	0.18	0.18	0.000
Scientific field:			
Engineering & Technology	0.35	0.35	0.000
Life sciences	0.19	0.19	0.000
Natural sciences & Mathematics	0.46	0.46	0.000
Social sciences & Humanities	0.00	0.00	0.000
Shanghai ranking, top-100 university	0.42	0.42	0.000
Observations	1,623	1,623	

Table C2: Summary Statistics, matched sample in 2007 with 90th percentile of exposure

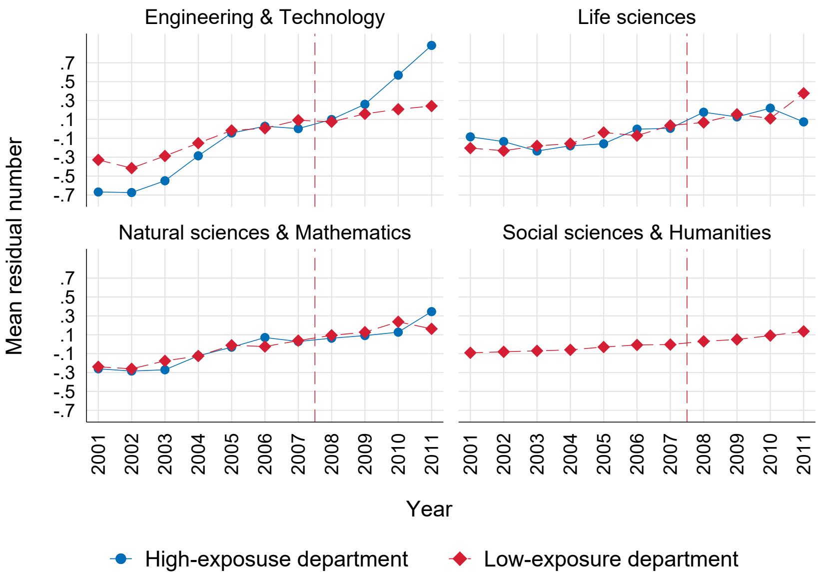
	Control	Treated	Difference
Scientific output:			
No. publications 2001–07	5.45	6.33	-0.883
No. citations 2001–07	25.59	29.04	-3.452
Age:			
19–31	0.35	0.39	-0.033
32–39	0.21	0.20	0.013
40–47	0.18	0.17	0.010
48–56	0.13	0.12	0.012
57–	0.13	0.13	-0.003
Male	0.69	0.73	-0.033
Faculty position:			
PhD-student	0.45	0.45	0.000
Postdoc	0.05	0.05	0.000
Tenured	0.33	0.33	0.000
Other	0.17	0.17	0.000
Scientific field:			
Engineering & Technology	0.72	0.72	0.000
Life sciences	0.00	0.00	0.000
Natural sciences & Mathematics	0.28	0.28	0.000
Social sciences & Humanities	0.00	0.00	0.000
Shanghai ranking, top-100 university	0.30	0.30	0.000
Observations	675	675	

Figure C1: Number of papers published annually by type of department and field, matched sample with above 75th percentile treatment



Note: Residual mean number of papers published annually by incumbents in high- and low-exposure departments, where the residual is defined as the difference between the actual number of papers published and the average number of papers published annually before 2008. A high-exposure department is a university-field combination with above the 75th-percentile presence of Asian-born researchers in 2001. Swedish researchers are classified into high- or low-exposure departments based on 2007 employment.

Figure C2: Number of papers published annually by type of department and field, matched sample with above 90th percentile treatment



Note: Residual mean number of papers published annually by incumbents in high- and low-exposure departments, where the residual is defined as the difference between the actual number of papers published and the average number of papers published annually before 2008. A high-exposure department is a university-field combination with above the 90th-percentile presence of Asian-born researchers in 2001. Swedish researchers are classified into high- or low-exposure departments based on 2007 employment.

CHAPTER II

Does mobility across universities raise scientific productivity?

with Olof Ejermo and Claudio Fassio

Abstract

Using a highly comprehensive new dataset on Swedish researchers, we investigate the effects of inter-university mobility on researcher productivity. Our study suggests substantial gains from mobility on scientific output. The empirical analysis addresses selection using inverse probability treatment censoring weights. We find that mobility induces a long-lasting increase in a researcher's publications by 32% and citations by 63%. Such mobility effects are not explained by promotions taking place jointly with a move. Positive effects are found among individuals who move between universities and not for those who move to or from university colleges. Moreover, we find that the positive effect of moving only applies to researchers in medicine, natural sciences and engineering and technology, with no effect of mobility found in the social sciences and in the humanities.

Keywords: Economics of science, mobility, scientific productivity, university
JEL Classification: O31, I23, J24

This paper has benefitted from comments by participants at the "Workshop on Scientists' Careers Inside and Outside the University" in Kassel, Germany; the "4th VPDE-BRICK Workshop in Economics of Innovation, Complexity and Knowledge", in Turin, Italy; the "11th Workshop on The Organisation, Economics and Policy of Scientific Research", in Turin, Italy; "DRUID17", New York, USA; and, seminars in Lund Univeristy and Epfl Lausanne. Moreover, the authors wish to thank Patrick Gaule, Thomas Åstebro, Otto Toivanen, Marco Guerzoni, Cristiano Antonelli and Anupama Phene for helpful comments on earlier versions of this paper.

1 Introduction

The lack of inter-university mobility among researchers has attracted substantial interest from policymakers because it is often claimed that low mobility across academic institutions hampers the diffusion of ideas, leading to an inefficient allocation of human capital and to intellectual inbreeding (OECD, 2008; Hynes et al., 2012). Many European countries might suffer from such problems, as indicated by the high share of faculty members who received their PhD at the same institution in which they work (Horta et al., 2010). This suggests that Europe, in particular, would benefit from a more integrated academic labor market, as less efficient knowledge production also implies negative consequences for economic growth, given the important role that universities may play in economic development (Foray & Lissoni, 2010).

In this paper, we investigate the importance of mobility by examining academic researchers' mobility across Swedish universities to see whether they become more scientifically productive and under what conditions. Sweden is an interesting empirical case, as its academic system is high performing, with several universities ranked among the top-100 worldwide and relatively high levels of public funding (Shanghai Jiao Tong University, 2017; Times Higher Education, 2017). At the same time, the system has historically displayed low levels of mobility across institutions. This might leave room for upgrading efficiency through researcher mobility.

We examine the effects of mobility on scientific productivity in terms of both publication output and the quality of scientific output, gauged through citation-weighted publication output. We analyze which factors are likely to have an impact on the overall effect of mobility, including the initial level of productivity of researchers, the interaction between mobility and promotion, and the importance of the status of the university of destination. Lastly, we investigate if the specific disciplinary field influences the impact of mobility on productivity.

To our knowledge, no solid empirical evidence exists yet on the scientific returns to mobility for academic researchers. A challenge in the measurement of such an effect is related to the fact that highly productive researchers and/or researchers who have better networks, are also more likely to be mobile (Zucker et al., 2002). This means that mobility can have a positive effect on productivity but also that the reverse effect—from productivity to mobility—cannot be ruled out (Hoisl, 2007).

Other factors are also likely to influence the impact of mobility on scientific productivity. In many cases, individuals move to a new university to obtain a promotion that would not be available at their current institution. Because researchers' promotions to a higher rank in the academic system can also have an impact on their productivity, the effect of mobility should ideally be separated from the effect of promotion. Furthermore, the status of universities plays a role in mobility patterns, because researchers prefer to move to universities that are more prestigious. Higher prestige tends to be associated with more resources, which in turn might lead to better research outcomes.

In addition, mobility effects can also be expected to differ by discipline. In some fields, especially in the hard sciences, moving to an environment with better research facilities, a better team of collaborators, or higher investment in labs and equipment can substantially increase performance. By contrast, this is less likely to be relevant in the social sciences or the humanities and researchers' output can be expected to be less dependent on the specific context in which it is undertaken.

We conduct our empirical analyses using a highly comprehensive new database on publications and citations of more than 35,000 Swedish university researchers who were active in the period 2002-2012, which is based on employer-employee and university registers at Statistics Sweden. These data allow us to follow individuals over time and to observe a large range of individual characteristics, controlling for confounding factors and analyzing individual heterogeneity. The data also enable us to identify researcher moves between academic institutions as well as publication output.

In our analysis, we address endogeneity through the method of inverse probability treatment censoring weights (IPTCW; Robins, 2000). The method assigns each individual a weight equal to the inverse probability of being treated, utilizing an individual's history of observables (Azoulay et al., 2009; Buenstorf, 2009), thereby allowing us to estimate the mobility effect over the full sample of researchers. In robustness estimations, we adopt instead a matched sample approach, in which we use a nonparametric coarsened exact matching (CEM) algorithm to identify an appropriate control group. Both methods rely on an unusually rich set of observables, e.g., whether researchers are married or have children, to account for selection into treatment. To the extent that these characteristics account for selection, our results have a causal interpretation.

Our estimations (both IPTCW and CEM) reveal that mobility leads to a lasting increase in both the quantity and quality of publication output of researchers who move (here called 'movers'), compared with those who remain (here called 'stayers'). In our preferred estimations, we find that the publication rate increases approximately by 32% and citations increase by 63%. The effect is not significantly different across the productivity distribution, i.e. both highly productive and less productive researchers benefit from it similarly. We also analyze the role of promotion in connection with mobility: some promotions may allow for more research time, whereas in other cases promotion may be associated with a heavier administrative burden or higher teaching load. We find that promotion associated with mobility does not contribute to the positive mobility effect, even though promotion is much more common among movers than among stayers. Moreover, using the university status (full research universities vs. university colleges) of the destination institution as a proxy for host institution quality, we find that only moves between full research universities have an impact, which is in line with an earlier study on the UK (Fernández-Zubieta et al., 2015a). Finally, when we distinguish among different disciplines, we find that moving has a positive effect for researchers in all disciplines but the humanities and the social sciences.

The paper contributes to the existing literature in several ways. First, we provide the first country-level analysis of the effect of researchers' mobility on productivity, encompassing most academic researchers in the national academic system, instead of focusing on a specific sample or discipline. This adds to the generalizability of our results. Second, the longitudinal dimension of our dataset and the richness of individual characteristics, coupled with our chosen estimation strategy, allows us to separate the effect of mobility from various confounding factors and to take selection issues into account. We also shed light on how specific factors, such as promotion, university hierarchy and initial productivity influence the mobility effect for individual researchers. Lastly, the comprehensiveness of our dataset allows us to check for differentiated effects of mobility across disciplinary fields.

2 Job mobility and researcher productivity

Although the effect of productivity on mobility (especially for highly productive individuals) is well established, the mobility effect on scientific productivity is still a matter of dispute, and evidence remains scarce. Some studies, mostly focused on the US academic system, find that job mobility increases the publishing performance of researchers. For example, Azoulay et al. (2011) find a positive effect of mobility on productivity and citations of scientists in the life sciences, suggesting that mobility fosters the diffusion of knowledge, as indicated by a higher number of citations that mobile scientists obtain from colleagues at recipient institutions. Similarly, Dubois et al. (2014), using a large sample of mathematicians active all over the world, find a positive effect of mobility on research productivity. Using data on the top 100 scientists in terms of publications in seven different disciplines, Halevi et al. (2016) also find that mobility between departments generally has a positive effect on publications and citations. Similarly, Di Lorenzo & Tartari (2014) in a study of 80 research-active academics working in UK life science departments find that mobility has a positive effect on scientific productivity.

The impact of mobility on productivity depends on institutional factors related to the specific workings of national academic labor markets. Although most results focused on the Anglo-Saxon context find a positive effect of mobility on academic performance, existing evidence from continental Europe shows a different picture. For instance, Bolli & Schläpfer (2015) find that inter-institutional mobility has no effect on the publication outcomes of economists in Austria, Germany, and Switzerland, in 2006-2008.

According to Fernández-Zubieta et al. (2015b) and Stephan (2012), most mobility within an academic system comes from nontenured staff (typically postdoctoral researchers) who have been unable to secure a permanent position and are often required to change institution at the end of their contracts. In a study on German postdoctoral researchers in economics and management, Bäker (2015) finds that mobility induces a negative short-term effect on scientific productivity, especially when researchers' social capital is strongly linked to the department with which they were originally affiliated (measured as number of co-authors and staff). Also in the Anglo-Saxon context, some studies find no evidence of a positive effect of mobility on research outcome. In a longitudinal study of academic careers among UK academics, Fernández-Zubieta et al. (2015a) find no evidence of a positive effect of mobility on academic performance. By contrast, they find that downward mobility (mobility to less-prestigious institutions) can reduce researcher productivity. A study by Kim et al. (2009) finds that also in the US, upward mobility, as proxied by becoming affiliated to a top 25 university, does not have a substantial effect on the productivity of academics in economics and finance. More precisely the effect was positive in the 1970s, but gradually disappeared in the 1980s and the 1990s, possibly because of the lower importance of physical proximity for teamwork, allowed by information and communication technologies.

At the same time, many European academic systems, including the Swedish system, do not require researchers to be mobile, for example, by leaving their university after receiving a PhD.[1] This can lead to aca-

[1] Germany is an important exception, where researchers must change university after completing PhD studies to attain a professorship (Bäker, 2015).

demic inbreeding.[2] Again, the empirical evidence is mixed. As shown by Cruz-Castro & Sanz-Menéndez (2010) using a sample of Spanish scientists, researchers with a PhD from the same institution at which they are currently active do not perform worse than PhD holders who are not ‘inbred.’ Other research suggests that concern over inbreeding is not limited to Europe. Horta et al. (2010) focus on the impact of hiring PhDs trained in the department that granted the person that PhD on academic performance in Mexico. The empirical results show that inbred academic researchers generally show lower performance in terms of scholarly output than academics who change their affiliation at least once over the course of their career. Morichika & Shibayama (2015) find similar results among Japanese researchers.

3 Data

We constructed a dataset of Swedish academics following a three-step procedure. In the first step, we sent requests to all Swedish universities and university colleges (defined below) asking for lists of staff whose position involves research—that is, professors, associate professors, postdocs, and PhD students—going as far back in time as they could.[3] These staff lists are public documents available in Sweden, and any government body (most universities in Sweden are state run) are required by law to release them to researchers upon request. Each listed individual has a social security number, giving each researcher a unique identifier even if the individual appears on several universities’ staff lists. The universities also supplied information about first and last names, affiliations, and e-mail addresses of the researchers to enable us to match them with publication records. In the end, out of the 28 universities and university colleges we contacted, 25 responded and sent us the requested information.

[2]The academic inbreeding perspective considers the social embeddedness of a researcher within his alma mater (his social capital) as an inhibiting factor, which may decrease the researcher’s ability to produce independent and original research. Academic inbreeding may also lead to hiring practices that reward relational capital more than academic merits. It must be stressed that some studies, such as the one by Bäker (2015), consider social (or relational) capital to be a positive factor that may instead increase publication productivity.

[3]A large group of staff is labelled ’Other researchers and teaching staff’. Even though those in this group do not have a formal academic position, many have a PhD and publish with an academic affiliation, so we include them in our analysis. The staff category consists of, for example, lab assistants and research engineers.

In the Swedish system, a distinction exists between full research universities (*universitet*) and university colleges (*högskolor*), from here on referred to as universities and university colleges.[4] All major scientific fields are represented at universities, whereas generally not all fields are represented at university colleges. Moreover, university colleges can only examine PhDs after a rigorous evaluation process. Universities are provided with more resources in general and in particular for PhD education and research. As a result, academics have on average more research time at universities.[5] However, some university colleges called university college are de facto universities as they have PhD examination rights within broad fields such as technology or medicine and are also allocated more resources. The exceptions are three specialized technical and medical university colleges: The Royal Institute of Technology (KTH), Chalmers University of Technology, and the Karolinska Institute (KI). Appendix Table A1 lists the academic divisions in our sample and whether they are considered universities or as university colleges in our study.

In the second step, with the assistance of Fraunhofer ISI in Germany, we match the lists to author IDs as given in the Scopus database, using combinations of their names, affiliations, and e-mail addresses that appear on publications and the staff lists. The Scopus database starts in 1996, and the staff lists cover different periods (see Appendix Table A1 for the years covered by each responding university's staff lists), leaving us with the period of analysis 2002-2012. For this period, the staff lists cover 70,202 unique researchers, of which we can match 35%, or 25,020 individuals, to an author ID in Scopus. The match rate might seem low, but this sample of researchers accounts for around 85% of publications (fractionalized by number of authors) associated with a Swedish author ID in Scopus.[6] For our matched researchers, Fraunhofer ISI added in-

[4] We use this term in this paper although readers who may look for them should bear in mind that most university colleges use the label universities when they translate their Swedish name to English.

[5] According to the latest available figures, an average employee at universities had 44% research time in 2011, whereas staff working at university colleges with/without any PhD education had 21/25% research time (SCB, 2017). The slight difference among the latter is probably explained by the fact that PhD training counts as teaching.

[6] There are two likely reasons for why only 35 percent of staff are matched. First, staff lists cover all disciplines at Swedish universities. However, coverage of publication activity varies by discipline. For example, in the social sciences and humanities researchers are more likely to publish in national journals or books such as monographs, which are less well covered by Scopus. Second, it is well-known that a small number of researchers account for a disproportionate number of publications and a large share of researchers do not publish (Azoulay et al., 2010; Lotka, 1926; Stephan, 1996).

formation on the number of Scopus publications by year. We count entire publications, rather than divide publication counts by the number of authors, and use the number of citations in a three-year window following the publication of an article as our measure of publication quality.

In the last step, we used researchers' social security numbers to link them to employer-employee and ancillary university-employee lists available from Statistics Sweden. This connection to register data gives us access to a wide range of data on researchers, including demographic and job characteristics. The matching of yearly publication output with university registers and further Swedish Statistics data allows us to infer whether an individual has changed university. We define mobility as a change in a researcher's main university employer from one year to the next.

One challenge is that not all academic researchers show up in Scopus publications; for example, they might publish in publication types that are rarely covered by Scopus, such as books, anthologies, or journals that publish in the national language (e.g., Swedish). This issue is more pressing in the social sciences and humanities. To address this issue, we use data from Statistics Sweden to predict the probability of matching a Swedish researcher with a publication in the Scopus database based on a wide range of individual-level characteristics. Indeed, to make our sample representative, we also want to include researchers who have no publications, but we need to ensure that the only reason we could not match those researchers is that they did not publish in Scopus-listed journals (and not for other reasons, such as homonymy, which prevented matching). Hence, we included as 'non-publishing researchers' individuals who were not matched to any Scopus publication but, according to our estimation, have a very low predicted probability (< 0.2) of being matched using our procedure. These individuals are likely to be researchers who truly do not have any publications listed in Scopus.[7] This allows us to add an additional 10,341 researchers from the register data. The result is a sample of 35,361 researchers (i.e., about 50% of all university researchers). For further details on the construction of the database and an analysis showing that it constitutes a representative sample of Swedish researchers with publications, see Ejermo et al. (2016).

We add some further restrictions on the sample for analysis. First, we omit all PhD students from the sample, including those who transition to a PhD student position from another researcher job. It frequently happens that someone either during their PhD studies, or after gradu-

[7] We later check the robustness of our results by dropping non-publishing researchers from the sample.

ating, moves to another academic division at the end of their studies or leaves academia. Hence, mobility among PhD students involves a different, largely unobservable, selection process that follows a different rationale than it does at later career stages (cf. Nisticò, 2018). Importantly, in our empirical analysis we use our observation of various characteristics that determine mobility in our data, further arguing that it allows us to account for selection effects. It is crucial for us to observe past productivity in terms of publications and citations, which, as we discuss in more depth below, are important determinants of future mobility. However, PhD students are likely to lack such paper trails simply because remaining in academia after receiving a PhD involves a selection into a research career. This makes selection into mobility for a PhD student determined to a greater extent by unobserved factors (to the econometrician), such as promising 'papers in the pipeline,' which in turn implies a higher degree of omitted-variable bias for this group.

Second, we can neither observe the productivity of researchers when they reside outside the country as the Scopus link is based on Swedish staff lists, nor can we observe individual characteristics when they are abroad, because Statistics Sweden only measures individuals residing in Sweden, but we have data on if they are internationally mobile. Such individuals therefore had to be dropped. By doing this, we lose 6% of the observations in our sample period.

Third, we omit from our sample researchers who move more than once during the 2002–2012 period. We refer to them as 'multiple movers.' It is not obvious how to account for the effects of different moves, for example whether effects from second moves are qualitatively the same as those from first moves. Moreover, multiple movers might differ from their nonmobile and one-time mobile counterparts. Thus, including them in the analysis would introduce additional heterogeneity between movers and stayers and would make the interpretation of our findings more difficult. Multiple movers comprise a relatively small share of mobile researchers (about 10%); in fact, a robustness analysis reveals that including multiple movers does not change our main results in any qualitatively significant way.[8]

[8]Those results are based on an indicator in which multiple movers have a value of one for all observations after the first move. Results are available from the authors on request.

4 Empirical strategy

4.1 Econometric specification

As a first step to disentangle the direction of causality between mobility and scientific productivity, we specify a difference-in-differences model using fixed-effects (Quasi-ML) Poisson regressions. Our main estimated equation relates researcher i's scientific productivity in year $t$to mobility in the following way:

$$E\left[y_{i,t}|PostMob_{i,t}, \boldsymbol{X}_{i,t}, \gamma_i, \delta_t\right] = Exp[\beta_1 PostMob_{i,t} + \boldsymbol{\beta}_2' \boldsymbol{X}_{i,t} + \gamma_i + \delta_t], \quad (1)$$

where $y_{i,t}$ is the dependent variable measuring the quantity or quality of publications measured as raw publication counts and citations, respectively; $PostMob_{i,t}$ is a typical difference-in-differences indicator that becomes 1 after a move; $\boldsymbol{X}_{i,t}$ is a vector of time-varying characteristics including the number of children each year, whether s/he marries (cf. Azoulay et al., 2017), number of years since graduation (experience); γ_i and δ_t are individual- and time-fixed effects, respectively. We include individual-level fixed effects, γ_i, to capture time-invariant heterogeneity, such as intrinsic differences in ability or motivation, that affect publishing productivity and the likelihood of mobility. Lastly, including calendar-year-fixed effects, δ_t, captures general time trends that affect publishing. Our main coefficient of interest is β_1, which is the change in researcher i's productivity following a move. To interpret the coefficient as an elasticity we exponentiate it and subtract one: $Exp\left[\beta_1\right] - 1$.

Our dependent variable is the number of publications or, alternatively, the number of citations each year. An advantage of this approach is that the fixed-effects (Quasi-ML) Poisson model takes the nonnegative count-data nature of our data into account (Wooldridge, 1999).[9] In contrast to OLS, the main advantage is that a Poisson model allows modelling of highly skewed distributions of nonnegative outcomes. We use (robust) standard errors clustered at the individual level, thus ensuring that the standard errors are consistent even if the underlying data-generating process is not Poisson (Cameron & Trivedi, 2010). Moreover, this estimator addresses another potentially important issue, that there may be risks of serial correlation in the error term (Bertrand et al., 2004). In fact, the Quasi-ML, FE Poisson is also robust to arbitrary pat-

[9]We also ran our specification using standard ordinary least squares (OLS) to check the robustness of the estimates. The results (available from the authors on request) show that they are very stable.

terns of serial correlation (Azoulay et al., 2010; Silva & Tenreyro, 2006; Wooldridge, 2010).

4.2 Threats to identification

In our empirical investigation, the threats to identification of the effect of mobility on publication and citation rates stem from the confounding effects of selection and treatment (mobility). The literature on mobility has shown that higher individual research productivity makes academics and other knowledge workers more likely to move (Azoulay et al., 2017; Hoisl, 2007). To the extent that such positive selection varies over time for an individual, it is likely to bias fixed-effects estimations.

In Table 1, we show summary statistics of 'stayers' (researchers who remain at the same university after graduation) and 'movers' (individuals who change their main university of employment once) in our sample. We find 1,270 movers, who each account for one move event by construction. This amounts to approximately 6% in our sample. By comparison, Azoulay et al. (2011) find among a sample of 9,000 elite scientists in the US a mobility rate of 30%. Bäker (2015) finds in a sample of 1,000 German academics in economics and management a mobility rate also of 30%. Similar rates are presented by Cruz-Castro & Sanz-Menéndez (2010) for a sample of 1,500 researchers in 'hard' sciences in Spain. Bolli & Schläpfer (2015) instead find a much lower 5% mobility rate among a sample of 1,000 economists working at Austrian, German, and Swiss universities. While in our Swedish case the rate of mobility is low, it is possibly explained by the fact that the sample is more representative of the population and includes many less productive researchers. Indeed, the table shows that movers have a more productive track record. Movers are associated with 0.53 more publications per year on average in the year of moving and have 3.05 more citations per year on average compared to stayers. They also more frequently hold tenured positions (3 percentage points more likely to be professors and 8 percentage points more likely to be associate professors), although a shorter period has elapsed since their degree (on average, 1.69 fewer years since graduating). Mobile researchers are also more likely to have children (85% vs 67%) and to be married or be cohabitating with a partner (65% vs 58%).

We employ two strategies to deal with the potentially endogenous relationship between publication and mobility. The first relies on the so-called inverse probability of treatment censoring weights (IPTCW), while the second is based on coarsened exact matching (CEM). In order to save space and because we consider IPTCW to be more representative of a general effect, we present only the IPTCW analysis in the main

Table 1: Descriptive statistics

	Movers (N = 1,270)					Stayers (N = 20,551)					
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Diff. in mean
No. of publications	1.74	1	2.85	0	34	1.20	0	2.40	0	50	0.53***
No. of citations	10.02	2	36.59	0	746	6.97	0	25.94	0	944	3.05***
Cum. no. of publications	10.32	4	18.39	0	192	7.59	1	18.91	0	452	2.73***
Cum. no. of citations	58.80	9	168.56	0	2626	44.56	1	174.80	0	7818	14.24***
Age	43.49	42	9.89	25	73	43.47	42	12.56	18	84	0.02
Years since degree	7.84	6	7.09	0	37	9.53	6	9.97	0	50	-1.69***
Male	0.56	1	0.49	0	1	0.59	1	0.49	0	1	-0.02
Children	0.85	1	0.95	0	5	0.67	0	0.92	0	8	0.18***
Married/Cohabitating	0.65	1	0.43	0	1	0.58	1	0.47	0	1	0.06***
Full Professors	0.18	0	0.33	0	1	0.15	0	0.33	0	1	0.03***
Associate Professors	0.35	0	0.40	0	1	0.27	0	0.41	0	1	0.08***
Postdocs	0.03	0	0.12	0	1	0.09	0	0.28	0	1	-0.06***
Guest researchers	0.33	0	0.39	0	1	0.42	0	0.46	0	1	-0.09***
‘Other’ researchers	0.11	0	0.22	0	1	0.07	0	0.22	0	1	0.04***

Note: Each individual is averaged as one observation in the table. PhD students are excluded. The category ‘other’ research staff mainly consists of PhDs working in academia and doing research but who lack an academic position. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

text. Appendix B discusses the second method, CEM, and Appendix C presents results using CEM.

4.3 Inverse probability of treatment censoring weights

Lacking a natural experiment, the standard way to account for selection into mobility would be to find an instrumental variable (IV) that explains mobility but is unrelated to a researcher's own productivity. However, in the absence of such an instrument, our primary choice to deal with self-selection into mobility is the inverse probability of treatment censoring weights (IPTCW; Robins, 2000). The basic idea behind this method is to give 'unexpected' movers—individuals whose move predicted on observables is *less* likely—a larger weight in the ensuing regressions than that of individuals whose move is more likely. Unlike fixed-effects estimations, IPTCW allows us to recover causal effects even in the presence of time-varying selection into treatment, that is, selection that correlates with past and future values of the dependent variable and other confounding variables described below. To the extent that we can account for selection by means of observables, the strategy mimics an IV approach by creating a 'pseudo-population,' in which an unexpected mover is more likely to move, without making any strong assumption on the functional form of the relationship.

We follow the method of Azoulay et al. (2009) and Buenstorf (2009) in constructing stabilized weights that take into account both endogenous selection into treatment and censoring (exiting the sample). In short, the treatment weights are defined as:

$$sw_{it} = \prod_{\tau=0}^{t} \frac{Prob\left(Mobility_{i\tau} \mid \boldsymbol{X}_{i\tau}\right)}{Prob\left(Mobility_{i\tau} \mid \boldsymbol{Z}_{i\tau-1}, \boldsymbol{X}_{i\tau}\right)}$$

And the censoring weights are defined as:

$$sw^{*}_{it} = \prod_{\tau=0}^{t} \frac{Prob(Exit_{i\tau} \mid \boldsymbol{X}_{i\tau})}{Prob(Exit_{i\tau} \mid \boldsymbol{Z}_{i\tau-1}, \boldsymbol{X}_{i\tau})},$$

where $\boldsymbol{X}_{i\tau}$ and $\boldsymbol{Z}_{i\tau-1}$ are matrices containing confounders associated with mobility. Following Azoulay et al. (2009), $\boldsymbol{X}_{i\tau}$_include various covariates while $\boldsymbol{Z}_{i\tau-1}$ contain the main selectors into mobility. In our case, $\boldsymbol{X}_{i\tau}$ comprises of age, age squared, gender, number of children, and marital status, as well as discipline, university, staff category, and year-fixed effects. As the main selectors into mobility, $\boldsymbol{Z}_{i\tau-1}$, we consider publications, citations and co-authors in $t-1$ as well as cumulative publications, cumulative citations and cumulative number of co-authors in $t-2$. As mentioned, scientists that are more productive are more likely to move. Therefore, including past publications and citations is natural. By including both the cumulative stock of publications and citations in $t-2$ as well as the flow of these variables in $t-1$, we are able to account for both differences in pre-mobility trends and the possibility that mobility is directly preceded by some dynamics in the outcome variables.[10] However, past productivity is not all. Many jobs are found through connections (see e.g., Calvó-Armengol & Jackson, 2004), so that researchers with a larger social network will get more job offers at other universities. Thus, to proxy for the social network of the researcher we also included co-authors at other Swedish universities as one of our main selectors into mobility.

Weighting by sw_{it} creates a 'pseudo-population' in which $\boldsymbol{Z}_{i\tau-1}$ no longer determines selection into treatment and the causal impact of mobility is the same as in the original population. The numerators and denominators are estimated by means of logit regressions.[11] The estimation of the denominator (using both $\boldsymbol{X}_{i\tau}$ and $\boldsymbol{Z}_{i\tau-1}$) and the numerator (just $\boldsymbol{X}_{i\tau}$) of sw_{it} for scientist i in year t is $\prod_{\tau=0}^{t}(1-p_{i\tau})$ if scientist i did not move by year t, and $\prod_{\tau=0}^{t-1}(1-p_{i\tau}) \times p_{it}$ if scientist i moved in t, in which $p_{i\tau}$ and p_{it} are the predicted probability obtained from logit estimations. Estimation of sw^*_{it} proceeds in the same fashion. The final IPTCW weights are defined as $sw_{it} \times sw^*_{it}$. Note that because the weights vary over time, they cannot be combined with fixed effects.

As mentioned, the main advantage of the IPTCW approach is that it will produce unbiased estimates if no unobserved time-varying confounders affect the propensity to be mobile. However, this is a strong assumption. For instance, as we discussed above, the hiring department is likely to observe 'papers in the pipeline' among potential hires, unobserved to us but whose existence will affect hiring chances and hence the likelihood of being mobile. However, as suggested by Azoulay et al. (2009), the method performs well when (i) the estimations are based on a large set of observables (and to the extent that these observables are correlated with unobservable confounders); (ii) subjects are drawn from the same labor market; and (iii) the dependent variable is measured similarly in the control and treatment groups. All three of these conditions are fulfilled in the current context.

[10]For example, looking at the publications rates in Figure A1 reveals that the mover has higher publication output in the year preceding the move. Consequently, the fixed effects estimator might underestimate the effect of mobility on productivity.

[11]We report these estimations in Appendix D.

5 Results

We analyze the effects of moving in three sections. First, we include all moves in our analyses (i.e., moves including any type of academic position, except moves to/from a PhD student position, as discussed above). We refer to this as our main effect of mobility. Next, we investigate the heterogeneous effects of mobility, differentiating between moves with and without promotions, moves up and down the university hierarchy (up = from a university college to a university and down = from a university to a university college), as well as how the effect of mobility differs depending on the past productivity of the researcher. Lastly, we investigate differences across disciplinary fields.

5.1 Main effect of mobility

Table 2 presents the results of the main analysis. Columns (3) and (6) presents the results of the IPTCW estimations on, respectively, the number of publications and of citations; columns (2) and (4) presents the results obtained with a simple fixed-effects (Quasi-ML) Poisson model on the same dependent variables. As a benchmark, we also report the results obtained with a simple pooled model without fixed effects in columns (1) and (4). We find positive effects from mobility on the quantity of publications and the quality of publications in all models. The IPTCW estimates are markedly higher than in the fixed-effects regressions, indicating an increase in publications and citations of 32% and 63%, respectively, for movers. For the matched sample, shown in Appendix C, the increases are slightly lower, 22% and 56% respectively, but still strongly significant. We also conducted our IPTCW analyses by dropping individuals who had zero articles published. This did not change those results in any substantive way. However, dropping these individuals brings the estimated IPTCW coefficients closer to the fixed effects estimates. This is intuitive, because individuals with zero articles published or cited are automatically dropped from the individual-fixed effects specifications.[12]

These results tell us that mobility has a positive effect on productivity for individual researchers who move from one university to another in the Swedish academic system.

We also run our main specifications using leading and lagged indicators of the move, which help us discern whether movers are publishing more or less than stayers before they move. The lagged indicators allow us to explore the duration and time pattern of the mobility effect.

[12]These results are available from the authors on request.

Table 2: Main Effect of mobility

	(1)	(2)	(3)	(4)	(5)	(6)
	Publications			Citations		
	Pooled	Fixed effects	Pooled IPTCW	Pooled	Fixed effects	Pooled IPTCW
PostMob	0.235***	0.165***	0.275***	0.440***	0.389***	0.488***
	(0.0727)	(0.0482)	(0.0855)	(0.165)	(0.146)	(0.157)
Observations	56,438	56,438	56,438	50,221	50,221	50,221
No. of scientists	9,121	9,121	9,121	8,081	8,081	8,081
Individual FE	No	Yes	No	No	Yes	No
IPTCW	No	No	Yes	No	No	Yes

Note: Baseline category is researchers who never move. All models include time-varying controls; these are years of experience, number of children, an indicator if an individual is married/cohabiting and year-fixed effects. Pooled models also include time-fixed controls; these are discipline dummies, year of birth and indicator of sex. To interpret coefficient as elasticities we take $[\exp(PostMob) - 1]$. Robust (QML) standard errors clustered at the individual level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

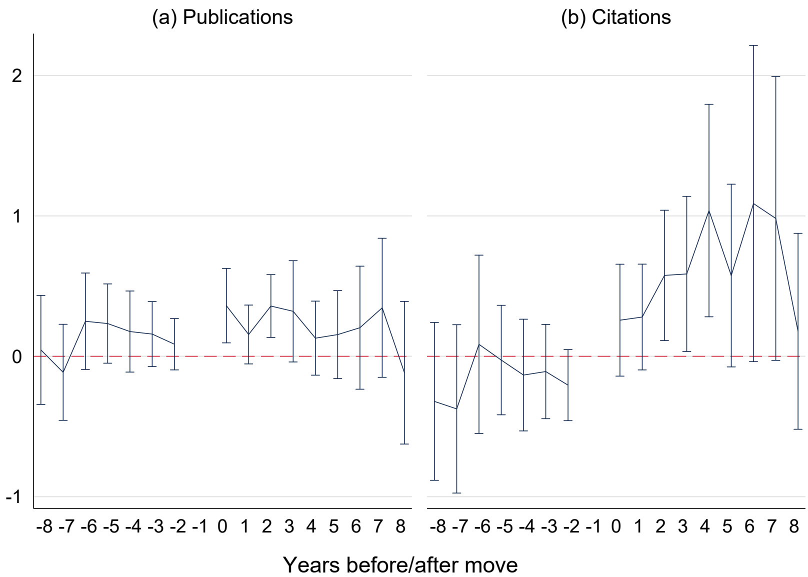
We plot the point estimates of the leading and lagged indicators from Poisson estimations with IPTCW in Figure 1.[13] Generally, we note that our preferred specification, using the IPTCW method, is successful in removing pre-treatment differences. In addition, each method reveals that the effects of mobility last up to eight years after the move, suggesting that mobility has a long-lasting effect on productivity.

5.2 Heterogeneous effects of mobility

We now turn to how the main effects of mobility differ by the type of move. First, we investigate the role of career transitions (promotions) in the effects of mobility, looking specifically at the effect of moving up or down the career ladder in academia. Second, we investigate whether mobility up the Swedish university hierarchy is important in determining the size effect of mobility. Previous studies indicate that a move has a positive impact on scientific output, especially if it involves upward mobility in the university hierarchy (Fernández-Zubieta et al., 2015a), which suggests that moving to better research environments has a positive impact on researcher productivity. Finally, we check whether the effects of mobility vary depending on the past productivity of researchers.

[13]We also ran these estimations using Poisson fixed effects with and without using the matched sample. These results are available in Appendices A and C , respectively.

Figure 1: Main effect of mobility with leads and lags using pooled Poisson with IPTC weights



Note: This figure plots point estimates for leading and lagging indicators for mobility. The omitted category is one year prior to the move. All specifications include controls for years of experience, number of children, an indicator if an individual is married/cohabiting, discipline dummies, year of birth, indicator of sex and year-fixed effects. Vertical bars correspond to 95% confidence intervals with robust (QML) standard errors clustered at the individual level.

In Appendix C, we present the same set of analyses using the CEM matched sample: all results are in line with those displayed in the paper.

Career ladders and mobility.— Earlier literature has highlighted how research opportunities vary by position (e.g., Sabatier, 2012). One of our hypotheses is that the mobility-productivity effect is closely tied to promotions. We consider this especially important in systems with low mobility rates, as in the Swedish case, in which promotions are likely to be a driver of mobility. Indeed, Appendix Table A2 , which tabulates position transitions subdivided by movers and stayers, confirms that the probability of a change in position is more common among movers than among stayers.

To analyze the general role of moving up or down the academic career hierarchy, we construct indicators: (i) 'Promotion,' which equals 1 the first time (and afterward) a researcher is observed to have moved up the career ladder; and (ii) 'Demotion,' which equals 1 the first time (and afterward) an academic is observed to have moved down the career ladder. Promotion is defined to take place from one year to the next if: an associate professor becomes a full professor; a postdoc becomes an associate professor or a full professor; and for other researchers if they become postdoc, associate or full professor. Demotions are defined as taking place for professors that change to any other type of position; for associate professors if they become a postdoc or 'other researcher'; for postdocs if they transit to the 'other researcher' category. In this context, 'other researchers,' by definition, cannot be demoted, nor can professors be promoted.[14]

In Table 3, we investigate the role of moving up or down the academic career hierarchy and mobility. To be able to distinguish the role of mobility on productivity from that of career transitions, we introduce a new indicator, 'Stay,' which equals one for researchers who never move. This indicator is used, in combination with the 'PostMob' indicator and the 'Promotion' and 'Demotion' dummies, to create four new indicators: 'PostMob & No Promotion/Demotion,' 'PostMob & Promotion,' 'Stay & Promotion,' 'PostMob & Demotion,' and 'Stay & Demotion.' The indicator 'PostMob & No Promotion/Demotion' equals one each year after a researcher moves if that move is not associated with a change in career position; otherwise, it is zero. Conversely, 'PostMob & Promotion' equals one each year after a mobility event when it is associated with a promotion, and zero otherwise. 'Stay & Promotion' equals one for each year after the first observed promotion for researchers who never move. Correspondingly, 'PostMob & Demotion' equals one each year after the first observed demotion if associated with a move in the same year and 'Stay & Demotion' equals one for researchers who are demoted and do not move. The baseline category is not moving and not being promoted. An individual is allowed, in our coding, to change status only once to any of the other categories.[15]

Again, we display the estimates obtained using (i) the IPTCW method (the preferred analysis method) and (ii) fixed effects. The effect of mobility per se is still positive and significantly different from zero in all specifications, and the magnitude of the coefficients is in line with the

[14]This coding is done such that a researcher is never demoted after a promotion and never promoted after a demotion.

[15]Thus, an individual who first moves and is promoted later will only be coded as 'PostMob & No Promotion/Demotion.'

Table 3: Effect of career transitions and moving

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Fixed effects	Pooled IPTCW	Fixed effects	Pooled IPTCW
PostMob & No Prom/Demo	0.185***	0.286**	0.353**	0.458**
	(0.0690)	(0.126)	(0.149)	(0.210)
PostMob & Promotion	0.200***	0.461***	0.341**	0.676**
	(0.0548)	(0.156)	(0.148)	(0.293)
Stay & Promotion	0.0801***	0.230***	0.120**	0.143
	(0.0208)	(0.0533)	(0.0524)	(0.108)
PostMob & Demotion	0.189	0.155	0.682	0.364
	(0.171)	(0.164)	(0.481)	(0.361)
Stay & Demotion	0.0730*	0.0658	0.0804	-0.0811
	(0.0392)	(0.0810)	(0.0845)	(0.104)
Observations	56,438	56,438	50,221	50,221
No. of scientists	9,121	9,121	8,081	8,081
Individual FE	Yes	No	Yes	No
IPTCW	No	Yes	No	Yes

Note: Baseline category is researchers who stay and are not promoted or demoted. See Table 2 for other notes.

results in Table 2. Moreover, the results indicate that a move coupled with a promotion leads to a higher impact on productivity, but further chi-squared tests on the equality of coefficients reveal that the effect is not significantly different from the effect of mobility without promotion. This is true for both publications and citations. We also find that researchers who remain and are promoted experience a positive and significant effect on publications and citations. The magnitude of the coefficients is equal or slightly lower than the coefficient for mobility without promotion, while it is always lower than the effect of mobility coupled with promotion. However, tests on the equality of coefficients indicate that the difference is significantly different only in column (2). By contrast, whenever a demotion is involved, the effect on productivity is almost never significantly different from zero.

In sum, we find that movers who are simultaneously promoted do not show a statistically significant different publication rate than other movers. This suggests that it is mainly the move itself that explains the effect, not promotion. Moreover, the results suggest that individuals who do not move but are promoted are equally able to increase their productivity over time, although a slightly lower effect is seen on citations. This last result is in line with previous findings by Cruz-Castro & Sanz-Menéndez (2010) in the Spanish context, according to which careers at the same academic institution are not necessarily associated with lower publication outcomes.

Moving up the university hierarchy.— An additional factor pointed out in the literature is that mobility can have a positive impact on scientific productivity by allowing the individual to gain access to a better research environment. This is often labeled as moving up the university hierarchy, that is, moving to a university with a higher 'ranking' (Dubois et al., 2014; Fernández-Zubieta et al., 2015a).

In Sweden, universities and university colleges (see Section 3) provide a natural division for analyzing hierarchical moves. Appendix Table A3 illustrates the significance of this distinction, providing the average number of publications for a researcher–year observation by university, ranked by the average number of publications. Eight universities and two university colleges, the latter ranked nine and ten, are among the top ten performers. From below, among the ten least publishing institutions are eight university colleges, seven ranked at the bottom and one ranked ninth from the bottom. These descriptive statistics are suggestive of the importance of university status for research performance, although they do not standardize results by discipline, composition of researchers, and so forth.

In Table 4, we show regression results in which we distinguish among different types of moves by introducing four different dummy variables: (i) 'University to university,' which equals one when a researcher moves from one university to another university; (ii) 'University college to university college,' which equals one when a researcher moves from one university college to another university college; (iii) 'University college to university,' which equals one when a researcher moves from a university college to a university; and (iv) 'University to university college,' which equals one when the opposite type of move occurs. The baseline is to remain, that is, *not* to move. The results of the estimation, displayed in Table 4, point unequivocally to the important role of moves between universities as the only type of moves that boost researchers' productivity. While this kind of move leads to a statistically significant increase of 49% in publication counts (22% in the fixed-effects estimates) and to a 78% increase in citations (55% in the fixed-effects estimates), other types of moves never lead to an increase in researcher productivity. In the case of moves from a university college to a university or to another university college, the moves even have a negative effect in terms of publications and citations.[16] A possible interpretation of these results is that the moves of researchers from one research-based university to another may be more motivated by the search for better research environments (which eventually lead to better productivity). By contrast, other types of moves may have other rationales such as family reasons or commuting, that have little or nothing to do with academic productivity.

Table 4: Effect of moving to/between universities

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Fixed effects	Pooled IPTCW	Fixed effects	Pooled IPTCW
PostMob & Uni to Uni.	0.202***	0.398***	0.437***	0.576***
	(0.0566)	(0.0915)	(0.164)	(0.157)
PostMob & Coll. to Coll.	0.0948	-0.870***	-0.391	-1.747***
	(0.172)	(0.248)	(0.303)	(0.384)
PostMob & Coll. to Uni.	0.0695	-0.314**	0.0830	-0.492
	(0.111)	(0.138)	(0.266)	(0.389)
PostMob & Uni to Coll.	0.0564	-0.262	0.169	-0.356
	(0.100)	(0.220)	(0.223)	(0.388)
Observations	56,438	56,438	50,221	50,221
No. of scientists	9,121	9,121	8,081	8,081
Individual FE	Yes	No	Yes	No
IPTCW	No	Yes	No	Yes

Note: Baseline category is researchers who never move. See Table 2 for other notes.

Earlier productivity and mobility.— We also want to investigate how the productivity effect of mobility varies depending on past productivity. As indicated above, we have constructed a matched sample as an alternative estimation method. This sample is now useful to investigate differences in pre-mobility as, differently from the IPTCW-method, it identifies a point in time in which researchers could be characterized as high or low in productivity. We first split the matched sample based on the median number of cumulative publications in the year of match and create an indicator variable which takes the value one for above-median productivity individuals in the year of match and which remains constant afterwards. This variable is interacted with all previously included covariates and these newly created variables added to the

[16] The matched sample approach in Appendix Table C4 shows similar results, with only moves among universities leading to improved publication and citation outputs.

specification to flexibly estimate the effect of mobility on highly productive individuals, our main variable of interest. This allows us to check for statistically significant differences of the mobility effect between the individuals above and below the median presented in Table 5.

Table 5: Main effect of mobility over the productivity distribution including interactions of indicator if researcher is above median productivity in year of match, matched sample

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Pooled	Fixed effects	Pooled	Fixed effects
PostMob	0.329**	0.336***	0.417**	0.570**
	(0.137)	(0.111)	(0.180)	(0.232)
Above median	28.99	.	-4.207	.
	(27.65)	(.)	(42.67)	(.)
PostMob × Above median	0.00949	-0.127	0.440*	-0.119
	(0.159)	(0.123)	(0.266)	(0.268)
Observations	8,483	8,483	7,547	7,547
Number of scientists	1,051	1,051	926	926
Individual FE	No	Yes	No	Yes
IPTCW Weights	No	No	No	No

Notes: The variable 'Above median' is an indicator value taking the value of one for all individuals above the median number of cumulative publications in the year of match. All covariates have been interacted with the 'Above median'-indicator. See Table 2 for other notes.

The table indicates that mobility does not have different effects depending on productivity. On the other hand, researchers above the median in terms of publications before a move get a higher boost in terms of citations, although this effect is only significant in the pooled model and not in the individual fixed effects models. It seems natural to conclude that it is rather underlying characteristics such as 'fame' or social networks, which boost citations, rather than mobility per se.

Mobility and disciplinary fields.— Another relevant dimension is the role of different disciplinary fields. In Table 6, we display results obtained by running the model in equation (1) separately for researchers in four different disciplinary fields. Our data allow us to identify the discipline of each researcher, distinguishing among social sciences and humanities, engineering and technology, medicine, and the natural sciences.[17] When we run our models separately by discipline, we find that the positive effect obtained at the aggregate level is confirmed for medicine and the natural sciences (only for what concerns publications) for engineering and technology (only for citations). On the contrary, we find no mobility effect among researchers active in the social sciences and the humanities, neither on publications nor on citations.

We interpret these results considering the different preconditions for research in the various disciplines. For disciplines such as engineering and medicine, the availability of a (highly costly) research infrastructure can make a difference in the overall quality of research performed. Having access to better equipment, labs, or teams, can substantially increase the chances of achieving better research outcomes in engineering as well as in medicine. However, this is not necessarily the case for research in the humanities and the social sciences, in which most research is conducted without the need for costly equipment. It is likely, then, that when researchers in the hard sciences move, they tend to move to academic environments with better endowments in terms of research infrastructure, which complements and augments their research productivity. In the social sciences and the humanities, however, moves may more frequently be related to other motives that are less directly linked to productivity prospects.

6 Concluding discussion

This paper contributes to the understanding of the effect of mobility of researchers on their individual quantity and quality of scientific output by adopting a new highly encompassing dataset on the universe of Swedish researchers and their publication activities. The existing literature argues that mobility might increase academics' individual productivity if it improves research opportunities, for instance, if a move takes the researcher to a more research-intensive environment. We find that productive researchers are more likely to be mobile. In other words, the outcome of interest also gives rise to selection into treatment. To deal with these issues, we employ a weighting technique (IPTCW) that coun-

[17] The disciplinary fields correspond to the first-digit level in the Organization for Economic Cooperation and Development's Field of Science and Technology (FOS) Classification in the Frascati Manual. Because all research in the agricultural and veterinary sciences is confined to the Swedish University of Agricultural Sciences, for which we lack information, this field is excluded from the analysis.

Table 6: Effect of mobility within disciplines

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Fixed effects	Pooled IPTCW	Fixed effects	Pooled IPTCW
	Panel A. Social science			
PostMob	-0.142	0.176	0.0181	0.282
	(0.102)	(0.195)	(0.305)	(0.365)
Observations	13,187	13,187	9,612	9,612
No. of scientists	1,981	1,981	1,438	1,438
	Panel B. Medicine			
PostMob	0.200***	0.276**	0.273**	0.290
	(0.0740)	(0.122)	(0.130)	(0.193)
Observations	17,690	17,690	17,219	17,219
No. of scientists	3,164	3,164	3,057	3,057
	Panel C. Natural sciences			
PostMob	-0.0970	0.297**	0.0595	0.0201
	(0.0965)	(0.147)	(0.171)	(0.199)
Observations	9,242	9,242	8,740	8,740
No. of scientists	1,493	1,493	1,396	1,396
	Panel D. Engineering/Technology			
PostMob	0.268***	0.303	0.535	1.403***
	(0.0834)	(0.185)	(0.384)	(0.362)
Observations	16,213	16,213	14,529	14,529
No. of scientists	2,541	2,541	2,239	2,239
Individual FE	Yes	No	Yes	No
IPTCW	No	Yes	No	Yes

Notes: Baseline is researchers within discipline who do not move. See Table 2 for other notes.

teracts endogeneity in moves. In robustness checks, we also use non-parametric matching techniques to create a sample of similar scientists who do and do not move.

In our preferred specifications, the estimated gain following a move is 32% in the publishing rate and 62% in the citation rate. Estimations including both leading and lagged effects suggest that they are not short-term effects, as they last up to eight years after a move. Moreover, the results are not driven mainly by individuals who are promoted at the same time as a move, suggesting that mobility per se explains a large part of productivity increases. We also find that the positive effect of mobility on both publications and citations is only due to moves between universities, as opposed to other types of moves involving more teaching-intense academic institutions, such as university colleges. These results suggest considerable sorting effects in the university system where more research-intensive departments (in this context, a university) will both hire better researchers and get a bigger 'bang for their buck' in terms of publication by their hires. Finally, the results point to a stronger effect of mobility for researchers active in medicine and the natural sciences and — only in terms of citations counts — in engineering and technology, while no effect is found among researchers in the social sciences and the humanities.

One should note that our results are limited to the personal productivity gains obtained. Other studies have investigated social gains such as spillovers in the new environment on others' publication activity, which are typically positive and sizable (e.g., Azoulay et al., 2010; Borjas & Doran, 2012; Moser et al., 2014). In addition, social effects of mobility may well extend beyond publication to innovation (Azoulay et al., 2009; Buenstorf, 2009). These effects may also extend to the sending organization (Kaiser et al., 2015).

Lastly, our results show that the effect of mobility differs significantly between different disciplines. This suggests that future research efforts should be devoted to investigating the underlying reasons (importance of teamwork, access to better equipment and labs, etc.) behind these differences. Understanding why the effect of mobility differs across disciplinary fields would also allow us to understand better why mobility may or may not lead to increased academic productivity. Hence, cross-disciplinary databases like the one used in this paper may be extremely useful. Such research should also acknowledge that studying the effect of mobility in a specific disciplinary field might be informative about the mechanisms that drive mobility in that specific context, but one should be careful in generalizing the findings to academia as a whole.

Results from this study can be of relevance also for other academic systems. However, some caution should be taken when generalizing results found in a relatively small academic system like the Swedish one to larger countries in Europe and worldwide. Potentially, the development of a more integrated European academic labor market could facilitate better matching and (in the long-run) compensate for lacking academic mobility at the national level.

References

Azoulay, P., Ding, W., & Stuart, T. (2009). The impact of academic patenting on the rate, quality and direction of (public) research output. *The Journal Of Industrial Economics*, *57*(4), 637–676.

Azoulay, P., Ganguli, I., & Zivin, J. G. (2017). The mobility of elite life scientists: Professional and personal determinants. *Research Policy*, *46*(3), 573–590.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal Of Economics*, *125*(2), 549–589.

Azoulay, P., Zivin, J. S. G., & Sampat, B. N. (2011). The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. In J. Lerner & S. Stern (Eds.), *The Rate and Direction of Inventive Activity Revisited* (pp. 107–155). University of Chicago Press.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). cem: Coarsened Exact Matching in Stata. *Stata Journal*, *9*(4), 524.

Bolli, T. & Schläpfer, J. (2015). Job mobility, peer effects, and research productivity in economics. *Scientometrics*, *104*(3), 629–650.

Borjas, G. J. & Doran, K. B. (2012). The collapse of the Soviet Union and the productivity of American mathematicians. *The Quarterly Journal Of Economics*, *127*(3), 1143–1203.

Buenstorf, G. (2009). Is commercialization good or bad for science? Individual-level evidence from the Max Planck Society. *Research Policy*, *38*(2), 281–292.

Bäker, A. (2015). Non-tenured post-doctoral researchers' job mobility and research output: An analysis of the role of research discipline, department size, and coauthors. *Research Policy*, *44*(3), 634–650.

Calvó-Armengol, A. & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American Economic Review*, *94*(3), 426–454.

Cameron, A. C. & Trivedi, P. K. (2010). *Microeconometrics Using Stata*, volume 2. Stata Press College Station, Tx.

Cruz-Castro, L. & Sanz-Menéndez, L. (2010). Mobility versus job stability: Assessing tenure and productivity outcomes. *Research Policy*, *39*(1), 27–38.

Di Lorenzo, F. & Tartari, V. (2014). Should I stay or should I go? How mobility explains individual scientific performance. In *Academy Of Management Proceedings*, volume 2014. Academy Of Management.

Dubois, P., Rochet, J.-C., & Schlenker, J.-M. (2014). Productivity and mobility in academic research: Evidence from mathematicians. *Scientometrics*, *98*(3), 1669–1701.

Ejermo, O., Alder, C., Fassio, C., & Källström, J. (2016). Publications of Academic Researchers In Sweden (PARIS). Technical report, CIRCLE, Lund University. `http://paris.circle.lu.se`.

Fernández-Zubieta, A., Geuna, A., & Lawson, C. (2015a). Productivity pay-offs from academic mobility: should I stay or should I go? *Industrial and Corporate Change*, *25*(1), 91–114.

Fernández-Zubieta, A., Geuna, A., & Lawson, C. (2015b). What do we know of the mobility of research scientists and impact on scientific production. In A. Geuna (Ed.), *Global Mobility of Research Scientists* (pp. 1–33). Elsevier.

Foray, D. & Lissoni, F. (2010). University research and public–private interaction. In B. Hall & N. Rosenberg (Eds.), *Handbook Of The Economics Of Innovation*, volume Volume 1 (pp. 275–314). North-Holland.

Halevi, G., Moed, H. F., & Bar-Ilan, J. (2016). Researchers' mobility, productivity and impact: Case of top producing authors in seven disciplines. *Publishing Research Quarterly*, *32*(1), 22–37.

Hoisl, K. (2007). Tracing mobile inventors—the causality between inventor mobility and inventor productivity. *Research Policy*, *36*(5), 619–636.

Horta, H., Veloso, F. M., & Grediaga, R. (2010). Navel gazing: Academic inbreeding and scientific productivity. *Management Science*, *56*(3), 414–429.

Hynes, M., Ackers, L., Avveduto, S., Basri, E., Boudard, E., Castro, L. C., Langer, J., Linkova, M., Scholz, B., Veugelers, R., et al. (2012). Excellence, equality and entrepreneurialism building sustainable research careers in the European Research Area. In *Report of the Expert Group* *on the Research Profession*. Brussels: European Commission, Directorate General for Research and Innovation.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review Of Economics And Statistics*, *86*(1), 4–29.

Kaiser, U., Kongsted, H. C., & Rønde, T. (2015). Does the mobility of R&D labor increase innovation? *Journal Of Economic Behavior & Organization*, *110*, 91–105.

Kim, E. H., Morse, A., & Zingales, L. (2009). Are elite universities losing their competitive edge? *Journal Of Financial Economics*, *93*(3), 353–381.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal Of Washington Academy Sciences*, *16*(12), 317–323.

Moffitt, R. A. (2004). Introduction to the symposium on the econometrics of matching. *Review Of Economics And Statistics*, *86*(1), 1–3.

Morichika, N. & Shibayama, S. (2015). Impact of inbreeding on scientific productivity: A case study of a Japanese university department. *Research Evaluation*, *24*(2), 146–157.

Moser, P., Voena, A., & Waldinger, F. (2014). German Jewish Émigrés and US invention. *American Economic Review*, *104*(10), 3222–3255.

Nisticò, R. (2018). The effect of PhD funding on postdegree research career and publication productivity. *Oxford Bulletin of Economics and Statistics*, *80*(5), 931–950.

OECD (2008). *The Global Competition For Talent: Mobility Of The Highly Skilled*. Paris: Organization for Economic Cooperation and Development.

Robins, J. M. (2000). Marginal structural models versus structural nested models as tools for causal inference. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials* (pp. 95–133). Springer New York.

Sabatier, M. (2012). Does productivity decline after promotion? The case of French academia. *Oxford Bulletin Of Economics And Statistics*, *74*(6), 886–902.

SCB (2017). Andel av total arbetstid för anställda vid universitet och högskolor efter kön, lärosäte, arbetsuppgifter och vartannat år. http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__UF__UF0304__UF0304B/UoHIndArsvRelLaro/table/tableViewLayout1/?rxid=561eb25b-5bd4-4075-b2fe-6197067d78c0. Accessed: 27 Jan, 2017.

Shanghai Jiao Tong University (2017). Academic ranking of world universities. http://www.shanghairanking.com. Accessed: 15 jan, 2017.

Silva, J. M. C. S. & Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, *88*(4), 641–658.

Stephan, P. E. (1996). The economics of science. *Journal Of Economic Literature*, *34*(3), 1199–1235.

Stephan, P. E. (2012). *How Economics Shapes Science*, volume 1. Harvard University Press Cambridge, Ma.

Times Higher Education (2017). World university rankings. www.timeshighereducation.com. Accessed: 15 jan, 2017.

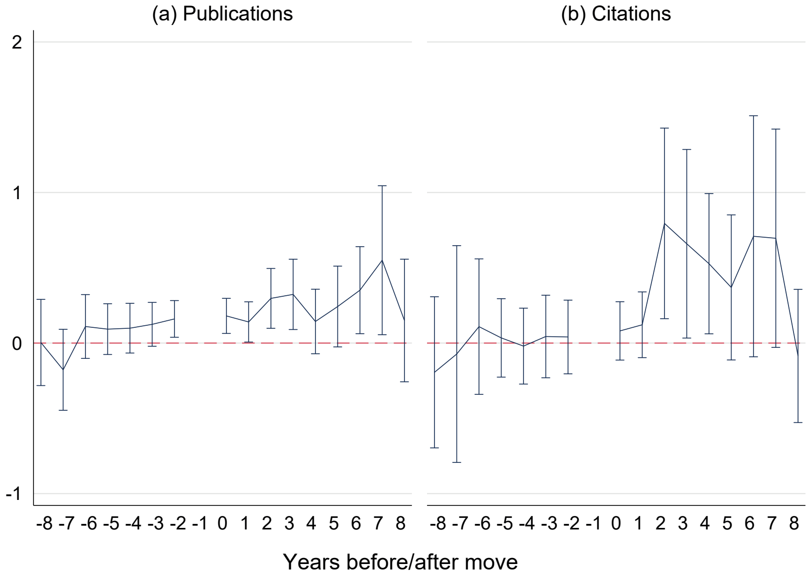
Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal Of Econometrics*, *90*(1), 77–97.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Zucker, L., Darby, M. R., & Torero, M. (2002). Labor mobility from academe to commerce. *Journal Of Labor Economics*, *20*(3), 629–660.

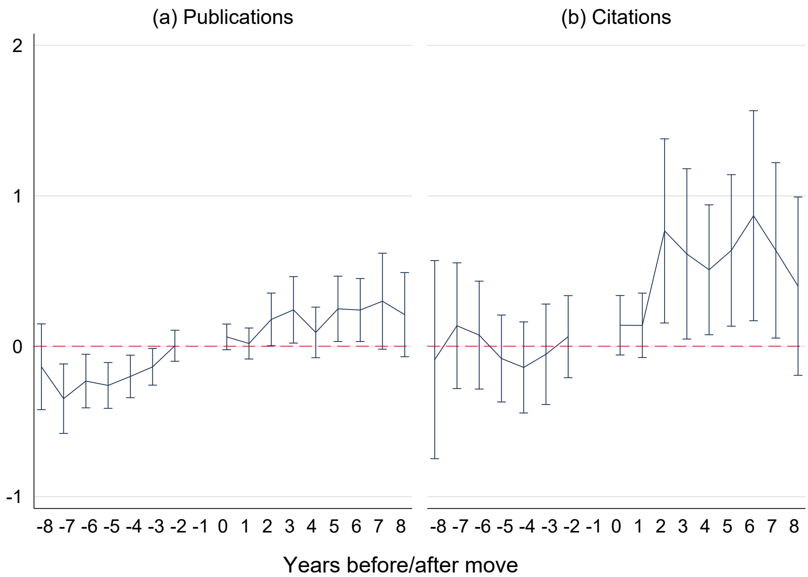
A Tables and graphs

Figure A1: Main effect of mobility with leads and lags using pooled Poisson



Note: This figure plots point estimates for leading and lagging indicators for mobility. The omitted category is one year prior to the move. All specifications include controls for years of experience, number of children, an indicator if an individual is married/cohabiting, discipline dummies, year of birth, indicator of sex and year-fixed effects. Vertical bars correspond to 95% confidence intervals with robust (QML) standard errors clustered at the individual level.

Figure A2: Main effect of mobility with leads and lags using fixed effects Poisson



Note: This figure plots point estimates for leading and lagging indicators for mobility. The omitted category is one year prior to the move. All specifications include controls for years of experience, number of children, an indicator if an individual is married/cohabiting, discipline dummies, year of birth, indicator of sex and year-fixed effects. Vertical bars correspond to 95% confidence intervals with robust (QML) standard errors clustered at the individual level.

Table A1: Years covered by staff lists from responding universities in relation to years covered by the Scopus database

Years	*BTH*	***CTH***	***GU***	*HB*	*HDA*	*HH*	*HJ*	*HKR*	*HS*	*HV*	***KAU***	***KI***	***KTH***	***LIU***	***LNU***	***LTU***	***LU***	*MAU*	*MDH*	***MIU***	*ORU*	*SH*	***SU***	***UMU***	***UU***
1996			**X**		X			X							**X**		**X**			**X**	X			**X**	**X**
1997			**X**		X		X	X			**X**				**X**		**X**			**X**	X	X		**X**	**X**
1998	X	**X**	**X**		X		X	X			**X**				**X**		**X**	X		**X**	X	X		**X**	**X**
1999	X	**X**	**X**		X		X	X			**X**		**X**		**X**		**X**	X		**X**	X	X		**X**	**X**
2000	X	**X**	**X**		X		X	X			**X**		**X**		**X**		**X**	X		**X**	X	X		**X**	**X**
2001	X	**X**	**X**		X		X	X			**X**	**X**	**X**	**X**	**X**		**X**	X	X	**X**	X	X		**X**	**X**
2002	X	**X**	**X**		X		X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X		**X**	**X**
2003	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X		**X**	**X**
2004	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2005	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2006	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2007	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2008	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2009	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2010	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2011	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2012	X	**X**	**X**		X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2013	X	**X**	**X**	X	X	X	X	X	X	X	**X**	**X**	**X**	**X**	**X**	**X**	**X**	X	X	**X**	X	X	**X**	**X**	**X**
2014																	**X**				X				

Note: BTH=Blekinge Institute of Technology, CTH=Chalmers University of Technology, GU=Gothenburg University, HB=University of Borås, HDA=Dalarna University, HH=Halmstad University, HJ=Jönköping University, HKR=Kristianstad University, HS=University of Skövde, HV=University West, KAU=Karlstad University, KI=Karolinska Institute, KTH=Royal Institute of Technology, LIU=Linköping University, LNU=Linnaeus University, LTU=Luleå Technical University, MAU=Malmö University, MDH=Mälardalen University, MIU=Mid-Sweden University, ORU=Örebro University, SH=Södertörn University, SU=Stockholm University, UMU=Umeå University, UU=Uppsala University. Universities are marked in boldface. Preferred time period of analysis between the dotted lines.

Table A2: Positional transition probabilities (%)

Position in t-1	*Full professor*			*Associate professor*			*Postdoc*			*Guest researchers*			*Other researchers*			*Total*		
Position in t	*All*	*Movers*	*Stayers*	*All*	*Movers*	*Stayers*	*All*	*Movers*	*Stayers*	*All*	*Movers*	*Stayers*	*All*	*Movers*	*Stayers*	*All*	*Movers*	*Stayers*
Full professor	**93.0**	**87.8**	**93.5**	0.1	0.4	0.0	0.0	0.0	0.0	2.5	6.7	2.2	1.2	1.3	1.2	28.2	24.8	28.5
Associate professor	5.3	8.1	5.0	**93.5**	**88.4**	**94.0**	4.1	7.4	3.6	6.8	20.1	5.9	2.8	6.0	2.4	40.4	39.9	40.4
Postdoc	0.2	0.8	0.1	2.0	3.6	1.9	**75.0**	**69.4**	**75.8**	0.5	1.3	0.5	2.9	4.2	2.8	6.7	9.2	6.4
Guest researchers	0.2	0.9	0.2	1.0	2.1	0.8	0.6	1.2	0.5	**82.3**	**53.0**	**84.4**	1.8	1.6	1.9	3.1	2.6	3.2
Other researchers	1.3	2.4	1.2	3.5	5.4	3.3	20.3	22.0	20.1	7.8	18.8	7.0	**91.3**	**86.9**	**91.7**	21.7	23.5	21.5
Total transitions (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Total transitions (N)	24,509	2,100	22,409	33,103	3,035	30,068	5,692	732	4,960	2236	149	2,087	16,428	1,591	14,837	81,968	7,607	74,361

Note: Numbers in boldface denotes the share of individuals, by category, that do not change position.

Table A3: Average number of publications and citations per year and researcher, by university and university status

University name	University status(University=1/University college=0)	Average numberof publications	Average numberof citations
Karolinska Institute	1	3.05	27.42
Royal Institute of Technology	1	2.88	10.85
Chalmers University of Technology	1	2.81	9.49
Lund University	1	2.33	14.82
Uppsala University	1	2.11	14.08
Linköping University	1	1.95	7.57
Gothenburg University	1	1.58	10.57
Luleå University of Technology	1	1.46	2.91
Blekinge Institute of Technology	0	1.38	1.94
Jönköping University	0	1.36	4.25
Stockholm University	1	1.19	8.54
Umeå University	1	1.12	6.8
Örebro University	1	1.07	4.85
University of Borås	0	1.01	3.18
Mid-Sweden University	1	0.95	2.22
Karlstad University	1	0.9	2.33
Mälardalen University	0	0.88	2.38
Linnaeus University	1	0.81	2.33
Halmstad University	0	0.74	1.45
Malmö University	0	0.71	2.15
Kristianstad University	0	0.68	2.22
University of Skövde	0	0.68	1.23
University College West	0	0.46	0.83
Dalarna University College	0	0.45	0.85
Södertörn University	0	0.37	1.83

Note: University of Borås only observed for one year.

B Construction of our matched sample

As an alternative approach, we use an algorithm (coarsened exact matching, CEM) to match mobile researchers with stayers with similar publication trends and other characteristics at the same career stage before the move. This reduces selection problems, in which scientists who are more productive are also more likely to be mobile, with different prospects for mobility at different stages of their career. An advantage of the matched sample approach is that it removes concerns regarding functional form, such as a nonlinear relationship between the dependent and the independent variables (Imbens, 2004; Moffitt, 2004).

A trade-off exists between finding matches for as many of the treated individuals as possible, which increases the generalizability of our results, against the precision of the match. To obtain high precision in matching, we can increase the number of matching variables, and the exactness of the categories by which we match the individuals, but at the cost of losing matched observations, which reduces generalizability. We use these data to construct our matched sample of 650 mobile researchers and 650 immobile 'twin' researchers implementing the nonparametric matching method coarsened exact matching (CEM) (Blackwell et al., 2009) to create a matched sample, matching one control for each mobile treated individual. In line with previous studies (Azoulay et al., 2009), we match researchers on (a) publication flow in year t-1, (b) cumulative publication stock in t-2, (c) cumulative citation count in t-1, (d) age in t-1, (e) position in t-1, and (f) calendar year in t-1.

Considering that what constitutes normal publication and citation rates varies among disciplines, we coarsen the joint distribution of these variables by discipline (social sciences/humanities, engineering, medicine, or natural sciences), into several strata. The distribution of publication flows is coarsened into four strata (the three bottom quartiles; the 75th to 95th percentiles and above the 95th percentile). The stock of cumulative publications is coarsened into eight strata (0 to 10th percentile; 10th to 25th percentile; 25th to 50th percentile; 50th to 75th percentile; 75th to 90th percentile; 90th to 95th percentile; 95th to 99th percentiles; and above the 99th percentile). Cumulative citations are coarsened into seven strata (the bottom quartile; 25th to 50th percentile; 50th to 75th percentile; 75th to 90th percentile; 90th to 95th percentiles; 95th to 99th percentiles; and above the 99th percentile). The distribution of age is coarsened into quartiles, which captures much of life-cycle patterns and is critical to eliminating differences in pretreatment values over time. We exact match on position and on calendar year, rather than coarsen these variables. Exact matching on calendar year is necessary, since otherwise we run the risk that matched publication rates capture, e.g., trends in publication rates, distorting the interpretation.

Using these matching criteria and with the restriction that a treated individual cannot also be a control individual, we match one-to-one without replacement. In cases of ties, CEM randomly chooses one match. Moreover, it is possible for the same individual to be assigned as a control for several treated. In these cases, we randomly select one matched pair. In the end, we find a control researcher for 650 movers. These movers (treated) and stayers (controls) hold the same position in t-1 and share similar career trajectories, the only (measured) difference being that in year *t* one moves, whereas the other does not. Note that while the matched sample may improve the internal validity, there is a trade-off in terms of external validity.

In Appendix C, we include further details of our matching procedure. Appendix Figure C1 shows the trend of mean publication and citation rates for treated and control researchers. The graph indicates that the matching procedure reduces pretreatment differences between the treated and the control group in terms of outcomes. It also indicates that researchers who move experience a positive effect from mobility on both publication output and citations, although the graph does not account for other confounding characteristics. Appendix Table C1 shows the balancing properties of the matched sample and verifies that pretreatment characteristics are similar for treated and controls. Especially important is that for the outcome variables— publications and citations—both the flow and the cumulative sums are similar. On average, an individual in our matched sample published 1.5-1.6 publications and the publications during that year received 7-8 citations on average within a three-year window.

Unlike the IPTCW method, matched samples can also be combined with a fixed-effect difference-in-differences approach. Still, to the extent that matching does not remove differences in unobservable variation, selection may influence the results. Moreover, by substantially reducing our sample, the matching procedure reduces the generalizability of our results. These reasons make the IPTCW approach preferable, while here we use the matched sample approach as a further robustness analysis.

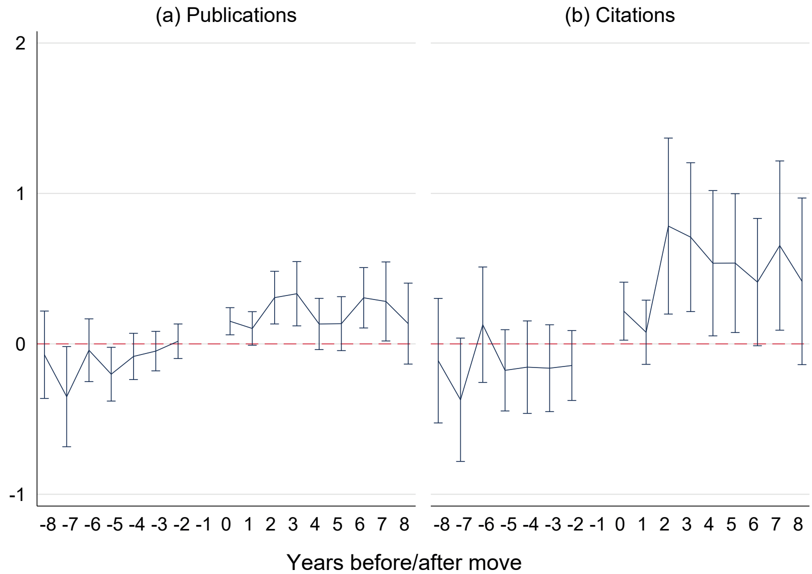
References

Azoulay, P., Graff Zivin, J. S. and Wang, J. (2010). Superstar Extinction, *The Quarterly Journal of Economics, 125(2)*, pp. 549–589.

Blackwell, M., Iacus, S., King, G. and Porro, G. (2009). cem: Coarsened exact matching in Stata, *Stata Journal, 9(4)*, pp. 524.
Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review, *Review of Economics and Statistics, 86(1)*, pp. 4–29.
Moffitt, R. A. (2004). Introduction to the Symposium on the Econometrics of Matching, *Review of Economics and Statistics, 86(1)*, pp. 1–3.

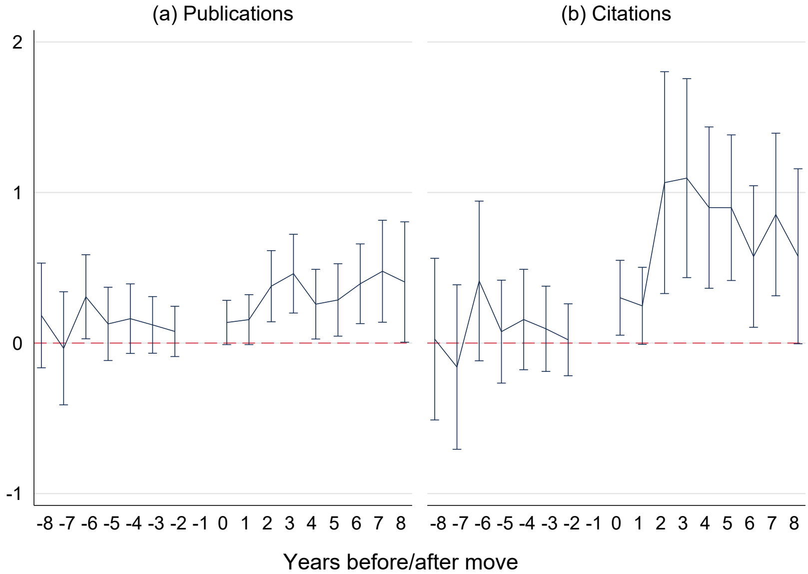
C Matched sample results

Figure C1: Main effect of mobility with leads and lags using fixed effects Poisson, matched sample



Note: This figure plots point estimates for leading and lagging indicators for mobility. The omitted category is one year prior to the move. All specifications include controls for years of experience, number of children, an indicator if an individual is married/cohabiting, discipline dummies, year of birth, indicator of sex and year-fixed effects. Vertical bars correspond to 95% confidence intervals with robust (QML) standard errors clustered at the individual level.

Figure C2: Main effect of mobility with leads and lags using pooled Poisson, matched sample



Note: This figure plots point estimates for leading and lagging indicators for mobility. The omitted category is one year prior to the move. All specifications include controls for years of experience, number of children, an indicator if an individual is married/cohabiting, discipline dummies, year of birth, indicator of sex and year-fixed effects. Vertical bars correspond to 95% confidence intervals with robust (QML) standard errors clustered at the individual level.

Table C1: Descriptive statistics, matched sample in year of match

	Movers (N = 650)					Stayers (N = 650)					
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Diff. in mean
No. of publications	1.52	1	2.47	0	24	1.60	1	3.00	0	46	-0.08
No. of citations	7.40	0	18.30	0	220	8.02	0	22.04	0	260	-0.62
Cum. no. of publications	10.34	4	18.20	0	211	10.46	4	18.40	0	214	-0.11
Cum. no. of citations	53.92	7	128.61	0	1507	54.69	7	133.53	0	1559	-0.76
Age	45.31	44	10.47	23	69	44.86	43	10.07	23	73	0.45
Years since receiving degree	10.50	8	9.21	0	38	9.21	7	7.85	0	37	1.29***
Male	0.61	1	0.49	0	1	0.61	1	0.49	0	1	0.00
Children	0.89	0	1.05	0	4	0.80	0	1.02	0	5	0.08
Married/Cohabitating	0.68	1	0.47	0	1	0.67	1	0.47	0	1	0.01
Full Professors	0.22	0	0.41	0	1	0.22	0	0.41	0	1	0.00
Associate Professors	0.30	0	0.46	0	1	0.30	0	0.46	0	1	0.00
Postdocs	0.04	0	0.19	0	1	0.04	0	0.19	0	1	0.00
Guest researchers	0.36	0	0.48	0	1	0.36	0	0.48	0	1	0.00
‘Other’ researchers	0.08	0	0.28	0	1	0.08	0	0.28	0	1	0.00

Note: Each individual is averaged as one observation in the table. PhD students are excluded. The category ‘other’ research staff mainly consists of PhDs working in academia and doing research but who lack an academic position. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C2: Main effect of mobility, matched sample

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Pooled	Fixed effects	Pooled	Fixed effects
PostMob	0.269***	0.196***	0.729***	0.443***
	(0.0720)	(0.0491)	(0.177)	(0.129)
Observations	8,484	8,483	7,548	7,547
Number of Scientists	1,051	1,051	926	926
Individual FE	No	Yes	No	Yes
IPTCW	No	No	No	No

Note: Baseline category is researchers who never move. All models include time-varying controls; these are years of experience, number of children, an indicator if an individual is married/cohabiting and year-fixed effects. Pooled models also include time-fixed controls; these are discipline dummies, year of birth and indicator of sex. To interpret coefficient as elasticities we take $[\exp(PostMob) - 1]$. Robust (QML) standard errors clustered at the individual level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C3: Effect of career transitions and moving, matched sample

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Fixed effects	Pooled	Fixed effects	Pooled
PostMob & No Prom/Demo	0.242**	0.236***	0.704***	0.489***
	(0.101)	(0.0721)	(0.251)	(0.139)
PostMob & Promotion	0.402***	0.194***	0.528***	0.261**
	(0.114)	(0.0691)	(0.147)	(0.127)
Stay & Promotion	0.213***	0.0962*	0.00592	0.129
	(0.0815)	(0.0579)	(0.135)	(0.115)
PostMob & Demotion	0.438**	0.275	1.165**	0.909*
	(0.198)	(0.172)	(0.542)	(0.477)
Stay & Demotion	0.100	0.101	0.0873	0.503*
	(0.107)	(0.127)	(0.208)	(0.262)
Observations	8,484	8,483	7,548	7,547
Number of scientists	1,051	1,051	926	926
Individual FE	No	Yes	Yes	Yes
IPTCW	No	No	No	No

Note: Baseline category is researchers who stay and are not promoted or demoted. See Table C2 for other notes.

Table C4: Effect of moving and university hierarchies, matched sample

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Fixed effects	Pooled	Fixed effects	Pooled
PostMob & Uni to Uni.	0.339***	0.223***	0.822***	0.515***
	(0.0827)	(0.0597)	(0.191)	0.150)
PostMob & Coll. to Coll.	-0.361	0.168	-1.108***	-0.0710
	(0.256)	(0.147)	(0.294)	(0.284)
PostMob & Coll. to Uni.	0.0263	0.159*	-0.0957	-0.102
	(0.137)	(0.0956)	(0.313)	(0.124)
PostMob & Uni to Coll.	0.104	0.167	0.154	0.267
	(0.208)	(0.104)	(0.464)	(0.265)
Observations	8,484	8,483	7,548	7,547
Number of scientists	1,051	1,051	926	926
Individual FE	No	Yes	No	Yes
IPTCW	No	No	No	No

Note: Baseline category is researchers who never move. See Table C2 for other notes.

Table C5: Effect of mobility within disciplines

	(1)	(2)	(3)	(4)
	Publications		Citations	
	Fixed effects	Pooled	Fixed effects	Pooled
	Panel A. Social science			
PostMob	-0.121	-0.376***	0.0648	-0.793***
	(0.142)	(0.126)	(0.237)	(0.294)
Observations	2,286	2,286	1,742	1,742
Number of scientists	274	274	208	208
	Panel B. Medicine			
PostMob	0.415***	0.207***	0.543***	0.256**
	(0.0938)	(0.0698)	(0.147)	(0.110)
Observations	2,468	2,468	2,413	2,413
Number of scientists	332	332	320	320
	Panel C. Natural sciences			
PostMob	0.260	0.211*	0.158	0.442***
	(0.164)	(0.117)	(0.193)	(0.158)
Observations	1,472	1,472	1,401	1,401
Number of scientists	219	219	205	205
	Panel D. Engineering/Technology			
PostMob	0.325*	0.237**	1.685***	0.550**
	(0.183)	(0.0956)	(0.449)	(0.225)
Observations	2,159	2,159	1,878	1,878
Number of scientists	284	284	242	242
Individual FE	Yes	No	Yes	No
IPTCW	No	No	No	No

Notes: Baseline is researchers within discipline who do not move. See Table C2 for other notes.

D Estimations of IPTC weights

Table D1: Logit estimation of probability of move

	(1) Denominator of treatment weight	(2) Numerator of treatment weight
Publications, t-1	0.0830***	
	(0.0169)	
Citations, t-1	-0.00186	
	(0.00159)	
Cumulative Publications, t-2	-0.00996**	
	(0.00425)	
Cumulative Citations, t-2	-7.29e-05	
	(0.000423)	
Co-authors, t-1	-0.0646**	
	(0.0261)	
Cumulative co-authors, t-2	0.00171	
	(0.00784)	
Age	0.219***	0.212***
	(0.0309)	(0.0304)
Age squared	-0.00263***	-0.00257***
	(0.000332)	(0.000327)
Male	-0.0874	-0.0947
	(0.0680)	(0.0677)
Number of Children	-0.0332	-0.0373
	(0.0391)	(0.0391)
Married	-0.0193	-0.0129
	(0.0801)	(0.0802)
Observations	92,436	92,436
Field FE	Yes	Yes
University FE	Yes	Yes
Position FE	Yes	Yes
Year FE	Yes	Yes

caption*Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table D2: Logit estimation of probability of exit

	(1) Denominator of treatment weight	(2) Numerator of treatment weight
Publications, t-1	-0.110***	
	(0.00959)	
Citations, t-1	0.00178***	
	(0.000499)	
Cumulative Publications, t-2	-0.0187***	
	(0.00199)	
Cumulative Citations, t-2	0.000867***	
	(0.000156)	
Co-authors, t-1	0.000293	
	(0.0131)	
Cumulative co-authors, t-2	-0.0155***	
	(0.00525)	
Age	-0.476***	-0.504***
	(0.00770)	(0.00759)
Age squared	0.00472***	0.00499***
	(8.28e-05)	(8.16e-05)
Male	-0.297***	-0.422***
	(0.0216)	(0.0212)
Number of Children	0.0256*	0.0234*
	(0.0141)	(0.0140)
Married	-0.0775***	-0.111***
	(0.0264)	(0.0263)
Observations	92,403	92,403
Field FE	Yes	Yes
University FE	Yes	Yes
Position FE	Yes	Yes
Year FE	Yes	Yes

caption*Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Chapter III

On the social origins of scientists: How intergenerational (im-)mobility shapes science

Abstract

I use register data on over 1 million Swedish secondary school graduates observed over the years 1986–2012 linked to biological parents to investigate the social background of PhD-graduates. I first establish that family environment is crucial to a future career as a scientist. Of particular importance is whether the parent also holds a PhD. Next, I investigate how the existence of intergenerational spillovers in the decision to obtain a PhD affect career outcomes by contrasting the inventive and publishing performance of PhD graduates who have PhD-educated parents to those without. Results are mediated by gender of the child and parent, suggesting some instances were intergenerational correlations may lead to less scientific output.

Keywords: Intergenerational mobility, scientists, education, parental background, innovation, scientific productivty
JEL Classification: O31, I23, J62

The author wish to thank Patrick Gaule for helpful comments on earlier versions of this paper.

1 Introduction

What motivates someone to pursue a career as a researcher? Earlier work has highlighted how scientists seem to have a *taste for science* and may be willing to forgo higher wages for jobs with more research time (Sauermann & Roach, 2010; Stern, 2004). When are such preferences formed? Is scientific aptitude given from birth or something that is learned growing up? In this paper, I address the question whether and to what extent socio-economic background determines the chances of becoming a scientist and what consequences such associations will have on future career outcomes. Specifically, I study the intergenerational correlation in the decision to choose a PhD-level education and the implications that such spillovers have for careers as inventors and academics.

Anecdotally, many famous scientists are born into *scientific families* in the sense that one or both of their parents worked with research in some capacity. Some prominent examples of scientists born to such scientific families include Henri Becquerel, Niels Bohr, Irene-Joliot Curie, Alexander Graham Bell, Stephen Hawking, Werner Heisenberg, Robert C. Merton, and Max Planck. Although these are all distinguished researchers who made important contributions and discoveries in their respective field, a strong intergenerational correlation in the choice to become a scientist raises the question whether these scientists would have gone on to careers as researchers had their parents not been scientists.

Recent work has shown how early-life expressions of talent or ability (measured, for example, in terms of test scores at a young age, winning mathematics competitions, or scoring high on IQ tests) are important determinants of future productivity of academics and inventors (Agarwal & Gaule, 2018; Aghion et al., 2017; Bell et al., 2018). However, the same line of literature also shows that the socio-economic background of children plays an intricate part in the choice of entering these careers in the first place. One potential explanation for this is that the differences reflect inherited differences or preferences for specific careers across children from different backgrounds. An alternative explanation is that the differences instead mirror differences in childhood environments. In particular, that children from more socio-economic advantageous backgrounds, in terms of parental resources, careers, or education, have access to specific human capital, or role models from parents or environment, which may enhance inherent talent (or compensate for the lack of such talent) and help them pursue specific careers. Such career-specific human capital will raise the expected rate of return to the choice of one's parent's occupation as the child's career compared to the expected return in that same occupation for the son/daughter whose parent was not in that occupation (Lentz & Laband, 1989). In research, scientist parents could pass along valuable human capital to their children, the ultimate effect of which would be to: i) motivate them to become scientists; and ii) better prepare them to be successful in applying for, and complete, a postgraduate education. However, at issue is also whether a strong intergenerational link implies favoritism or nepotism. That is, if career-specific human capital lowers the barriers to enter a scientific career for some groups based, not on their ability (by nature or nurture), but on access to networks and family connections. Since slots on postgraduate educations are fairly fixed, the worry in such a case would be that society is missing out on talented scientist due to them lacking exposure to career-specific human capital or networks.

Considering the importance that research — as a source of ideas and inventions — plays in economic growth (e.g., Aghion & Howitt, 1992; Mokyr, 2002; Romer, 1990), differences in the propensity to pursue careers in science across socio-economic groups could have wide-ranging implications for society. For example, assuming that innate talent is independently distributed, differences in propensity to pursue careers as scientists or inventor due to social factors could imply that society is missing talented individuals who would have made high-impact contributions in these fields. Indeed, some recent evidence support that lowering entry barriers for disfranchised groups into the scientific workforce could have a large impact on the innovative activity in the economy. For instance, Celik (2015) estimates that such *misallocation of talent* might cost the US up to a 10 percent lower rate of innovation and Bell et al. (2018) show how social and demographic characteristics at birth affect the choice to enter a career as an inventor, but do not determine the impact of patents measured by forward citations.

Despite the widespread importance assigned to science and scientists as drivers of economic growth and societal development, studies on what factors determine entry into a career in research remain scarce. Even less is known about whether factors that determine entry into a scientific career also will have an impact on the level and direction of scientific output.

In this paper, I use population-wide register data on over a million Swedish upper-secondary school graduates born between 1971 and 1985 linked to their biological parents to paint a portrait of the social background of people who pursue PhDs. Moreover, I study the implications that the social background has for career outcomes. I leverage the detailed data on children and their parents to control for a wide range of demographic, geographic, school, and family characteristics that otherwise could confound associations.

In the first part of the analysis, I document some facts about the importance of the parent-child link for obtaining a PhD. First, I confirm that it indeed is more likely for a child to graduate with a PhD if at least one of the parents also holds a PhD. For instance, having a PhD-trained father increases the probability to pursue a PhD by more than twice the mean for both boys and girls, which is equivalent to moving from the 5th to the 9th decile in the GPA distribution. This intergenerational correlation remains when controlling for family characteristics, parental income and children's secondary school grades. Furthermore, consistent with a strong role model effect, the boy-father and daughter-mother relationships are of particular importance. Moreover, I find that it is more likely for children with PhD-educated parents to obtain a PhD-level education in the exact same narrowly defined field as their parent compared to PhD graduates whose parents hold a lesser university degree. The results taken together provide suggestive evidence that the observed intergenerational correlations are not driven mainly by inherited differences in ability across children. Instead, the results are indicative of the importance of the childhood environment (c.f., Bell et al., 2018).

Finally, I contrast the exposure to research careers children get from their parents to that from other sources in the childhood environment. Although the results show that exposure to PhD jobs or education is a significant factor explaining the decision to obtain PhD-level education, the magnitudes are small compared to the parental channel. Thus, I conclude that most of the observed intergenerational association is driven by environmental factors *within* the family.

In the second part of the paper, I investigate implications of the parent-child spillovers for inventive and scientific output in the later career. Using patents and scientific publications as additional outcomes, I contrast the performance of PhD graduates who have PhD-educated parents to those without in these areas. The analysis reveals a complex relationship: while PhD-educated men with a PhD-educated parent do not perform differently from other PhDs, I find that PhD-educated women are less likely to patent if the mother is PhD-educated and are less likely to publish if the father is PhD-educated. Thus, depending on the specific relationship, the implications of having a PhD-educated parent may differ widely for the overall allocation of human capital and for scientific and inventive activity.

Related literature.— The paper relates to several lines of literature. First, it adds to the literature on the supply of scientists and inventors (Ehrenberg, 1992; Goolsbee, 1998; Romer, 2001).[1] In regard to this literature, however, this paper mostly relates to a nascent line of the studies focusing on the social background of inventors by linking patent records to historical and current US census data (Akcigit et al., 2017; Bell et al., 2018), as well as using Finnish register-based data (Aghion et al., 2017). A main takeaway from these papers is that — similar to the finding in this paper — inventors tend to come from high-income and highly educated families, although for inventors, parental income seems to be a stronger predictor. The contribution that this paper makes is to make a first effort to investigate the importance of socio-economic factors for obtaining a PhD education and to provide some first evidence on the consequences of such selection for later-life outcomes in patenting and publishing. In contrast to studies on inventors, the focus on the choice to obtain a doctorate allows me to also include the direct effect of mothers' characteristics for daughters' careers in research. This effect is missing in earlier work on inventors either due to few women patenting (Bell et al., 2018) or due to employing conscription data only covering the male population (Aghion et al., 2017).

Furthermore, the paper relates to the literature on how scientists and inventors contribute to overall innovation (e.g., Akcigit et al., 2016; Bloom et al., 2013), in particular, to a line of literature interested in how the allocation of human capital across occupations and firms affects aggregate innovation and growth in the economy (Acemoglu et al., 2018; Celik, 2015; Hsieh et al., 2013; Murphy et al., 1991). Although the paper does not provide direct estimates of the cost of human-capital *misallocation* for the economy, the evidence provided in the paper shows that a strong intergenerational link in the choice to pursue a PhD could result in a lower level of inventive and scientific activity.

A final line of research related to this paper is the literature on intergenerational transmission of human capital. The existence of positive correlations in terms of economic, educational, social and occupational outcomes between parents and children has received considerable attention in recent years.[2] With regards to this literature, this paper primarily relates to the literature on intergenerational correlations in educational attainment and occupations (see e.g., Björklund & Salvanes,

[1] This line of literature has mainly been concerned with the impact of educational, science or tax policy on the choice to enter a researcher career (e.g., Freeman et al., 2009; Freeman & Van Reenen, 2009; Moretti & Wilson, 2017; Toivanen & Väänänen, 2016) and not the role of social background. An exception is Arcidiacono et al. (2016), who study minority graduates in STEM fields from the University of California.

[2] For a survey of the wider literature on intergenerational mobility, see Black & Devereux (2011).

2011; Holmlund et al., 2011; Laband & Lentz, 1983; Lentz & Laband, 1989; Lindquist et al., 2015).

While most literature in this field focuses on general intergenerational transmissions of education in years of schooling, this paper contributes by specifically studying the parent-child correlation in top-level education, as well as careers as researchers. Related is a study by Clark (2014), which finds much stronger long-run intergenerational persistence among elite occupations in Sweden compared to what is found in papers focusing on correlations in years of schooling (see e.g., Holmlund et al., 2011). By focusing on the top of the educational distribution, this paper makes a first foray into investigating non-linearities in the transmission of human capital.[3]

2 Postgraduate education in Sweden

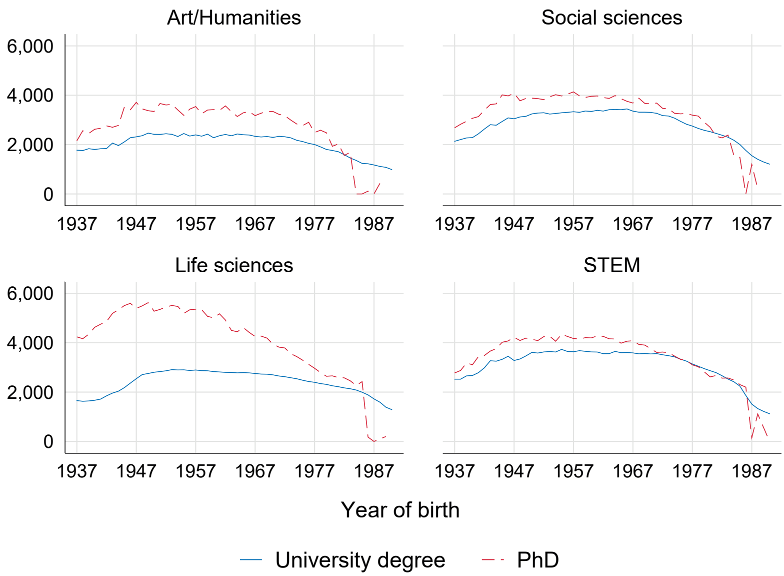
Education in Sweden — up to and including the university level — is essentially free of charge for students, being paid partly by direct government grants and partly by loans. Doctoral students, however, are typically employed by the university and the position is usually financed through government block grants, or by stipends and grants held by senior faculty. Thus, nowadays self-financing of a PhD is very uncommon (Karlsson et al., 2006). Entry into postgraduate PhD-programs is usually competitive and requires at least a three-year undergraduate (bachelor's) degree, but usually a four- or five-year graduate (master's) degree is necessary for admission. A PhD-degree is normally completed after four years of full-time study including writing and publicly defending a doctoral thesis. The thesis is assessed by a three-person committee and graded as either passed or failed (Kyvik & Tvede, 1998). Besides the thesis and course work, postgraduate studies usually involve some teaching and administrative work. However, the extent will vary across university departments.

For most of the post-war era, the aim of Swedish postgraduate education policy has been to increase throughput and reduce the time to graduation. In 1969, Swedish authorities made the first major steps towards reforming postgraduate education in this regard. Modelled on the American PhD-degree, a new doctoral degree was introduced (*Dok-*

[3]There is related literature in the income mobility literature concerned with intergenerational correlations in the top of the income and earnings distributions (e.g., Björklund et al., 2012; Corak & Heisz, 1999; Majlesi et al., 2019). In general, these papers find much higher correlations at the top of the income/wealth distribution compared to studies focused on the general transmission of income and earnings.

torsexamen), replacing the older more extensive doctorate (*Doktorsgrad*) (Andrén, 2013). However, time to completion was still usually long and graduation rates in the 1970s and first half of the 1980s were still quite constant or even slightly decreasing. During the period 1975–1985, only about 800 PhD students graduated each year (Karlsson et al., 2006).

Figure 1: Median disposable income by year of birth, educational level and field in 2012 (in hundreds of SEK)



Note: The figure plots median disposable income in 2012 by year of birth for all people born between 1937 and 1990 subdivided by educational level and field. University degree is a degree with least three years of university education (i.e. bachelor or master's degree). Disposable income is the yearly net-of-tax labor and capital income as well transfers.

From the mid-1980s, the government made further efforts to improve funding for PhD students by introducing a specific employment contract for doctoral students (*Doktorandanställning*), making most doctoral students university employees. Moreover, a shorter doctoral degree, *Licentieexamen,*, which had been abolished in the 1969 reform, was re-introduced. This degree corresponds to two years full-time studies and is most common in engineering disciplines (Kyvik & Tvede, 1998). In subsequent years government efforts to increase the number of PhD graduates continued. In the 1993 research bill, the political slogan became "2,000 doctors a year in 2000" and already by 1998, the number of PhD-graduates had reached over 2,500 graduates per year (Viberg & Bengtsson, 2018). During the same time, average graduation times went down to about four years, although there are still large differences across subjects (Karlsson et al., 2006).

Looking to the labor market outcomes for Swedish PhD graduates, the employment rates are generally high. According to a recent report from the Swedish Higher Education Authority, about 80% of PhD graduates are employed within three years of graduating and almost 90 percent within eight years (Viberg & Bengtsson, 2018). However, there are large differences across fields with PhD graduates in STEM fields having the highest employment rates and PhD graduates in the arts and humanities the lowest. Even so, the employment rates are generally high compared to the average working age population, in which about 75 percent were employed in 1990–2012.

Most PhD graduates in Sweden work in the higher education sector, followed by health and research and development sectors. The most common occupations are university teacher and researcher or medical doctor (Viberg & Bengtsson, 2018). Figure 1 shows the income premium for obtaining a PhD education by plotting a cross-section of the median disposable income for PhD and university graduates in 2012 for birth cohorts 1937 to 1990. Although these are raw figures not accounting for selection, we see that there is a positive income premium for individuals obtaining a PhD education over the life cycle. The premium varies across fields and is especially large for the life sciences. For other fields the premium is less pronounced.

3 Data

Data sources.— For the analysis, I combine several registry databases provided by Statistics Sweden. The basis of the data is *Registerbaserad arbetsmarknadsstatistik*, or RAMS ('Labour statistics based on administrative sources') covering the years 1985–1989 and the *Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier* or LISA ('Longitudinal integration database for health insurance and labor market studies') covering the years 1990–2012. Combined, the RAMS and LISA data comprise rich longitudinal records for the whole working age population (16 and older) of Sweden from 1985 to 2012 registered as residents in Sweden as of December 31 for each year. The data combine various administrative registers and integrate existing data from tax authorities, as well as the labor market, and educational and social sectors. For our purposes, the data contain demographic information (sex, year-of-birth, birth county) and socio-economic data (earnings, as well as the level and field of education). Level and field of education is classified by a three-digit code according to International Standard Classification of Education 1997 (ISCED97)[4] so that fields of education are clearly delineated.

Importantly, the data include personal identifiers, which allows me to merge the data with other registers available from Statistics Sweden. Using the multi-generational registry, which links individuals living in Sweden from 1961 onwards and born after 1932 to their biological parents, I can identify and add information on parents and siblings.

Next, I add information on grade point average (GPA), graduation year and municipality of the school from the secondary schooling registry. Data are available for all upper-secondary school (Swedish, *Gymnasium*) graduates 1973–2012. For graduates from 1997 onwards there is also an indicator for the school and for the educational track. In 1997, the grading system was changed from a 1–5 scale to a 20-point scale. This means that I cannot compare GPA for graduates after this year to those of earlier cohorts. Instead, I assign each graduate to his/her percentile rank based on relative position in the grade distribution of their corresponding graduation cohort.[5]

Sample construction and definitions.— Since a prerequisite for university education (and subsequently a PhD degree) is to finish secondary education, I delimited the sample to those that graduated from secondary school. I focus on children born in 1971–1985 to allow for consistent measurement of completed PhD-education (only 1.52% of PhD-holders graduated before the age of 27 in my sample). These cohorts of graduates account for 1,203,236 who can be matched to at least one of their biological parents. I measure children's highest level of education completed by 2012. I have to drop 58,493 additional observations due to missing information on birth county. For parents, I measure highest level of completed education when the child is 19 years old. I drop 57,598 additional observations due to missing education information on the parents. Moreover, I also measure parents' income (deflated to 1990-levels) when the child is 19 years old. I further drop 24,711 observation due to missing data on parents' income. Finally, I drop 2,314 parent-

[4]See, UNESCO United Nations Educational & Organization (2003).

[5]Note that this assumes that the distribution of grades was not affected by the reform. In further robustness analysis, I checked if the changed grading system affected my analysis. I found that it did not have any impact.

Table 1: Summary statistics

	Mean	Std.Dev.	Min.	Max.
A. Child characteristics:				
Boy	0.505	0.500	0	1
Birth order	1.386	0.610	1	11
Birth year	1977.523	4.312	1971	1985
GPA, percentile	51.763	27.693	1	100
B. Family characteristics:				
No. of bio. siblings	0.772	0.762	0	10
Mother PhD	0.005	0.069	0	1
Father PhD	0.016	0.127	0	1
Mother birth year	1950.036	5.934	1922	1971
Father birth year	1947.406	6.413	1911	1969
Mother annual income (100s of SEK)	1,255.931	793.153	0	44,623
Father annual income (100s of SEK)	1,845.700	1,546.493	0	171,358
C. Child outcomes:				
PhD	0.011	0.106	0	1
Inventor	0.004	0.061	0	1
Publish	0.008	0.088	0	1
Observations	1,060,120			

child combinations where both parents have a PhD since including these in the analysis would prevent me from disentangling the separate effects of fathers and mothers.

Thus, in the final sample I have 1,060,120 individuals linked to their parents ("children") observed 1985–2012. Our main variable of interest is whether the child completed a PhD-level education, operationalized as a 0/1-indicator. In later analysis, I also consider two additional career outcomes. First, I investigate whether children of parents with PhDs are likely to invent, measured as ever filing for a patent to the European Patent Office (EPO). Second, I consider whether a "PhD-child" is likely to contribute to knowledge production, measured as publishing with a Swedish academic affiliation. The patent data were collected in a project matching all inventors in the EPO with a Swedish address with their social security number to be able to match them to administrative registers. The publication data were collected in a separate project to match staff directories of Swedish universities to publication records from the Scopus database. Both databases link to register data and are described extensively in Jung & Ejermo (2014) and Ejermo et al. (2016), respectively.

The rich data allow me to create a host of variables capturing important child and family characteristics. In regressions, I include the following base control variables: dummy for gender (1=boy), indicators for birth year, and gender-birth year interactions; indicator for birth order; indicator for number of siblings; mother's and father's age-at-birth (linear and quadratic terms); 25 dummies for birth county; and, 290 dummies for the municipality of children's secondary school when graduating.

Summary statistics.— Table 1 presents summary statistics for the main outcome variables and covariates. We can note some patterns. First, it is rare to have a "PhD-parent", only about 0.5 percent of mothers and about 1.6 percent of fathers have a PhD-level education. Fathers are, on average, slightly older than mothers and, looking at the annual income, they earn substantially more. Second, looking at Panel C and the outcomes for the children, we see that it has not become more common to earn a PhD over the generations. The average share of PhDs among the children in our sample is about 1.1 percent, corresponding to about 12,721 individuals, which more or less is the same as the average share of mothers and fathers with a PhD-education. The variable *inventor* is an indicator equal to one if the child ever applied for or was granted a patent, and zero otherwise. With only 0.4 percent, or about 4,240 individuals, ever filing for a patent we see that this constitutes an even rarer event. Finally, the variable *publish* is a 0/1-indicator if the child ever published in academic journals with a Swedish university affiliation. This represents a larger proportion of PhDs compared to patenting and corresponds to about 8,480 individuals.

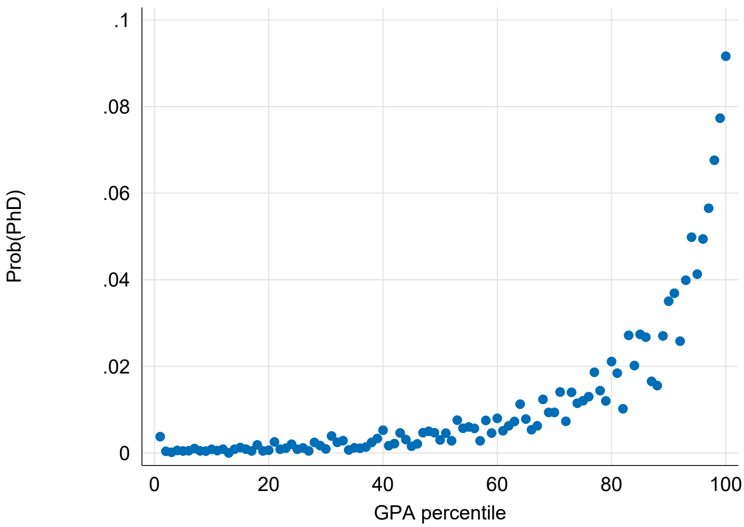
4 Obtaining a PhD

In this section, I study the determinants for obtaining a PhD. First, I start by showing some descriptive statistics on the social background of PhD holders. Second, I establish the importance of parental characteristics for the choice to obtain a PhD-level education in linear probability regressions. After having established the relative importance of parental characteristics, the second goal is to investigate the nature of the associations. Specifically, I focus on how the associations in PhD-education vary across gender and types of educational field of parents and children, as well as compare parental characteristics to other sources of childhood exposure. The objective is to shed some light on possible causal mechanisms to explain the observed associations.

4.1 Descriptive statistics

I begin the analysis of the social background of PhD holders by plotting the estimated probability of obtaining a PhD against the percentile ranks of secondary school GPA (see Figure 2). From the figure, we that the probability of acquiring a PhD has an increasing and convex relationship to students' position in the GPA distribution. In fact, comparing someone in the middle to someone at the very right tail of grade distribution reveals that the latter is almost ten times as likely to graduate with a PhD. We can further note that the figure bares a strong resemblance to the relationship between measures of visuo-spatial IQ and chances of becoming an inventor, found by Aghion et al. (2017).

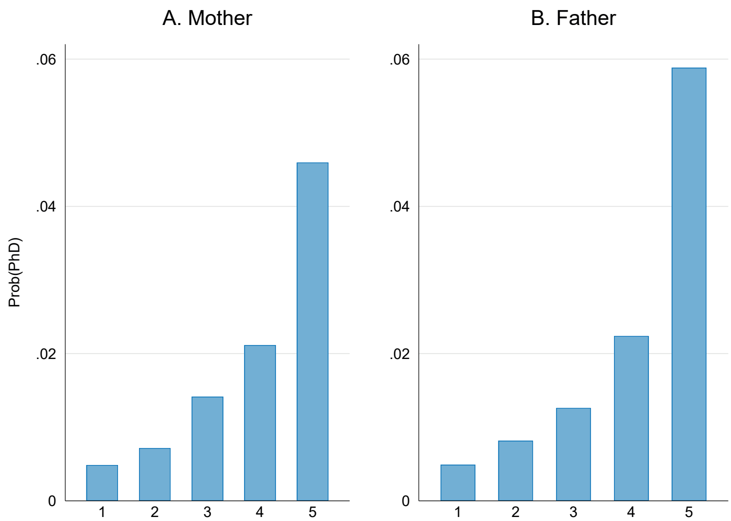
Figure 2: GPA and becoming a PhD



Note: The figure plots the estimated probability of graduating from a PhD-level education conditional on secondary school GPA. GPA percentiles are calculated based on the grade distribution of each seperate cohort 1971–1985.

To make a first effort to investigate an intergenerational association in PhD education, Figure 3 displays histograms with the estimated probability of obtaining a PhD conditioned on highest education level of the parents, separately for mothers and fathers. I measure parents' education level by five indicators: primary (9 years of schooling); secondary schooling; some university education; three years or more of university education; and PhD-level education. The figure makes it clear that having a mother or father with a PhD increases the chances of obtaining a PhD. Compared to having a father with a bachelor or master's degree (three years or more university education), having a father with a PhD increases the probability of obtainiing a PhD by a factor of three.

Figure 3: Parents' education and becoming a PhD

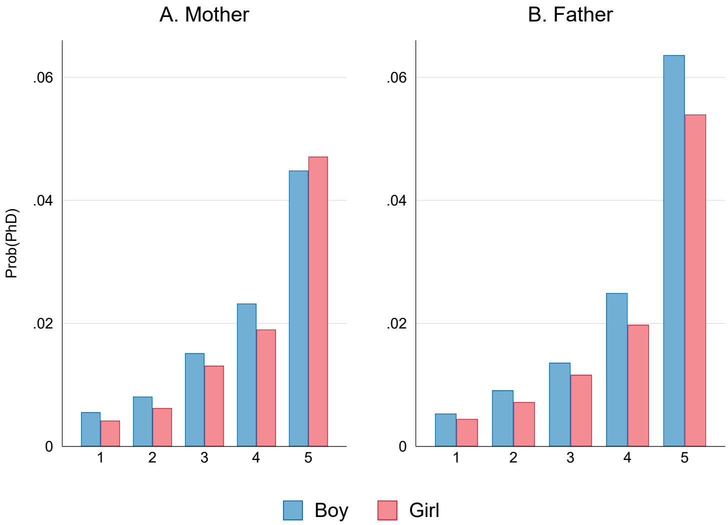


Note: The figure displays the estimated probability of graduating from a PhD-level education conditional on education of (A) mother and (B) father. Parents are divided into five education groups: 1 = primary (9 years of education); 2 = secondary; 3 = less than three years tertiary; 4 = three years or more tertiary; and, 5 = PhD-level education.

To investigate how the associations vary with the gender of child and parent, I decompose the probabilities of obtaining a PhD across mother's and father's education levels separately by the gender of the child. The results are displayed in Figure 4. From the figure, we can see that boys are more likely to obtain a PhD across most of parents' education levels. However, at the PhD-level the association is especially strong for the father-son and mother-daughter correlations. In fact, having a PhD-educated mother more than compensates for the observed gender differences in obtaining a PhD for girls.

Another potential aspect that may matter for the choice of following your parents' PhD education is the field of their education. In Figure 5,

Figure 4: Parental education and becoming a PhD, by gender of child

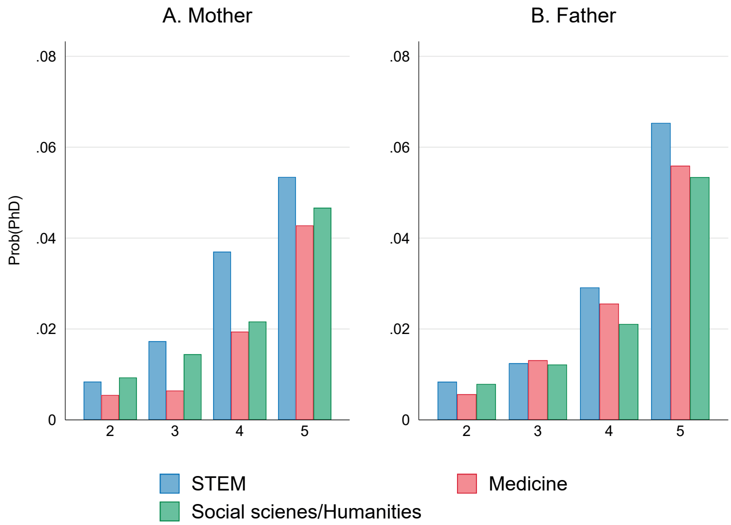


Note: The figure displays the estimated probability of graduating from a PhD-level education conditional on education of (A) mother and (B) father and gender of child. Parents are divided into five education groups: 1 = primary (9 years of education); 2 = secondary; 3 = less than three years tertiary; 4 = three years or more tertiary; and, 5 = PhD-level education.

I introduce indicators for three fields of education based on the 3-digit ISCED classification: science, technology, engineering and mathematics (STEM); medicine; and, social science or arts and humanities. Educational fields that fall outside these categories are assigned to a separate category. Since this "other" category is somewhat diffuse, it is excluded from the figure. Note that specific tracks for primary education do not exist; thus, the primary education level is excluded.

Looking at Figure 5, we see that having a STEM-educated parent is a strong predictor for obtaining a PhD. The correlation is relatively strong if the mother has a PhD-degree in a STEM-field (or more than three years of STEM university education). In Appendix A, Figure A1 shows how having a STEM-educated mother makes it more probable for a child to obtain a PhD compared to mothers with the same educational level but different field regardless of the gender of the child. Thus, the gender differences across children are somewhat attenuated by field-

Figure 5: Parental education and becoming a PhD, by parental level and field of education

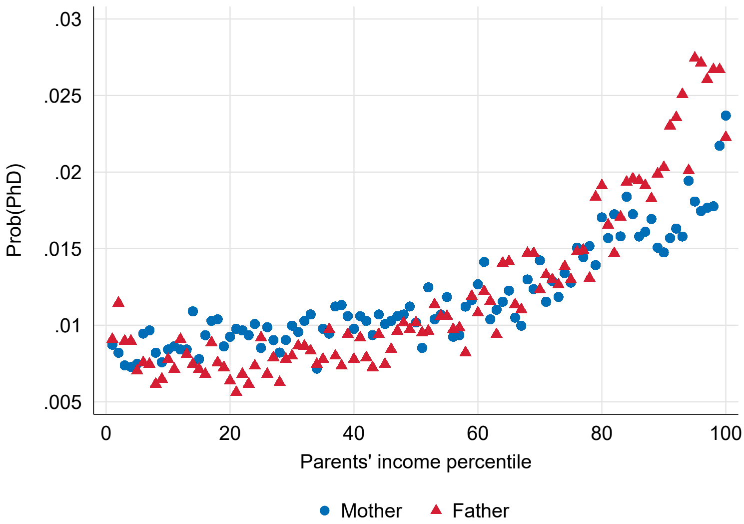


Note: The figure displays the estimated probability of graduating from a PhD-level education conditional on educational level and field of (A) mother and (B) father. Parents are divided into five education groups: 1 = primary (9 years of education); 2 = secondary; 3 = less than three years tertiary; 4 = three years or more tertiary; and, 5 = PhD-level education. Fields of education are STEM, Medicine or Social science/Humanities according to the 1-digit ISCED97 code. Primary education is missing since there is no specific field for education at this level.

specific effects.

Next, I turn to the importance of parental income for obtaining a PhD. Specifically, Figure 6 plots the relationship between mother's and father's income percentile and the probability of obtaining a PhD. We see a positive increasing convex relationship for both mother's and father's income. Holding parents' income percentile fixed, there are no large differences between father's and mother's income in the lower tail of the income distribution, although the effect of mother's income is slightly larger for incomes from about the 20th to the 50th percentile. However, roughly from the 70th percentile and upwards, the effect of the father's income increases at a faster pace relative to that of the mother. Above the 80th percentile, the effect of father's income has outpaced that of the mother's income. Thus, having a high-income parent makes it more probable to pursue a PhD, but the effect is larger for having a high-income father.

Figure 6: Parental income and becoming a PhD



Note: The figures plot the estimated probability of graduating from a PhD-education conditional on mother and father's income percentile. Parental income is measured when the child is 19 years old. Percentile ranks are calculated using residuals from a regression of log income on year-of-birth dummies. Income is deflated to 1990-year's level.

Finally, GPA and parents' education and income all correlate. In Figures A2–A4 in Appendix A, I display correlations between all three variables. The figures reveal the expected pattern: children of high-income and educated parents have above average-GPA and high-income parents are, on average, more educated.

To summarize the discussion above, Figures 2–6 show how the choice of obtaining a PhD is associated with own grades and gender, but also with socio-economic characteristics, such as parental education and income. Furthermore, these relationships are increasing and nonlinear, in the sense that the probability of obtaining a PhD increases at a faster pace in the upper tails of the distributions and that GPA, parental education and income are all correlated.

4.2 Regression specification

In this section, the goal is to disentangle how own and parental characteristics affect the probability of obtaining a PhD-level education. Of particular interest is the correlation of parents' PhD education to the PhD education of their children. To be able to estimate these associations, I specify the following linear probability model:

$$y_i = \alpha + \beta_1 GPA_i + \sum_m \beta_m X_{i,m}^{Mother} + \sum_f \beta_f X_{i,f}^{Father} + \sum_c \beta_c X_{i,c}^{Child} + \epsilon_i \quad (1)$$

where: y_i is a dummy equal to one if individual i is a PhD/invent/publish; GPA_i is a vector including a measure of individual i's GPA; $X_{i,m}^{Mother}$ and $X_{i,f}^{Father}$ are vectors of characteristics pertaining to i's mother and father, respectively; $X_{i,c}^{Child}$ is a vector of control variables specific to individual i; α and βs are parameters to be estimated; and, ϵ_i is an idiosyncratic error term.

Own GPA is introduced via indicators for decile rank with the top decile split into two, the 91st–95th and 96th–100th percentiles, to capture the nonlinear relationship between grades and probability of obtaining a PhD at the top of the grade distribution. The excluded category is the lowest decile. Parents' education is included as an indicator equal to one if the mother or father has completed PhD education and zero otherwise. Moreover, indicators for the 1-digit ISCED field of study is included for all parents regardless of whether the parent has a PhD-level education. General education or primary education is the base category for the field indicators. For parental income, I include indicators for quintile ranks for both mother and father. Again, I split the top decile into two groups, 91st–95th and 96th–100th percentiles. Income deciles are based on the residual from a Mincer-type regression of (log) wage on year-of-birth dummies. I also include base controls for family- and child-level characteristics, as specified above.

4.3 Baseline results

The results from the regressions corresponding to Equation 1 are shown in Table 2 below. Since it is an extensive regression with many covariates, I only include a selection of independent variables of interest to save space. Table A1 in Appendix A shows the corresponding table with all independent variables (base controls excluded).

In Column (1), I regress a dummy for having obtained PhD-level education on controls and own GPA. From Column (1), we can clearly see

Table 2: Determinants of a PhD-education

	(1)	(2)	(3)	(4)	(5)
GPA 91-95	0.0416***	0.0406***	0.0399***	0.0393***	0.0386***
	(0.000850)	(0.000848)	(0.000846)	(0.000845)	(0.000845)
GPA 96-100	0.0704***	0.0691***	0.0678***	0.0669***	0.0660***
	(0.00115)	(0.00115)	(0.00115)	(0.00115)	(0.00115)
Mother PhD		0.0202***		0.0210***	0.0199***
		(0.00297)		(0.00297)	(0.00298)
Mother soc sci./hum.		0.00495***		0.00377***	0.00299***
		(0.000257)		(0.000260)	(0.000263)
Mother medicine		0.00348***		0.00243***	0.00220***
		(0.000244)		(0.000243)	(0.000246)
Mother STEM		0.00661***		0.00550***	0.00508***
		(0.000589)		(0.000588)	(0.000585)
Father PhD			0.0306***	0.0304***	0.0292***
			(0.00180)	(0.00180)	(0.00181)
Father soc sci./hum.			0.00417***	0.00324***	0.00222***
			(0.000319)	(0.000323)	(0.000329)
Father medicine			0.00794***	0.00705***	0.00595***
			(0.000729)	(0.000729)	(0.000720)
Father STEM			0.00418***	0.00368***	0.00307***
			(0.000217)	(0.000218)	(0.000218)
Mother income 91-95					0.00394***
					(0.000643)
Mother income 96-100					0.00242***
					(0.000695)
Father income 91-95					0.00498***
					(0.000740)
Father income 96-100					0.00282***
					(0.000775)
Mean of dependent variable	0.0115	0.0115	0.0115	0.0115	0.0115
SD of dependent variable	0.106	0.106	0.106	0.106	0.106
Observations	1,060,112	1,060,112	1,060,112	1,060,112	1,060,112

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed PhD education (0/1 indicator). All regressions also include controls for parents' age-at-birth (linear and quadratic terms), boy-birth cohort dummies and interactions, as well as indicators for family size, birth order, county-of-birth, and municipality of secondary school (indicator, based on school municipality in the year of graduating).

that being in either of the two top deciles of the grade distribution, 91–95 or 96–100, has positive and significant correlation to PhD education. Compared to the sample mean of about 1.1 percent, the estimates are sizeable. Moving up to the top percentile of the grade distribution has an economically important effect, as students in the top percentile are 7 percentage points more likely to pursue a PhD compared to the average.

Next, Columns (2) and (3) introduce covariates measuring the education of mother and father, respectively. First, we can note that the coefficients for both parents' PhD education are positive and significant, indicating that the intergenerational correlation in PhD-level education holds also when holding grades and controls fixed. The coefficients are sizeable, but only about one-third to one-half the size of having a GPA in the top percentile. We see that parents' education attenuates part of the GPA effect. Second, we can note that the coefficients for parents' field of education are significant and positive, indicating that the educational subject of parents also matters for the choice of pursuing a PhD. Having a mother with a STEM-education increases the probability of obtaining a PhD by almost 0.7 percentage points and having a father with an education in medicine increases the probability by about 0.8 percentage points.

To specifically test for heterogeneity in the effect of having a PhD-educated parent across fields, I interact field indicators with the PhD status of parents. The results are available in Table A2 in Appendix A. The interaction effects are small in size and not significant in any of the specifications across Columns (1) to (3). Thus, I conclude that the effect of parents' PhD status does not differ between fields.

Column (4) shows the result when the education of both parents is included in the regression. Controlling for the other parent's education reduces the size of the PhD coefficients for both the mother and the father, but both are still positive and significant and have economically important magnitudes. Considering both parents' education shows that the father's PhD-education is slightly more important relative to the mother's education, having a father with a PhD-level education makes you about 3 percentage points more likely to obtain a PhD when also considering the mother's education; whereas having a mother with a PhD makes you about 2 percentage points more likely when controlling for father's PhD status. Looking at the field indicators, we see that having a father with a degree in medicine or a mother with a STEM-degree is still the most important field for explaining a PhD-level education.

In Column (5), I include indicators for mother's and father's income. We see that having either the father or the mother belonging to the second-highest or the highest income bracket has a positive and significant association with PhD education. Interestingly, the coefficient for the highest income bracket, 96–100, is smaller than the second highest, 91–95, indicating that the probability of PhD education does not monotonically increase with parents' income, which is in contrast with earlier findings for inventors (Aghion et al., 2017; Bell et al., 2018). Moreover, while father's income is relatively more important in the 91–95 percentile of income, the coefficients for mother's and father's income in top incomes are in parity. By including parents' income, the size of the other covariates is reduced somewhat, but top-GPA is still the most powerful determinant for obtaining a PhD, followed by a PhD-educated mother or father.

Robustness analysis.— The results are robust to splitting the sample by the grade reform in 1997 and running the regression separately in the before and after period (see Table B1 in Appendix B), although the size of coefficients is reduced somewhat in the later period possibly indicating that the role of socio-econnomic background diminishes over time. Alternatively, including a dummy indicating if the individual graduated before or after 1997 does not change the results (available upon request). Moreover, including secondary school- and track-fixed effects (for the period from 1997 onwards when these are available) does not change the results in any qualitative way (see Table B2 in Appendix B). Coefficients on GPA and parents' PhD education remain significant and are similar in size compared to regressions using the same sample without the fixed effects reported in Table B1 Column (2).

To summarize, although the effects of grades, and parental education and income are smaller in the regressions than what Figures 2–6 suggest, they are all important for the choice to obtain a PhD. Among the considered determinants, I find that grades and parents' PhD-level of education to be the most important for explaining PhD-education. The results suggest that an intergenerational association in PhD-education does seem to exist, also when controlling for GPA, parental income and other family and individual characteristics. Concisely, having a mother or father with a PhD makes it 2–3 percentage points more likely to have a PhD-level education, holding other socio-economic factors and grades constant. These estimates are sizeable and have important implications. To illustrate, comparing an individual with a PhD educated father to one without, the latter would have to move from the 5th to the 9th decile of the grade distribution to fully compensate for lacking a PhD-educated father (looking at Table A1).

4.4 Differences across genders

By including a gender dummy, the results presented above conditioned the effects by the gender of the child. However, the descriptive analysis of PhD graduates suggested that the role of parents for PhD education differs by the gender of the child and parent. Since children inherit equal amounts of genes from both parents, if the parent-child association varies in a strong gender-specific manner, this is consistent with the interpretation that the associations are largely driven by differences in childhood environment, such as strong role model effects (Bell et al., 2018). In Table 3, I investigate gender-specific effects by splitting the sample by the sex of the child.

In Columns (1)–(3) of Table 3, the results for the sub-sample of girls are displayed. We see that, holding other covariates constant, there are no large differences between the effects of mother's or father's PhD education on the likelihood of obtaining a PhD. In fact, when including both mother's and father's PhD education, the effect of father's education is slightly larger. Further F-tests of equality of coefficients (not included) reveal that this is indeed the case. Following own GPA and parents' PhD status, having a STEM-educated mother or a parent in the top-income bracket is the most important determinant explaining PhD-education for girls.

Next, Columns (4)–(6) of Table 3 include the results for the sample of boys. For boys, the size of the coefficient of father's PhD status is larger than the coefficient on mother's education, indicating a stronger intergenerational father-son correlation. F-tests of equality of coefficient reject the null hypothesis that the coefficients are the same for both mothers and fathers with $p < .005$. In contrast to girls, we can further note that having a father or mother in the very top income bracket does not increase the probability of pursuing a PhD for boys. Moreover, while having a STEM-educated mother is important for pursuing a PhD for both boys and girls, the most important field for explaining a PhD education for boys is if the father has training in medicine.

Comparing the results for boys and girls, we can note that being in the top-grade percentiles, 91–95 or 96–100, is a stronger predictor for boys getting a PhD-level education compared to girls. Moreover, the mother's PhD education matters more for girls when taking both parents' education into account — see Columns (3) and (6). The results in Table 3 confirm the descriptive results from above — that gender-specific intergenerational correlation in terms of parents' PhD status is more pronounced for boys, but a PhD-educated mother is relatively more important for girls than boys. In other words, a PhD-educated mother will compensate for the lower overall likelihood to pursue a PhD among

Table 3: Determinants of a PhD-education, by gender of child

	Girls			Boys		
	(1)	(2)	(3)	(4)	(5)	(6)
GPA 91-95	0.0286***	0.0282***	0.0274***	0.0534***	0.0525***	0.0518***
	(0.000966)	(0.000966)	(0.000964)	(0.00155)	(0.00155)	(0.00155)
GPA 96-100	0.0483***	0.0475***	0.0462***	0.0955***	0.0937***	0.0927***
	(0.00128)	(0.00127)	(0.00127)	(0.00211)	(0.00211)	(0.00211)
Mother PhD	0.0343***		0.0265***	0.0267***		0.0184***
	(0.00403)		(0.00402)	(0.00385)		(0.00385)
Mother soc sci./hum.	0.00362***		0.00264***	0.00457***		0.00343***
	(0.000351)		(0.000354)	(0.000384)		(0.000388)
Mother medicine	0.00254***		0.00167***	0.00379***		0.00272***
	(0.000329)		(0.000328)	(0.000364)		(0.000364)
Mother STEM	0.00699***		0.00595***	0.00559***		0.00449***
	(0.000825)		(0.000823)	(0.000846)		(0.000845)
Mother income 91-95	0.00600***		0.00498***	0.00495***		0.00361***
	(0.000881)		(0.000881)	(0.000947)		(0.000949)
Mother income 96-100	0.00634***		0.00428***	0.00284***		0.000489
	(0.000978)		(0.000977)	(0.00102)		(0.00102)
Father PhD		0.0324***	0.0292***		0.0321***	0.0299***
		(0.00241)	(0.00241)		(0.00252)	(0.00253)
Father soc sci./hum.		0.00242***	0.00128***		0.00408***	0.00300***
		(0.000438)	(0.000442)		(0.000480)	(0.000485)
Father medicine		0.00484***	0.00354***		0.00861***	0.00744***
		(0.000952)	(0.000951)		(0.00108)	(0.00108)
Father STEM		0.00324***	0.00274***		0.00390***	0.00341***
		(0.000294)	(0.000293)		(0.000319)	(0.000320)
Father income 91-95		0.00590***	0.00502***		0.00598***	0.00540***
		(0.00102)	(0.00102)		(0.00108)	(0.00109)
Father income 96-100		0.00519***	0.00436***		0.00113	0.000582
		(0.00108)	(0.00108)		(0.00113)	(0.00114)
Mean of dependent variable	0.0102	0.0102	0.0102	0.0127	0.0127	0.0127
SD of dependent variable	0.100	0.100	0.100	0.112	0.112	0.112
Observations	525,990	525,990	525,990	536,445	536,445	536,445

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed PhD education (0/1 indicator). All regressions also include controls for parents' age-at-birth (linear and quadratic terms), boy-birth cohort dummies and interactions, as well as indicators for family size, birth order, county-of-birth, and municipality of secondary school (indicator, based on school municipality in the year of graduating).

girls.

Furthermore, we see from the table that parental income matters more for girls compared to boys. Grades are also a stronger predictor that boys will pursue a PhD compared to girls. Considering that it is less likely for girls to obtain a PhD in our sample, this suggest that parental characteristics are relatively more important for girls than boys, making girls more likely to pursue a PhD based on socio-economic background factors.

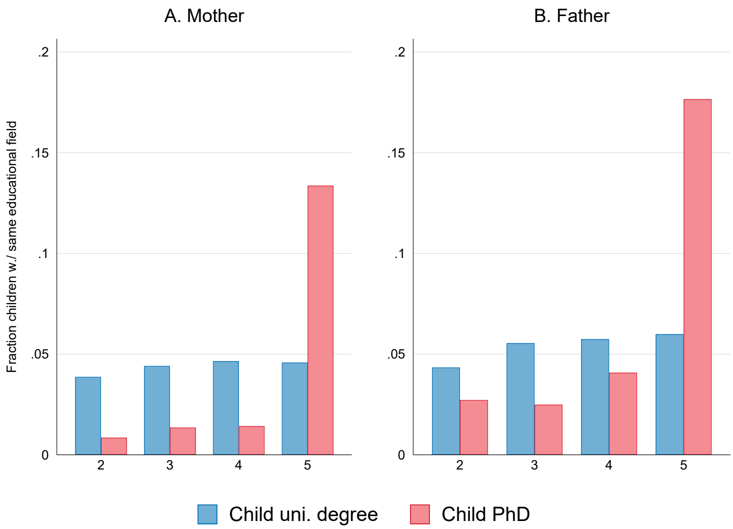
4.5 The choice of educational field for PhD

So far, the analysis has revealed that there seems to be a strong correlation between parents and children both having a PhD-level education. Moreover, these associations differ across children and parents in a gender-specific manner, which may suggest that part of the correlations are explained by the different exposure that girls and boys get from mothers and fathers, respectively. In this subsection, I make further efforts to distinguish between what kinds of exposure children get from their parents. I do this by investigating whether PhD students are likely to share parent's field of education if the parent also holds a PhD. If children are more likely to obtain a PhD in the same narrowly defined discipline as a PhD-educated parent, this might indicate that the PhD-educated parents are able to help or transmit career-specific human capital that helps their children complete an education in their own field. Conversely, if children do PhDs in other disciplines, this might instead suggest that what is transmitted to children is general encouragement to acquire knowledge or preference for research in general.

To begin, Figure 7 plots the proportions of university and PhD graduates who share the same 3-digit ISCED field classification as their parents by their parents' educational level. From the figure, it is more common for PhD-educated children to be educated in the same narrowly defined field as their PhD-educated parents. While around 5 percent of university-educated children with PhD-educated parents have the same exact field of education as their mother or father, about 15 percent of PhD-educated children share the same educational field as their PhD-educated parents. In fact, the proportion of university-educated children who share the same education as their parents is larger for all levels of parents' education, except when parents hold a PhD.

To be able to test the differences in the choice of field, I define a distance metric measuring the distance between children's and parents' education. The metric is calculated in the following way: a distance of "0" is the same 3-digit ISCED code; a distance of "1" is the same 2-digit

Figure 7: Children's and parents' educational field, by educational levels



Note: The figures plot the fraction of university and PhD-educated children with the same educational field as their parents by the parents and children's educational level. Educational field is defined according to 3-digit ISCED97 codes. Parents are divided into five education groups: (1) primary (9 years of education); (2) secondary; (3) less than three years tertiary; (4) three years or more tertiary; and, (5) PhD-level education. Primary education is missing since there is no specific field for education at this level. Child uni. degree is equivalent to three years or more tertiary education.

ISCED code, but not the same 3-digit code; a distance of "2" is the same 1-digit ISCED code, but not the same 2- or 3-level code; and a distance of "2 <" is a different ISCED code altogether.

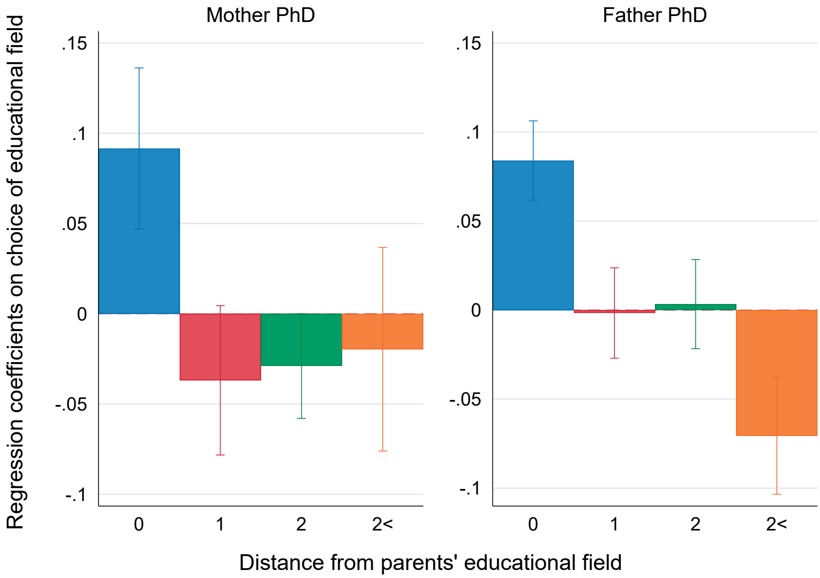
Figure 8 plots coefficients from separate regressions where the outcome in each regression is a dummy equal to one if the distance is equal to "0", "1", "2" or "2 <", respectively, and zero otherwise. I regress the distance dummies on an indicator for parents' PhD status and controls using the sample of all PhD graduates. To facilitate comparison between fields, I also require parents to have at least a three-year university degree. Thus, the regressions answer the question whether PhD students who have PhD-educated parents are more likely to follow their parents' exact, or more or less general, educational orientation compared to other PhD students with less-educated parents.

From Figure 8, it seems clear that what distinguishes PhD holders who have PhD-educated parents is that they are almost 10 percentage points more likely to have the same exact education as their parents compared to other PhD graduates. There is no difference between the groups when distance from parent's education is "1" or "2", while having a father with a PhD makes it less likely to have a PhD in a completely different area. This result suggest that it is transmission or exposure to some specific human or social capital from parents that explains the intergenerational correlations rather than a general interest in research.

The interpretation of field choice as an outcome of environmental exposure relies on there not being any inherited genetic differences across children explaining the choice of educational field. However, a worry might be that inherent abilities make it easier to pursue specific education. For instance, the proclivity for learning mathematics or languages is likely to be, at least partly, inherited. Such inherited differences could affect the choice of pursuing a STEM education as opposed to a education in the social sciences. To test this empirically, I run similar regressions as in Figure 8 with "0" distance but introduce indicators for the 1-digit ISCED field to capture differences in inborn ability for broader fields of education across children. The regression now gives the probability of obtaining a PhD in the same field as parents compared to other PhDs within the same broad field.

Table A3 in Appendix. A presents the results for choice of field. Note that Columns (1) and (2) are identical to the regressions underlying the first bar of Figure 8. Table A3 Columns (4)–(6), introduce the field-fixed effects for the broader fields of education. Neither the size nor the significance of coefficients is affected. Thus, it seems likely that it is parents' PhD education and not inborn proclivity for specific education, that largely explains the choice of field among PhDs with PhD-educated parents.

Figure 8: The effect of having a PhD-educated parent on the choice of educational field of PhD-educated children, by distance from parents' educational field



Note: The figures plots OLS coefficients from separate regressions of a dummy equal to one depending on the distance between children and parents' educational field is 0, 1, 2 or 2 < on an 0/1-indicator of PhD-status of parents. Regressions uses the sample of all PhD-educated children with parents with at least a 3-year university degree. The distance metric is defined into four categories based on 3-digit ISCED97 codes of parents and children. The metric is calculated in the following way: a distance of 0 is the same 3-digit ISCED code; a distance of 1 is the same 2-digit ISCED code, but not the same 3-digit code; a distance of 2 is the same 1-digit ISCED code, but not the same 2- or 3-level code; and a distance of 2 < is a different ISCED code altogether. Regressions include controls for parents' income quintile, 1-digit field of education, age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, and municipality of high school (indicator, based on school municipality in the year of graduating). Vertical bars correspond to 95%-level confidence intervals. Regressions use robust standard errors clustered across siblings.

4.6 Alternative sources of exposure to PhD education and careers

Although the results above are useful in establishing that the exposure children get through parents is important to the choice whether to pursue a PhD education, they likely only represent one potential source of exposure, or they pick up on the environmental factor correlated with a PhD-educated family and with pursuing a PhD. Such correlated factors could be the quality of schools in PhD-dense neighborhoods, growing up close to a university, etc. While we were able to rule out some of the confounding factors in our analysis (for example, differences across schools), other sources of exposure to PhD education or careers, such as the availability of jobs, might confound the analysis. Moreover, replicating the level of exposure one gets through a parent is likely difficult from a policy point of view. Thus, it is interesting to also investigate other sources of exposure.

In this subsection, I investigate how important exposure to PhD-trained individuals in the childhood environment outside the family is for pursuing PhD education. In this part of my analysis, I closely follow that of Bell et al. (2018). To be able to distinguish the exposure one gets from the general environment to that of parents, I turn to the sub-sample of children without a PhD-educated parent. As the general childhood environment, I will be considering the commuting zone containing the municipality of residence of the child at the age of 16 (i.e. the first year an individual shows up in the data). In turn, municipalities are assigned into commuting zones based on the classification of Swedish municipalities into 75 local labor market regions in 2009 (*Lokal Arbetsmarknad*, LA2009) by Statistics Sweden based on commuting patterns. By focusing on childhood commuting zones, I capture the broader sources of environmental exposure outside the family, such as labor markets, neighborhoods or peers.

In Table 4 Column (1), I regress the proportion of children in a commuting zone who get a PhD on the proportion of PhDs in their childhood commuting zone, as well as indicators for birth year to capture differences across cohorts in the propensity to pursue a PhD. The coefficient of 0.42 implies that a one standard deviation increase (0.24 percentage point) in the proportion of PhDs in a commuting zone is associated with an increase of about 0.1 percentage points ($\sim 12\%$) more individuals pursuing a PhD, seemingly a quite sizeable effect. However, considering that the mean share of PhDs in a commuting zone is 0.2 percent a one standard deviation increase is equivalent to a 130 percent increase in the share of PhDs. Besides the establishment of a new university, it is difficult to think of any policy that could have that kind of effect on the number of PhDs in a region.

Table 4: Exposure to PhDs in childhood commuting zone and the fraction of children with a PhD-level education

	(1)	(2)
Fraction PhDs in childhood CZ	0.421***	0.262**
	(0.0828)	(0.109)
Unit of observation	Childhood CZ by birth year	Childhood CZ by birth year by current CZ
Fixed effects	Birth year	Birth year, current CZ
Mean of dependent variable	0.00859	0.0130
SD of dependent variable	0.00813	0.0760
Mean of independent variable	0.00186	0.00225
SD of independent variable	0.00241	0.00295
Observations	1,125	30,972

Note: The table show OLS coefficients from separate regressions of the fraction of children with a PhD-education on the fraction PhDs in his/her childhood commuting zone (CZ). The sample consist of all children without PhD-educated parents. Each child is assigned to a childhood CZ based on the municipality of residence at age 16, the first year the child shows up in the data. CZ are defined according to the local labor market region in 2009 (LA2009) by Statistics Sweden. Robust standard errors clustered at the childhood CZ in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A worry might be that the results in Column (1) of Table 4 are driven by children who remain in the area where they grew up and that the increase in the number of PhDs is driven by the supply of jobs in that region. In that case, children will get PhDs not because of exposure to science or PhD careers, but because that is the kind of jobs available in their home region. To alleviate concerns about the effect being mechanically driven by the supply of jobs, I include fixed effects for the current commuting zone (as of 2012) of children in Column (2) of Table 4. The identifying variation now comes from the children whose current commuting zone differs from their childhood commuting zone. That is, children who at some point after the age of 16 move. Although reduced in size, thee coefficient remains positive and is significant at the 5% level. Thus, while the types of jobs available in the childhood commuting zone drive some of the results, there is still a sizeable environmental effect when including the current commuting zone. Yet the magnitude of the exposure effects a child gets from the general environment is still small in comparison to the one they get through parents.

5 Scientific and inventive productivity and social background

The analysis has so far focused on how parental characteristics or the general environment influences the decision whether to obtain a PhD or not. However, an additional aspect is whether having a PhD-educated parent enhances (or compensates for the lack of) productivity in subsequent careers. In this section, I investigate how the existence of intergenerational spillovers affect inventive and scientific output at the individual level. In particular, I study how the social background of PhD graduates affects patenting and publishing activity. Evaluating differences in later-career outcomes are of importance if we want to gauge if the intergenerational associations give rise to a misallocation of talent and lower levels of innovation or knowledge production overall.

To investigate this matter, I run regressions similar to Equation 1 using patenting (*inventor*) or publishing (*publish*) as dependent variables and focusing on individuals who have obtained a PhD-level education. Thus, the regressions now compare the differences in background characteristics between PhD holders who patent and/or publish to those that do not. Moreover, when considering patenting as an outcome I restrict the sample to PhD holders in STEM and medicine fields since these are the disciplines where patenting can be considered an outcome of their training and work. In fact, PhD holders with STEM and medical training represent about 99 percent of all inventors in the sample of PhD holders. Table 5 displays the results.

In ancillary regressions, I also consider differences in number of patents and publications and their respective societal impact, using the number of citations accrued to those patents and publications as dependent variables. The results for these additional regressions on the *intensive margin* of innovation and publishing are available in Tables C1 and C2 in Appendix C.

Table 5: Probability of becoming an inventor or publishing given parents' education, all PhD graduates

	Inventor (0/1)			Publish (0/1)		
	(1)	(2)	(3)	(4)	(5)	(6)
GPA 91-95	0.0195	0.0204	0.0201	-0.0831	-0.0801	-0.0806
	(0.0374)	(0.0375)	(0.0374)	(0.0623)	(0.0619)	(0.0622)
GPA 96-100	0.0163	0.0170	0.0169	-0.0799	-0.0754	-0.0773
	(0.0371)	(0.0372)	(0.0371)	(0.0620)	(0.0616)	(0.0619)
Mother PhD	-0.0362**		-0.0369**	0.0291		0.0229
	(0.0162)		(0.0163)	(0.0341)		(0.0343)
Mother soc sci./hum.	-0.0000716		-0.000604	0.00223		0.00383
	(0.00881)		(0.00890)	(0.0154)		(0.0156)
Mother medicine	0.00166		0.00172	0.0111		0.0106
	(0.00929)		(0.00933)	(0.0160)		(0.0162)
Mother STEM	0.0174		0.0167	0.0494**		0.0525**
	(0.0135)		(0.0136)	(0.0236)		(0.0237)
Mother income 91-95	0.00550		0.00686	-0.00294		-0.00333
	(0.0127)		(0.0126)	(0.0218)		(0.0219)
Mother income 96-100	0.00383		0.00529	-0.0527**		-0.0528**
	(0.0121)		(0.0123)	(0.0211)		(0.0214)
Father PhD		-0.00451	-0.00598		-0.0361**	-0.0351**
		(0.0106)	(0.0106)		(0.0178)	(0.0179)
Father soc sci./hum.		0.00715	0.00861		-0.0103	-0.00628
		(0.00903)	(0.00911)		(0.0154)	(0.0157)
Father medicine		-0.00122	0.000627		0.0164	0.0213
		(0.0115)	(0.0116)		(0.0206)	(0.0208)
Father STEM		0.0154*	0.0155*		-0.0109	-0.0106

– Table continued on next page –

– Table continued from previous page –

		(0.00837)	(0.00839)		(0.0145)	(0.0146)
Father income 91-95		-0.00201	-0.00227		0.0172	0.0185
		(0.0122)	(0.0122)		(0.0208)	(0.0208)
Father income 96-100		0.00441	0.00433		-0.0305	-0.0291
		(0.0127)	(0.0128)		(0.0209)	(0.0209)
Mean of dependent variable	0.0766	0.0766	0.0766	0.464	0.464	0.464
SD of dependent variable	0.266	0.266	0.266	0.499	0.499	0.499
Observations	9,938	9,938	9,938	12,152	12,152	12,152

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcomes are a 0/1-dummy for inventor or not (column (1)–(3)) and a 0/1-dummy for publishing or not (columns (4)–(6)) (0/1 indicators). Inventor status is equal to one if the individual ever filed and/or was granted a patent. Publish is equal to one if the individual ever published with a Swedish academic affiliation. Regressions uses the sample of all PhD-graduates. Regressions in columns (1)–(3) further restrict the sample to PhDs in STEM and Medicine fields. All regressions include controls for parents' age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Inventing.— Columns (1)–(3) of Table 5 display the results using an indicator for becoming an inventor or not as the dependent variable. When comparing PhDs, GPA does not seem to matter for patenting. However, note that there is little variation in GPA among PhD-educated individuals. Turning to the importance of parental characteristics, we can first note that the only significant covariates explaining *inventing* among PhD holders are mother's PhD status and father's STEM education, the former being negatively associated with becoming an inventor and significant at the 5% level. The size of the coefficient implies that PhD holders with a PhD-educated mother are about 3.6 percentage points less likely to patent compared to other PhD holders. Considering that the mean of the dependent variable is about 7.6 percent of PhD's patenting, the effect is substantial and corresponds to about a 50 percent decline in the probability to patent. The coefficient of father's PhD status is small and insignificant. Considering father's STEM education, the estimate implies that the probability to invent increases by about 1.5 percentage points (~20%); the coefficient is significant at the 10% level.

The patterns reported in Table 5 hold when considering patent count and citations as dependent variables (see Table C1 in Appendix C). Thus, a PhD graduate having a PhD-educated mother is associated with a lower likelihood of patenting and with a lower output as well as impact of that output.

Publishing.— Next, Columns (4)–(5) of Table 5 consider the likelihood of publishing among PhD holders. First, looking at the mother's characteristics, the coefficient on mother's PhD-status is positive but insignificant. However, having a STEM-educated mother makes it more likely to publish. When considering both parents' characteristics in Column (6), having a STEM-educated mother makes ii about 5 percentage points (~ 11%) more likely to publish with an academic affiliation, significant at the 5% level. Moreover, having a mother in the top-income bracket is associated with a decrease in the probability of publishing by about 5.3 percentage points (~ 12%), significant at the 5% level. Turning to the fathers' characteristics, we see that father's PhD status is negatively associated with an academic career and publishing; the coefficient is significant at the 5% level. The size of the coefficient implies that having a PhD-educated father is associated with about 3.5 percentage points (~ 7.5%) lower likelihood of publishing.

Besides the effect of having a mother in the top-income bracket, the findings for publishing and parental characteristics do not extend to the case were I instead use the number of publication or citations as the dependent variables (see Table C2 in Appendix C). Thus, although the *extensive margin* — publishing or not publishing — seem affected by some parental characteristics, neither the *intensive margin* — the quantity of papers published — nor the quality of publications seem affected by the covariates included.

Differences across genders.—To attempt to investigate the underlying mechanisms driving these results, I re-run the regressions in Table 5 split by the gender of the PhDs. Table C3 in Appendix C displays the results.

Columns (1) and (2) of Table C3 display the effect on becoming an inventor subdivided by the gender of the child. We see that women who hold PhDs drive the negative effect of mother's PhD status on patenting reported above. The coefficient is imprecisely measured, significant only at the 10% level, but sizeable in magnitude: having a PhD-educated mother makes girls 2.6 percentage points (~ 62%) less likely to invent. Although, the correlation between men's propensity to patent and mother's PhD education is also negative, it is not significant.

Next, looking at publishing as the outcome in Columns (3) and (4) of Table C3, we see that women who hold PhDs drive the negative effect of father's PhD staatus on the likelihood of publishing. The coefficient is sizeable and significant at the 1% level. The coefficient implies that a PhD-educated woman with a PhD-educated father has about a 7.1 percentage point (~ 16%) lower likelihood of publishing. Moreover, while having a STEM-educated mother increases chances to enter an academic career for girls, a STEM-educated father reduces the propensity to publish. Parental characteristics matter less for boys.

The results in this section have two implications. First, having a PhD-educated parent will have an impact on whether PhDs enter an inventive or publishing career. Second, this effect differs across the genders of child and parent. Taken together with the evidence on the role of PhD-educated parents for the choice of education, this is consistent with children inheriting, or being exposed to, different tastes, preferences or specific human capital for specific educational fields or careers depending on whether the mother or the father holds a PhD. To some extent, such differences likely reflect the type of career or education that the parents have and the role models provided to children. The results on the differing role of PhD-educated parents for patenting and publishing across children's genders speak to this interpretation. Furthermore, the results could explain some differences found in the propensity to invent across genders (Ding et al., 2006; Jung & Ejermo, 2014). For instance, if girls are likely to do PhDs in the same fields as their PhD-educated mothers, but these fields, or in the careers that they pursue after graduating, patent less, this will perpetuate gender differences in inventing across generations.

6 Conclusion

In this paper, I use Swedish register data to investigate the role of GPA and social background for the probability of obtaining a PhD-level education. In particular, the paper studies the implications of having a PhD-educated parent for obtaining a PhD-level education. Additionally, the paper considers how such intergenerational spillovers affect career outcomes in invention and publishing.

The main findings can be summarized as follows. For the determinants considered in my analysis, grades are the most important predictor for obtaining a PhD but the socio-economic background of graduates attenuate the effect of grades. In particular, I find intergenerational correlations in the probability of obtaining a PhD, where children of PhD-educated parents are more likely to obtain a PhD-level education. Having a PhD-trained parent more than doubles the chances of obtaining a PhD and is equivalent in magnitude to moving from the middle of the grade distribution to the very top. These associations show a heterogeneous impact across the gender of child and parent. For boys, the association to father's PhD education is stronger compared to the mother's, indicating a gender-specific component to the father–son association. However, for girls such a component is less pronounced, although having a mother with a PhD is relatively more important for girls and compensates for the lower overall probability of girls seeking a PhD education. Moreover, a PhD student with a PhD-trained parent is more likely to follow the parent's choice of a narrowly defined field compared to a PhD student without a PhD-educated parent. Furthermore, additional analysis shows that while other sources of exposure to PhD education and careers outside the family can be important to the decision to obtain a PhD, the magnitude of such effects is small in comparison to the exposure children get through their family.

In the final part of the paper, I study how the existence of intergenerational spillovers affect inventive and scientific production in terms patenting and publishing. The results suggest that parents' PhD education has a heterogeneous impact on PhD graduates' propensity to invent and publish. While PhD-educated men with a PhD-educated parent do not perform differently from other PhD holders, PhD-educated women are less likely to patent if the mother is PhD educated and are less likely to publish if the father is PhD educated. Parents' influence over the choice of field could be one possible explanations. However, distinguishing whether the effect is due to that women of lower ability is drawn into obtaining a PhD-level education by their PhD-educated parents, that they do not need to perform as well in patenting and publishing activities due to the connection with their PhD educated parents, or that it has to do with what job tasks allotted to women, is beyond the current scope of the paper.

Taken together, the findings in this paper point to the importance of the childhood environment for the choice of entering PhD education and for the choice of subsequent career. Overall, the results are consistent transmission of career-specific human capital, networks, or role model effects as potential mechanisms to explain the associations. Moreover, result indicate that exposure to PhD education and researcher careers in childhood commuting zones has relatively small impact, thus, most of the associations in PhD education seem to be driven by exposure within the family

However, identifying the exact mechanisms to explain the results is beyond the scope of the current paper and would require further study. Indeed, future research should be devoted to investigating the importance of different types of childhood exposure for future careers as scientists. A promising avenue could be using exogenous variation in parental exposure due to, for example, parental deaths (e.g., Adda et al., 2011; Kalil et al., 2016; Lang & Zagorsky, 2001). Such a study would ideally use between-sibling variation in exposure to net out family-specific confounders. Seeing as both PhD education and parental death at a young age are rare events (especially in the case of highly educated parents), such a study would require data covering many cohorts of PhD graduates and their parents.

Furthermore, the analysis has interesting policy implications. In particular, it suggests that parents having a PhD education is a major potential determinant for becoming a scientist. Although it is unlikely that the exposure children get from their parents could be replicated by policy, the analysis suggests alternative policies could leverage large social pay-offs. For example, seeing as women tend to patent less and that the intergenerational spillovers imply that their daughters will also patent less (potentially due to the choice of field or career), providing young girls with female role models in science or simply exposing children to the possibility of careers in science and research has the potential to have large social effects.

In sum, the paper attempts to disentangle the reasons for the supply of scientists and to provide some evidence on the social background of the PhD educated. The findings add to evidence that the social origins of scientists matter even in relatively egalitarian welfare states such as Sweden or Finland (Aghion et al., 2017) where education, up to and including university education, is free or heavily subsidized.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., & Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, *108*(11), 3450–91.

Adda, J., Bjorklund, A., & Holmlund, H. (2011). The role of mothers and fathers in providing skills: Evidence from parental deaths. IZA discussion paper 5425, IZA.

Agarwal, R. & Gaule, P. (2018). Invisible geniuses: Could the knowledge frontier advance faster? IMF Working Paper No. 18/268, International Monetary Fund.

Aghion, P., Akcigit, U., Hyytinen, A., & Toivanen, O. (2017). The social origins of inventors. NBER Working Paper 24110, National Bureau of Economic Research.

Aghion, P. & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2).

Akcigit, U., Celik, M. A., & Greenwood, J. (2016). Buy, keep, or sell: Economic growth and the market for ideas. *Econometrica*, *84*(3), 943–984.

Akcigit, U., Grigsby, J., & Nicholas, T. (2017). The rise of American ingenuity: Innovation and inventors of the golden age. NBER working paper No. 23047, National Bureau of Economic Research.

Andrén, C.-G. (2013). *Visioner, vägval och verkligheter: Svenska universitetens utveckling efter 1940*. Nordic Academic Press.

Arcidiacono, P., Aucejo, E. M., & Hotz, V. J. (2016). University differences in the graduation of minorities in stem fields: Evidence from california. *American Economic Review*, *106*(3), 525–562.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2018). Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*, *134*(2), 647–713.

Björklund, A., Roine, J., & Waldenström, D. (2012). Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? *Journal of Public Economics*, *96*(5-6), 474–484.

Björklund, A. & Salvanes, K. G. (2011). Education and family background: Mechanisms and policies. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education*, volume 3 (pp. 201–247). Elsevier.

Black, S. & Devereux, P. (2011). Recent developments in intergenerational mobility. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (1 ed.)., volume 4 chapter 16, (pp. 1487–1541). Elsevier.

Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, *81*(4), 1347–1393.

Celik, M. A. (2015). Does the cream always rise to the top? The misallocation of talent and innovation. Unpublished, University of Pennsylvania.

Clark, G. (2014). *The son also rises: Surnames and the history of social mobility*. Princeton University Press.

Corak, M. & Heisz, A. (1999). The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data. *Journal of Human Resources*, *34*(3), 504–507.

Ding, W. W., Murray, F., & Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *Science*, *313*(5787), 665–667.

Ehrenberg, R. G. (1992). The flow of new doctorates. *Journal Of Economic Literature*, *30*(2), 830–875.

Ejermo, O., Alder, C., Fassio, C., & Källström, J. (2016). Publications of Academic Researchers In Sweden (PARIS). Technical report, CIRCLE, Lund University. `http://paris.circle.lu.se`.

Freeman, R., Chang, T., & Chiang, H. (2009). Supporting "the best and brightest" in science and engineering: NSF graduate research fellowships. In R. B. Freeman & D. L. Goroff (Eds.), *Science and Engineering Careers in the United States: An Analysis of Markets and Employment* (pp. 19–57). National Bureau of Economic Research, Inc.

Freeman, R. & Van Reenen, J. (2009). What if congress doubled R&D spending on the physical sciences? In J. Lerner & S. Stern (Eds.), *Innovation Policy And The Economy*, volume Volume 9 (pp. 1–38). University Of Chicago Press.

Goolsbee, A. (1998). Does government R&D policy mainly benefit scientists and engineers? *American Economic Review*, *88*(2), 298–302.

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, *49*(3), 615–51.

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2013). The allocation of talent and US economic growth. NBER working paper No. 18693, National Bureau of Economic Research.

Jung, T. & Ejermo, O. (2014). Demographic patterns and trends in patenting: Gender, age, and education of inventors. *Technological Forecasting And Social Change*, *86*, 110–124.

Kalil, A., Mogstad, M., Rege, M., & Votruba, M. E. (2016). Father presence and the intergenerational transmission of educational attainment. *Journal of Human Resources*, *51*(4), 869–899.

Karlsson, S., Hyenstrand, P., & Lundberg, E. (2006). *Forskarutbildningen i Sverige: variation i volym, effektivitet och kostnader sedan tidigt 1990-tal*. Vetenskapsrådet.

Kyvik, S. & Tvede, O. (1998). The doctorate in the Nordic countries. *Comparative Education*, *34*(1), 9–25.

Laband, D. N. & Lentz, B. F. (1983). Like father, like son: Toward an economic theory of occupational following. *Southern Economic Journal*, *50*(2), 474–493.

Lang, K. & Zagorsky, J. L. (2001). Does growing up with a parent absent really hurt? *Journal of Human Resources*, *36*(2), 253–273.

Lentz, B. F. & Laband, D. N. (1989). Why so many children of doctors become doctors: Nepotism vs. human capital transfers. *Journal of Human Resources*, *24*(3), 396–413.

Lindquist, M. J., Sol, J., & Van Praag, M. (2015). Why do entrepreneurial parents have entrepreneurial children? *Journal of Labor Economics*, *33*(2), 269–296.

Majlesi, K., Lundborg, P., Black, S., & Devereux, P. (forthcoming 2019). Poor little rich kids? The role of nature versus nurture in wealth and other economic outcomes and behaviors. *Review of Economic Studies*.

Mokyr, J. (2002). *The Gifts Of Athena: Historical Origins Of The Knowledge Economy*. Princeton University Press.

Moretti, E. & Wilson, D. J. (2017). The effect of state taxes on the geographical location of top earners: Evidence from star scientists. *American Economic Review*, *107*(7), 1858–1903.

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics*, *106*(2), 503–530.

Romer, P. M. (1990). Endogenous technological change. *Journal Of Political Economy*, *98*(5 Pt 2).

Romer, P. M. (2001). Should the government subsidize supply or demand in the market for scientists and engineers? In A. B. Jaffe, J. Lerner, & S. Stern (Eds.), *Innovation Policy And The Economy*, volume 1 (pp. 221–252). MIT Press.

Sauermann, H. & Roach, M. (2010). A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. *Research Policy*, *39*(3), 422–434.

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, *50*(6), 835–853.

Toivanen, O. & Väänänen, L. (2016). Education and invention. *Review Of Economics And Statistics*, *98*(2), 382–396.

UNESCO United Nations Educational, S. & Organization, C. (2003). International standard classification of education, ISCED 1997. *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, 195–220.

Viberg, A. & Bengtsson, A. (2018). *Doktorsexaminerades etablering på arbetsmarknaden*. Swedish Higher Education Authority.

A Supplementary Materials for Section 4

Table A1: Determinants of a PhD-education with all covariates reported

	(1)	(2)	(3)	(4)	(5)
GPA 11-20	0.00202***	0.00196***	0.00199***	0.00195***	0.00190***
	(0.000169)	(0.000169)	(0.000169)	(0.000170)	(0.000170)
GPA 21-30	0.00255***	0.00241***	0.00242***	0.00234***	0.00220***
	(0.000186)	(0.000186)	(0.000186)	(0.000186)	(0.000187)
GPA 31-40	0.00404***	0.00381***	0.00381***	0.00367***	0.00348***
	(0.000216)	(0.000216)	(0.000216)	(0.000216)	(0.000217)
GPA 41-50	0.00491***	0.00461***	0.00460***	0.00442***	0.00421***
	(0.000227)	(0.000227)	(0.000227)	(0.000227)	(0.000228)
GPA 51-60	0.00784***	0.00745***	0.00738***	0.00716***	0.00688***
	(0.000277)	(0.000277)	(0.000276)	(0.000276)	(0.000276)
GPA 61-70	0.0111***	0.0106***	0.0105***	0.0102***	0.00979***
	(0.000327)	(0.000326)	(0.000326)	(0.000326)	(0.000326)
GPA 71-80	0.0165***	0.0158***	0.0156***	0.0152***	0.0148***
	(0.000392)	(0.000390)	(0.000390)	(0.000389)	(0.000389)
GPA 81-90	0.0265***	0.0256***	0.0252***	0.0247***	0.0241***
	(0.000505)	(0.000503)	(0.000503)	(0.000501)	(0.000501)
GPA 91-95	0.0422***	0.0410***	0.0403***	0.0397***	0.0390***
	(0.000854)	(0.000851)	(0.000849)	(0.000848)	(0.000848)
GPA 96-100	0.0714***	0.0696***	0.0684***	0.0674***	0.0665***
	(0.00116)	(0.00115)	(0.00115)	(0.00115)	(0.00115)
Mother PhD		0.0324***		0.0234***	0.0225***
		(0.00280)		(0.00280)	(0.00280)
Mother soc sci./hum.		0.00493***		0.00377***	0.00298***
		(0.000257)		(0.000260)	(0.000264)

– Table continued on next page –

– Table continued from previous page –

Mother medicine	0.00345***		0.00240***	0.00216***
	(0.000244)		(0.000244)	(0.000247)
Mother STEM	0.00687***		0.00569***	0.00527***
	(0.000598)		(0.000597)	(0.000594)
Mother other	0.000776**		0.000417	0.000534*
	(0.000304)		(0.000305)	(0.000305)
Father PhD		0.0339***	0.0311***	0.0299***
		(0.00175)	(0.00175)	(0.00176)
Father soc sci./hum.		0.00412***	0.00319***	0.00215***
		(0.000320)	(0.000324)	(0.000330)
Father medicine		0.00770***	0.00674***	0.00564***
		(0.000736)	(0.000736)	(0.000726)
Father STEM		0.00420***	0.00371***	0.00309***
		(0.000218)	(0.000218)	(0.000218)
Father other		0.00230***	0.00193***	0.00147***
		(0.000337)	(0.000338)	(0.000339)
Mother income 21-40				-0.000677**
				(0.000307)
Mother income 41-60				0.000846***
				(0.000317)
Mother income 61-80				0.00243***
				(0.000340)
Mother income 81-90				0.00391***
				(0.000471)
Mother income 91-95				0.00421***
				(0.000651)

– Table continued on next page –

– Table continued from previous page –

Mother income 96-100					0.00235***
					(0.000713)
Father income 21-40					-0.000261
					(0.000289)
Father income 41-60					0.000834***
					(0.000307)
Father income 61-80					0.00271***
					(0.000342)
Father income 81-90					0.00370***
					(0.000494)
Father income 91-95					0.00521***
					(0.000752)
Father income 96-100					0.00273***
					(0.000784)
Mean of dependent variable	0.0115	0.0115	0.0115	0.0115	0.0115
SD of dependent variable	0.106	0.106	0.106	0.106	0.106
Observations	1,062,435	1,062,435	1,062,435	1,062,435	1,062,435

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed PhD education (0/1 indicator). All regressions also include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Table A2: Field-parent PhD-status interactions

	(1)	(2)	(3)
GPA 91-95	0.0405***	0.0398***	0.0390***
	(0.000849)	(0.000848)	(0.000846)
GPA 96-100	0.0690***	0.0676***	0.0664***
	(0.00114)	(0.00114)	(0.00114)
Mother PhD	0.0345**		0.0243
	(0.0165)		(0.0166)
Mother soc sci./hum.	0.00325***		0.00240***
	(0.000330)		(0.000331)
Mother medicine	0.00238***		0.00167***
	(0.000315)		(0.000314)
Mother STEM	0.00522***		0.00440***
	(0.000607)		(0.000606)
Mother PhD × mo soc sci./hum.	-0.00386		0.000240
	(0.0172)		(0.0172)
Mother PhD × mo medicine	-0.00809		-0.00566
	(0.0170)		(0.0170)
Mother PhD × mo STEM	0.00785		0.00605
	(0.0181)		(0.0182)
Mother income 91-95	0.00546***		0.00417***
	(0.000645)		(0.000646)
Mother income 96-100	0.00471***		0.00238***
	(0.000702)		(0.000705)
Father PhD		0.0318***	0.0285***
		(0.00873)	(0.00875)
Father soc sci./hum.		0.00143***	0.000750*
		(0.000421)	(0.000422)
Father medicine		0.00606***	0.00523***
		(0.000745)	(0.000746)
Father STEM		0.00152***	0.00145***
		(0.000349)	(0.000349)
Father PhD × fa soc sci./hum.		-0.00237	-0.00189
		(0.00935)	(0.00937)
Father PhD × fa medicine		-0.00713	-0.00604
		(0.00920)	(0.00923)
Father PhD × fa STEM		0.00924	0.00933
		(0.00920)	(0.00922)
Father income 91-95		0.00587***	0.00515***
		(0.000741)	(0.000742)

– Table continued on next page –

– Table continued from previous page –

Father income 96-100		0.00372***	0.00303***
		(0.000778)	(0.000779)
Mean of dependent variable	0.0115	0.0115	0.0115
SD of dependent variable	0.106	0.106	0.106
Observations	1,062,435	1,062,435	1,062,435

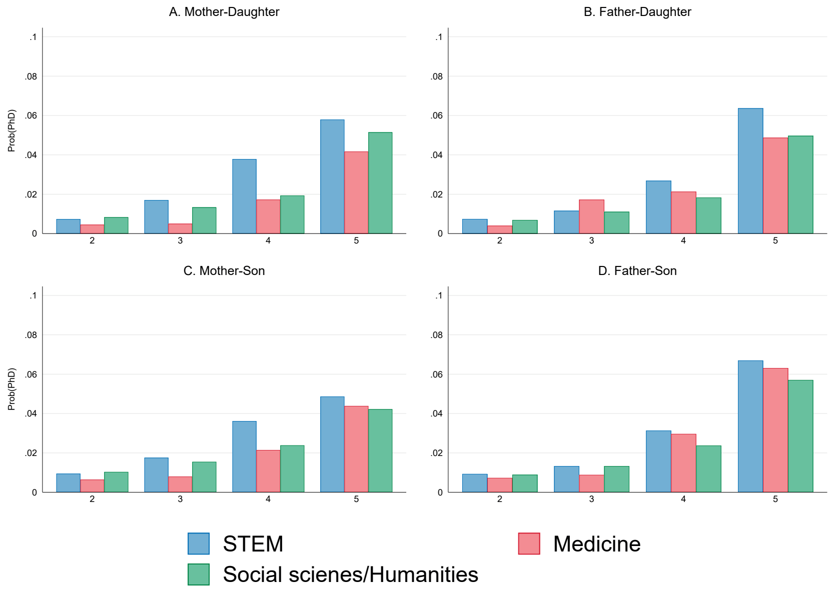
Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed PhD education (0/1 indicator). All regressions also include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Table A3: PhD in same field as parents

	Same field (0/1), mothers	Same field (0/1), fathers	Same field (0/1), either parent	Same field (0/1), mothers	Same field (0/1), fathers	Same field (0/1), either parent
	(1)	(2)	(3)	(4)	(5)	(6)
Mother PhD	0.0911***		0.0769***	0.0913***		0.0773***
	(0.0227)		(0.0242)	(0.0223)		(0.0232)
Father PhD		0.0865***	0.0846***		0.0840***	0.0823***
		(0.0116)	(0.0118)		(0.0112)	(0.0114)
Field-fixed effects	NO	NO	NO	YES	YES	YES
Mean of dependent variable	0.0180	0.0632	0.0614	0.0181	0.0632	0.0614
SD of dependent variable	0.133	0.243	0.240	0.133	0.243	0.240
Observations	7,204	6,206	8,506	7,202	6,203	8,503

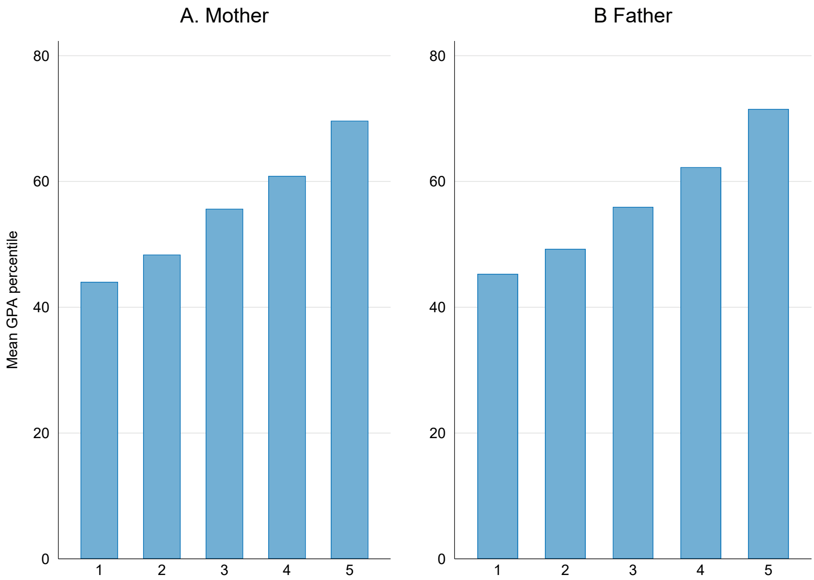
Notes: Table show OLS coefficients from a regression of dummy equal to one if parent and child share the same 3-digit ICSED97 educational code on an indicator equal to one if parent's have a PhD-level education. Sample is all PhD-educated children with at least one parent with university education. Robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include dummies for parents' 1-digit educational field and income percentile, parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Figure A1: Parental education and becoming a PhD, by parental field of education and gender of child



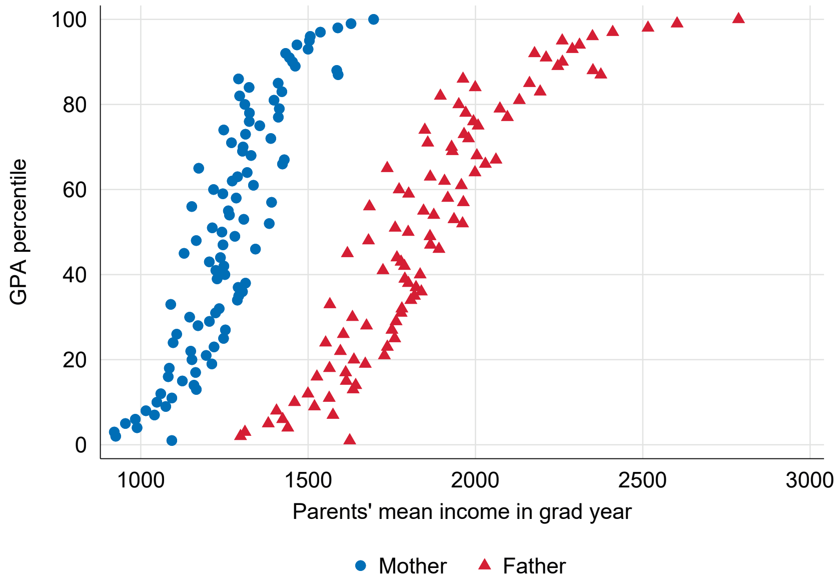
Note: The figure displays the probability of graduating from a PhD-level education conditional on educational level and field of (A) mother and (B) father by gender of child. Parents are divided into five education groups: (1) primary (9 years of education); (2) secondary; (3) less than three years tertiary; (4) three years or more tertiary; and, (5) PhD-level education. Fields of education is STEM, Medicine or Social science/Humanities. Primary education is missing since there is no specific field for education.

Figure A2: Mean GPA and parental education levels



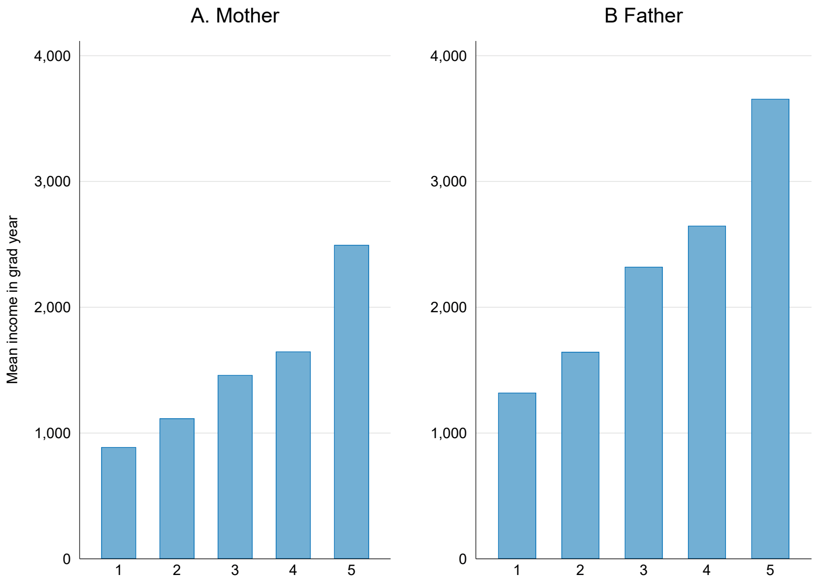
Note: The figure shows mean percentile rank conditional on parents' educational level. Parents are divided into five education groups: 1 = primary (9 years of education); 2 = secondary; 3 = less than three years tertiary; 4 = three years or more tertiary; and, 5 = PhD-level education.

Figure A3: GPA percentile and parental income



Note: The figure shows GPA percentile rank conditional on mother's and father's income. Parents' income is measured in the year the child turn 19. Income is deflated to 1990's level.

Figure A4: Parental mean income and education level



Note: The figure shows mother's and father's mean income by eduction level. Parents are divided into five education groups: 1 = primary (9 years of education); 2 = secondary; 3 = less than three years tertiary; 4 = three years or more tertiary; and, 5 = PhD-level education. Parents' income is measured in the year the child turn 19. Income is deflated to 1990's level.

B Robustness analysis for Section 4

Table B1: Before and after grade change in 1997

	Before 1997 (1)	After 1997 (2)
GPA 91-95	0.0464***	0.0247***
	(0.00127)	(0.00105)
GPA 96-100	0.0840***	0.0350***
	(0.00170)	(0.00133)
Mother PhD	0.0281***	0.0134***
	(0.00551)	(0.00305)
Father PhD	0.0416***	0.0150***
	(0.00305)	(0.00187)
Mother soc sci./hum.	0.00421***	0.000958***
	(0.000435)	(0.000289)
Mother medicine	0.00312***	0.000921***
	(0.000396)	(0.000275)
Mother STEM	0.00603***	0.00349***
	(0.000943)	(0.000664)
Father soc sci./hum.	0.00320***	0.000763**
	(0.000560)	(0.000352)
Father medicine	0.0107***	0.00218***
	(0.00136)	(0.000693)
Father STEM	0.00407***	0.00150***
	(0.000352)	(0.000241)
Mother income 91-95	0.0102***	0.00162***
	(0.00166)	(0.000565)
Mother income 96-100	0.00945***	0.00111*
	(0.00188)	(0.000610)
Father income 91-95	0.00738***	0.00243***
	(0.00133)	(0.000763)
Father income 96-100	0.00815***	-0.000937
	(0.00145)	(0.000757)
Mean of dependent variable	0.0161	0.00657
SD of dependent variable	0.126	0.0808
Observations	544,730	515,382

– Table continued on next page –

– Table continued from previous page –

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed PhD education (0/1 indicator). All regressions also include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Table B2: Determinants of a PhD-education post-1997, adding school- and track-fixed effects

	(1)	(2)	(3)
GPA 91-95	0.0220***	0.0216***	0.0216***
	(0.000993)	(0.000994)	(0.000997)
GPA 96-100	0.0307***	0.0300***	0.0301***
	(0.00130)	(0.00130)	(0.00131)
Mother PhD		0.0114***	0.0117***
		(0.00304)	(0.00304)
Father PhD		0.0124***	0.0129***
		(0.00186)	(0.00187)
Mother soc sci./hum.		0.000326	0.000225
		(0.000286)	(0.000289)
Mother medicine		0.000403	0.000411
		(0.000272)	(0.000274)
Mother STEM		0.00203***	0.00208***
		(0.000665)	(0.000661)
Father soc sci./hum.		0.000292	0.000321
		(0.000349)	(0.000354)
Father medicine		0.000918	0.00126*
		(0.000702)	(0.000694)
Father STEM		0.000331	0.000348
		(0.000240)	(0.000240)
Mother income 91-95			0.000647
			(0.000564)
Mother income 96-100			-0.000196
			(0.000612)
Father income 91-95			0.000560
			(0.000762)
Father income 96-100			-0.00297***
			(0.000766)
Mean of dependent variable	0.00654	0.00654	0.00654
SD of dependent variable	0.0806	0.0806	0.0806
Observations	513,545	513,545	513,545

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed PhD education (0/1 indicator). All regressions include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

C Supplementary Materials for Section 5

Table C1: Patent count and citations regressions, all PhD graduates

	No. of patents			No. of patent citations		
	(1)	(2)	(3)	(4)	(5)	(6)
GPA 91-95	0.146	0.150	0.148	0.159**	0.163**	0.162**
	(0.120)	(0.120)	(0.120)	(0.0732)	(0.0736)	(0.0741)
GAP 96-100	0.208	0.203	0.206	0.211***	0.208***	0.209***
	(0.127)	(0.126)	(0.127)	(0.0799)	(0.0790)	(0.0801)
Mother PhD	-0.247***		-0.237***	-0.187***		-0.184***
	(0.0486)		(0.0510)	(0.0407)		(0.0428)
Mother soc sci./hum.	0.0349		0.0226	0.0238		0.0151
	(0.0564)		(0.0549)	(0.0468)		(0.0461)
Mother medicine	0.0725		0.0627	0.0517		0.0453
	(0.0600)		(0.0593)	(0.0494)		(0.0488)
Mother STEM	0.0954		0.0726	0.0773		0.0615
	(0.0742)		(0.0739)	(0.0644)		(0.0643)
Mother income 91-95	-0.0895		-0.0945	-0.0735		-0.0768
	(0.0712)		(0.0721)	(0.0583)		(0.0588)
Mother income 96-100	0.0131		0.0135	-0.00720		-0.00499
	(0.0819)		(0.0805)	(0.0625)		(0.0620)
Father PhD		0.0357	0.0276		-0.00214	-0.00860
		(0.108)	(0.109)		(0.0756)	(0.0759)
Father soc sci./hum.		0.00975	0.0193		0.00644	0.0139
		(0.0394)	(0.0397)		(0.0341)	(0.0344)

– Continued on next page –

– Table continued from previous page –

Father medicine		0.0370	0.0367		0.0308	0.0324
		(0.0650)	(0.0658)		(0.0532)	(0.0536)
Father STEM		0.131***	0.133***		0.101***	0.103***
		(0.0451)	(0.0447)		(0.0375)	(0.0371)
Father income 91-95		0.0932	0.0930		0.0296	0.0301
		(0.105)	(0.104)		(0.0755)	(0.0751)
Father income 96-100		-0.00716	-0.00694		-0.00271	-0.00245
		(0.0730)	(0.0737)		(0.0631)	(0.0640)
Mean of dependent variable	0.267	0.267	0.267	0.211	0.211	0.211
SD of dependent variable	1.948	1.948	1.948	1.580	1.580	1.580
Observations	9,938	9,938	9,938	9,938	9,938	9,938

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Outcome is completed is number of patents applied for (columns (1)–(3) and forward-citations of those patents measured in 5-year window (columns (4)–(6)). Regressions uses the sample of all PhD-graduates and restricts the sample to PhDs in STEM and Medicine fields. All regressions also include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Table C2: Publication count and citations regressions, all PhD graduates

	No. of publications			No. of citations		
	(1)	(2)	(3)	(4)	(5)	(6)
GPA 91-95	-0.399	-0.329	-0.340	-1.440	-1.822	-1.683
	(1.454)	(1.456)	(1.459)	(9.073)	(8.945)	(9.084)
GPA 96-100	0.838	0.904	0.880	12.87	12.46	12.49
	(1.456)	(1.462)	(1.461)	(9.029)	(8.957)	(9.028)
Mother PhD	0.933		0.947	16.92		17.46
	(1.734)		(1.751)	(22.25)		(22.46)
Mother soc sci./hum.	-0.131		-0.194	-6.278		-6.399
	(0.483)		(0.490)	(8.280)		(7.969)
Mother medicine	-0.114		-0.160	-7.632		-7.746
	(0.474)		(0.481)	(7.932)		(7.695)
Mother STEM	0.0638		0.00756	-5.480		-6.103
	(0.674)		(0.678)	(10.80)		(10.80)
Mother income 91-95	-0.432		-0.469	7.604		7.141
	(0.596)		(0.605)	(11.59)		(11.63)
Mother income 96-100	-0.830*		-0.868*	-1.619		-1.940
	(0.498)		(0.509)	(5.108)		(5.146)
Father PhD		-0.154	-0.102		0.509	1.394
		(0.454)	(0.459)		(5.237)	(5.306)
Father soc sci./hum.		0.0682	0.113		-3.793	-3.511
		(0.392)	(0.402)		(6.055)	(5.771)
Father medicine		-0.255	-0.174		-3.752	-3.026
		(0.434)	(0.421)		(6.239)	(5.693)
Father STEM		0.0956	0.125		-2.709	-2.306
		(0.337)	(0.325)		(5.717)	(5.249)
Father income 91-95		0.204	0.251		1.870	2.112
		(0.579)	(0.593)		(4.911)	(5.002)
Father income 96-100		-0.00848	0.0221		1.705	1.977
		(0.554)	(0.556)		(4.290)	(4.363)
Mean dependent variable	4.854	4.854	4.854	25.78	25.78	25.78
SD dependent variable	13.31	13.31	13.31	149.6	149.6	149.6
Observations	12,148	12,148	12,148	12,148	12,148	12,148

– Continued on next page –

– *Table continued from previous page* –

Notes: Table show OLS coefficients with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is completed is number of publication (columns (1)–(3) and citations of those publications measured in a 3-year window (columns (4)–(6)). Regressions uses the sample of all PhD-graduates. All regressions also include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Table C3: Probability of becoming an inventor or academic given parents' education, by gender of child

	Invent (0/1)		Publish (0/1)	
	Girls (1)	Boys (2)	Girls (3)	Boys (4)
GPA 91-95	-0.0107	0.0350	-0.169*	-0.0118
	(0.0501)	(0.0517)	(0.102)	(0.0770)
GPA 96-100	-0.0204	0.0332	-0.158	-0.0140
	(0.0495)	(0.0515)	(0.102)	(0.0765)
Mother PhD	-0.0263*	-0.0426	0.0142	0.0315
	(0.0142)	(0.0284)	(0.0491)	(0.0489)
Father PhD	0.00339	-0.0106	-0.0772***	0.00163
	(0.0140)	(0.0159)	(0.0270)	(0.0246)
Mother soc sci./hum.	0.00420	-0.00357	0.0123	-0.000210
	(0.0105)	(0.0136)	(0.0234)	(0.0211)
Mother medicine	-0.00171	0.00570	0.00868	0.0140
	(0.0111)	(0.0142)	(0.0247)	(0.0218)
Mother STEM	0.00382	0.0269	0.0714**	0.0379
	(0.0147)	(0.0220)	(0.0350)	(0.0326)
Father soc sci./hum.	-0.00809	0.0218	-0.00757	-0.00222
	(0.0111)	(0.0138)	(0.0237)	(0.0212)
Father medicine	-0.0137	0.0148	0.0168	0.0333
	(0.0141)	(0.0175)	(0.0322)	(0.0280)
Father STEM	-0.000549	0.0276**	-0.0373*	0.0118
	(0.0104)	(0.0127)	(0.0221)	(0.0198)
Mother income 91-95	-0.00728	0.0140	0.0298	-0.0299

– *Continued on next page* –

– Table continued from previous page –

	(0.0135)	(0.0196)	(0.0333)	(0.0293)
Mother income 96-100	0.0178	-0.00810	-0.0498	-0.0526*
	(0.0155)	(0.0183)	(0.0323)	(0.0292)
Father income 91-95	-0.00101	-0.000735	0.0189	0.0234
	(0.0150)	(0.0186)	(0.0316)	(0.0283)
Father income 96-100	0.00230	0.00665	-0.0163	-0.0352
	(0.0149)	(0.0199)	(0.0314)	(0.0288)
Mean of dependent variable	0.0422	0.102	0.442	0.480
SD of dependent variable	0.201	0.303	0.497	0.500
Observations	4,218	5,720	5,338	6,814

Notes: Table show with robust standard errors clustered across siblings in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcomes are a 0/1-dummy for inventor or not, (column (1)–(2)) or a 0/1-dummy for publishing or not (columns (3)–(4)) (0/1 indicators). All regressions also include controls for parents age-at-birth (linear and quadratic terms), boy/birth cohort-interactions, as well as indicators for family size, birth order, twin status, last-born status, county-of-birth, and municipality of high school (indicator, based on school municipality in the year of graduating).

Lund Economic Studies

1. Guy Arvidsson	Bidrag till teorin för verkningarna av räntevariationer, 1962
2. Björn Thalberg	A Trade Cycle Analysis. Extensions of the Goodwin Model, 1966
3. Bengt Höglund	Modell och observationer. En studie av empirisk anknytning och aggregation för en linjär produktionsmodell, 1968
4. Alf Carling	Industrins struktur och konkurrensförhållanden, 1968
5. Tony Hagström	Kreditmarknadens struktur och funktionssätt, 1968
6. Göran Skogh	Straffrätt och samhällsekonomi, 1973
7. Ulf Jakobsson och Göran Norman	Inkomstbeskattningen i den ekonomiska politiken. En kvantitativ analys av systemet för personlig inkomstbeskattning 1952-71, 1974
8. Eskil Wadensjö	Immigration och samhällsekonomi. Immigrationens ekonomiska orsaker och effekter, 1973
9. Rögnvaldur Hannesson	Economics of Fisheries. Some Problems of Efficiency, 1974
10. Charles Stuart	Search and the Organization of Marketplaces, 1975
11. S Enone Metuge	An Input-Output Study of the Structure and Resource Use in the Cameroon Economy, 1976
12. Bengt Jönsson	Cost-Benefit Analysis in Public Health and Medical Care, 1976
13. Agneta Kruse och Ann-Charlotte Ståhlberg	Effekter av ATP - en samhällsekonomisk studie, 1977
14. Krister Hjalte	Sjörestaureringens ekonomi, 1977
15. Lars-Gunnar Svensson	Social Justice and Fair Distributions, 1977
16. Curt Wells	Optimal Fiscal and Monetary Policy - Experiments with an Econometric Model of Sweden, 1978
17. Karl Lidgren	Dryckesförpackningar och miljöpolitik - En studie av styrmedel, 1978
18. Mats Lundahl	Peasants and Poverty. A Study of Haiti, London, 1979
19. Inga Persson-Tanimura	Studier kring arbetsmarknad och information, 1980
20. Bengt Turner	Hyressättning på bostadsmarknaden - Från hyresreglering till bruksvärdesprövning, Stockholm 1979
21. Ingemar Hansson	Market Adjustment and Investment Determination. A Theoretical Analysis of the Firm and the Industry, Stockholm 1981
22. Daniel Boda Ndlela	Dualism in the Rhodesian Colonial Economy, 1981
23. Tom Alberts	Agrarian Reform and Rural Poverty: A Case Study of Peru, 1981
24. Björn Lindgren	Costs of Illness in Sweden 1964-75, 1981

25. Göte Hansson	Social Clauses and International Trade. An Economic Analysis of Labour Standards in Trade Policy, 1981
26. Noman Kanafani	Oil and Development. A Case Study of Iraq, 1982
27. Jan Ekberg	Inkomsteffekter av invandring, 1983
28. Stefan Hedlund	Crisis in Soviet Agriculture?, 1983
29. Ann-Marie Pålsson	Hushållen och kreditpolitiken. En studie av kreditrestriktioners effekt på hushållens konsumtion, sparande och konsumtionsmönster, 1983
30. Lennart Petersson	Svensk utrikeshandel, 1871-1980. En studie i den intraindustriella handelns framväxt, 1984
31. Bengt Assarsson	Inflation and Relative Prices in an Open Economy, 1984
32. Claudio Vedovato	Politics, Foreign Trade and Economic Development in the Dominican Republic, 1985
33. Knut Ödegaard	Cash Crop versus Food Crop Production in Tanzania: An Assessment of the Major Post-Colonial Trends, 1985
34. Vassilios Vlachos	Temporära lönesubventioner. En studie av ett arbetsmarknadspolitiskt medel, 1985
35. Stig Tegle	Part-Time Employment. An Economic Analysis of Weekly Working Hours in Sweden 1963-1982, 1985
36. Peter Stenkula	Tre studier över resursanvändningen i högskolan, 1985
37. Carl Hampus Lyttkens	Swedish Work Environment Policy. An Economic Analysis, 1985
38. Per-Olof Bjuggren	A Transaction Cost Approach to Vertical Integration: The Case of Swedish Pulp and Paper Industry, 1985
39. Jan Petersson	Erik Lindahl och Stockholmsskolans dynamiska metod, 1987
40. Yves Bourdet	International Integration, Market Structure and Prices. A Case Study of the West-European Passenger Car Industry, 1987
41. Krister Andersson and Erik Norrman	Capital Taxation and Neutrality. A study of tax wedges with special reference to Sweden, 1987
42. Tohmas Karlsson	A Macroeconomic Disequilibrium Model. An Econometric Study of the Swedish Business Sector 1970-84, 1987
43. Rosemary Vargas-Lundius	Peasants in Distress. Poverty and Unemployment in the Dominican Republic, 1989
44. Lena Ekelund Axelson	Structural Changes in the Swedish Marketing of Vegetables, 1991
45. Elias Kazarian	Finance and Economic Development: Islamic Banking in Egypt, 1991
46. Anders Danielson	Public Sector Expansion and Economic Development. The Sources and Consequences of Development Finance in Jamaica 1962-84, 1991

47. Johan Torstensson	Factor Endowments, Product Differentiation, and International Trade, 1992
48. Tarmo Haavisto	Money and Economic Activity in Finland, 1866-1985, 1992
49. Ulf Grönkvist	Economic Methodology. Patterns of Reasoning and the Structure of Theories, 1992
50. Evelyne Hangali Maje	Monetization, Financial Development and the Demand for Money, 1992
51. Michael Bergman	Essays on Economic Fluctuations, 1992
52. Flora Mndeme Musonda	Development Strategy and Manufactured Exports in Tanzania, 1992
53. Håkan J. Holm	Complexity in Economic Theory. An Automata Theoretical Approach, 1993
54. Klas Fregert	Wage Contracts, Policy Regimes and Business Cycles. A Contractual History of Sweden 1908-90, 1994
55. Per Frennberg	Essays on Stock Price Behaviour in Sweden, 1994
56. Lisbeth Hellvin	Trade and Specialization in Asia, 1994
57. Sören Höjgård	Long-term Unemployment in a Full Employment Economy, 1994
58. Karolina Ekholm	Multinational Production and Trade in Technological Knowledge, 1995
59. Fredrik Andersson	Essays in the Economics of Asymmetric Information, 1995
60. Rikard Althin	Essays on the Measurement of Producer Performance, 1995
61. Lars Nordén	Empirical Studies of the Market Microstructure on the Swedish Stock Exchange, 1996
62. Kristian Bolin	An Economic Analysis of Marriage and Divorce, 1996
63. Fredrik Sjöholm	R&D, International Spillovers and Productivity Growth, 1997
64. Hossein Asgharian	Essays on Capital Structure, 1997
65. Hans Falck	Aid and Economic Performance - The Case of Tanzania, 1997
66. Bengt Liljas	The Demand for Health and the Contingent Valuation Method, 1997
67. Lars Pålsson Syll	Utility Theory and Structural Analysis, 1997
68. Richard Henricsson	Time Varying Parameters in Exchange Rate Models, 1997
69. Peter Hördahl	Financial Volatility and Time-Varying Risk Premia, 1997
70. Lars Nilsson	Essays on North-South Trade, 1997
71. Fredrik Berggren	Essays on the Demand for Alcohol in Sweden - Review and Applied Demand Studies, 1998
72. Henrik Braconier	Essays on R&D, Technology and Growth, 1998
73. Jerker Lundbäck	Essays on Trade, Growth and Exchange Rates, 1998
74. Dan Anderberg	Essays on Pensions and Information, 1998

75. P. Göran T. Hägg	An Institutional Analysis of Insurance Regulation – The Case of Sweden, 1998
76. Hans-Peter Bermin	Essays on Lookback and Barrier Options - A Malliavin Calculus Approach, 1998
77. Kristian Nilsson	Essays on Exchange Rates, Exports and Growth in Developing Countries, 1998
78. Peter Jochumzen	Essays on Econometric Theory, 1998
79. Lars Behrenz	Essays on the Employment Service and Employers' Recruitment Behaviour, 1998
80. Paul Nystedt	Economic Aspects of Ageing, 1998
81. Rasha M. Torstensson	Empirical Studies in Trade, Integration and Growth, 1999
82. Mattias Ganslandt	Games and Markets - Essays on Communication, Coordination and Multi-Market Competition, 1999
83. Carl-Johan Belfrage	Essays on Interest Groups and Trade Policy, 1999
84. Dan-Olof Rooth	Refugee Immigrants in Sweden - Educational Investments and Labour Market Integration, 1999
85. Karin Olofsdotter	Market Structure and Integration: Essays on Trade, Specialisation and Foreign Direct Investment, 1999
86. Katarina Steen Carlsson	Equality of Access in Health Care, 1999
87. Peter Martinsson	Stated preference methods and empirical analyses of equity in health, 2000
88. Klas Bergenheim	Essays on Pharmaceutical R&D, 2000
89. Hanna Norberg	Empirical Essays on Regional Specialization and Trade in Sweden, 2000
90. Åsa Hansson	Limits of Tax Policy, 2000
91. Hans Byström	Essays on Financial Markets, 2000
92. Henrik Amilon	Essays on Financial Models, 2000
93. Mattias Lundbäck	Asymmetric Information and The Production of Health, 2000
94. Jesper Hansson	Macroeconometric Studies of Private Consumption, Government Debt and Real Exchange Rates, 2001
95. Jonas Månsson	Essays on: Application of Cross Sectional Efficiency Analysis, 2001
96. Mattias Persson	Portfolio Selection and the Analysis of Risk and Time Diversification, 2001
97. Pontus Hansson	Economic Growth and Fiscal Policy, 2002
98. Joakim Gullstrand	Splitting and Measuring Intra-Industry Trade, 2002
99. Birger Nilsson	International Asset Pricing, Diversification and Links between National Stock Markets, 2002
100. Andreas Graflund	Financial Applications of Markov Chain Monte Carlo Methods, 2002

101. Therése Hindman Persson	Economic Analyses of Drinking Water and Sanitation in Developing Countries, 2002
102. Göran Hjelm	Macroeconomic Studies on Fiscal Policy and Real Exchange Rates, 2002
103. Klas Rikner	Sickness Insurance: Design and Behavior, 2002
104. Thomas Ericson	Essays on the Acquisition of Skills in Teams, 2002
105. Thomas Elger	Empirical Studies on the Demand for Monetary Services in the UK, 2002
106. Helena Johansson	International Competition, Productivity and Regional Spillovers, 2003
107. Fredrik Gallo	Explorations in the New Economic Geography, 2003
108. Susanna Thede	Essays on Endogenous Trade Policies, 2003
109. Fredrik CA Andersson	Interest Groups and Government Policy, A Political Economy Analysis, 2003
110. Petter Lundborg	Risky Health Behaviour among Adolescents, 2003
111. Martin W Johansson	Essays on Empirical Macroeconomics, 2003
112. Joakim Ekstrand	Currency Markets - Equilibrium and Expectations, 2003
113. Ingemar Bengtsson	Central bank power: a matter of coordination rather than money supply, 2003
114. Lars Pira	Staples, Institutions and Growth: Competitiveness of Guatemalan Exports 1524-1945, 2003
115. Andreas Bergh	Distributive Justice and the Welfare State, 2003
116. Staffan Waldo	Efficiency in Education - A Multilevel Analysis, 2003
117. Mikael Stenkula	Essays on Network Effects and Money, 2004
118. Catharina Hjortsberg	Health care utilisation in a developing country -the case of Zambia, 2004
119. Henrik Degrér	Empirical Essays on Financial Economics, 2004
120. Mårten Wallette	Temporary Jobs in Sweden: Incidence, Exit, and On-the-Job Training, 2004
121. Tommy Andersson	Essays on Nonlinear Pricing and Welfare, 2004
122. Kristian Sundström	Moral Hazard and Insurance: Optimality, Risk and Preferences, 2004
123. Pär Torstensson	Essays on Bargaining and Social Choice, 2004
124. Frederik Lundtofte	Essays on Incomplete Information in Financial Markets, 2005
125. Kristian Jönsson	Essays on Fiscal Policy, Private Consumption and Non-Stationary Panel Data, 2005
126. Henrik Andersson	Willingness to Pay for a Reduction in Road Mortality Risk: Evidence from Sweden, 2005

127. Björn Ekman	Essays on International Health Economics: The Role of Health Insurance in Health Care Financing in Low- and Middle-Income Countries, 2005
128. Ulf G Erlandsson	Markov Regime Switching in Economic Time Series, 2005
129. Joakim Westerlund	Essays on Panel Cointegration, 2005
130. Lena Hiselius	External costs of transports imposed on neighbours and fellow road users, 2005
131. Ludvig Söderling	Essays on African Growth, Productivity, and Trade, 2005
132. Åsa Eriksson	Testing and Applying Cointegration Analysis in Macroeconomics, 2005
133. Fredrik Hansen	Explorations in Behavioral Economics: Realism, Ontology and Experiments, 2006
134. Fadi Zaher	Evaluating Asset-Pricing Models in International Financial Markets, 2006
135. Christoffer Bengtsson	Applications of Bayesian Econometrics to Financial Economics, 2006
136. Alfredo Schclarek Curutchet	Essays on Fiscal Policy, Public Debt and Financial Development, 2006
137. Fredrik Wilhelmsson	Trade, Competition and Productivity, 2006
138. Ola Jönsson	Option Pricing and Bayesian Learning, 2007
139. Ola Larsson	Essays on Risk in International Financial Markets, 2007
140. Anna Meyer	Studies on the Swedish Parental Insurance, 2007
141. Martin Nordin	Studies in Human Capital, Ability and Migration, 2007
142. Bolor Naranhuu	Studies on Poverty in Mongolia, 2007
143. Margareta Ekbladh	Essays on Sickness Insurance, Absence Certification and Social Norms, 2007
144. Erik Wengström	Communication in Games and Decision Making under Risk, 2007
145. Robin Rander	Essays on Auctions, 2008
146. Ola Andersson	Bargaining and Communication in Games, 2008
147. Marcus Larson	Essays on Realized Volatility and Jumps, 2008
148. Per Hjertstrand	Testing for Rationality, Separability and Efficiency, 2008
149. Fredrik NG Andersson	Wavelet Analysis of Economic Time Series, 2008
150. Sonnie Karlsson	Empirical studies of financial asset returns, 2009
151. Maria Persson	From Trade Preferences to Trade Facilitation, 2009
152. Eric Rehn	Social Insurance, Organization and Hospital Care, 2009
153. Peter Karpestam	Economics of Migration, 2009
154. Marcus Nossman	Essays on Stochastic Volatility, 2009
155. Erik Jonasson	Labor Markets in Transformation: Case Studies of Latin America, 2009

156. Karl Larsson	Analytical Approximation of Contingent Claims, 2009
157. Therese Nilsson	Inequality, Globalization and Health, 2009
158. Rikard Green	Essays on Financial Risks and Derivatives with Applications to Electricity Markets and Credit Markets, 2009
159. Christian Jörgensen	Deepening Integration in the Food Industry – Prices, Productivity and Export, 2010
160. Wolfgang Hess	The Analysis of Duration and Panel Data in Economics, 2010
161. Pernilla Johansson	From debt crisis to debt relief: A study of debt determinants, aid composition and debt relief effectiveness, 2010
162. Nils Janlöv	Measuring Efficiency in the Swedish Health Care Sector, 2010
163. Ai Jun Hou	Essays on Financial Markets Volatility, 2011
164. Alexander Reffgen	Essays on Strategy-proof Social Choice, 2011
165. Johan Blomquist	Testing homogeneity and unit root restrictions in panels, 2012
166. Karin Bergman	The Organization of R&D - Sourcing Strategy, Financing and Relation to Trade, 2012
167. Lu Liu	Essays on Financial Market Interdependence, 2012
168. Bujar Huskaj	Essays on VIX Futures and Options, 2012
169. Åsa Ljungvall	Economic perspectives on the obesity epidemic, 2012
170. Emma Svensson	Experimenting with Focal Points and Monetary Policy, 2012
171. Jens Dietrichson	Designing Public Organizations and Institutions: Essays on Coordination and Incentives, 2013
172. Thomas Eriksson	Empirical Essays on Health and Human Capital, 2013
173. Lina Maria Ellegård	Political Conflicts over Public Policy in Local Governments, 2013
174. Andreas Hatzigeorgiou	Information, Networks and Trust in the Global Economy - Essays on International Trade and Migration, 2013
175. Gustav Kjellsson	Inequality, Health, and Smoking, 2014
176. Richard Desjardins	Rewards to skill supply, skill demand and skill match-mismatch: Studies using the Adult Literacy and Lifeskills survey, 2014
177. Viroj Jienwatcharamongkhol	What Drives Exports? Empirical Evidence at the Firm Level, 2014
178. Anton Nilsson	Health, Skills and Labor Market Success, 2014
179. Albin Erlanson	Essays on Mechanism Design, 2014
180. Daniel Ekeblom	Essays in Empirical Expectations, 2014

181. Sofie Gustafsson	Essays on Human Capital Investments: Pharmaceuticals and Education, 2014
182. Katarzyna Burzynska	Essays on Finance, Networks and Institutions, 2015
183. Mingfa Ding	Corporate Ownership and Liquidity in China's Stock Markets, 2015
184. Anna Andersson	Vertical Trade, 2015
185. Cecilia Hammarlund	Fish and Trips in the Baltic Sea - Prices, Management and Labor Supply, 2015
186. Hilda Ralsmark	Family, Friend, or Foe? Essays in Empirical Microeconomics, 2015
187. Jens Gudmundsson	Making Pairs, 2015
188. Emanuel Alfranseder	Essays on Financial Risks and the Subprime Crisis, 2015
189. Ida Lovén	Education, Health, and Earnings – Type 1 Diabetes in Children and Young Adults, 2015
190. Caren Yinxia Nielsen	Essays on Credit Risk, 2015
191. Usman Khalid	Essays on Institutions and Institutional change, 2016
192. Ross Wilson	Essays in Empirical Institutional Analysis, 2016
193. Milda Norkute	A Factor Analytical Approach to Dynamic Panel Data Models, 2016
194. Valeriia Dzhamalova	Essays on Firms' Financing and Investment Decisions, 2016
195. Claes Ek	Behavioral Spillovers across Prosocial Alternatives, 2016
196. Graeme Cokayne	Networks, Information and Economic Volatility, 2016
197. Björn Thor Arnarson	Exports and Externalities, 2016
198. Veronika Lunina	Multivariate Modelling of Energy Markets, 2017
199. Patrik Karlsson	Essays in Quantitative Finance, 2017
200. Hassan Sabzevari	Essays on systemic risk in European banking, 2017
201. Margaret Samahita	Self-Image and Economic Behavior, 2017
202. Aron Berg	Essays on informational asymmetries in mergers and acquisitions, 2017
203. Simon Reese	Estimation and Testing in Panel Data with Cross-Section Dependence, 2017
204. Karl McShane	Essays on Social Norms and Economic Change, 2017
205. Elvira Andersson	From Cradle to Grave: Empirical Essays on Health and Economic Outcomes, 2017
206. Yana Pryymachenko	Heavy Metal Exposure in Early Life - Health and Labour Market Perspectives, 2017
207. Alemu Tulu Chala	Essays on Banking and Corporate Finance, 2017
208. Jim Ingebretsen Carlson	Essays on economic behavior, focusing and auctions, 2018
209. Jörgen Kratz	Essays on Matching, 2018

210. Anna Welander Tärneberg	Essays on Health in Developing countries, 2018
211. Osmis Areda Habte	Essays on competition and consumer choice, 2018
212. Thomas Hofmarcher	Essays in Empirical Labor Economics, 2019
213. Hjördis Hardardottir	Time and inequality – A study of individual preferences, 2019
214. Erik Grenestam	Essays in Applied Microeconomics, 2019
215. Sara Moricz	Institutions, Inequality and Societal Transformations, 2019
216. John Källström	Mobility in Science, 2019