ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
UNIVERSITY OF GOTHENBURG
254

Choices among Doctors, Students and Primary Care Providers: Empirical Evidence from Sweden

Andrea Berggren

UNIVERSITY OF GOTHENBURG

ISBN 978-91-88199-67-6 (printed)
ISBN 978-91-88199-68-3 (pdf)
ISSN 1651-4289 (printed)
ISSN 1651-4297 (online)

Printed in Sweden,
Stema 2023

SVANENMÄRKET
Trycksak
3041 0234

ACKNOWLEDGEMENTS

The process of writing this thesis has been a long and winding road. First of all, I would like to express my gratitude to my two supervisors: Mikael Lindahl and Gustav Kjellsson. Mikael, thanks for your clear guidance and econometric expertise. Our meetings has always resulted in a clear path forward. Gustav, thanks for always being available and open for discussion. There have been many. I am glad that we could work together and I have learnt a lot from doing so. I have appreciated the moral and academic support that you both have provided in good and bad times. I would also like to thank my final seminar opponents, Ylenia Brilli and Lina Maria Ellegård for great comments and dicussions, that truly improved this thesis. I would also like to acknowledge Jens Dietrichson who has been an excellent co-author of my third chapter.

This journey has also led me to meet many wonderful colleagues and friends. The environment at the Department of Economics has been warm and friendly and offices have always been open for random discussions (of course, until the pandemic literally shut everything down). Thank you to all the people I have had the privilege of getting to know.

I am very grateful for all the support given by Randi Hjalmarsson, both in terms of academic advise and as a director of the PhD-program. An extra big thank you to my fellow PhD-cohort: Lina Andersson, Jakob Enlund, Louise Jeppsson, Lisa Norrgren and Ruijie Tian. I will never forget the long and dark hours we spent toghether during the first years of the program. And to Louise - thanks for being a superb coauthor. I am so happy that my first paper was written with you. This paper also benefited greatly from a mentorship with Mårten Palme, Elin Molin and Paula Rooth.

The thesis has really improved from comments and discussions from seminars at the Department of Economics at the University of Gothenburg. While there too many people to list here, I would especially like to mention Anna Bindler, Andreea Mitrut, Eva Ranehill, and Joseph Vecci. This journey has also required support from the administrative staff, thank you Maria Siirak, Ann-Christin Räätäri

Nyström and Katarina Forsberg. I am also grateful to Hans Bjurek and Pelle Ahlerup for guidance and support regarding the development of my pedagogical skills.

I gratefully acknowledge financial support from Adlerbertska Stipendiestiftelsen, Donationsnämnden and Stiftelsen Kjellbergska Flickskolans Donationer.

Finally, I would like to thank my friends and family for always being supportive and helpful during this, sometimes very stressful, journey. Thank you to my parents, Anne and Ulf Berggren, and to my extended family: Lila Alborz and Eva Ersdal. This thesis would not have been written without your love, support and dedication to ensure the finalization of this project. Above everything, Sharam Alborz, you have been my rock throughout this time. You, Lovis and Eira are my world and I dedicate this thesis to you.

Billdal, January 2023
Andrea Berggren

Contents

Introduction

Microeconomics is fundamentally the study of individual choices under scarce resources. This dissertation is comprised of three self-contained chapters which each studies the consequences of individual choices in two domains: education and primary care. Choices can be made either directly, for example through the choice of coursework in high school in chapter two or choice of primary care provider in chapter three. Choices can also be made indirectly, via a physician's choice to prescribe antibiotics to a child which is the focus of the first chapter.

The three chapters share an empirical approach which applies econometric methods explicitly targeted to pinpoint causal effects. Identifying causal effects is of extra importance from a policy-perspective as understanding the full impact of a policy requires separating causal effects from changes driven by correlated covariates. Another common theme in this dissertation is the use of large micro-level data sets from Sweden. The results presented in this thesis provide important evidence from a policy-perspective which could be helpful to understand how public policy could exacerbate or ameliorate social inequality within health and education, in particular among adolescents and children.

In Chapter One, I investigate the consequences of a physician's decision to prescribe antibiotics on childrens' health care utilization. Efforts to combat the rising problem with antimicrobial resistance have led governments to impose restrictions on antibiotics use globally. While it has been shown that increased prudence decrease antibiotics prescribing, see for example Oliveira et al. (2020) for a systematic overview, less is known about the individual consequences of the restricted access to antibiotics. Studying the causal consequences of antibiotics consumption in primary care is challenging since individuals choose their health care provider which likely introduces a bias if this choice is correlated with prescribing practices. To overcome this limitation, I leverage detailed register-data from one region in Sweden, Scania (*Skåne*), which contains physician identifiers. Using this information, I construct a measure of each physicians propensity to prescribe antibiotics at a primary care visit when the child is aged 0-5 years.

I show that there is a large variation in physicians´ propensity to prescribe antibiotics, and that this measure strongly predicts the probability of actually obtaining an antibiotics prescription. Children are conditionally randomly assigned to physicians since the measure is unrelated to predetermined background characteristics, which allows for a causal interpretation. The findings show that being prescribed antibiotics at an index visit leads to an increased probability of more interactions with the health care system in the short term. I estimate a precise increase in the probability of having at least one additional visit within 30 and 90 days, to both in- and out-of-hours primary care centers, but no effect on either outpatient emergency visits or hospitalizations. Going from the lowest- to the highest-prescribing physician leads to an increase in the probability of having a revisit within 30 days of 3.3 percentage points, or 14.3% of the mean, a revisit to in-hours and out-of-hours PCC by 1.6 (9.5% of the mean) and 0.9 (47% of the mean) percentage points, respectively.

The results are robust to an extensive number of specification checks. I further corroborate the results by restricting the visits to those with a diagnosis of respiratory tract infection, a common condition in children, and show that the effects in this subsample are remarkably similar to those found for the full sample of visits.

Overall, my results show that being prescribed antibiotics causes an increase in health care utilization in both the short and medium term. The short-term increase in visits should be taken into account by policy-makers, as the decision whether to prescribe antibiotics affects the work burden in an already capacity-constrained sector.

In Chapter Two, The Impact of Upper Secondary School Flexibility on Sorting and Educational Outcomes, co-authored with Louise Jeppsson, we estimate the causal impact of an upper secondary curriculum reform in Sweden. The reform increased students' course-taking flexibility and was implemented in year 2000. In the most popular upper secondary program, it led to a significant decrease in mandatory mathematics requirements. Using administrative Swedish data, we estimate the causal impact of the reform on tertiary education outcomes and expected earnings using a differences-in-discontinuity identification strategy. The method compares students born immediately before and after a cutoff date which dictates whether the students were exposed to the reform or not. Since the reform was implemented in year 2000, cohorts born in 1983 started school in the old curriculum and those born in 1984 under the new, reformed, curriculum. For the

main part of the analysis, we compare students born in October-December 1983 to students born in January-March 1984. To disentangle the school starting age effect from the unconfounded effect of the reform we subtract similar comparisons between students born in neighboring non-reform cutoff years.

We find a positive effect of the reform on students' probability of ever enrolling in tertiary education, an increase of 3 percent. The positive impact on Social Science students' enrollment in tertiary education translates into an increase in the probability of students exiting tertiary education with a degree. Estimating the effect by gender shows that the positive impact on the probability of earning a degree was driven by a large and positive impact for females. Interestingly, we find a marginally significant positive effect for women and no impact for men on the probability of having the highest degree in a relatively mathematics-intensive field. The reform does not affect the speed of students entering into tertiary education after graduating from upper secondary school, on average. However, the average outcome masks the distributional effects of the reform.

The heterogeneity analysis reveals that relatively disadvantaged students (measured along a socio-economic status index) were not negatively affected by the curriculum reform. Rather, students in the lowest SES quartile seem to have benefited the most from the more flexible curriculum and have a large increase of 19 percent in the probability of entering a mathematics intensive program. On the other hand, the most advantaged students had a reduced probability of attending the same program as well as a lower speed to enter tertiary education. To the extent that majors in Business and Economics give relatively higher earnings, this group were harmed by the reform.

Our results are informative for policy makers speculating about the optimal level of flexibility and mathematics content. Increasing flexibility had a positive impact on academic outcomes. The decline in mathematics attainment lead relatively more disadvantaged students in particular to choose more advanced programs than their peers. In particular, the most advantaged students were negatively affected by the reform in terms of chosen programs in higher education. As such, the reform possibly lead to a dismantling of the socio-demographic gradient in educational attainment.

In Chapter Three, Local Media Information and Choice of Primary Health Care Provider, co-authored with Jens Dietrichson and Gustav Kjellsson, we investigate how local media information affects the choice of primary health care provider. In this chapter, we study if local media reporting affects the number of enrolled patients among primary care providers that are mentioned in the news. We compare the difference in list size between treated primary care centers, that are subject to a publication, and control primary care centers that are never exposed in local media. Since publications occur at different points in time, we employ a method that can manage a staggered implementation of the treatment.

A first result of the study is that we find differentiated pre-treatment trends for treated and control providers for both positive and negative articles. This pattern indicates that at least some patients have and act on quality information *before* the articles come out which is interesting in relation to our research question. However, the identification strategy employed in this chapter requires treatment and control units to follow parallel trends before the onset of treatment. While informative of individuals behavior, the differentiated pre-trends are thus problematic for the identification of causal effects. We address the problem by using the 12 months before the publication date to estimate a differentiated pre-trend and then we study how the treated group deviates from the extrapolated trend after the publication date.

The main analysis is not able to detect any significant effect of either positive or negative coverage. While the effect of positive articles is close to zero, the event study suggests that there is a trend break following the treatment. However, the effects are small and insignificant. We test heterogeneity between different groups of articles by categorizing both positive and negative articles into those more or less likely to affect patients' enrollment, depending on the content of the article. We find a more pronounced effect among articles classified as strongly negative or positive, but the estimates are still small and insignificant. When splitting the data between providers located in different types of markets, namely in urban and rural towns, we record a stronger, yet insignificant effect among rural providers both with regards to positive and negative news.

Overall, the small or absent effects of media coverage are of interest when designing these patient choice markets. Unless the information reported in the local newspaper is already known to the public, these results suggest that patients do not turn away from low quality providers - even in case of reports of mistreatment. One major explanation for the lack of quality improvements exercised by patient choice and provider competition may still be that patients either do not have, or do not act on, information on provider quality. These results are of particular interest in the primary care context where patients to a large degree are left alone to make decisions of where to enrol - without guidance by other medical expertise.

Chapter 1

Antibiotic consumption and health care utilization in children

Abstract

As the global threat of antibiotic resistance grows more urgent, guidelines on the prudent use of antibiotics for human consumption are becoming commonplace worldwide, and access to antibiotics is increasingly restricted. This paper seeks to answer how obtaining an antibiotic prescription in primary care affects children's health care utilization, using the propensity of general practitioners (GPs) to prescribe antibiotics at an index visit for children ages 0–5. I show that GP behavior is unrelated to predetermined child characteristics, which allows for a causal interpretation of my results. The results show that being prescribed antibiotics at an index visit increases the probability of more interactions with the health care system in the short term. I estimate a precise increase in the probability of having at least one additional visit within 30 and 90 days, to both in- and out-of-hours primary care centers (PCCs), but no robust effect on either outpatient emergency visits nor hospitalizations. The results are robust to an extensive number of specification checks. I further corroborate the results by restricting the visits to those with a diagnosis of respiratory tract infection, a common condition in children, and show that the effects in this subsample are remarkably similar to those found for the full sample of visits. The results in this paper indicate that the GP's prescription decision has a significant effect on the downstream use of health care services.

Ethics Approval has been obtained from the Swedish Ethical Review Authority Dnr: 068-18.

1.1 Introduction

Antibiotics are the most commonly prescribed drugs to children in the Western world (Youngster et al., 2017). They can save lives, but antibiotic consumption in humans is considered the main driver of antibiotic resistance, a huge global public health concern (Adda, 2020). According to WHO (2022), common infections such as pneumonia, salmonella, and blood poisoning are becoming more difficult and more expensive to treat because bacteria are becoming more resistant. As a consequence, policies and guidelines to reduce antibiotic consumption have become increasingly commonplace, resulting in a downward trend in the aggregate consumption of antibiotics in children. From a societal point of view, reducing the consumption of antibiotics is a desirable outcome. However, less is known about individual outcomes of more prudent antibiotic use.

In this paper, I investigate the causal effect of antibiotic consumption in children on downstream health care utilization. I study outcomes immediately related to an index visit in the short term: the probability of having recurring visits (within 10, 30, and 90 days), where they occur (at in- or out-of-hours primary care centers, at emergency units, or during hospitalizations), and whether there is a change of health care provider.[1] To capture the effect of antibiotics on health care utilization for the entire sampling period, I also investigate the impact of antibiotics on the total number of health visits, where they occur, and the probability of obtaining diagnoses such as asthma, eczema, and respiratory tract infection (RTI).

The setting for the study is the primary care sector in Sweden, where, as in the United States, access to antibiotics requires a prescription. The largest share of antibiotics is prescribed by primary care physicians. Patients are free to choose their primary health care center, which makes obtaining causal evidence of antibiotic consumption challenging, since the choice induces correlations between individual and family characteristics, antibiotics, and health care consumption. I address this challenge by exploring a unique data set containing the full universe of visits to primary care centers (PCCs) in one of the largest regions in Sweden, Scania (Skåne). To leverage causal estimates, I explore possibly exogenous variation coming from the supply of antibiotics. I use the propensity to prescribe antibiotics, at the physician level, as an instrument for the probability of obtaining an antibiotic prescription at an index visit in the primary care sector. An undesired feature of the Swedish primary care sector is a discontinuous relationship between patients and physicians (Vård och Omsorgsanalys, 2021a). One

[1] Index visits are defined as those with no prior visits within 180 days.

reason for this is a shortage of general practitioners, which makes long-standing relationships between patients and physicians difficult. Besides the GP shortage, PCCs have problems with high staff turnover and physicians with temporary short-term contracts (Riksrevisionen, 2014).[2] While an undesired outcome for society, it enables an identification strategy that assumes quasi-random assignment of patients to physicians.

I show that physicians' antibiotic prescribing behavior strongly predicts the probability of obtaining a prescription for antibiotics at a visit but is unrelated to predetermined background characteristics, which allows for a causal interpretation of the reduced form estimate. Causal interpretation of the IV requires that the exclusion restriction is satisfied - which in this case implies that there are no other characteristics of physicians with higher propensities to prescribe antibiotics that might directly affect subsequent healthcare utilization.[3] The first-stage estimates show that meeting with a physician who is 10 percentage points more prescription prone significantly increases the probability of being prescribed antibiotics by 5.7 percentage points. Moreover, I find that being prescribed antibiotics at an index visit leads to an increased probability of more interactions with the health care system in the short term. I estimate a precise increase in the probability of having at least one additional visit within 30 and 90 days, to both in- and out-of-hours PCCs, but no effect on either outpatient emergency visits or hospitalizations. Being assigned to a physician who is 10 percentage points more prescription prone increases the probability of a revisit within 30 days by 0.663 percentage points. An alternative interpretation of the reduced-form estimate is that going from the lowest- to the highest-prescribing physician leads to an increase in the probability of having a revisit within 30 days of 3.3 percentage points, or 14.3% of the mean, a revisit to in-hours and out-of-hours PCC by 1.6 (9.5% of the mean) and 0.9 (47% of the mean) percentage points, respectively. The instrumental variables (IV) estimate has a larger magnitude: antibiotics increase the probability of any revisit within 30 days by 11.8 percentage point, or 56% of the mean. The results are robust to a number of tests, including changing the index visit definition, placebo tests, controls for co-treatments, and alternative specifications of the instrument.

To discern whether the short-term effects are driven by case mix—that is, whether high-prescribing physicians prescribe more antibiotics because they meet

[2] I explain the institutional details at greater length in section 1.2.

[3] While inherently untestable, I test the plausibility of this assumption in section 1.5.6 in two ways: First, I perform a placebo analysis by regressing the impact of the instrument on my main outcomes for a subset of visits that very rarely are prescribed antibiotics. Second, I test the sensitivity of the results to controlling for physician's prescribing behavior for other prescription drugs.

with more sick children—I restrict the sample to children who are diagnosed with RTIs and thus compare those with more or less the same symptoms, which is possibly the most relevant comparison. A drawback to this comparison, and the reason that this is not the main analysis, is that the probability of being diagnosed could be affected by the antibiotic prescription decision and is therefore a bad control in the terminology by Angrist and Pischke (2008). I find no positive effect of antibiotics on the probability of revisit within 10 days. However, the positive effects on revisits within 30 and 90 days are remarkably similar between the full and restricted samples. The main difference between the full and restricted samples is a significant reduction in the probability of hospitalization in the latter sample.[4] In sum, the effects on health care utilization within 1 and 1.5 months are not driven by differences in case mix. However, the short-term effects on emergency visits and hospitalizations are more sensitive to choice of sample and should be interpreted with more caution. However, even for the latter set of the results, the confidence intervals overlap for almost all point estimates from the two samples. Since the location of the PCCs is an important factor in the choice of PCC providers, and because the location often correlates with sociodemographic characteristics, I also conduct a heterogeneity analysis, which reveals that distance to PCC is a source of heterogeneity but not very important because the treatment effects are similar across the subsamples.

In the medium term, I show that physicians' antibiotic prescribing practices have a significant positive effect on the total number of both in- and out-of-hours PCC visits after the initial index visit. Being prescribed antibiotics at the index visit increases the total number of visits by 0.56, which is large relative to the mean of 1.5 PCC visits half a year after the index visit. The evidence in this paper strongly shows that the physicians' prescribing practice increases interactions with the primary sector of the health care system. I present evidence that this is not a supply-side effect, as it is not the physicians who reschedule meetings with patients for whom they prescribed antibiotics.[5] Rather, it seems to be driven by an increased demand for care, but I cannot distinguish whether this is because antibiotics increase infection susceptibility or because patients develop a preference for physicians.

The main contribution of this paper is to provide evidence of individual consequences of antibiotic consumption. To the best of my knowledge, no other

[4]The restriction to children with RTIs is conditional on an outcome, since the probability of being diagnosed can be affected by physicians' prescription propensity. Nevertheless, it is interesting to see the similarities and differences between the two samples, as they shed light on possible mechanisms for the results.

[5]Table 1.12 shows that the probability of having a subsequent visit with the same doctor as on the index visit is actually negatively affected by physicians' prescribing behavior.

paper has investigated the causal effect of antibiotics on short-term health care utilization, which is an important outcome, particularly in a setting where access to antibiotics is becoming more restricted. If prescribing antibiotics is associated with a higher utilization of health care services, then efforts to combat antibiotic resistance can have spillover effects on the work burden and patient inflow for primary care physicians.

The most closely related work at this point is that by Sievertsen et al. (2021), who study the cumulative effect of antibiotic consumption on children's cognitive outcomes measured as test scores at age 10. They employ a similar identification strategy, which uses the prescription propensity at the mothers' PCCs, at the PCC level, and find a negative impact of consuming more antibiotics at ages 0–5 on test scores in school. My paper differs in several ways. First and foremost, we study different research questions. The focus in this paper is on how physicians' prescribing behavior affect downstream use of health care. This captures both health and behavioral responses by patients to gauge the full effect of obtaining an antibiotic prescription at a primary care visit. Sievertsen et al. (2021) focus on the effect of cumulative childhood consumption of antibiotics on cognitive skills. Moreover, this paper is specifically focused on prescribing behavior in the primary sector, while Sievertsen et al. (2021) aggregates total consumption of antibiotics from all sectors of the health care system. Second, we use different sources of variation. While they explore differences in prescribing propensities between PCCs, my data allow me to use variation between physicians within primary care centers. The papers complement each other well. Sievertsen et al. (2021) captures in part the effect found in my paper, to the extent that I estimate a positive effect of antibiotics on subsequent health care utilization, which affects the probability of obtaining more antibiotics. This, in turn, has the possibility of affecting cumulative childhood consumption.

A second contribution of this paper is to provide evidence of how physicians' behavior causally affects the prescribing of antibiotics. This is important from a policy perspective, since it informs policy-makers about where to target efforts to combat antibiotic resistance. Huang and Ullrich (2021) explicitly focus on the supply side of antibiotic prescribing, using a different identification strategy in which they explore variation in antibiotic prescription styles related to physician exits from general practice clinics. Their main focus is how a large share of the supply of antibiotics can be attributed to differences in physician prescription style. They find that 53% to 56% of between-clinic differences in all antibiotic consumption is due to physician practice styles. My paper can be viewed as a complement to that of Huang and Ullrich (2021), verifying the importance of supply-side antibiotic prescriptions using a different setup while focusing more on the health care utilization effects of such differences in practices.[6]

Finally, a third contribution is that the detailed data with access to geographic proximity allow for investigating treatment heterogeneity. Access to PCCs is a highly debated topic and, indeed, one of the rationales for the introduction of telemedicine (Ellegård et al., 2021).[7]

Prior evidence on the individual consequences of reduced antibiotic consumption is scarce. On the one hand, antibiotic consumption is linked to negative health outcomes such as obesity, asthma, and eczema, so reducing antibiotic use may be beneficial for both individuals and society (Bejaoui et al., 2020). But in some cases, antibiotics can speed up recovery. As a consequence, parents often demand antibiotics, and surveys show that they have overoptimistic expectations (Coxeter et al., 2017). Empirical work is mainly centered around the effects of payment schemes (Ellegård et al., 2018; Currie et al., 2014) or institutional context (Fogelberg, 2013) rather than the individual consequences of obtaining antibiotics. Two exceptions, as already discussed, are Huang and Ullrich (2021) and Sievertsen et al. (2021). With regard to methodology, this paper adds to the health economics literature that studies variation in physician treatment styles at within- and between-region hospitals or PCC providers, such as Dalsgaard et al. (2014); Chandra and Staiger (2007) and Currie et al. (2016), which use differences in treatment styles as a source of exogenous variation. The identification strategy explored in this paper is closely linked to the literature on judge fixed effects, such as Kling (2006), Doyle (2007), and more recent work by Dahl et al. (2014); Dobbie et al. (2018) and Bhuller et al. (2020). These papers use the within-court random assignment of judges to cases to estimate the effects of incarceration (or incarceration of parents) on a wide range of outcomes. More generally, this gatekeeper fixed effects approach has also been used in a health-related setting, such as by Maestas et al. (2013), who studies the effect of disability benefits on labor market supply using variation between examiners within offices, and by Bakx et al. (2020), who examines the effect of nursing home eligibility on mortality and health.

The remainder of the paper is structured as follows: section 1.2 describes the institutional setup, section 1.3 provides the data and some descriptive evidence,

[6]Huang and Ullrich (2021) also studies one adverse health outcome: preventable hospitalizations due to infections. They find little evidence that differences in practice styles adversely affect health, except for an increase in hospitalizations for a subclass of antibiotic drugs, penicillin. This paper does not include this outcome because the majority of symptoms for this class of hospitalizations are experienced by adults.

[7]The use of digital services really took off in 2018 and onward, whereas the period covered in this paper ends at 2017.

and section 1.4 discusses the empirical strategy and the details of the instrument. The results are presented in section 1.5. Section 1.6 tests the robustness of the results, and section 1.7 concludes.

1.2 Background

The primary care sector

The health care sector in Sweden has universal coverage and is funded by taxes. Individuals are automatically enrolled in the health care system, and only 6% have private insurance (Glenngård, 2015). Visits for children are free from fees. The share of health care expenditures constitutes 11% of the GDP, slightly higher than in other OECD countries, where the share is 9.8% on average, but lower than in the United States, where health care spending makes up approximately 17% of the GDP.

The primary care sector is at the front line of the health care system. It is a decentralized system, under the responsibility of the 21 regions, and is organized via group practices, through several primary health care centers (PCCs). PCCs employ nurses and physicians and often serve as gatekeepers to more specialized care through referrals. The primary care sector is responsible for a large share of total antibiotic consumption, approximately 60%. The remaining share consists of antibiotics prescribed in open specialized care, dental care, and inpatient care (Nord et al., 2013).

Choice of health care providers

Patients have the freedom to register with any PCC they like and also to change providers whenever they like. There are no restrictions on the number of times an individual can change providers, and the providers cannot decline a registration. This choice was introduced in 2009 by the Act of Free Choice (SFS, 2008:962). The reform was designed to improve the patient's ability to register with the provider that best suits their needs. However, if patients do not actively make a choice, they will be assigned to the geographically closest clinic. For 80% of inhabitants, it is less than a five-minute drive to the second-closest PCC (Glenngård, 2015). According to the law, patients should be able to register with a regular, personal, GP though this system has been functioning very poorly. The main reason is that there is a shortage of specialized physicians. A prerequisite for a doctor to serve as a regular GP is that he or she must be a general practitioner or a specialist within another field (Riksrevisionen, 2014). To become a GP, the physician must first have obtained a medical doctor (MD) certification and have done 24 months of clinical training (AT, "*allmäntjänsgöring*") as part of the degree. Thereafter, the physician undergoes specialized training (ST, "*specialisttjänstgöring*"), which typically takes 5 years. During this period, physicians are usually called residents. Due to the shortage of GPs, many PCCs rely heavily on interns, residents, and rental doctors, which is expensive and comes at the cost of physician-patient relationships and a high staff turnover (Riksrevisionen, 2014).

Children in primary care

The front-line health system for children is divided into two parts: child health care centers (CHCs, "*barnavårdscentraler*") which are preventative and standard primary care centers (PCCs) which are curative. The CHCs are responsible for children's general well-being and track weight, height, and physical and physiological development from birth to school age. This is done through a standardized program offering visits at different ages. The participation rate is almost 100%. The CHCs typically employ nurses and one or a few pediatricians. Standard checkups must be performed by a doctor; otherwise, the child regularly meets with a nurse. The pediatrician can also write referrals if the child needs specialized care. Thus the main objective of the CHCs is preventative care (National Board of Health and Welfare, 2014). The focus of this paper is the treatment of sick children. In this regard, they offer little help to families of sick children but rather refer them to PCCs. The main difference is therefore that CHCs typically meet health children with a preventative focus, while sick children are treated in PCCs by GPs. The focus of this paper is the treatment of *sick* children, so data on the PCCs are more useful for this study.

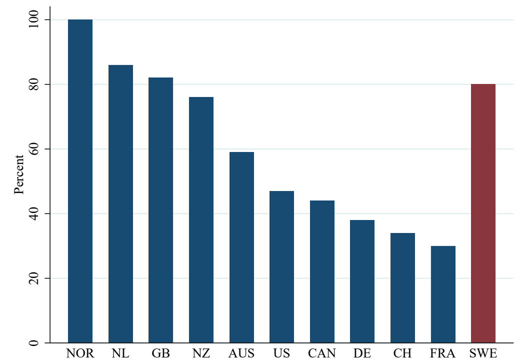
The allocation of physicians to children in PCCs is typically subject to a triage system, in which a nurse must first determine that seeing a doctor is necessary. One potential issue is whether some parents have strong preferences for certain physicians, which could possibly correlate with prescribing behavior. While I cannot empirically assess this, I have corresponded with physicians working in Scania, who, independently of one another, state that it is rare for parents to have such preferences. More commonly, they want to see the first available physician to get help for their sick child as soon as possible. The types of visits observed in this paper are to PCCs, where it is not likely that a child would be allocated to a GP specializing in children, since PCCs typically do not employ specialized physicians. This also mitigates the concern about parents selecting GPs based on their children's specific needs. Children with certain types of health requirements

are typically referred to specialized care by CHCs. Neither PCC nor CHC Visits for children are subject to fees.

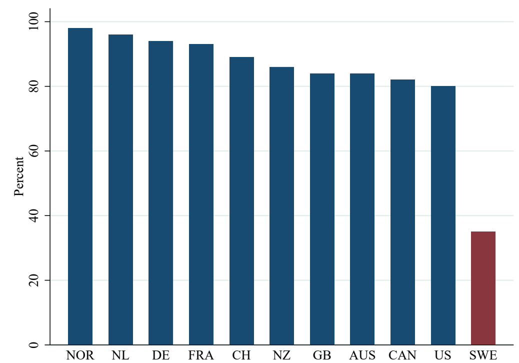
The physician-patient relationship

A continuous relationship between the patient and physician is crucial for well-functioning primary care and leads to improved health for the patient (Cabana and Jee, 2004). Relative to other comparable countries, Sweden has a low share of patients with regular, personal, GPs even though the share of patients registered at a their regular center level is high.

Figure 1.1: Physician-patient continuity



(a) Share with regular PCC



(b) Share with regular physician or nurse

Notes: This figure shows the shares of individuals reporting that they have a regular PCC and/or a regular, personal GP or nurse, in Sweden and 10 other countries. The data comes from International Health Policy Survey (IHP) and was accessed from Vård och Omsorgsanalys (2021b).

Figure 1.1 shows that only 35% of the population in Sweden have a regular and personal contact with either a nurse or a physician, compared with around 80–98% for the other countries (Vård och Omsorgsanalys, 2021b). Most of the fixed GP contacts are with a physician. Approximately 25–27% of the Swedish population have a regular physician, and this share has been steady (or only slightly declining) at least from 2013 onward (Vård och Omsorgsanalys, 2021a). Moreover, the share of individuals with regular physicians is measured on the population level; it is likely that the share is even lower for the children in my sample. In Scania, as in many other regions, a a personal and regular physician contact is prioritized for elderly, chronically ill, and multimorbid patients (Vård och Omsorgsanalys, 2021b). These patients are unlikely to appear in my sample, since I focus on young children with no prior visits within 6 months.

This a particular feature of the organization of the primary care sector in Sweden, in contrast to the organization of primary care in other countries, such as the United States. While not ideal for the patients, it is ideal for the purpose of the identification strategy employed in this paper, which assumes that physicians are conditionally assigned similar patients.

Prescribing of antibiotics

The procedure of prescribing antibiotics is subject to guidelines, but in practice, adherence to treatment guidelines is often poor (Finkelstein et al., 2021). This leads to unnecessary prescribing, such as prescribing antibiotics for a cold even though antibiotics are ineffective against cold viruses. Norms, parental expectations, and lack of diagnostic tools are commonly cited reasons for poor adherence (Abaluck et al., 2020).[8] Given that adherence to treatment guidelines is poor, it may also be that some parents (wealthier or more educated) have a higher probability of getting antibiotics for their children, which could possibly exacerbate health inequalities (if recovery is sped up) or ameliorate differences (if antibiotics are negative for health outcomes).

For the prudent use of antibiotics, it is typically recommended to prescribe a narrow-spectrum drug rather than a broad-spectrum drug, as the latter category has a larger effect on antibiotic resistance. Interestingly, Finkelstein et al. (2021) use Swedish data to study guideline adherence and find that the largest adherence gap between medical experts (parents or individuals being physicians) and

[8] There are relevant changes in the guidelines for two common child illnesses, acute otitis media and acute pharyngotonsillitis. I describe these changes in more detail in Appendix D. While difficult to assess empirically, extensive literature specifically addresses poor adherence to the acute otitis media guideline (see, for example, Célind et al. (2014)) and guidelines for common conditions in primary care, including pharyngotonsillitis (see, for example, Nord et al. (2013)).

nonexperts is for guidelines governing antibiotic treatment and conclude that this could be pointing to a possible conflict when considering antibiotics prescribing, as they conclude that the guidelines are more designed to promote public health than focused on the more narrow interest of the patient.
"*the association is most negative for guidelines regarding appropriate use of antibiotics which are designed to promote public health rather than the narrow interest if the patient*" (p. 4).
Thus this poor adherence to guidelines can be considered a potential conflict between the individual and the societal objective of reducing antibiotic consumption, which underscores the importance of carefully examining the impact of prescribing antibiotics on health care utilization.

1.3 Data and Descriptive Statistics

The backbone of the data is individual-level data on health care utilization in the primary care sector, containing information about the type of visit, date, diagnosis codes from the International Classification of Diseases (ICD-10), and name of the PCC. Uniquely for Scania, and crucial for the identification strategy, I have access to physician identifiers. A physician identifier consists of three letters. The identifiers are not necessarily unique across PCCs, for example, identifier ABC could exist at several PCCs throughout the sampling period. With the data at hand, it is impossible to uncover if this physician is the same individual, working at different practices or several individuals with the same 3-letter combination. However, within each PCC there can only exist one physician per identifier. To ensure that only one individual is associated with a physician identifier, I construct the instrument using physician IDs only within PCCs as these definitely are unique.[9] As a robustness check, I also construct the instrument using physician identifiers across the sample, as this will allow for movement of physicians between PCCs. The results are presented in the Appendix, Table A6, they have a slightly larger magnitude but are overall very similar to the main results found in this paper.

The data in this project cover the time period 2010–2017 in one of the largest regions in Sweden, Scania. Scania is the third-largest region, with a mix of rural and urban areas, and has 1.4 million inhabitants (SCB 2022). I define the base population to be children ages 0–5 (born in 2010–2017) at the time of the visit.[10] I

[9] While the physician identifiers are unique to each physician within PCCs, they cannot be linked to any other data, such as data on physician characteristics.

[10] Children born in 2005–2009 are not included because my data on hospitalizations start in 2009, and I need to be able to define the time and conditions around the birth.

link the children to a primary care physician identifier for each physical physician visit at a PCC and make several sample restrictions to ensure that the sample is relevant. Since antibiotics are prescribed by physicians and, in general, not over the phone, this restriction ensures that the visits included in the sample are actually potential antibiotic visits. For the same reason, I also exclude physician contacts at CHCs.[11]

To identify antibiotic treatment, I match children to Sweden's National Prescribed Drug Register, which contains information on all dispensed antibiotic prescriptions, including dosage and ATC code (ATC-code=J01). I include only those prescriptions that can be matched to a PCC encounter.[12] Prescriptions for are recorded only if they are filled, potentially leading to an underestimation of the frequency of (antibiotic) prescribing. Thereafter, the sample of primary care visits and corresponding antibiotic information is matched to the registry data provided by Statistics Sweden, which links several administrative registers by personal identification numbers. Linking the children to parents using the Multi-Generational register, I can obtain information about the child's date of birth, gender, siblings, birth order, and parents' background characteristics, such as marital status, origin, occupation, and education. I also link the children to the inpatient register provided by the National Board of Health and Welfare (Socialstyrelsen) to record hospitalizations and background variables at the time of birth. Finally, I add data on the distances from the home address to the 10 closest PCCs. Distance measures are straight distance, distance by road, and distance in duration (in minutes, routed along the closest road). I use this last part of the data for a heterogeneity analysis. I restrict the sample to keep only those children for which I have complete information on the background characteristics. The sample contains 346 864 visits by 76,183 unique individuals.

For the main portion of the analysis, I restrict visits to those more than 180 days apart and refer to them as "index visits."[13] This restriction causes a significant sample size reduction. Finally, if individuals have multiple index visits with

[11]These visits have an antibiotic prescription rate equal to 0.3%, since visits to CHCs follows a national program where the visits are preventative, routinely scheduled checkups, as described in detail in section 1.2.

[12]The matching is by a unique individual identifier and the exact date on which the primary care visit was made and the prescription was dispensed.

[13]There is no consensus in the literature about how long the time window between visits should be. In the epidemiological literature, studies use 30 days between visits (see e.g., Sabbatini et al., 2016), and Finkelstein et al. (2021) include only those observations with no use of antibiotics within the preceding 2 years (though they do not have access to primary care data). I use 180 days following the definition in Milos Nymberg et al. (2021). This is a trade-off between keeping as good as independent visits while not reducing the sample size too much. Keeping only the first visit will lead to the problem that children will be very young (mean age 0.7 years), and for this group, physicians are very prone to prescribe antibiotics, as they should be. Note that the first visit to a PCC for the child is not automatically classified as an index visit, if it was preceded by a visit to an emergency unit,

the same physician, I keep only the first occurrence. The final sample contains 72,245 index visits by 50,951 individuals.

1.3.1 Variables

Treatment

The treatment is defined as the child being prescribed any type of antibiotic (i.e., either narrow or broad spectrum) at an index visit and the antibiotic being dispensed at a pharmacy. The variable is an indicator variable equal to 1 if the drug was prescribed and 0 if not.

Outcomes

The outcomes studied in this paper can broadly be divided into two categories: short term and medium term. The variables in the short term are the probability of having an additional visit, to the same or a different physician, within 10, 30, or 90 days. The revisit can occur at an in-hours PCC (in general, referred to as only PCC in this paper), out-of-hours PCC (*närakut*), or emergency unit, or as a hospitalization. The out-of-hours PCCs are open exclusively on weekends and evenings. The hospital visits include only those requiring inpatient care, which typically are for more severe conditions. Visits to emergency units include revisits to hospitals from the outpatient register, that are registered at an emergency unit. All these variables are binary and equal to 1 if the child had at least one subsequent visit within the specified time frame or to the specific type of facility.

A concern in Sievertsen et al. (2021) is that parents may switch providers as a response to the decision of whether to prescribe antibiotics. An advantage of this study is that I can test this channel directly by constructing indicator variables, switch_{t+1} and switch_{t+3}, with the former equal to 1 if the next visit to a PCC was at a different unit than the index visit, and the latter equal to 1 if the patient switched to a different unit for one or more of the three subsequent visits.[14]

In the medium term, I focus on two types of variables to capture the effect of antibiotics on the child's general health. The first is the total number of visits in the 6 months following the initial index visit, both overall and at the different types of facilities.[15] The second is the effect on the probability of being diagnosed

out-of -hours PCC or hospitalized within 180 days. In section 1.6, I test the robustness of the results to shrinking the no-visit window to 90 days and expanding it to 365.

[14]Note that the switch is defined based on actual visits. An alternative is to define a switch as equal to 1 if there was an active change in the listed unit, but actual visits are more relevant.

[15]Medium term is defined as the 6 months following the initial index visit. The difference between short and medium term in this paper is that the short-term outcomes are directly related to index

with asthma, eczema, or RTI. See Appendix C2–C4 for a detailed list of the ICD-10 codes for each diagnosis class. According to the epidemiological literature, excessive antibiotic consumption increases the risk of asthma and eczema. The mechanism is the disruption in gut microbiota caused by antibiotics, and children are more susceptible to this, as their biota is still developing (Jernberg et al., 2010; Schwartz et al., 2020). RTIs are included to test whether antibiotics have an effect on the probability of becoming ill, such as through a reduction in the efficacy of the immune system, in which case we would expect a positive impact on RTIs. Moreover, if antibiotics treat the child's condition better than no or alternative treatment, we would expect a positive impact, since the next time the patient acquired respiratory symptoms, the parents would be more likely to turn to the health care sector for help for their child.

Background Variables

I include an extensive set of background variables, as these are an important tool to check for conditional random assignment. I broadly divide the background characteristics into variables defined for the individual, the individual parents, and the PCC. All parental background characteristics are measured the year before childbirth. A more detailed description of the control variables can be found in the Appendix, Table C1.

visits, while the medium-term outcomes measure total visits in the half-year following the index visit.

Table 1.1: Summary Statistics

	Mean(AB=1) (1)	Mean(AB=0) (2)	Diff. (3)	Std. Error (4)
Individual Characteristics				
Age	2.91	2.62	-0.293***	0.012
Female	0.49	0.48	-0.009*	0.005
Number of siblings	2.23	2.16	-0.078***	0.009
Born Jan-March	0.27	0.26	-0.001	0.004
Born April-June	0.27	0.27	0.007*	0.004
Born July-Sept	0.26	0.26	-0.002	0.004
Born Oct-Dec	0.21	0.20	-0.004	0.004
Number of index visits	1.86	1.87	0.013	0.008
Total visits	7.91	7.81	-0.096*	0.051
Individual Health at birth				
Hospital length (days)	2.63	2.68	0.055**	0.022
Caesarean Birth	0.15	0.15	0.000	0.003
Twin birth	0.03	0.02	-0.006***	0.002
Birth order	1.20	1.20	-0.005	0.004
Pre-term	0.04	0.04	-0.002	0.002
Birth weight$\leq 1000g$	0.01	0.00	-0.000	0.001
Birth weight$1001-2500g$	0.03	0.03	-0.001	0.002
Complications at birth	0.08	0.08	0.004	0.002
Parents Characteristics				
Both parents born abroad	0.18	0.18	-0.004	0.004
Age at birth$_m$	30.98	30.90	-0.074	0.047
Age at birth$_f$	33.80	33.72	-0.082	0.057
Mother married at birth	0.46	0.46	-0.004	0.005
Years of educ$_m$	13.00	13.06	0.053***	0.020
Years of educ$_f$	12.56	12.64	0.081***	0.020
Not working$_f$	0.09	0.10	0.011***	0.003
Not working$_m$	0.13	0.15	0.014***	0.003
Family income	365 939	361 207	-4 732**	2 143
Parents Highest Education level				
Elementary school$_f$	0.13	0.12	-0.007**	0.003
Elementary school$_m$	0.11	0.11	-0.009***	0.003
High school$_f$	0.47	0.46	-0.008*	0.005
High school$_m$	0.37	0.37	0.001	0.005
University$_f$	0.40	0.42	0.015***	0.005
University$_m$	0.52	0.52	0.008*	0.005
Health care characteristics				
Distance to closest PCC (km)	1.93	1.89	-0.042*	0.022
Number of listed patients, PCC	9767	9766	0.172	39

Notes: The final sample contains 72 745 observations. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As is evident in Table 1.1, there are significant differences between the groups of visits that lead to antibiotics in column 1 and those that did not in column 2. I add a t-test of the difference in means between the two groups. With regard to the individual characteristics, visits leading to antibiotics are by patients who are slightly older, female children from marginally larger families. The magnitudes are small; for example, 49% of the patients who got antibiotics were female compared with 48% in the no-antibiotics group. With regard to the health characteristics, there are small but very significant differences between the groups in the length of hospital stay at birth and the share of twins. The largest difference between the groups is with regard to the level of parents' education. Overall, in terms of both years and level of education, parents of children in the antibiotics group have a lower mean. On the other hand, they have a slightly higher annual disposable income. As these characteristics are constant for each individual, the differences can occur for two reasons: (i) the children are, on average, different in terms of characteristics; or (ii) the children are equal, but the share of individuals having multiple index visits differs with parents' background characteristics.

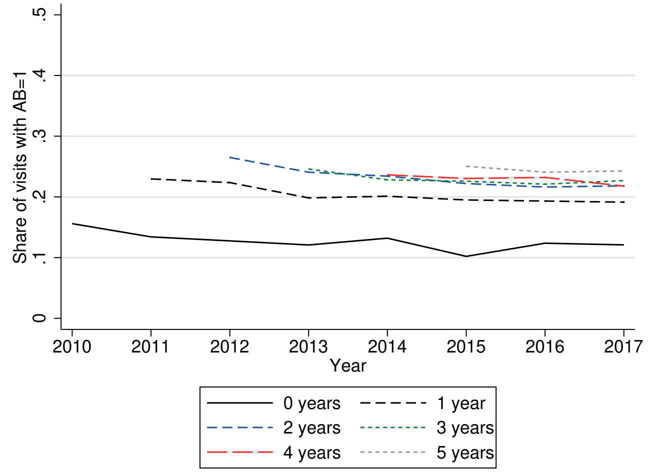
1.3.2 Descriptive Evidence

Prescribing of antibiotics for children, on the aggregate level, has been downward trending since the 1990s in Sweden. In this paper, it is difficult to present statistics for the aggregate number of prescriptions in the age group 0–5, since the age in my sample is skewed.[16] To account for this, I present the share of in-hours PCC visits that were linked to an antibiotic prescription per age cohort. From Figure 1.2, it is clear that the share of PCC visit that resulted in an antibiotics prescription is at different levels for different age cohorts but that the level of prescribing within cohorts has remained stable across the sample. The prescribing rate for 0-year old children is lowest, followed by the rate to 1-year old children. From 2 years and onwards the rates are more similar, this is line with the fact that conditions that often require antibiotics are more common when the child have started preschool, usually around 1-2 years of age (Daysal et al., 2021).

To get a better understanding of the types of conditions that cause physicians to prescribe antibiotics for children, I plot the most common diagnosis groups for which antibiotics are prescribed.

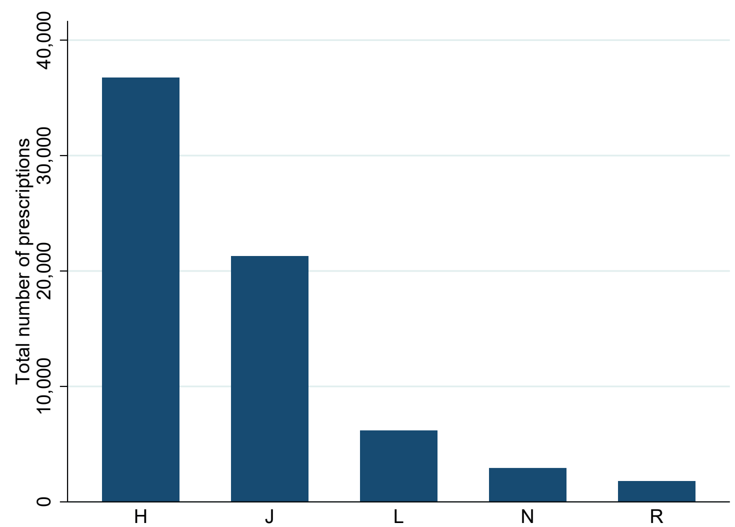
[16] I include children ages 0–5 at the time of the visit during 2010–2017. This implies that I have only three cohorts (those born in 2010, 2011, and 2012) for which I can observe the full range of antibiotic prescriptions for this time period.

Figure 1.2: Share of PCC visits that prescribe antibiotics, per visit year and age



Notes: The y-axis represent the share of PCC visits that resulted in an antibiotics prescription age cohort and the x-axis shows the year of the visit

Figure 1.3: Number of prescriptions per diagnosis class



Notes: The y-axis represent the total number of prescriptions throughout the sample period per diagnosis class. The diagnosis classes are as follows: *H = ear and eye, J = respiratory system, L = skin and subcutaneous tissue, N = genitourinary system, and R = symptoms, signs, and abnormal clinical and laboratory findings, not elsewhere classified.*

From Figure 1.3, it is clear that the most common diagnoses for which antibiotics are being prescribed are ear and eye-related conditions. This group is dominated by ear-related conditions, mainly acute otitis media (AOM). This is expected, since AOM is one of the main reasons children visit PCCs (Lundborg Ander and Eggertsen, 2004).

1.4 Empirical Strategy

The aim of this paper is to estimate the causal effect of being prescribed antibiotics in childhood at an index visit at a PCC on downstream health behavior. The estimating equation is

$$y_{iv} = \alpha + \beta AB_{iv} + \mathbf{X}_{iv}\delta + \epsilon_{iv} \tag{1.1}$$

where y_{iv} represents the outcome variables for individual i related to index visit v. In the short term, these are recurring visit (within 10, 30, or 90 days), type of revisit (in-hours, out-of-hours, emergency, or hospitalization), and change of health care provider. In the medium term, these are total number of health visits and the probability of having diagnoses such as asthma, eczema, or upper respiratory tract infection (RTI). AB_{iv} is an indicator variable equal to 1 if the individual i receives an antibiotic prescription at a PCC. $\mathbf{X}_{iv}$ includes the observed characteristics of the patient, described in Table 1.1. Equation 1.1 can be estimated with a simple OLS model. However, β will be biased if there are omitted variables affecting both the consumption of antibiotics and the health outcome. In this context, the fact that individuals choose their health care providers is a source of omitted variable bias. For example, certain family characteristics, such as attitudes and expectations, may affect both health behavior and the probability of obtaining antibiotics. Moreover, omitted variable bias will also stem from unobserved (by the econometrician) characteristics of the patient, such as severity of symptoms and the general health condition, which affect both the probability of being prescribed antibiotics and subsequent health care behavior. To formalize this idea, I closely follow Maestas et al. (2013) and outline the thought experiment and how this can be translated into an instrumental variables approach. If we rewrite equation 1.1 as

$$y_{iv} = \alpha + \beta AB_{iv} - s_i + \mathbf{X}_{iv}\delta + \epsilon_{iv}$$

where s_i is the unobserved severity of the illness, we can think of this as the unobserved share of the change in outcomes associated with an antibiotic prescription. From a societal perspective, the optimal use of antibiotics follows the prescribing guidelines, and thus a physician should prescribe antibiotics to patients sick enough to exceed a prescription threshold specified in the guidelines.

I call this the guideline prescription threshold (GPT):

$$GPT > \mathbf{X}_{iv}\delta - s_i + \epsilon_{iv}$$

In reality, patients in general and children in particular may exhibit symptoms in different ways, and physicians may or may not use diagnostic tools or tests for bacterial infections. This implies that physicians have imperfect information about the severity of their patients' diseases, so the prescription rule becomes

$$GPT > \mathbf{X}_{iv}\delta - \hat{s}_{ij} + \epsilon_{iv}$$

where $\hat{s}_{ij}$ is physician j's estimate of the severity of patient i's illness. Importantly, the physician observes this more accurately than the econometrician, through patients' journals, test results, and general condition. It is also a function of the characteristics of the physician: previous research has shown that physician characteristics such as age, gender, experience, and taste for medication affect prescribing (Cadieux et al., 2007)). We can decompose $\hat{s}_{ij}$ into patient-specific and physician-specific parts:

$$\hat{s}_{ij} = s_i + \omega_j$$

where ω_j is a systematic part of physician j's judgment of the patient, which on average leads to systematic over- or underprescribing to patients. Substituting the above equation into the expression for the prescription rule yields the following:

$$GPT > \mathbf{X}_{iv}\delta - s_i - \omega_j + \epsilon_{iv} \tag{1.2}$$

This leads to the notion that the physician will prescribe antibiotics to patient i only if the patient is sufficiently ill:

$$AB_{iv} = 1(s_i > \mathbf{X}_{iv}\delta - GPT - \omega_j) \tag{1.3}$$

As shown in equations 1.2 and 1.3, physicians with high ω_j will overestimate the severity of symptoms, leading to a lower prescription threshold and a higher propensity to prescribe antibiotics. ω_j can also reflect other time-invariant physician characteristics, such as his or her views on antibiotic resistance in society, to the extent that they influence the interpretation of the patient's illness and thereby appropriate treatment. The variation in physicians' ω_j gives rise to a natural source of variation that can lead to a possibly exogenous change in the supply of antibiotics. The change can be exogenous, conditional on an extensive set of fixed effects described in more detail later in this section, for two reasons. First, the way the primary care system is set up in Sweden does not, in general, allow for long-standing relationships between the physician and the patient. Second, if more than one interaction occurred with the same physician, I keep only the first interaction (i.e., the original assignment in the terminology of Cunningham (2021)).

I construct a physician-visit specific measure of physicians' prescription propensity, defined as the share of PCC visits leading to a prescription, crucially excluding the focal visit. For the construction of the instrument, I follow Dahl et al. (2014) and Dobbie et al. (2018) and include all visits, not just index visits. The upside of using a larger sample is that the instrument will be measured with more precision, which is particularly important in smaller PCCs with fewer visits by children.[17] I begin by constructing the leave-out mean:

$$Propensity_{jk} = \frac{\#AB_k - 1(AB_k = 1)}{\#visits_j - 1} \tag{1.4}$$

In other words, $Propensity_{jk}$ measures the prescription rate of physician j for visit k, for all visits except the focal visit. I regress $Propensity_{jk}$ on an extensive set of fixed effects—visit month, age, year, and PCC fixed effect—to reflect the notion that random assignment of index visits to physicians occurs within these cells:

$$Propensity_{jk} = \alpha + \pi_{1visitmonth} + \pi_{2age} + \pi_{3year} + \pi_{4PCC} + e_{jk} \tag{1.5}$$

After controlling for the fixed effects, I predict the residual:

$$\hat{e}_{jk} = Propensity_{jk} - \widehat{Propensity}_{jk} \tag{1.6}$$

This residualized measure of prescription propensity is then used as an instrument for the probability of getting antibiotics, and the set of estimating equation becomes:

$$AB_{iv} = \alpha + \beta_1 \hat{e}_{jk} + \beta_2 \mathbf{X}_{iv} + \epsilon_{iv} \tag{1.7}$$

$$y_{iv} = \alpha + \beta AB_{iv} + \delta \mathbf{X}_{iv} + \epsilon_{iv} \tag{1.8}$$

where AB_{iv} is instrumented by the residualized measure of physicians' prescription propensity (PPP), as defined in equation 1.6. By instrumenting antibiotic prescription with PPP, I identify the local average treatment effect (LATE) of antibiotics for children on the margin. These children are the group for which studying the prescribing of antibiotics is meaningful, since this group of patients consists of

[17] I also perform a robustness check using only the index visits to construct the instrument. The result is presented in the Appendix, Table A3

those who receive an antibiotic prescription because of being assigned to a high-prescribing (high ω_j) physician but would not if assigned to a low-prescribing (low ω_j) physician. In the terminology of the causal inference framework, these are the compliers (Angrist and Pischke, 2008).

For causal interpretation of the antibiotics estimates, the independence assumption must be satisfied, ensuring that children really are randomly assigned to PCCs, conditional on an extensive list of fixed effects (instrument validity). The instrument must also satisfy the exclusion restriction, which in this application means that physicians cannot affect the children's future outcomes by means other than through their prescription of antibiotics. I elaborate on the plausibility of the exclusion restriction in section 1.5.6. It is important to note that even if the exclusion restriction fails, the reduced-form estimates can, under the assumption of random assignment to physicians, still be given a causal interpretation of the effect of being assigned to a more or less lenient, in terms of antibiotic prescribing practice, physician on future outcomes. The instrument must also be relevant, which implies prescribing behavior should be correlated with antibiotic prescription only if there is a systematic, underlying physician-specific threshold for antibiotic prescribing decisions. The underlying variation in this instrument may come from between (within PCCs) or within physicians. In Appendix B, I explore and decompose the underlying variation, from which I can conclude that the main share of the variation in this instrument comes from differences in antibiotic prescribing between physicians relative to within. Finally, the last assumption for the LATE that needs to be defined is monotonicity. Monotonicity assumes that if a patient receives antibiotics from a low-prescribing physician during an index visit, the patient would also have received antibiotics if assigned to a high-prescribing physician. I test this assumption in section 1.5.5.

1.4.1 Instrument Validity

For physicians' propensity to be a valid instrument, I must first assume that patients are as good as randomly assigned to physicians. This implies that the assignment must be uncorrelated with unobserved characteristics conditional on observed characteristics. In this paper, the assumption stipulates that there is conditional random assignment. As shown in equation 1.5, I add an extensive set of fixed effects, which effectively means, for example, that physicians are allowed to specialize in certain age groups, or that high or low prescribers work at different times of the study period, but they cannot decide which patients they should meet. As described in section 1.2, individuals are free to choose the PCC unit, so conditioning on PCC fixed effects is crucial. Since I exclude PCCs with only one physician and include only the first interaction between a patient and a particular physician, I assume that the assignment of patients for an index visit, at a certain PCC, in a particular year, month, and cohort, is as good as random. I investigate the plausibility of this assumption in two ways: First, I regress both the probability of being prescribed antibiotics and the instrument PPP on an extensive set of patient characteristics. Table 1.2 shows the results of this analysis. Second, when evaluating the relevance of the instrument in the first stage, I add covariates sequentially and check the consistency of the point estimate. This is an indirect test of random assignment, as only covariates that are correlated with PPP will affect the point estimate.

Table 1.2: Randomization Test

	AB	Physician Propensity
Female	0.00369	0.00039
	(0.00280)	(0.00073)
Number of siblings	0.00839***	0.00024
	(0.00170)	(0.00040)
Born abroad	0.04100	0.01809
	(0.12842)	(0.02404)
Both parents born abroad	-0.00541	0.00162
	(0.00534)	(0.00115)
Age at birth$_m$	0.00075*	-0.00003
	(0.00041)	(0.00010)
Age at birth$_f$	-0.00002	-0.00002
	(0.00038)	(0.00007)
Married at birth$_m$	-0.02956	0.00108
	(0.01798)	(0.00399)
Married at birth$_f$	0.02003	-0.00041
	(0.01778)	(0.00394)
Years of education$_m$	-0.00150	0.00008
	(0.00161)	(0.00038)
Years of education$_f$	-0.00250**	0.00023
	(0.00116)	(0.00036)
Parents unemployed	-0.00448	-0.00080
	(0.00483)	(0.00124)
Family disposable income quintile 2	0.00881*	-0.00056
	(0.00467)	(0.00117)
Family disposable income quintile 3	0.01560***	-0.00164
	(0.00491)	(0.00118)
Family disposable income quintile 4	0.02013***	-0.00192
	(0.00536)	(0.00130)
Family disposable income quintile 5	0.02474***	-0.00037
	(0.00596)	(0.00137)
University^f	0.00586	-0.00179
	(0.00523)	(0.00151)
University^m	0.00658	-0.00100
	(0.00592)	(0.00138)
Size of the PCC	0.00000	0.00000
	(0.00000)	(0.00000)
Hospital length at birth	-0.00141**	-0.00025*
	(0.00069)	(0.00014)
Twin birth	0.03685***	-0.00397
	(0.01047)	(0.00245)
Birth order	0.01087**	-0.00140
	(0.00425)	(0.00088)
Pre-term	0.01081	-0.00111
	(0.01745)	(0.00378)
Birth weight$<$ 1000g	0.01517	-0.00126
	(0.02260)	(0.00447)
Birth weight $1001-2500$g	-0.00453	0.00410
	(0.01795)	(0.00394)
Complications at birth	-0.00655	0.00197
	(0.00626)	(0.00142)
Born Jan-March	0.01386***	0.00196*
	(0.00422)	(0.00109)
Born April-June	0.00688*	0.00075
	(0.00397)	(0.00110)
Born July-Sept	0.00469	0.00104
	(0.00410)	(0.00111)
F-stat	6.576	1.306
p-value	0.0000	0.1581
Obs	72745	72745

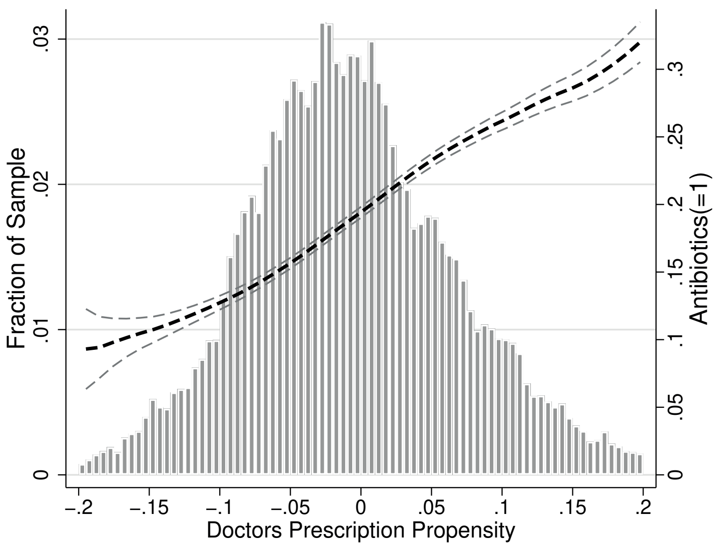
Notes: Variables are defined in Appendix C1. The coefficients are net of fixed effects, which include PCC (160), age (6), visit month (12) and year (8). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The output in Table 1.2 is consistent with the proposition that patients are conditionally randomly assigned to a certain physician. In contrast to the importance of the covariates on antibiotics, very few of the covariates are significantly related to the instrument. The null hypothesis that the parameters are jointly different from zero cannot be rejected in column 2. In other words, physicians meet with similar children, so changes in the outcomes associated with physician behavior does not come from differences in the patient pool. For a more graphical overview of the importance of each covariate, I plot the coefficients in Appendix Figures A1 and A2.

1.4.2 Instrument Relevance

Figure 1.4 plots the distribution of the residualized leave-out measure of physicians' antibiotic prescription propensity.

Figure 1.4: Effect of Prescribing Behavior on Antibiotics



Notes: The figure shows the histogram captures the tendency among GPs to prescribe antibiotics, excluding the focal visit. The prescribing propensity and the probability of obtaining antibiotics are residualized using the full set of control variables and fixed effects, described in Table 1.1. Dashed lines represent 95 percent confidence intervals. Standard errors are two-way clustered at the PCC and individual level

Figure 1.4 is a visual representation of the identification strategy. The histogram captures the average antibiotic prescription rate in all other visits for a physician (residualized). Overlaid is a local polynomial estimation of the first stage. Physicians' residualized tendency to prescribe antibiotics ranges from -0.19 at the 1st percentile to 0.27 at the 99th percentile, with a standard deviation of 0.09. Antibiotic prescription is monotonically increasing in prescription propensity and is highly linear. For each 0.01 increase in the instrument, the corresponding likelihood of antibiotic prescription increases by approximately 0.57.

Table 1.3: First Stage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Physician Propensity	0.5811***	0.5632***	0.5639***	0.5638***	0.5636***	0.5642***	0.5647***
	(0.0239)	(0.0227)	(0.0227)	(0.0226)	(0.0227)	(0.0228)	(0.0227)
mean of dep. var.	.194	.194	.194	.194	.194	.194	.194
F-stat	591	304	152	92.9	78.7	62	55.8
Obs	72745	72745	72745	72745	72745	72745	72745
PCC FE	✓	✓	✓	✓	✓	✓	✓
Age FE		✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓
Visit month FE				✓	✓	✓	✓
Child Background					✓	✓	✓
Parents background						✓	✓
Child health at birth							✓

Notes: The outcome is the probability of obtaining antibiotics at an index visit. The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using the 72 745 index visits, i.e., visits at least 6 months apart. The baseline control variables are described in Table 1.1. Parents' variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. Fixed effects include PCC (160), age (6), visit month (12) and year (8). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1.3 shows the linear estimates from the first-stage regression in which the dependent variable is equal to 1 if the child received an antibiotic prescription during the index visit and 0 otherwise. As an indirect test of whether children are quasirandomly assigned to physicians, I add the fixed effects and control variables sequentially. Only covariates that are correlated with physicians' prescription propensity will change the coefficient. In the first column, only the PCC fixed effects are included. The coefficient in column 1 is 0.58, which implies that when the PPP increases by 10 percentage points, the probability that the child receives antibiotics increases by 5.8 percentage points. The point coefficient changes slightly when adding age fixed effects but is barely changed by the inclusion of control variables in the subsequent columns. Importantly, the point estimate is barely affected when including control variables for children's health at birth. This is reassuring and consistent with the evidence presented in Table 1.2.

1.5 Results

1.5.1 Short-Term Effects

To capture the effect of antibiotics on short-term outcomes, I begin by investigating the impact of getting a prescription on the probability of having one or many visits within 10, 30, or 90 days relative to the index visit in column 1 of Table 1.4. Columns 2–5 investigate the impact of antibiotics on the probability of having a subsequent visit to any of the four different types of facilities.

Table 1.4: Types of revisits

	Panel A: Type of revisit $\leq$ *10 days*				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
OLS: Antibiotics, no X	0.0098***	0.0160***	0.0017	-0.0047***	-0.0032***
	(0.0035)	(0.0030)	(0.0011)	(0.0006)	(0.0006)
OLS: Antibiotics	0.0091***	0.0148***	0.0010	-0.0047***	-0.0020***
	(0.0033)	(0.0029)	(0.0011)	(0.0006)	(0.0006)
RF: Physician Propensity	0.0224	0.0047	0.0084*	0.0078	0.0015
	(0.0161)	(0.0136)	(0.0046)	(0.0047)	(0.0030)
IV: Any Antibiotics	0.0396	0.0083	0.0149*	0.0137*	0.0027
	(0.0281)	(0.0240)	(0.0081)	(0.0082)	(0.0052)
mean of dep. var.	.123	.0981	.0108	.00786	.00582
	Panel B: Type of revisit $\leq$ *30 days*				
OLS: Antibiotics, no X	0.0338***	0.0329***	0.0097***	-0.0050***	-0.0038***
	(0.0049)	(0.0044)	(0.0015)	(0.0008)	(0.0007)
OLS: Antibiotics	0.0345***	0.0326***	0.0087***	-0.0049***	-0.0019**
	(0.0046)	(0.0041)	(0.0015)	(0.0008)	(0.0007)
RF: Prescription Propensity	0.0663***	0.0355*	0.0203***	0.0093*	0.0011
	(0.0204)	(0.0181)	(0.0056)	(0.0055)	(0.0037)
IV: Any Antibiotics	0.1175***	0.0629**	0.0360***	0.0165*	0.0020
	(0.0350)	(0.0316)	(0.0101)	(0.0095)	(0.0064)
mean of dep. var.	.21	.171	.0197	.0116	.00805
	Panel C: Type of revisit $\leq$ *90 days*				
OLS: Antibiotics, no X	0.0454***	0.0430***	0.0136***	-0.0067***	-0.0045***
	(0.0053)	(0.0049)	(0.0021)	(0.0011)	(0.0009)
OLS: Antibiotics	0.0491***	0.0451***	0.0121***	-0.0064***	-0.0018**
	(0.0049)	(0.0045)	(0.0021)	(0.0011)	(0.0009)
RF: Prescription Propensity	0.1038***	0.0691***	0.0289***	0.0103	-0.0046
	(0.0256)	(0.0233)	(0.0089)	(0.0065)	(0.0046)
IV: Any Antibiotics	0.1837***	0.1224***	0.0512***	0.0182	-0.0081
	(0.0447)	(0.0409)	(0.0159)	(0.0113)	(0.0082)
mean of dep. var.	.354	.285	.0375	.0194	.0124
Observations	72745	72745	72745	72745	72745

Notes: Fixed effects include PCC (160), age (6), visit month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using the 72 245 index visits, i.e., visits at least 6 months apart. The baseline controls are added to all rows except row 1 and contain all the variables described in Table 1.1 and defined in Appendix C1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

I begin by focusing on the reduced-form estimates. It is important to note that the reduced-form impact reflects the causal impact of physician behavior, under the assumption of quasi-random assignment, on the outcomes. For all three time frames, physician prescribing behavior is associated with an increase in the probability of the child having at least one subsequent visit. For the shortest time frame, this is driven by visits to out-of-hours PCCs, and for 30 and 90 days, it is driven by visits to both in- and out-of-hours PCCs. There is no effect on visits to emergency units and hospitalizations. With regard to magnitudes, being assigned to a physician who is 10 percentage points more prescription prone increases the probability of a revisit within 30 days by 0.663 percentage points. An alternative interpretation of the reduced form is that going from the least prescribing (a change equal to 46 percentage points: 1st percentile, -0.19, to 99th percentile, 0.27) to the most leads to an increase in having a revisit within 30 days and 90 days of 3.0 percentage points (14.5% of the mean) and 4.8 percentage points (13.5% of the mean), respectively.

The positive impact on revisits is driven primarily by visits to primary care units and to a lesser extent by visits to hospitals, either to an emergency unit or for hospitalizations. The estimates on out-of-hours PCCs are large with respect to the low mean; however, it should be noted that they have fairly wide confidence intervals, which makes it difficult to obtain a precise estimate for this rare outcome. For example, the lower bound of the confidence interval at the 95% level implies a reduced-form point estimate of 0.09 percentage points (45.7% of the mean) when focusing on visits within 30 days and 1.1 percentage points (30.4% of the mean) for 90 days. Some of the estimated point estimates in the robustness section also imply slightly smaller point estimates, though similar, positive, and statistically significant. We can interpret the reduced-form estimate as the intent-to-treat (ITT) effect (Angrist and Pischke, 2008).

Turning to the OLS and IV estimates, it is informative to compare OLS with and without control variables. They are very similar, but the raw OLS estimates are slightly larger in absolute numbers. There is a positive correlation between OLS and primary care visits and a negative correlation with hospital visits. Shifting focus to the IV estimates, which use quasirandom variation in the supply of antibiotics via PPP, the coefficient is consistently larger than with the OLS estimates. For example, in column 1 of Table 4, the OLS coefficient with controls shows that getting an antibiotic prescription increases the probability of a revisit within 30 days by 3.5 percentage points (16.4% of the mean), while the corresponding IV estimate has a magnitude of 11.8 percentage points (56% of the mean). The difference between the OLS and IV could be due to the OLS being biased. Another explanation could be that there is treatment heterogeneity with particularly large effects of antibiotics for individuals at the margin so that the ATE from the OLS estimation is different from the LATE.

A second reason that IV $>$ OLS may be that there is something special about the instrument that is driving the results (for example, if high-prescribing physicians are also more likely to schedule revisits or simply are worse physicians). If this is the case, it would violate the exclusion restriction, which is crucial for the interpretation of the IV as a causal estimate. If high-prescribing physicians are behaving differently from low-prescribing physicians, and this is a channel that affects the outcomes, the exclusion restriction is violated. However, to the extent that high-prescribing physicians are positively correlated with both antibiotics and short-term outcomes, the IV should be upward biased. In that case, the IV is an upper bound. I explore the plausibility of the exclusion restriction in section 1.5.6.

Consistently throughout Table 1.4, it is clear that being prescribed antibiotics, when taking differences in unobservables into account via the instrumentation, leads to a statistically significant increase in the probability of revisits to both in- and out-of-hours PCCs while not having a pronounced impact on the rare events of emergency visits and hospitalizations. This is an interesting result that lines up with the idea that getting antibiotics leads to a preference for specific doctors. There may be multiple explanations for why being prescribed antibiotics increases the probability of revisits. From a policy perspective, it is important to discern whether revisits by children who received antibiotics were because they became more sick.[18] On the other hand, getting access to antibiotics in a restrictive setting may lead the child or parent to experience the help they received from the health care system differently. Obtaining a prescription may increase the feeling of getting "real help." If this is the explanation for the increase in revisits, it calls for different types of interventions relative to children becoming sick.

Next, I ask whether antibiotics have an impact on changing health care providers. Does the prescription decision by the physician change the patient's behavior? Table 1.5 shows the results on the probability of switching PCCs.

[18] Antibiotics could have side effects that would require reexamination by a physician or a change in the type of antibiotic drug, such as from narrow to broad spectrum.

Table 1.5: Change of providers

	(1) Visit unit$_t$ $\neq$ visit_unit$_{t+1}$	(2) Visit unit_t$\neq$ visit_unit$_{t+3}$
OLS: Any Antibiotics, no X	0.0079**	0.0130***
	(0.0033)	(0.0045)
OLS: Any Antibiotics	0.0094***	0.0237***
	(0.0031)	(0.0037)
RF: Prescription Propensity	0.0287	0.0242
	(0.0193)	(0.0239)
IV: Any Antibiotics	0.0507	0.0429
	(0.0343)	(0.0422)
Observations	72745	72745
mean of dep. var.	.125	.227

Notes: The outcome is a binary variable equal to 1 if the first post-index visit is to a different unit relative to the index visit (column 1) or any of the three following visits (column 2). Fixed effects include PCC (160), age (6), visit month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using the 72 745 index visits, i.e., visits at least 6 months apart. The baseline controls are added to all rows except row 1 and contain all the variables described in Table 1.1 and defined in Table C1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In the first column, the dependent variable is equal to 1 if the first visit following the index visit is at a different unit (this includes out-of-hours PCCs and hospitals). The dependent variable in the second column is equal to 1 if any of the three following visits are at a unit different from the index visit unit. There is a positive correlation between antibiotics and the probability of changing PCCs when focusing on the OLS estimates, but this goes away when taking causality into account and focusing on the estimates using physician prescribing behavior as the source of variation and the IV estimates. Thus there is no evidence of antibiotic-shopping behavior.

1.5.2 Medium-Term Effects

To capture the effect of consuming antibiotics on the health of the child more broadly, I investigate the effect of the treatment on total number of visits after the initial index visit. The maximum length of time here is 6 months.

Table 1.6: Total visits within 6 months

	(1) Any	(2) In hours	(3) Out hours	(4) Emergency	(5) Hospital
OLS: Antibiotics, no X	0.1344***	0.1008***	0.0234***	-0.0092***	-0.0073***
	(0.0211)	(0.0176)	(0.0039)	(0.0019)	(0.0018)
OLS: Antibiotics	0.1948***	0.1424***	0.0226***	-0.0046**	-0.0068***
	(0.0200)	(0.0167)	(0.0038)	(0.0018)	(0.0018)
RF: Physician Propensity	0.3151***	0.1861**	0.0467***	-0.0157	0.0186*
	(0.1045)	(0.0833)	(0.0140)	(0.0095)	(0.0111)
IV: Any Antibiotics	0.5580***	0.3295**	0.0826***	-0.0278	0.0330*
	(0.1823)	(0.1468)	(0.0248)	(0.0170)	(0.0196)
Observations	72745	72745	72745	72745	72745
mean of dep. var.	1.45	1.02	.0706	.0226	.0337

Notes: The outcome is the total number of visits 6 months after the index visit. Fixed effects include PCC (160), age (6), visit month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The baseline controls are added to all rows except row 1 and contain all the variables described in Table 1.1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The OLS reveals a positive correlation between number of visits, overall and for both in- and out-of-hours PCCs, irrespective of controlling for covariates. There is a significant negative correlation between antibiotics and emergency and inpatient care. However, both the reduced-form and IV estimates suggest a significant increase in the total number of visits to PCCs. There is also a marginally significant increase in number of hospitalizations. This could be a consequence of more involvement with the health care system in general.[19]

Finally, I test whether an antibiotic prescription affects the probability of obtaining a certain diagnosis at a post-index visit, at any of the following visits within 6 months.[20]

[19] PCC physicians write referrals to hospitals and more specialized care.

[20] See Appendix C2–C4 for a detailed list of the ICD-10 codes included for each diagnosis class.

Table 1.7: Diagnoses

	(1) Eczema	(2) Asthma	(3) RTI
	Panel A: Excluding diagnosis at index visit		
OLS: Any antibiotics	0.0001	-0.0015**	0.0250***
	(0.0005)	(0.0006)	(0.0023)
RF: Physician Propensity	-0.0009	0.0016	0.0184*
	(0.0024)	(0.0025)	(0.0099)
IV: Any antibiotics	-0.0017	0.0023	0.0321*
	(0.0036)	(0.0053)	(0.0193)
Observations	72745	72745	72745
mean of dep. var.	.00339	.00387	.0621

Notes: The outcome is the probability of being diagnosed with respective diagnosis at at least one follow-up visit. Fixed effects include PCC (160), age (6), visit month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using the 72 245 index visits, i.e., visits at least 6 months apart. The baseline controls are added to all rows and contain all the variables described in Table 1.. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1.7 reveals that there is a positive correlation between OLS and the rare diagnosis of asthma.[21] This goes away when focusing on the reduced-form and IV estimates. However, there is a large, albeit only marginally statistically significant effect of getting antibiotics on the probability of being diagnosed with a condition within the RTI group. This could be because high-prescribing physicians tend to see children with more respiratory symptoms or because treatment with antibiotics has a negative effect on susceptibility to infections, which would make patients more prone to further illnesses. I investigate this more thoroughly in the next section.

1.5.3 Conditional on Respiratory Tract Infections

Next, I restrict the analysis to include only those visits for which the child was diagnosed with an RTI at the index visit. RTIs are the most common reason for children to visit a primary care center (Hedin, 2015). The diagnoses included in this group are defined in Appendix C2.[22] The high incidence of RTIs, combined

[21] As these include only diagnoses given after the initial index visits, the share of diagnoses here is lower than the incidence of these diagnoses in the sample population.

[22] The most common diagnoses within this group in the data are acute upper respiratory infections (J00–J06), followed by otitis media and mastoiditis (H65–H70).

with the fact that the largest share of antibiotics is prescribed for this category, makes the restriction to this group relevant for the purposes of this paper.

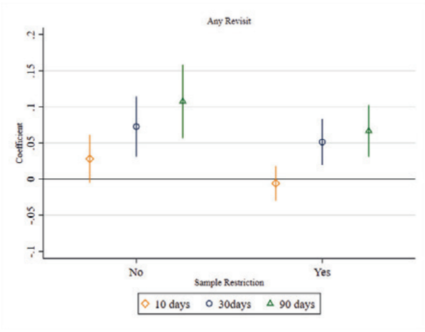
The positive impact of antibiotics on short-term revisits in Table 1.4 can be driven by several things. As shown in Table 1.2, the patient mix of a physician is not likely to be an issue, since we can see that physician behavior is unrelated to predetermined covariates. However, it could still be that the case mix (i.e., mix of symptoms, severity of illness) differs among physicians. For example, if more experienced physicians always receive more children with more symptoms, then both their prescription propensity and the positive impact of antibiotics can be driven by physicians' case mix. By including the fixed effects, I adjust for the case mix based on time, age, and season, but I have not taken diagnoses into account thus far. By restricting the sample to children having the same diagnosis, I limit the concern of case mix driving the results, since patients within this group are expected to display more or less similar symptoms. To the extent that case mix is an important mechanism, defining the instrument in this subsample is more relevant.

Importantly, though, the downside of making this restriction, and the reason for why this is not the main analysis, is that the probability of a physician giving a specific diagnosis to a patient may be correlated with the choice of whether to prescribe antibiotics. In the causal inference framework, this type of variable is called a bad control (Angrist and Pischke, 2008). This implies that there can be children who do not get a diagnosis within this group even though they display similar symptoms as patients who are diagnosed with RTIs, which results in a selected sample. This should be kept in mind when interpreting the results.[23]

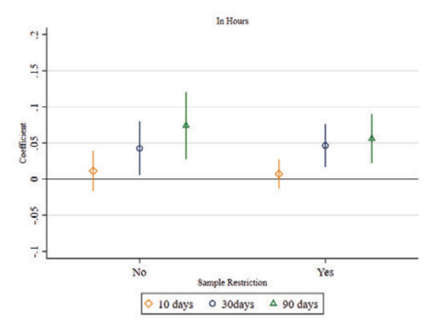
To facilitate a comparison between Table 1.4, the short-term outcomes using the full sample, and the same outcomes for the restricted sample, I present the coefficients from the reduced-form estimation in Figure 1.5. The full corresponding table can be found in Appendix Table A1. There I also present the reestimated first stage. The first stage in this subsample is precisely estimated and has a coefficient of approximately 0.6.

[23] In Appendix Table A4, I also test the main results to include all diagnosis fixed effects as baseline controls. The diagnoses are grouped into 117 chapters defined by WHO, and the reduced-form estimates are remarkably similar to those presented in Tables 1.3 and 1.4. The IV estimates are slightly larger, since the first-stage coefficient is smaller in the sample conditioned on diagnosis chapters, 0.38 versus 0.56 in the unrestricted sample.

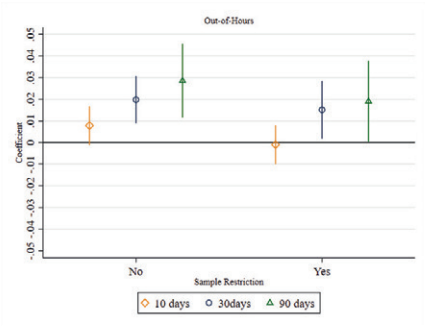
Figure 1.5: Comparison between unconditional and conditional samples



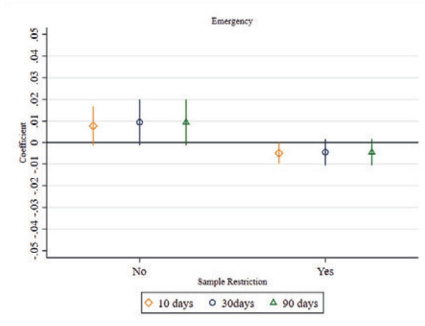
((a))



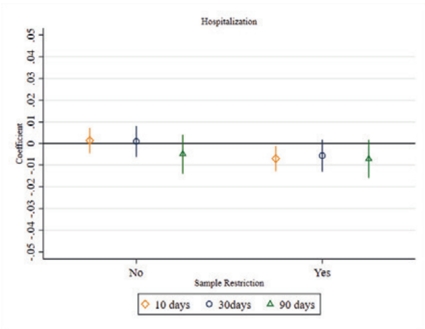
((b))



((c))



((d))



((e))

Notes: The figure shows the reduced-form estimates from regressing physician propensity on the health care outcomes, where the coefficient is displayed on the y-axis. Sample restriction = yes if the sample is restricted to those with an RTI diagnosis. The confidence intervals are at the 95% level of significance.

The overall takeaway from the plots in Figure 1.5 is that the point estimates are overall remarkably similar, though closer to zero for the restricted sample. One should keep in mind that the confidence bands for the restricted sample are tighter because the instrument is more precisely estimated as the share of antibiotic prescriptions is higher (36% versus 19% in the unrestricted sample).

The sets of estimates are not statistically different from each other, as the confidence intervals overlap for almost every coefficient. With regard to 10, 30 and 90 days, the point estimates are remarkably similar and larger than the full-sample point estimate for visits overall and to in-hours and out-of-hours PCCs. The largest difference between the two sets of coefficient is when studying the impact of antibiotics on having a visit to an emergency unit, where restricting the sample yields statistically significant negative effects of a re-visit within 10 days, while the full sample showed positive and insignificant estimates. The restricted sample also displays a reduction in hospitalizations within 10 days which is not found for the full sample of index visits.

Table 1.7 showed that antibiotics did have a sizable impact on the probability of being diagnosed with an RTI infection at a recurring visit. I replicate that table in Table 1.8 using the restricted sample.

Table 1.8: Diagnoses, restricted sample

	(1) Eczema	(2) Asthma	(3) RTI
	Panel A: Excluding diagnosis at index visit		
OLS: Any antibiotics, no **X**	0.0007	-0.0014*	0.0161***
	(0.0006)	(0.0008)	(0.0035)
OLS: Any antibiotics	0.0007	-0.0014*	0.0161***
	(0.0006)	(0.0008)	(0.0035)
RF: Physician Propensity	-0.0023	-0.0027	-0.0059
	(0.0020)	(0.0025)	(0.0110)
IV: Any antibiotics	-0.0037	-0.0043	-0.0083
	(0.0030)	(0.0042)	(0.0183)
Observations	28158	28158	28158
mean of dep. var.	.00262	.00417	.0887

Notes: The outcome is the probability of being diagnosed with any of the diagnoses presented in the columns at at least one of the follow-ups visits. The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using 28 158 patients diagnosed with RTI at the index visit, i.e., visits at least 6 months apart. The baseline controls are added to all rows except row 1 and contain all the variables described in Table 1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In the restricted sample, where all children are diagnosed with RTI at the index visit, I do not find that physician behavior significantly impacts the probability of getting this diagnosis. To sum up, the similarity between the results from the full sample and those from the restricted sample is assertive in that it is not the case mix that is driving the main results. There is some indication however, that the null, or even slightly increased probability, of hospitalizations and emergency visits in the very short-term is sensitive to how the sample is constructed and should therefore be interpreted with more caution.

1.5.4 Treatment Heterogeneity

One of the distinct features of my data set is that the access to data on distance to PCCs can be used to study whether the short-term treatment effect depends on the distance to the PCC. The barrier to a physical revisit can, for example, be affected by the distance to the PCC. This is interesting from an organizational perspective. Indeed, one of the aims of the Free Choice Act was to increase the availability and quality of medical care within the primary care sector. Government reports show that more PCCs opened up following the reform, but that these were located in areas with higher income, while areas with lower average income experienced closures of public PCCs (Riksrevisionen, 2014). Thus investigating heterogeneity from a geographic perspective is important both because individuals undecided about making a physician visit could be affected by proximity and because distance to a PCC likely reflect underlying differences in sociodemographic backgrounds. The mean distance to the closest PCC is 1.9 km, with a minimum of 6 meters and a maximum of 17 km. The majority of patients thus live relatively close to a PCC.[24] Since distance is a continuous measure, I divide the sample into quartiles. I present the results in Appendix Figure A3–A6. There is no clear pattern with regard to treatment based on distance to the PCC.

In Appendix Figure A7–A8, I present treatment heterogeneity based on the type of antibiotic drug prescribed.[25] The effects on short-term revisits are slightly larger for the second-line (broad-spectrum) antibiotic group, probably because these drugs have a greater impact on gut microbiota and are associated with more side effects (Jourdan et al., 2020). However, due to the low incidence of second-line drugs in the sample, the confidence intervals are very large.

[24]Note that the closest PCC may or may not be the unit of visit at either the index or follow-up visits.

[25]Following Huang and Ullrich (2021), I group the drugs by ATC 3, where J01C are penicillins, J01D (cephalosporins), J01E (sulfonamides and trimethoprim), J01F (macrolides, lincosamides, and streptogramins), and J01M (quinolones) are considered second-line drugs. These are typically associated with poorer practice.

1.5.5 Monotonicity

An assumption in the IV estimation, as discussed in section 1.4, is monotonicity. Monotonicity requires the instrument to work the same way across all visits. If a patient is prescribed antibiotics at an index visit, then that patient must also be prescribed antibiotics by an even higher-prescribing physician. Monotonicity is required for the IV estimates to be interpreted as local average treatment effects. A testable implication of this assumption is that the first-stage coefficients should be positive for all subsamples (Bhuller et al., 2020). I implement this test by continuing to construct the instrument using the full sample of visits, but I run a regression separately for each subgroup using only index visits. The sample mean is the share in that subgroup that gets antibiotics at the index visit. The first-stage estimates are presented in Table 1.9.

Table 1.9: First stage in subgroups

	Sample Mean	Any Antibiotics	Any Antibiotics
Female	0.197	0.586	0.587
		0.031	0.031
Male	0.191	0.546	0.547
		0.029	0.029
Born Abroad	0.198	0.570	0.571
		0.045	0.046
Born in Sweden	0.193	0.587	0.562
		0.025	0.025
Unemployed^p	0.180	0.558	0.560
		0.042	0.042
Employed^p	0.197	0.564	0.564
		0.025	0.025
Married^m	0.195	0.563	0.564
		0.030	0.030
Not Married^m	0.193	0.562	0.563
		0.028	0.028
Immigrant^p	0.198	0.570	0.571
		0.045	0.046
Native	0.193	0.561	0.562
		0.025	0.025
Income$>$ 50%	0.199	0.552	0.552
		0.030	0.030
Income$\leq$ 50%	0.189	0.574	0.574
		0.028	0.028
All	0.194	0.564	0.565
		0.023	0.023
Controls			✓

Notes: Fixed effects include PCC (160), age (6), birth month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The baseline controls, added in column 2, are described in Table 1. The control variables are measured the year prior to childbirth. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The pairwise difference across the columns is the inclusion of the control variables. The coefficient is similar, and positive, for all the subgroups, which is consistent with the monotonicity assumption. The test of monotonicity also has some bearing on the characteristics of compliers. If the first-stage coefficient is larger (smaller) for a certain subgroup, this implies that the concentration of compliers is higher (lower) in that subgroup (Norris et al., 2021). Parents with income below the median, children born abroad, and females are groups with a relatively larger share of compliers, but the differences are small.

1.5.6 Exclusion Restriction

For the IV estimates to have causal interpretation, the health care utilization outcomes are only allowed to be affected by antibiotics via physicians' prescription behavior. If physicians with certain antibiotic prescribing behavior treat patients differently in other aspects as well, it would violate the exclusion restriction. This assumption is inherently untestable, but I will do three tests to shed some light on its plausibility and provide some justification of its validity. First, I conduct a placebo test wherein I test the effect of the instrument on the health care outcomes from Table 1.4, using a restricted subsample treated for conditions that rarely require antibiotics. I rank the diagnosis groups (based on the first letter of the ICD code) by treatment incidence. I restrict the sample to keep only those diagnosis groups for which antibiotics are prescribed in less than 5% of the visits. The mean value of antibiotic prescriptions in this subsample is 2.6%.

Table 1.10: Placebo visits

	Panel A: Type of revisit $\leq$ 10 *days*				
	(1) Any	(2) In hours	(3) Out hours	(4) Emergency	(5) Hospital
RF: Physician Propensity	-0.0085	-0.0338	0.0062	0.0190*	0.0002
	(0.0271)	(0.0234)	(0.0083)	(0.0106)	(0.0058)
mean of dep. var.	.141	.106	.0102	.0173	.00791
	Panel B: Type of revisit $\leq$ 30 *days*				
RF: Prescription Propensity	-0.0062	-0.0320	0.0062	0.0236*	-0.0040
	(0.0356)	(0.0320)	(0.0105)	(0.0120)	(0.0065)
mean of dep. var.	.227	.179	.0173	.0211	.00964
	Panel C: Type of revisit $\leq$ 90 *days*				
RF: Prescription Propensity	0.0140	-0.0153	0.0093	0.0273**	-0.0073
	(0.0389)	(0.0369)	(0.0147)	(0.0137)	(0.0086)
mean of dep. var.	.358	.282	.0343	.0283	.0134
Observations	21328	21328	21328	21328	21328

Notes: The outcomes is the probability of revisits for a restricted subsample of index visits, namely those with diagnoses for which antibiotics is prescribed in less than 5%. Control variables are included. The control variables are measured the year prior to childbirth. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As shown in Table 1.10, physicians' antibiotic prescribing behavior has little effect on the probability of having an additional visit to all types of health care facilities. The exception is that the instrument has a small, marginally statistically significant, effect on the probability of having a revisit to an emergency unit within 30 and 90 days. This could indicate that high prescribers are relatively worse physicians in the sense that they do not provide the best possible patient care. However, in Appendix Table A2, I present the placebo for the total number of visits, and there is no positive effect on the total number of emergency visits, albeit on the total number of hospitalizations.

Second, I follow (Sievertsen et al., 2021) and include physicians' prescribing behavior for other groups of prescription drugs as control variables. As initially noted by Mueller-Smith (2015), one way to assess the exclusion restriction is if one dimension of judges' behavior (sentencing stringency) is affected by controlling for additional dimensions of judges' behavior. Both the first-stage coefficient and the short-term effects are very robust to controlling for physicians' prescribing behavior for other drugs. The first column in Panel I and the first row in Panel II-A-II.C replicates the results from Table 1.4. Treatment with drugs other than antibiotics are not driving the short-term effects.

Table 1.11: Controls for co-treatment

	Panel I: First Stage			
	(1)	(2)	(3)	(4)
Prescribing behavior, AB	0.5647***	0.5585***	0.5648***	0.5600***
	(0.0227)	(0.0231)	(0.0227)	(0.0227)
mean of dep. var.	.194	.194	.194	.194
F-stat	55.8	56.1	54.8	54.6
Prescribing behavior, respiratory		✓		
Prescribing behavior, skin			✓	
Prescribing behavior, eyes				✓

	Panel II.A: Type of revisit $\leq$ *10 days*				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
RF: Physician Propensity (PP)	0.0224	0.0047	0.0084*	0.0078	0.0015
	(0.0161)	(0.0136)	(0.0046)	(0.0047)	(0.0030)
PP Inc respiratory drugs	0.0218	0.0038	0.0085*	0.0078	0.0017
	(0.0160)	(0.0135)	(0.0047)	(0.0047)	(0.0030)
PP Inc skin drugs	0.0226	0.0049	0.0085*	0.0077	0.0015
	(0.0161)	(0.0136)	(0.0046)	(0.0047)	(0.0030)
PP Inc eyes and ears drugs	0.0200	0.0035	0.0078*	0.0079	0.0008
	(0.0162)	(0.0136)	(0.0046)	(0.0048)	(0.0029)
mean of dep. var.	.123	.0981	.0108	.00786	.00582

	Panel II.B: Type of revisit $\leq$ *30 days*				
RF: Physician Propensity (PP)	0.0663***	0.0355*	0.0203***	0.0093*	0.0011
	(0.0204)	(0.0181)	(0.0056)	(0.0055)	(0.0037)
PP Inc respiratory drugs	0.0656***	0.0348*	0.0205***	0.0091*	0.0011
	(0.0202)	(0.0180)	(0.0057)	(0.0055)	(0.0037)
PP Inc skin drugs	0.0663***	0.0355*	0.0204***	0.0093*	0.0011
	(0.0205)	(0.0182)	(0.0057)	(0.0055)	(0.0037)
PP Inc eyes and ears drugs	0.0647***	0.0346*	0.0200***	0.0095*	0.0006
	(0.0202)	(0.0180)	(0.0056)	(0.0055)	(0.0036)
mean of dep. var.	.21	.171	.0197	.0116	.00805

	Panel II.C: Type of revisit $\leq$ *90 days*				
RF: Physician Propensity (PP)	0.1038***	0.0691***	0.0289***	0.0103	-0.0046
	(0.0256)	(0.0233)	(0.0089)	(0.0065)	(0.0046)
PP Inc respiratory drugs	0.1004***	0.0648***	0.0297***	0.0103	-0.0045
	(0.0251)	(0.0228)	(0.0089)	(0.0065)	(0.0047)
PP Inc skin drugs	0.1037***	0.0692***	0.0289***	0.0102	-0.0046
	(0.0257)	(0.0234)	(0.0089)	(0.0065)	(0.0046)
PP Inc eyes and ears drugs	0.1035***	0.0694***	0.0284***	0.0104	-0.0046
	(0.0256)	(0.0232)	(0.0089)	(0.0065)	(0.0046)
Observations	72745	72745	72745	72745	72745
mean of dep. var.	.354	.285	.0375	.0194	.0124

Notes: "Respiratory" drugs include all prescription drugs in ATC group R, excluding R02AB and R05X. "Skin" contains all prescription drugs in ATC group D. "Eyes and ears" contains all prescription drugs in ATC group S. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Third, I test whether the positive effect on revisits is mechanical—if, for example, a high-prescribing physician is more likely to schedule a follow-up visit. In Table 1.12, I test whether the subsequent visit following the index visit is to the same physician. Only 5.1% of consecutive visits are to the same physician.

Table 1.12: Physician continuity

	Physician-continuity		
	(1) OLS	(2) RF	(3) IV
Coefficient	-0.0004	-0.0322***	-0.0571***
	(0.0019)	(0.0083)	(0.0150)
Observations	72745	72745	72745
mean of dep. var.	.0506	.0506	.0506

Notes: The outcome is a binary variable equal to 1 if the next visit is to the same physician as the index visit. Fixed effects include PCC (160), age (6), birth month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The baseline controls are included and described in Table 1.1. The control variables are measured the year prior to childbirth. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

From the results in Table 1.12, I can conclude that it is not the case that high-prescribing physicians simply schedule more revisits. On the contrary, both the reduced form and the instrumental variable show a statistically significant negative effect on the probability of seeing the same physician.

1.6 Robustness

Throughout the paper, I have referred to the main sample consisting of index visits, which are defined as visits with no prior visits within 180 days. Here I elaborate on the definition of index visits and present results using visits occurring closer together, with no prior visits within 90 days, shown in Table 1.13, and farther apart, with at least 365 days between visits, shown in Table 1.14.

Table 1.13: Index visits $\geq$ 90 days

	(1) Any	(2) In hours	(3) Out hours	(4) Emergency	(5) Hospital
Panel A: Type of revisit $\leq$ 10 days					
OLS: Antibiotics	0.0072**	0.0134***	0.0006	-0.0050***	-0.0018***
	(0.0031)	(0.0027)	(0.0009)	(0.0005)	(0.0005)
RF: Physician Propensity	0.0171	0.0051	0.0080**	0.0060*	-0.0020
	(0.0123)	(0.0104)	(0.0037)	(0.0031)	(0.0020)
IV: Any Antibiotics	0.0334	0.0100	0.0156**	0.0118*	-0.0039
	(0.0238)	(0.0203)	(0.0072)	(0.0060)	(0.0039)
mean of dep. var.	.129	.104	.0117	.00729	.00549
Panel B: Type of revisit $\leq$ 30 days					
OLS: Antibiotics	0.0380***	0.0365***	0.0085***	-0.0052***	-0.0018***
	(0.0041)	(0.0039)	(0.0012)	(0.0006)	(0.0006)
RF: Physician Propensity	0.0506***	0.0288**	0.0179***	0.0071*	-0.0033
	(0.0150)	(0.0142)	(0.0046)	(0.0037)	(0.0024)
IV: Any Antibiotics	0.0989***	0.0564**	0.0350***	0.0140**	-0.0064
	(0.0283)	(0.0271)	(0.0090)	(0.0071)	(0.0046)
mean of dep. var.	.226	.185	.0221	.0112	.00795
Panel C: Type of revisit $\leq$ 90 days					
OLS: Antibiotics	0.0536***	0.0489***	0.0130***	-0.0067***	-0.0017**
	(0.0042)	(0.0040)	(0.0017)	(0.0009)	(0.0008)
RF: Physician Propensity	0.0858***	0.0633***	0.0238***	0.0060	-0.0073**
	(0.0188)	(0.0183)	(0.0069)	(0.0045)	(0.0032)
IV: Any Antibiotics	0.1679***	0.1237***	0.0466***	0.0118	-0.0142**
	(0.0359)	(0.0351)	(0.0135)	(0.0088)	(0.0063)
mean of dep. var.	.387	.312	.0423	.0197	.0124
Observations	111642	111642	111642	111642	111642

Notes: This table replicates the results presented in Table 1.4, using a washout period of 90 days instead of 180. Fixed effects include PCC (160), age (6), birth month (12) and year (8). Control variables are always included and described in Table 1.1 and Appendix C1. The control variables are measured the year prior to childbirth. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1.14: Index visits ≥ 365 days

	Panel A: Type of revisit≤ 10 days				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
OLS: Antibiotics	0.0180***	0.0232***	0.0021	-0.0059***	-0.0013
	(0.0046)	(0.0039)	(0.0018)	(0.0011)	(0.0010)
RF: Physician Propensity	0.0317	0.0164	0.0048	0.0057	0.0048
	(0.0213)	(0.0188)	(0.0067)	(0.0067)	(0.0057)
IV: Any Antibiotics	0.0654	0.0337	0.0099	0.0118	0.0100
	(0.0433)	(0.0386)	(0.0137)	(0.0138)	(0.0116)
mean of dep. var.	.112	.0861	.00944	.00851	.00786
	Panel B: Type of revisit≤ 30 days				
OLS: Antibiotics	0.0354***	0.0344***	0.0085***	-0.0065***	-0.0011
	(0.0064)	(0.0053)	(0.0026)	(0.0013)	(0.0012)
RF: Physician Propensity	0.0710***	0.0406*	0.0182**	0.0111	0.0011
	(0.0258)	(0.0232)	(0.0083)	(0.0072)	(0.0067)
IV: Any Antibiotics	0.1463***	0.0837*	0.0375**	0.0228	0.0023
	(0.0528)	(0.0478)	(0.0171)	(0.0148)	(0.0138)
mean of dep. var.	.189	.151	.0163	.0119	.0107
	Panel C: Type of revisit≤ 90 days				
OLS: Antibiotics	0.0448***	0.0422***	0.0116***	-0.0080***	-0.0010
	(0.0077)	(0.0065)	(0.0035)	(0.0017)	(0.0015)
RF: Physician Propensity	0.0870***	0.0604**	0.0307**	0.0067	-0.0108
	(0.0315)	(0.0273)	(0.0119)	(0.0085)	(0.0079)
IV: Any Antibiotics	0.1792***	0.1245**	0.0632**	0.0138	-0.0223
	(0.0650)	(0.0564)	(0.0247)	(0.0175)	(0.0166)
mean of dep. var.	.315	.248	.0307	.0198	.0161
Observations	32141	32141	32141	32141	32141

Notes: Notes: This table replicates the results presented in Table 1.4,, using a wash-out period of 365 days instead of 180. Fixed effects include PCC (180), age (6), birth month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. Control variables are always included and described in Table 1.1 and Appendix C1. The control variables are measured the year prior to childbirth. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

With regard to robustness, the results presented in Tables 1.13 and 1.14 verify that they are not sensitive to how the index visit is defined. The main difference between the tables in this section and Table 1.4 is the precision, with 90 days giving more precise estimates and 365 days giving less precise estimates, relative to 180 days. This is due to the difference in sample size. In terms of signs and magnitudes, they are very similar, and the main conclusions still hold.

I also test the sensitivity of the results to how the instrument is constructed. First, instead of using the full sample of primary care visits, I reconstruct the instrument using only the sample of index visits. The results are presented in Appendix Table A3. The first-stage coefficient is approximately half the size (0.26) but still strong and robust to the sequential inclusion of control variables. Relative to the instrument using all visits, the accuracy of prescribing propensity is lower using index visits. However, the reduced-form effects are also smaller, and as a consequence, scaling the reduced-form coefficient with a smaller first stage yields results very similar to those presented in Table 1.4. Second, I include diagnosis groups as baseline controls and the results are presented in the Appendix, Table A4. Third, I use only visits from 2012-2017 to account for my slightly skewed sample with regards to age. These results are found in table A5. The results are robust to all of these three modifications.

I also redefine physicians' prescription propensity by using physician identifiers across PCCs. This allows for the same physician to operate at several PCCs over the sampling period.[26] The results from this sensitivity check are presented in Appendix Table A6 and are very similar to those presented in Table 1.4. The first-stage coefficient is slightly increased, from 0.57 to 0.69. The reduced-form coefficients are marginally larger, and the IV estimates are very similar.

1.7 Conclusion

Antibiotics are the most commonly prescribed group of drugs to children but have been on a downward trend since the 1990s in Sweden, which is the setting for this paper. Little is known about the individual consequences of reducing antibiotic consumption. Numerous studies document the benefit of reducing antibiotic prescribing on the societal level (Adda, 2020), and a few document a link between antibiotics and the individual.[27] However, these are generally based on records kept in journals or case studies, which cannot differentiate between unobserved characteristics that affect both antibiotics and the outcomes.

In this paper, I investigate the causal effect of antibiotic consumption in children on downstream health care utilization. My research question is whether treatment with antibiotics affects subsequent health care utilization. The Swedish health care sector has two compelling factors that allow me to investigate this research question. First, antibiotic prescribing here is among the lowest in Europe, allowing me to study the consequences in a setting where access is constrained. Second, the primary sector is characterized by a shortage of specialist general practitioners, which leads to a very low share of residents who have a regular,

[26] Or different individuals that happen to share the same physician 3-letter identifier, see the discussion in Section 1.3

[27] See, for example, Groth et al. (2011), who find that reducing antibiotics as a first-line treatment for children diagnosed with acute otitis media did not have a significant impact on more adverse events, in this case acute mastoiditis, which is a more complicated diagnosis.

personal physician contact that enables an ongoing patient-physician relationship. This leads to the possibility of random assignment between the child and the physician.

I document significant differences between the groups of children who consume and do not consume antibiotics. I take these differences into account by using an instrumental variables approach and use idiosyncratic variation in physicians' antibiotic prescription propensities, within PCCs, as an instrument for the probability of being prescribed antibiotics at an index visit to the PCC.

I also document large differences in physicians' prescription propensities in the primary care sector, using a detailed data set from the third-largest region in Sweden, Scania. I show that the physicians are conditionally randomly assigned to patients, which allows for a causal interpretation. The instrument is a strong predictor of obtaining antibiotics or not. The first-stage estimates show that meeting with a physician who is 10 percentage points more prescription prone significantly increases the probability of antibiotics by 5.7 percentage points.

With regard to health care utilization, I estimate a precise increase in the probability of having at least one additional visit within 30 and 90 days, to both in- and out-of-hours PCCs, and a small, marginally significant increase on visits to emergency units or hospitalizations. Being assigned to a physician who is 10 percentage points more prescription prone increases the probability of a revisit within 30 days by 0.663 percentage points. An alternative interpretation of the reduced form is that going from the least prescribing leads to an increase in having a revisit within 30 days and 90 days of 3.0 percentage points (14.5% of the mean) and 4.8 percentage points (13.5% of the mean), respectively.

To test the plausibility that high-prescribing physicians simply prescribe more antibiotics, and have more revisits, because they meet with sicker children, I restrict the sample to children with the same diagnosis, respiratory tract infection. The immediate positive effect on the very short time window of 10 days goes away when restricting the sample to children, which indicates that the effect might have been due to differences in case mix with respect to symptoms and severity of illness. However, the remaining positive effects on 30 and 90 days are remarkably similar between the unrestricted and restricted samples; the main difference is a significant reduction in the probability of very short-term visits to emergency units in the restricted sample. An analysis of treatment heterogeneity by distance to the primary care center reveals that the effect differs very little along the distance dimension, probably because geographic distance to a PCC is a low barrier, since 80% (of the full population) have less than a five-minute drive to the second-closest PCC (Glenngård, 2015).

Also in the medium term, I show that physicians' antibiotic prescribing practices have a significant positive effect on the total number of both in- and out-of-hours PCC visits, post the initial index visit. Being prescribed antibiotics at the index visit increases the total number of visits by almost 0.56, which is large relative to the mean of 1.5 PCC visits half a year after the index visit. The evidence in this paper strongly shows that the physicians' prescribing practice increases interactions with the primary sector of the health care system. I provide evidence that this is not a supply-side effect, as it is not the physicians who reschedule meetings with patients for whom they prescribed antibiotics.[28] Rather, it seems to be driven by an increased demand for care, but I cannot distinguish whether this is because antibiotics increase infection susceptibility or because patients develop a preference for physicians.

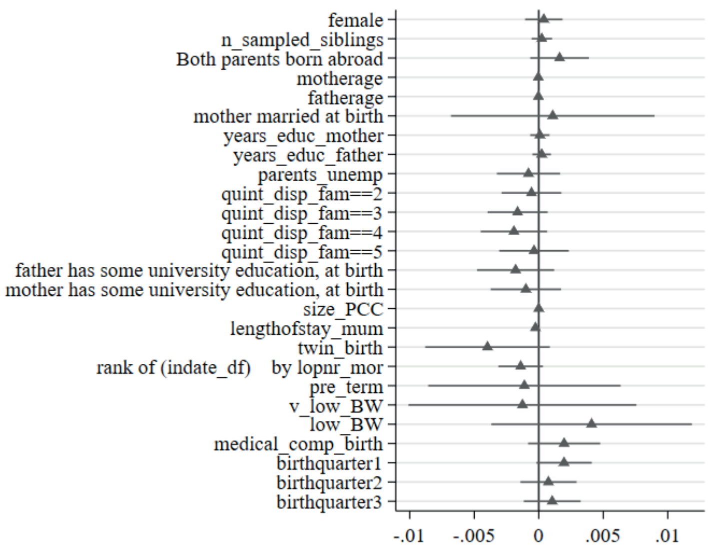
Overall, my results show that being prescribed antibiotics causes an increase in health care utilization in both the short and medium term. The short-term increase in visits should be taken into account by policy-makers, as the decision whether to prescribe antibiotics affects the work burden in an already capacity-constrained sector. Future research should ideally have a longer follow-up period and use more detailed health data (for example, health records or sick leave in school) to assess the causal long-term health implications of consuming antibiotics.

[28]I show in Table 1.5 that the probability of having the subsequent visit with the same doctor as the index visit is actually negatively affected by physicians' prescribing behavior.

Appendices 1

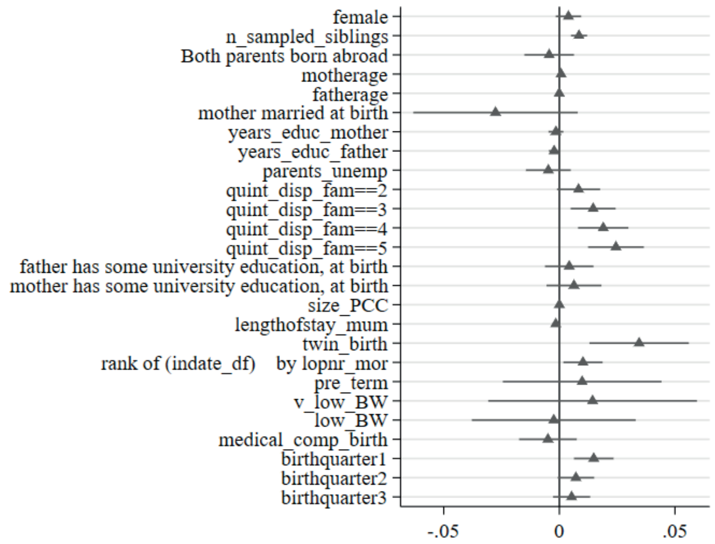
Figures

Figure A1: Coefficient plot on the instrument, randomization test



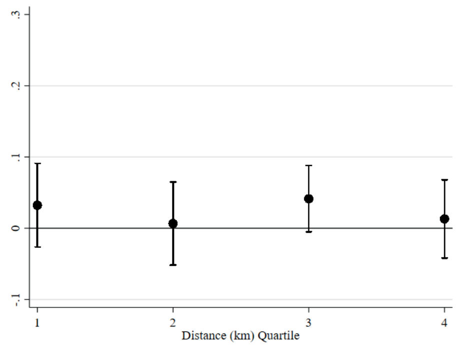
Note: The outcome is the residualized physician propensity to prescribe and the figure displays the coefficient of each control variable. Confidence intervals are at the 95% significance level two-way clustered at PCC and individual level

Figure A2: Coefficient plot on antibiotics

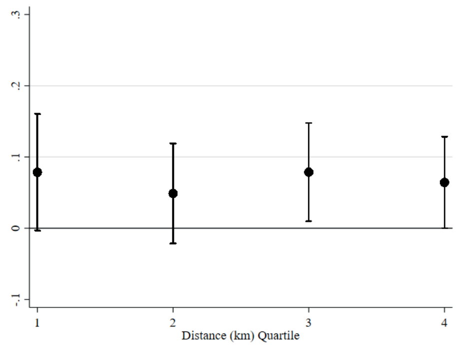


Note: The outcome is the probability of obtaining antibiotics at an index visit and the figure displays the coefficient of each control variable. Confidence intervals are at the 95% significance level two-way clustered at PCC and individual level

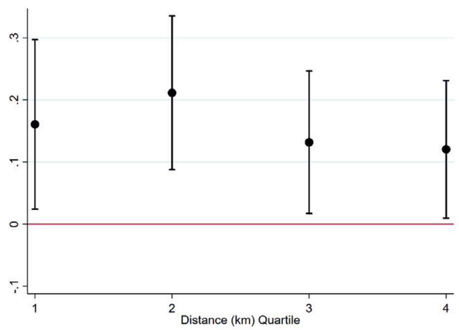
Figure A3: Health care utilization by distance (km) quartile



((a)) Any revisit $\leq$ 10 days



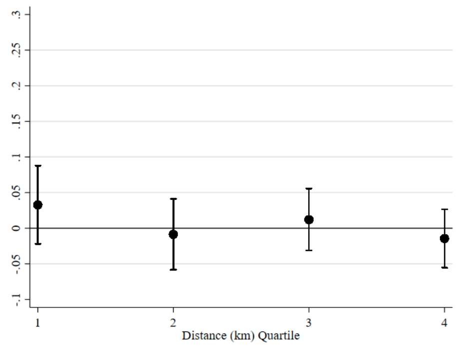
((b)) Any revisit $\leq$ 30 days



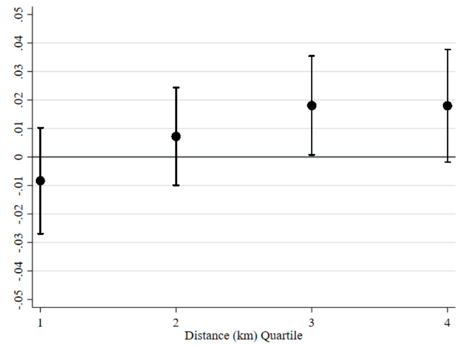
((c)) Any revisit $\leq$ 90 days

Notes: The figure shows the reduced-form coefficient per quartile. The upper and lower bounds are calculated at the 90% level of significance. The quartiles are based on the distribution of distance to the closest PCC.

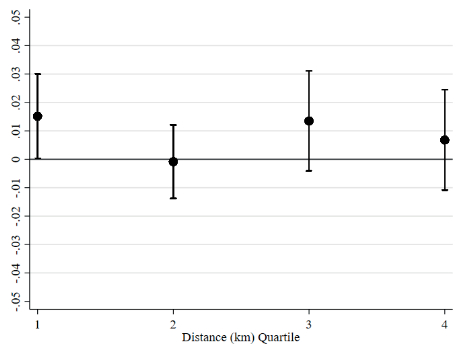
Figure A4: Health care utilization by distance (km) quartile in $\leq$ 10 days



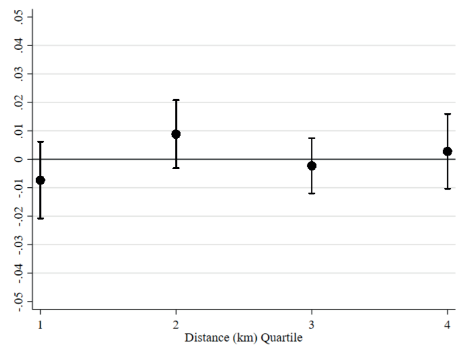
((a)) In-hours PCC



((b)) Out-of-hours PCC



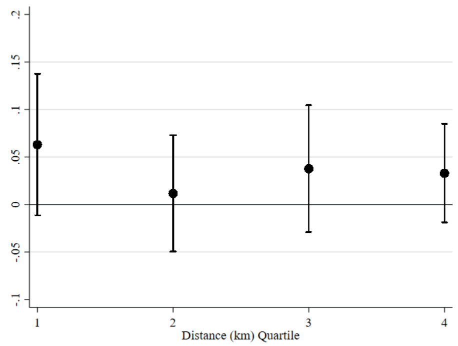
((c)) Emergency



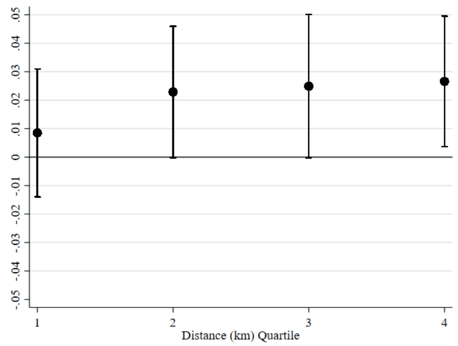
((d)) Hospitalization

Notes: The figure shows the reduced-form coefficient per quartile. The upper and lower bounds are calculated at the 90% level of significance. The quartiles are based on the distribution of distance to the closest PCC.

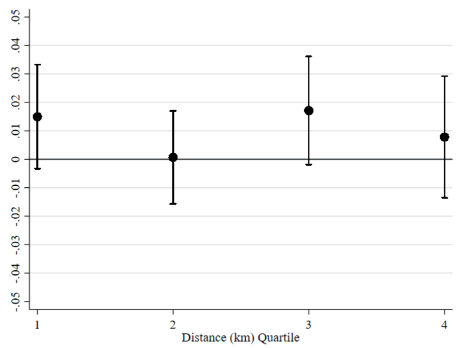
Figure A5: Health care utilization by distance (km) quartile in $\leq$ 30 days



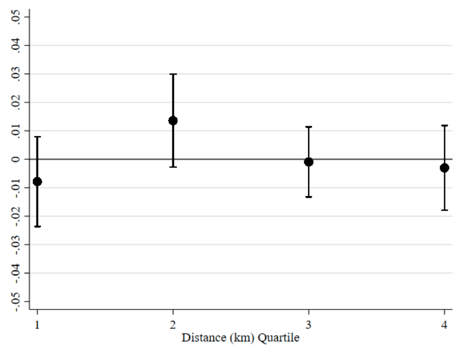
((a)) In-hours PCC



((b)) Out-of-hours PCC



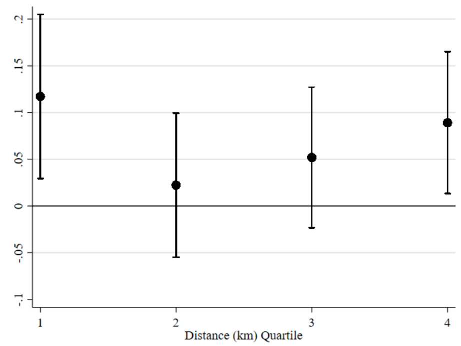
((c)) Emergency



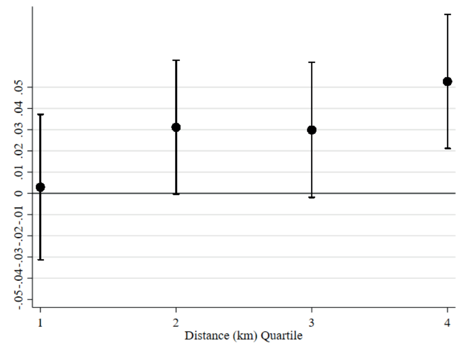
((d)) Hospitalization

Notes: The figure shows the reduced-form coefficient per quartile. The upper and lower bounds are calculated at the 90% level of significance. The quartiles are based on the distribution of distance to the closest PCC.

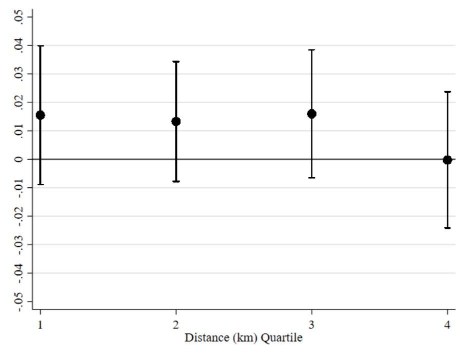
Figure A6: Health care utilization by distance (km) quartile in $\leq$ 90 days



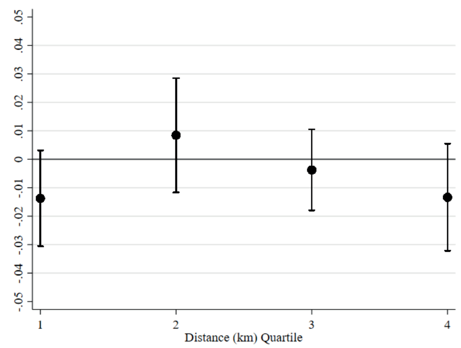
((a)) In-hours PCC



((b)) Out-of-hours PCC



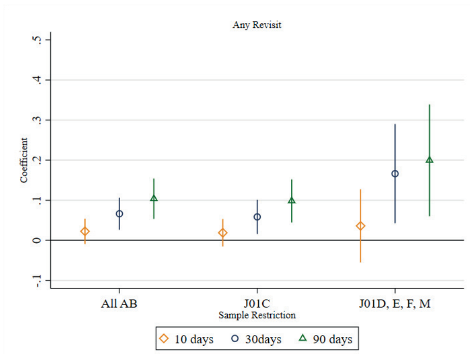
((c)) Emergency



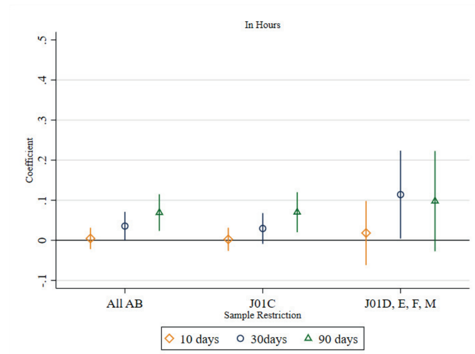
((d)) Hospitalization

Notes: The figure shows the reduced-form coefficient per quartile. The upper and lower bounds are calculated at the 90% level of significance. The quartiles are based on the distribution of distance to the closest PCC.

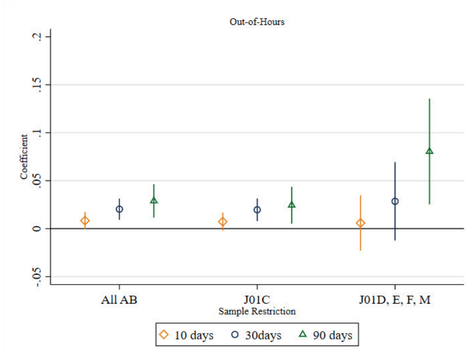
Figure A7: Short-term effects by type of antibiotic drug, PCC



((a)) Any revisit ≤ 10 days



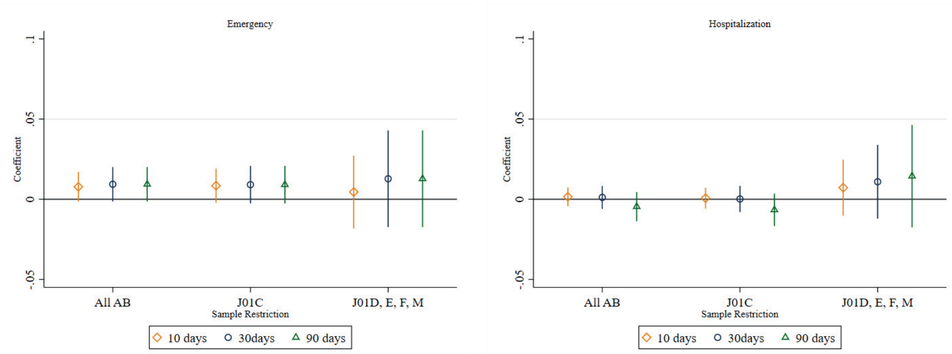
((b)) Any revisit ≤ 30 days



((c)) Any revisit ≤ 90 days

Notes: The figure shows the IV estimate where the drugs are classified into two groups and instrumented with physicians' prescription propensity. The upper and lower bounds are calculated at the 95% level of significance.

Figure A8: Short-term effects by type of antibiotic drug, emergency and hospitalizations



((a)) Any revisit ≤ 90 days

Notes: The figure shows the reduced-form coefficient per quartile. The upper and lower bounds are calculated at the 90% level of significance. The quartiles are based on the distribution of distance to the closest PCC.

Tables

Table A1: Short-term outcomes, restricted sample

	Panel I: First Stage				
	(1)	(2)	(3)		
Physician Propensity	0.6028***	0.6039***	0.6042***		
	(0.0228)	(0.0227)	(0.0227)		
mean of dep. var.	.348	.348	.348		
F-stat	68.7	58.7	53.2		
Obs	28157	28157	28157		
Child Background	✓	✓	✓		
Parents background		✓	✓		
Child health at birth			✓		
	Panel II.A: Type of revisit≤ 10 days				
	(1) Any	(2) In hours	(3) Out hours	(4) Emergency	(5) Hospital
OLS: Antibiotics, no X	-0.0000	0.0061	-0.0015	-0.0004	-0.0043***
	(0.0044)	(0.0039)	(0.0015)	(0.0006)	(0.0009)
OLS: Antibiotics	0.0037	0.0084**	-0.0015	-0.0002	-0.0031***
	(0.0045)	(0.0040)	(0.0016)	(0.0006)	(0.0008)
RF: Physician Propensity	-0.0059	0.0070	-0.0010	-0.0049**	-0.0070**
	(0.0122)	(0.0103)	(0.0046)	(0.0025)	(0.0030)
IV: Any Antibiotics	-0.0097	0.0116	-0.0017	-0.0081*	-0.0115**
	(0.0201)	(0.0170)	(0.0075)	(0.0041)	(0.0050)
mean of dep. var.	.122	.0986	.0145	.00311	.00562
	Panel II.B: Type of revisit≤ 30 days				
OLS: Antibiotics, no X	0.0320***	0.0308***	0.0068***	-0.0008	-0.0048***
	(0.0058)	(0.0055)	(0.0022)	(0.0011)	(0.0010)
OLS: Antibiotics	0.0414***	0.0377***	0.0073***	-0.0001	-0.0035***
	(0.0059)	(0.0055)	(0.0022)	(0.0010)	(0.0010)
RF: Prescription Propensity	0.0514***	0.0463***	0.0151**	-0.0045	-0.0056
	(0.0162)	(0.0151)	(0.0068)	(0.0031)	(0.0038)
IV: Any Antibiotics	0.0850***	0.0767***	0.0250**	-0.0074	-0.0093
	(0.0265)	(0.0247)	(0.0113)	(0.0052)	(0.0062)
mean of dep. var.	.212	.171	.0269	.007	.00721
	Panel II.C: Type of revisit≤ 90 days				
OLS: Antibiotics, no X	0.0381***	0.0362***	0.0101***	-0.0029*	-0.0053***
	(0.0062)	(0.0062)	(0.0032)	(0.0016)	(0.0013)
OLS: Antibiotics	0.0560***	0.0503***	0.0109***	-0.0018	-0.0033***
	(0.0064)	(0.0064)	(0.0033)	(0.0016)	(0.0012)
RF: Prescription Propensity	0.0668***	0.0559***	0.0190**	-0.0010	-0.0071
	(0.0182)	(0.0174)	(0.0096)	(0.0045)	(0.0045)
IV: Any Antibiotics	0.1106***	0.0926***	0.0315**	-0.0017	-0.0118
	(0.0301)	(0.0288)	(0.0158)	(0.0075)	(0.0074)
mean of dep. var.	.366	.292	.0485	.0144	.0114
Observations	28158	28158	28158	28158	28158

Notes: The tables shows the short-term effects for the restricted RTI-sample. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A2: Medium-term outcomes, placebo sample

	(1) Any	(2) In hours	(3) Out hours	(4) Emergency	(5) Hospital
Physician Propensity	0.0332	-0.0206	0.0147	-0.0150	0.0456*
	(0.1697)	(0.1363)	(0.0208)	(0.0172)	(0.0240)
Observations	21328	21328	21328	21328	21328
mean of dep. var.	1.39	.964	.0637	.0225	.0442

Notes: The outcomes is the total number of revisits within 6 months for a restricted subsample of index visits, namely those with diagnoses for which antibiotics is prescribed in less than 5%. Control variables are included. The control variables are measured the year prior to childbirth.Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A3: Alternative definition of physicians' prescription propensity

	Panel I: First Stage				
	(1)	(2)	(3)		
Physician Propensity	0.2628***	0.2626***	0.2628***	0.2630***	
	(0.0177)	(0.0177)	(0.0177)	(0.0177)	
mean of dep. var.	.194	.194	.194	.194	
F-stat	98.9	81.1	64.7	57.5	
Child Background		✓	✓	✓	
Parents background			✓	✓	
Child health at birth				✓	
	Panel II.A: Type of revisit≤ 10 days				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
RF: Physician Propensity (PP)	0.0224	0.0047	0.0084*	0.0078	0.0015
	(0.0161)	(0.0136)	(0.0046)	(0.0047)	(0.0030)
RF: PP_{index}	0.0176*	0.0106	0.0041	0.0005	0.0025
	(0.0102)	(0.0089)	(0.0030)	(0.0026)	(0.0021)
IV: Antibiotics	0.0396	0.0083	0.0149*	0.0137*	0.0027
	(0.0281)	(0.0240)	(0.0081)	(0.0082)	(0.0052)
IV: ($AB = PP_{index}$)	0.0670*	0.0403	0.0157	0.0017	0.0093
	(0.0382)	(0.0335)	(0.0116)	(0.0099)	(0.0080)
mean of dep. var.	.123	.0981	.0108	.00786	.00582
	Panel II.B: Type of revisit≤ 30 days				
RF: Physician Propensity (PP)	0.0663***	0.0355*	0.0203***	0.0093*	0.0011
	(0.0204)	(0.0181)	(0.0056)	(0.0055)	(0.0037)
RF: PP_{index}	0.0275**	0.0155	0.0082**	0.0009	0.0030
	(0.0135)	(0.0118)	(0.0039)	(0.0032)	(0.0026)
IV: Antibiotics	0.1175***	0.0629**	0.0360***	0.0165*	0.0020
	(0.0350)	(0.0316)	(0.0101)	(0.0095)	(0.0064)
IV: ($AB = PP_{index}$)	0.1046**	0.0588	0.0310**	0.0035	0.0112
	(0.0508)	(0.0444)	(0.0152)	(0.0120)	(0.0097)
mean of dep. var.	.21	.171	.0197	.0116	.00805
	Panel II.C: Type of revisit≤ 90 days				
RF: Physician Propensity (PP)	0.1038***	0.0691***	0.0289***	0.0103	-0.0046
	(0.0256)	(0.0233)	(0.0089)	(0.0065)	(0.0046)
RF: PP_{index}	0.0369**	0.0232	0.0159***	-0.0008	-0.0015
	(0.0155)	(0.0150)	(0.0054)	(0.0035)	(0.0031)
IV: Antibiotics	0.1837***	0.1224***	0.0512***	0.0182	-0.0081
	(0.0447)	(0.0409)	(0.0159)	(0.0113)	(0.0082)
IV: ($AB = PP_{index}$)	0.1403**	0.0883	0.0606***	-0.0029	-0.0057
	(0.0580)	(0.0562)	(0.0211)	(0.0132)	(0.0118)
mean of dep. var.	.354	.285	.0375	.0194	.0124
Observations	72745	72745	72745	72745	72745

Notes: The instrument is constructed using only index visits. The baseline controls are always included and described in Table 1.1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A4: Controlling for diagnosis chapters at baseline

	Panel I: First Stage				
	(1)	(2)	(3)	(4)	
Physician Propensity	0.5647***	0.3747***	0.3753***	0.3757***	
	(0.0227)	(0.0175)	(0.0176)	(0.0176)	
mean of dep. var.	.194	.194	.194	.194	
F-stat	55.8	37.3	29.6	30	
Obs	72745	72745	72745	72745	
Instrument	Standard	Diagnoses FE	Diagnoses FE	Diagnoses FE	
Child Background	✓	✓	✓	✓	
Parents background	✓		✓	✓	
Child health at birth	✓			✓	
	Panel II.A: Type of revisit≤ 10 days				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
RF: Physician Propensity	0.0190	0.0026	0.0067	0.0078	0.0019
	(0.0153)	(0.0130)	(0.0046)	(0.0048)	(0.0030)
IV: Any Antibiotics	0.0524	0.0042	0.0195	0.0227*	0.0060
	(0.0437)	(0.0372)	(0.0128)	(0.0130)	(0.0082)
mean of dep. var.	.123	.0981	.0108	.00786	.00582
	Panel II.B: Type of revisit≤ 30 days				
RF: Prescription Propensity	0.0607***	0.0334*	0.0158***	0.0092*	0.0022
	(0.0197)	(0.0177)	(0.0056)	(0.0056)	(0.0036)
IV: Any Antibiotics	0.1694***	0.0916*	0.0446***	0.0271*	0.0062
	(0.0544)	(0.0494)	(0.0160)	(0.0151)	(0.0100)
mean of dep. var.	.21	.171	.0197	.0116	.00805
	Panel II.C: Type of revisit≤ 90 days				
RF: Prescription Propensity	0.0940***	0.0642***	0.0232***	0.0103	-0.0037
	(0.0248)	(0.0228)	(0.0089)	(0.0067)	(0.0046)
IV: Any Antibiotics	0.3176***	0.2390***	0.0544**	0.0246	-0.0005
	(0.0767)	(0.0717)	(0.0221)	(0.0175)	(0.0134)
Observations	72745	72745	72745	72745	72745
mean of dep. var.	.354	.285	.0375	.0194	.0124

Notes: The instrument is re-estimated using diagnosis chapters as fixed effects. Fixed effects include PCC (160), age (6), visit month (12), year (8) and diagnoses (116). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using the 72,745 index visits, i.e., visits at least 6 months apart. Control variables are always included and described in Table 1.1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A5: Using only visits from 2012 to 2017

	Panel I: First Stage				
	(1)	(2)	(3)	(4)	
Physician Propensity	0.5619***	0.5620***	0.5625***	0.5630***	
	(0.0238)	(0.0239)	(0.0240)	(0.0239)	
mean of dep. var.	.199	.199	.199	.199	
F-stat	78.2	65.7	49.9	44.8	
Obs	69223	69223	69223	69223	
Child Background		✓	✓	✓	
Parents background			✓	✓	
Child health at birth				✓	
	Panel II.A: Type of revisit≤ 10 days				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
RF: Physician Propensity	0.0140	-0.0030	0.0076	0.0076*	0.0018
	(0.0165)	(0.0141)	(0.0048)	(0.0043)	(0.0027)
IV: Any Antibiotics	0.0249	-0.0053	0.0135	0.0136*	0.0032
	(0.0289)	(0.0250)	(0.0085)	(0.0075)	(0.0047)
mean of dep. var.	.124	.0995	.0112	.00795	.00518
	Panel II.B: Type of revisit≤ 30 days				
RF: Prescription Propensity	0.0542***	0.0227	0.0219***	0.0078	0.0019
	(0.0207)	(0.0181)	(0.0060)	(0.0050)	(0.0034)
IV: Any Antibiotics	0.0963***	0.0403	0.0388***	0.0138	0.0034
	(0.0355)	(0.0316)	(0.0109)	(0.0087)	(0.0060)
mean of dep. var.	.211	.172	.0202	.0117	.00707
	Panel II.C: Type of revisit≤ 90 days				
RF: Prescription Propensity	0.0952***	0.0592**	0.0306***	0.0086	-0.0032
	(0.0261)	(0.0238)	(0.0097)	(0.0063)	(0.0044)
IV: Any Antibiotics	0.1692***	0.1052**	0.0544***	0.0153	-0.0056
	(0.0455)	(0.0417)	(0.0174)	(0.0112)	(0.0078)
mean of dep. var.	.355	.286	.0381	.0194	.0112
Observations	69225	69225	69225	69225	69225

Notes: The outcome are the short-term outcomes and the estimation sample only uses visits from 2012-2017. Fixed effects include PCC (158), age (6), visit month(12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using 69 225 index visits, i.e., visits at least 6 months apart. Control variables are always included and described in Table 1.1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A6: Using physician identifiers across PCCs

	Panel I: First Stage				
	(1)	(2)	(3)	(4)	
Physician Propensity	0.6957***	0.6960***	0.6966***	0.6969***	
	(0.0282)	(0.0284)	(0.0285)	(0.0285)	
mean of dep. var.	.193	.193	.193	.193	
F-stat	110	90.3	70.7	64.1	
Obs	78819	78819	78819	78819	
Child Background		✓	✓	✓	
Parents background			✓	✓	
Child health at birth				✓	
	Panel II.A: Type of revisit≤ 10 days				
	(1)	(2)	(3)	(4)	(5)
	Any	In hours	Out hours	Emergency	Hospital
RF: Physician Propensity	0.0476**	0.0270	0.0115**	0.0066	0.0024
	(0.0196)	(0.0169)	(0.0057)	(0.0054)	(0.0038)
IV: Any Antibiotics	0.0683**	0.0388	0.0166**	0.0095	0.0034
	(0.0278)	(0.0243)	(0.0080)	(0.0076)	(0.0054)
mean of dep. var.	.134	.11	.0107	.00778	.00575
	Panel II.B: Type of revisit≤ 30 days				
RF: Prescription Propensity	0.0987***	0.0628***	0.0217***	0.0107	0.0035
	(0.0263)	(0.0234)	(0.0069)	(0.0066)	(0.0047)
IV: Any Antibiotics	0.1416***	0.0901***	0.0311***	0.0154*	0.0050
	(0.0373)	(0.0335)	(0.0098)	(0.0093)	(0.0068)
mean of dep. var.	.227	.188	.0194	.0115	.008
	Panel II.C: Type of revisit≤ 90 days				
RF: Prescription Propensity	0.1513***	0.1109***	0.0310***	0.0109	-0.0014
	(0.0329)	(0.0308)	(0.0101)	(0.0078)	(0.0060)
IV: Any Antibiotics	0.2172***	0.1591***	0.0444***	0.0157	-0.0021
	(0.0480)	(0.0448)	(0.0144)	(0.0111)	(0.0086)
mean of dep. var.	.373	.304	.037	.0191	.0123
Observations	78819	78819	78819	78819	78819

Notes: The instrument is re-estimated using physician identifiers across the sample (i.e. unconditional on PCC fixed effects). The difference in sample size is due to the fact that there are more identifiers when using within PCC IDs, and I only keep the first interaction with a new physician ID, as such this sample restriction reduces the sample size less since there are fewer IDs. Fixed effects include PCC (158), age (6), visit month (12) and year (8). The instrument is estimated using data from all visits to the same PCC in the same year, excluding the focal visit. The estimation sample is using the 78,819 index visits, i.e., visits at least 6 months apart. The baseline controls, added in column 2, are described in Table 1.1. The control variables are measured the year prior to childbirth. Standard errors are two-way clustered at the individual and PCC levels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix B

In the paper, the instrument is constructed as a physician-visit-specific measure, defined as the ratio of the physicians' total number of antibiotic prescriptions to his or her total number of visits, excluding the focal visit. This instrument thus varies both within and between physicians. The within variation comes from the fact that the instrument is calculated as the leave-out mean—that is, it excludes one visit. The between variation is the part of the variation in the instrument that comes from changes in prescription tendencies between physicians within PCCs.

$$Propensity_{jk} = \frac{\#AB_k - 1(AB_k = 1)}{\#visits_j - 1} \tag{1..9}$$

To elaborate further on the appropriateness of this instrument, and the source of variation used to estimate the causal impact of antibiotics on the outcomes listed in the main text, I decompose the instrument into variation coming from between and within. The decomposition exercise will be done for two different definitions of the instrument, the first one being that used in the main analysis and specified in equation 1.9. The second definition of the instrument will be allowed to vary across years as well, modifying equation 1.9 with a yearly subscript:

$$Propensity_{jky} = \frac{\#AB_{ky} - 1(AB_{ky} = 1)}{\#visits_{jy} - 1} \tag{1..10}$$

The decomposition of the variance is shown in Table B1.

Table B1: Decomposition of instrument variance

Instrument: Residualized $Propensity_{jk}$

	Mean	**Std. dev.**	**Min**	**max**	**Obs**
Overall	0.00	.0892608	-.6418511	.876839	N = 396663
Between		.1515799	-.3194716	.8663847	n = 6637
Within		.0202082	-.5070165	.5070165	T-bar = 60

Instrument: Residualized $Propensity_{jky}$

	Mean	**Std. dev.**	**Min**	**max**	**Obs**
Overall	0.00	.1054369	-.7054439	.9738747	N = 396138
Between		.1514936	-.3291142	.9036545	n = 6619
Within		.0604319	-.5066533	1.036464	T-bar = 60

The instrument is residualized, so the mean is 0. The first row, overall, simply displays the mean and the standard deviation for the full sample, effectively

ignoring the panel structure with respect to the physicians. Taking the panel structure into account enables me to decompose the variance. The row with the between variance shows the physician-level means (i.e., keeping the within variation constant). The number of unique physician IDs in the sample is 6637 for the main instrument and 6619 for the year-varying. One physician has 60 visits on average. The between standard deviation is is approximately 15 for both instruments. In other words, taking the time dimension into account by letting prescription pattern vary over time does not really affect the part of the variation coming from between physicians. The within part of the variation is much lower than the between, for both instruments. This exercise provides two insights: (i) the largest part of the variation in the instrument comes from between physicians rather than within, and (ii) allowing the instrument to vary each year (and thereby absorbing trends in antibiotic prescribing) increases variance, but only the within part. In other words, the variation in the propensity to prescribe antibiotics across physicians is larger than that observed over time. Conceptually, this implies that if you were to draw random physicians, the difference in prescribing propensity is larger than the difference for the same physician in two randomly selected visit days.

Appendix C

Table C1: Definition of control variables

Variable name	Definition	Included diagnosis codes
Individual Characteristics		
Age	Age at time of index visit	
Female	Gender of child, =1 if female, 0=male	
Number of siblings	Total number of siblings	
Born Jan-March	Born in quarter 1	
Born April-June	Born in quarter 2	
Born July-Sept	Born in quarter 3	
Born Oct-Dec	Born in quarter 4	
Number of index visits	Total number of index visits	
Total visits	Total number of visits throughout the sampling period	
Individual Health at birth[a]		
Hospital length (days)	Length of stay at the time of birth (outdate-indate)	
Caesarean Birth	=1 if the delivery was by caesarean	O80
Twin birth	=1 if twin birth	O84
Birth order	Rank of the birthdates by the same mother	
Pre-term	=1 born in less than 37 completed weeks of gestation	P072, P073
Birth weight$\leq 1000g$		P070
Birth weight$\leq 2500g$		P071
Complications at birth[b]	Newborn with BW & 2500g but with other significant problem or multiple problems	R69, P92, P70, P36, P21, P00, P59
Parents Characteristics[c]		
Both parents born abroad	=1 if both parents born outside of Sweden	
Age at birth$_m$	Years of age at birth	
Age at birth$_f$	Years of age at birth	
Mother married at birth	=1 if married at birth	
Years of educ$_m$	Years of education based on a 7 step education variable	
Years of educ$_f$	Years of education based on a 7 step education variable	
Not working$_f$	=1 if not working	
Not working$_m$	=1 if not working	
Family income	Yearly disposable income for the household unit, in SEK	
Parents Highest Education level		
Elementary school$_f$	=1 if Education category=1-2	
Elementary school$_m$	=1 if Education category=1-2	
High school$_f$	=1 if Education category=3-4	
High school$_m$	=1 if Education category=3-4	
University$_f$	=1 if Education category=5-7	
University$_m$	=1 if Education category=5-7	

[a]The health variables comes from registry held by the National Board for Health and Welfare
[b]Included diagnoses are grouped according to DRG-codes Q45 and Q55
[c]The data for these variables comes from Statistics Sweden

Table C2: ICD-10 diagnosis codes of relevant diagnoses for RTI

J00–J06	Acute upper respiratory infections
J10–J18	Influenza and pneumonia
J20	Acute bronchitis
J22	Unspecified acute lower respiratory infection
R05	Cough
R06	Dyspnoea
R50	Fever of other and unknown origin
B99	Other infectious disease
H65–H70	Otitis media and mastoiditis

Table C3: ICD-10 diagnosis codes of relevant diagnoses for asthma

J45	Asthma

Table C4: ICD-10 diagnosis codes of relevant diagnoses for eczema

L20	Atopic dermatitis
L30	Other and unspecified dermatitis

Appendix D

As mentioned in section 1.2 of the paper, there are two relevant guideline changes for prescribing of antibiotics. While changes in treatment guidelines to promote prudent use of antibiotics are commonplace, I have identified these two as being most relevant for the context of this paper. The reason is that they (i) treat illnesses experienced by children and (ii) are treated in the primary care sector.

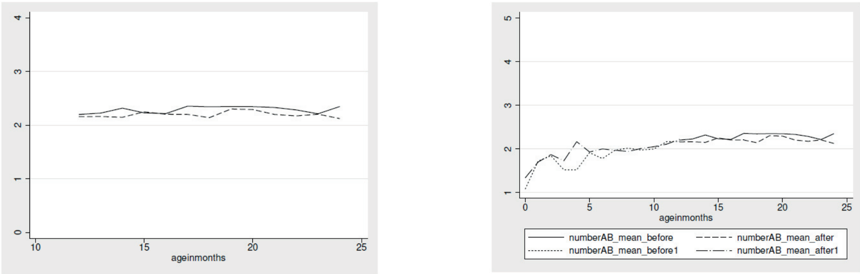
The first change is for treatment guidelines for a common child complication: acute otitis media (AOM). To ensure more prudent use of antibiotics, the Swedish Medical Products Agency (Läkemedelsverket) issued new treatment guidelines for AOM in October 2010 that effectively changed the age of children for whom antibiotics should be first-line treatment. Prior to the change in guidelines, antibiotics were recommended to children 0–2 years old. Post-change, this group was reduced to children 0–1 year old. Instead of antibiotics, "watchful waiting" was recommended for children 1–2 years old with AOM.

The second change in guidelines is for the treatment of acute pharyngotonsillitis (hereafter tonsillitis). New guidelines were issues in September 2012; however, there were only subtle changes relative to the preceding guideline. Prior to the change, antibiotic prescribing required two to four Centor Criteria to be fulfilled, as well as a positive rapid antigen detection test (RADT) for group A Streptococcus, which has long been considered the most important pathogen in tonsillitis, especially within children (Pallon, 2022).[29] If the test was negative, the physician could still offer the patient antibiotics but in a more restrictive manner. However, after the change, three criteria and a positive RADT test were required for antibiotics. If the RADT was negative, the physician was clearly instructed not to prescribe antibiotics (Erikson et al., 2014).

Prior research by Nord et al. (2013) clearly states that the adherence to treatment guidelines is especially poor for these two conditions, using aggregate data and patterns in antibiotic prescribing over time. However, it is very difficult to directly evaluate the effect of these two guideline changes, since adequately evaluating adherence requires access to more data, such as test results (for tonsillitis) and other symptoms that are not reflected by diagnosis codes directly (such as fever, pain, and single- versus double-sided infection with regard to AOM). Nevertheless, in Figure D1, I plot the average number of yearly antibiotic prescriptions per age in months at the time of the physician visit.

[29] The Centor score gives 1 point each for fever, swollen lymph glands, tonsillar exudates, and absence of a cough (Pallon, 2022).

Figure D1: Number of antibiotic prescriptions per month of age



((a)) 13-24 months ((b)) 0–24 months

Notes: Figure D1a shows the number of prescriptions for children ages 13–24 months at the time of the visit, as these children should have experienced the change in guidelines. The solid line represents the year prior to the guideline change (September 2009–September 2010) and the dashed line the year after (October 2010–October 2011). Figure D1b also includes the same number for children ages 0–12 months at the time of the visit; these can be thought of as a placebo group, since the recommendations in the guidelines were unchanged.

From Figure D1, it is clear that there was no dramatic change in antibiotic prescribing following the guideline change. Moreover, while I cannot assess the adherence to the tonsillitis guidelines due to a lack of data, I test the robustness of my main result using only data from 2012 in Appendix Table A5. The results are very similar to the main results presented in Table 1.4. Thus the AOM guideline change is not driving the results.

Chapter 2

The Impact of Upper Secondary School Flexibility on Sorting and Educational Outcomes

Co-authored with Louise Jeppsson and published in *Economics of Education Review* (2021)

Abstract

This paper estimates the causal impact of an upper secondary curriculum reform in Sweden that increased students' course-taking flexibility in year 2000. In the most popular upper secondary program, it led to a significant decrease in mandatory mathematics requirements. Using administrative Swedish data, we estimate the causal impact of the reform on tertiary education outcomes and expected earnings using a differences-in-discontinuity identification strategy. The method compares students born immediately before and after the cutoff date. The inclusion of students born in neighboring non-reform cutoff years enables us to disentangle the school starting age effect from the unconfounded effect of the reform. We find no negative effects of the reduced mathematics requirements. Rather, we find a positive effect of the reform on students' probability of enrolling in, and earning a degree from, tertiary education. Our heterogeneity analysis suggests that relatively disadvantaged students were not negatively affected by the reform.

Please cite as Berggren, A., Jeppsson, L. (2021). The impact of upper secondary school flexibility on sorting and educational outcomes. *Economics of Education Review*, 81, 102080.

2.1 Introduction

A well-educated labor force in science, technology, engineering, and math (STEM) offers a competitive edge in the global economy. Skills in mathematics and science have been shown to be positively associated with economic growth (Hanushek and Kimko, 2000). Policy makers in industrialized countries have shown great interest in improving the accumulation of such skills through curriculum reforms and better preparation of young individuals for tertiary education.[1] When determining the school curriculum, policy makers make choices regarding the overall time devoted to different subjects and what subjects should be compulsory. These choices reflect priorities and preferences concerning what knowledge and skills should be required and, consequently, there is substantial heterogeneity in curriculum priorities across countries (OECD, 2018).

Herein lies a potential trade-off for the policy maker. More advanced education leads to a higher human capital stock, but enforcing a too strict curriculum might also lead less able students to shy away from further investments in human capital. Critics of a rigid curriculum argue that restricting students' choices is undemocratic since mandating a fixed curriculum for all students deprives them of the opportunity to take courses they are interested in and comply with their personal aspirations (Noddings, 2011). On the other hand, under a flexible curriculum, students with potentially high returns to more advanced courses may opt out of those courses and hence reduce their tertiary education prospects. This paper sheds light onto the aforementioned trade-off by examining whether students' academic and labor market outcomes are affected by a reform that introduced a more flexible course system while simultaneously decreasing compulsory mathematics requirements which altered university eligibility.

There is quite little empirical evidence on the returns to different course choices. This is quite surprising since every student need to make these decisions and it has been subject to a vast amount of policy discussions(Altonji et al., 2012). When students have the flexibility to choose courses in line with their own interest and skills they may obtain higher grades. A higher GPA has been shown to have a positive impact on students' completion beliefs and aspirations to pursue post-secondary education (DesJardins et al., 2019; Kunz and Staub, 2016).

Opponents of flexibility would argue that the content of the curriculum dictates eligibility to academic subjects and programs. In US for example, a heavily influential report by Gardner (1983) called for more courses in academic subjects,

[1] See for example Görlitz and Gravert (2018) investigating reform changes in Germany; Ning (2014), Sosa (2016) and Goodman (2017) investigating reform changes in the U.S. and Joensen and Nielsen (2016) investigating curriculum reform changes in Denmark.

for example mathematics and science. Inspired by the report, there were many reforms increasing the strictness in course choices and academic content. This literature is centred around the importance of specific courses or subjects included in the course curriculum on subsequent outcomes, with a focus on mathematics and science (Altonji, 1995; Levine and Zimmerman, 1995; Rose and Betts, 2004; Joensen and Nielsen, 2016; Sosa, 2016; Goodman, 2017; Görlitz and Gravert, 2018; Ning, 2014). They commonly refer to reforms introducing a stricter curriculum and find that there are significant returns to advanced courses. Furthermore, as established in the literature on tracking there is a risk that students have problems in anticipating future educational performance (Brunello et al., 2007).

In this paper, we explore a reform that increased flexibility but did so at the cost of decreasing mathematics requirements. The cost of stricter curriculum is the loss of flexibility. A relevant question to ask is whether the positive impact of more mathematics knowledge established in the above cited literature is outweighed by the general increase in flexibility? In this respect the literature is scarce. The paper most closely related to ours is a working paper by Yu and Mocan (2018). They investigate the causal effect of increased upper secondary school curriculum flexibility on student outcomes. The authors exploit a curriculum reform in China launched in 2004 that increased students' freedom when selecting courses. The authors find a positive impact on both students' academic achievement at university level and their mental well-being. In contrast to Yu and Mocan (2018), who measure outcomes for a representative sample of students while still in tertiary education, we have access to data on the entire student population in Sweden and are able to follow them up to the age of 27. We also focus our analysis on the importance of less strict mathematics requirements. Our detailed data allows for estimation of distributional effects. From the point of view of the social planner, knowing *where* and *how* in the distribution students react to more flexibility is vital information to ensure equity in educational opportunities. Another contribution is our evaluation of a school reform that has never before been evaluated.

The reform was introduced in all Swedish upper secondary schools in year 2000 and the analysis first examines how it alters students' course-taking behavior. Second, we ask whether it has a causal impact on tertiary education outcomes and annual expected earnings. Finally, we examine the distributional effects along the dimension of parents' socio-economic status (SES). Leveraging detailed Swedish administrative data, containing the entire population of upper secondary school students, we estimate the causal effect in a regression discontinuity (RD) framework. The identification strategy exploits the discontinuity given by student's birth date since it decides whether he/she started upper secondary school after the new curriculum was introduced in autumn 2000. We compare students born in a 3 months window around the cutoff date, i.e., October 1983-March 1984. The RD estimations cannot disentangle the school starting age effect on outcomes from the true effect of the curriculum reform.[2] To tease out the unconfounded effect on outcomes we follow Carneiro et al. (2015) and Bertrand et al. (2019) and employ a difference-in-discontinuity (RD-DD) design where we augment the RD regression with students born in October–March in neighboring non-reform cutoff years.[3]

The decrease in mathematics requirements was most prominent on the most popular upper secondary school program in Sweden, the Social Science program.[4] In this program, 25 percent of the previously mandatory mathematics coursework was moved to a list of elective courses (GyVux 1994/97:16; GY2000:16). No such change occurred in any other Swedish upper secondary program. We will focus the empirical analysis within this program as it directly altered the pathways to university for the students due to the fact that the now voluntary mathematics course was pivotal for eligibility to, for example, Business and Economics.

The results show that, in line with the aim of the more flexible curriculum, the reform significantly altered students' course-taking behavior. We find a significant and large drop in mathematics attainment across males and females, by approximately 37 percent. The decrease was not offset by an increased enrollment in elective STEM-related courses. Rather, we find that students tend to substitute mathematics with non-STEM electives. However, the drop in mathematics does not lead to a significant impact on the probability of completing tertiary education in a field that requires the pivotal mathematics course. Splitting by gender show a marginally significant increase in mathematics-field for women. Nor do we find an effect on the speed at which students enter tertiary education after graduating from upper secondary school. Taken together, these results suggest that students' educational prospects, on average, were not limited by the choice to take less mathematics under the more flexible curriculum. On the contrary, our results suggest that the reform increased students' probability of enrolling in tertiary education by 3 percent. Furthermore, the reform led to an increase in

[2]In Sweden, a student's school starting year is based on his or her calendar year of birth. The school starting age effect implies that students born in December differ from students born in January regardless of whether the reform was in place or not since school-wise they are one year younger than their January born peers. See for example Black et al. (2011) and Fredriksson and Öckert (2014) regarding the importance of the school starting age effect.

[3]We include students born in 1982-1983, 1984-1985, 1985-1986, 1986-1987 and 1987-1988.

[4]For a more detailed explanation of the Swedish upper secondary program system, see Section 2.2.

the probability of exiting tertiary studies with a degree. Splitting the sample by gender shows that the overall effect was driven by a large and positive impact on females, for whom we estimate a 5.6 percent increase in the probability of earning a tertiary degree. Our results are robust to both the choice of bandwidth and other coinciding school reforms. As the students in our sample are too young to allow us to study actual earnings, we estimate the impact on expected earnings based on field of study and gender.[5] We find a small positive effect of the reform on females annual expected earnings. We propose a possible pathway to mediate the positive impact on tertiary education enrollment: an increase in GPA which increased due to the reform.

Treatment heterogeneity is analysed through the distributional impact of the reform along the dimension of parents' socio-economic status (SES). We find no evidence that relatively disadvantaged students were negatively affected by the reform. It rather seems that students in the lowest quartile benefited the most from the more flexible curriculum. However, they did so at the cost of the more advantaged students. This group had a decreased probability of attaining mathematics-related education combined with a lower speed to entering university. Taken together, the results suggest that the latter group did substitute a relatively difficult course for "fast and easy" education. The opposite is true for the disadvantaged students.

A potential challenge to the identification strategy is posed by the introduction of a new upper secondary program in Sweden, the Technology program, at the same time as the curriculum reform in autumn 2000. The new program could potentially induce a different sample of students to enter the Social Science program after the reform. We estimate the probability to enroll in the Social science program by gender.[6] The reform increased the probability of choosing the Social Science program among women. Importantly for the validity of the design, we show that pre-determined observable characteristics are balanced across the cutoff.

The rest of the paper is structured as follows: Section 2.2 describes the details of the school reform and the institutional framework of the educational system in Sweden, Section 2.3 presents the identification strategy, Section 2.4 describes the data, and Section 2.5 presents and discusses the main results and heterogeneity analysis. Section 2.6 concludes the paper.

[5]Students are 27 years old in the most recent data and the differential life cycle trajectories in earnings based on study choice are not yet materialized (Bhuller et al., 2017). Field of study is coded in detail and contains 116 education categories.

[6]90 percent of the students at the Technology program when first introduced were male, see Table A2

2.2 Institutional Background

Attending upper secondary school is not required by Swedish law. Nevertheless, after completing nine years of compulsory education in Sweden, most students choose to continue their education in the Swedish upper secondary school system. In 1999 and 2000, approximately 98 percent of all compulsory school graduates entered upper secondary school in the same year (Skolverket, 2000; Skolverket, 2001).[7] Without any grade retention or other discontinuities in prior education, students are expected to enter upper secondary school in the autumn semester of the year in which they turn 16 years old and then graduate after three years. Students apply for enrollment in specific upper secondary programs, either within preparatory or vocational tracks. Students are admitted based on their grades from lower secondary school.[8] In year 2000, the number of available national upper secondary school programs increased from 16 to 17 as a Technology program was officially introduced (Skolverket, 2000).

2.2.1 The Upper Secondary School Reform GY2000

From 1994 to 2011, the Swedish upper secondary school curriculum was regulated by Lpf 94, although an important revision of the existing program structure and curricula was made as part of the GY2000 reform, implemented in year 2000. A main objective of the reform was to increase the share of elective coursework and therefore also the students' course choice flexibility, in particular in the Natural Science and Social Science programs, the Swedish government thought that the course plans for these two programs were too rigid (Skolverket, 1998; Prop.1997/98:169).

The GY2000 reform increased upper secondary school students' course choice, to various degrees, on existing upper secondary school programs. The percentage of upper secondary school credits devoted to mandatory courses decreased while credits devoted to choice based coursework increased mainly through the introduction of a new package of elective courses from which students choose a number of courses to fill a quota of credits (GY2000:19; GyVux 1994/97:17).[9] While all Swedish upper secondary school programs were affected by the reform,

[7]Swedish compulsory education is divided into lower primary school (age 7-10), upper primary school (age 10-13) and lower secondary school (age 13-16). The majority of references in this section is available only in Swedish. The reference list for references in Swedish is found in Appendix B.

[8]The decision is made prior to lower secondary school graduation.

[9]There are specialization tracks within some of the vocational programs that experienced a small decrease in elective coursework. Choice based coursework within the 15 vocational programs made up 14.9-56.1 percent of total credits prior the reform and 22-52 percent after the reform.

this paper focuses on students enrolled in the Social Science program for the main analysis. The Social Science program is the most popular upper secondary program in Sweden and prior to the reform, the government raised concerns about the strict program curriculum. Before the reform social science students had a quota of 190 course credits, corresponding to 8.8 percent of the total credits, to obtain from individual course choices.[10] After the reform the quota of credits to be earned from choice based course work differed between program tracks, ranging from 18-24 percent of total credits. With the exception of the course Mathematics C, described below, each school was to decide what electives to offer.

The reason for focusing the analysis on students in the Social Science program is that one implication of the reform was that a full-year course in intermediate mathematics *Mathematics C*, was made elective as opposed to mandatory. That is, the course was moved from the mandatory course list to the package of elective courses (GY2000:16; GyVux 1994/97:16). Swedish media published articles informing about the increase in curriculum flexibility and the new Technology program, yet no information about the changes regarding the Mathematics C course seems to have been dispersed to the public.[11] If students and parents were poorly informed about this change the risk of student sorting based on changes in mathematics requirements at the Social Science program is attenuated. Prior to the reform, student were required to complete three mathematics courses, Mathematics A, B and C, corresponding to approximately 9.3 percent of the total amount of course credits.[12] After the reform, students were required to complete only the A- and B-level courses in mathematics, corresponding to 6 percent of the total amount of credits in the new curriculum. Although each upper secondary school was free to decide what electives to offer, Mathematics C was made an exception, so that after the reform all upper secondary schools were required to include this course in the elective course package offered to students in the Social Science program. The Swedish National Agency for Education (*Skolverket*) deemed mathematics as particularly important for tertiary education

[10]The same figure applies to the Natural Science program. Within Social science students were offered extra flexibility within two of the available specialization track; Business Administration and Humanities, but no flexibility within the Social science track (GyVux 1994/95:14).

[11]Tidningarnas Telegrambyrå (1999), "FAKTA: NYA GYMNASIESKOLAN",Tidningarnas Telegrambyrå, September 15; Anna Lena Wallström (1999), "Fler valmöjligheter för gymnasieelever", Borås Tidning, September 16, page 14; Inga-Lill Hagberg (1999), "GYMNASIEFÖRSLAG Teknik och miljö nya val", Svenska Dagbladet, September 16, page 4; Tidningarnas Telegrambyrå (1999), "BRÅTTOM ATT VÄLJA TILL FÖRÄNDRAT GYMNASIUM", Tidningarnas Telegrambyrå, November 4; Lena Hennel (1999), "Lärarkritik mot gymnasiereform", Svenska Dagbladet, November 5, page 5 ;Anna Asker(1999), "Nytt teknikprogram ska avhjälpa teknikerbristen", Svenska Dagbladet, December 7, page 30.

[12]Approximately 5.1 percent for Mathematics A, 1.9 percent for Mathematics B and 2.3 percent for Mathematics C (GyVux 1994/95:16).

since courses in mathematics is a common entry requirement for many university programs (Skolverket, 1998). For example, the intermediate mathematics course Mathematics C is an entry requirement for popular undergraduate programs in business and economics at Swedish universities as well as for other university programs such as those for future architects and real estate agents (UHR, 2016; SACO, 2018).

A second feature of the GY2000 reform was the introduction of a new higher education preparatory program, the Technology program. Prior to the reform, the Natural Science program offered a technical specialization track. The aim of the new Technology program was to increase the supply of available programs for students interested in the natural sciences and technology since the government at the time deemed that the technical orientation within the current Natural Science program was not sufficient to meet the demand from students interested in technology (Prop.1997/98:169). While we are not explicitly interested in the introduction of the Technology program, it may have induced a different sample of students entering the Social Science program after the reform. In Section 2.3, we discuss this challenge for identification more thoroughly and provide evidence in Section 2.5.1 that the introduction of the Technology program should not be of significant concern.

2.3 Empirical Strategy

This study estimates the causal average impact of an upper secondary school curriculum reform in autumn 2000 on students' course taking behavior, tertiary education outcomes and annual expected earnings. The identification explores the discontinuity given by students' birth dates as it dictates whether they started upper secondary school when intermediate mathematics was mandatory or not. We compare students born immediately to the right of the threshold, in January 1984 to students born precisely before, in December 1983. To capture the causal impact of the flexible curriculum, α_{RD}, in the limit, individuals born in December 1983 must be identical to children born in January 1984 such that the only difference comes from curriculum regime.

Effectively, we estimate two regressions, one on each side of the threshold:

$$y_i = \delta + \lambda R_{ic} + \gamma f(B_i - c) + \beta f(B_i - c) R_{ic} + \theta \mathbf{X}_{ic} + \pi \mathbf{W}^p_{ic} + \eta_m + v_{ic}. \quad (2.1)$$

Where y_i is the outcome for student i. Reform exposure, R_i, is and indicator variable equal to 1 if individual i was born in or after January 1984, c, and hence entered upper secondary school in year 2000 when the reform was implemented. Birth month and year, B_i, is normalized around the cutoff such that $c = 0$. α_{RD} is estimated as $\hat{\lambda}$. Split time trends $f(.)$ are included to allow for different slopes before and after the reform. We include a vector of control variables similar to those used in related work (Kirkeboen et al., 2016; Malamud and Pop-Eleches, 2010, 2011). Throughout the paper we show results both with and without controls and they are robust to the inclusion of controls.[13] η_m contains controls for municipality fixed effects, the level at which compulsory and upper secondary education is operated in Sweden. Equation 2.1 is estimated using a local polynomial regression with a first-order polynomial as suggested by Gelman and Imbens (2018). We use a bandwidth of three months on each side and a triangular kernel since it is shown to be boundary optimal (Cheng et al., 1997). In practice, the choice of kernel should not significantly alter the results (Lee and Lemieux, 2010).

One concern is that birth month of students is correlated with, for example, educational attainment. Previous research has shown substantial differences in educational achievements depending on month of birth.[14] To account for the effect of school starting age, we follow the identification strategy in Carneiro et al. (2015) and Bertrand et al. (2019) and include cohorts born in neighboring non-reform cutoff years, 1982-1983, 1984-1985, 1985-1986, 1986-1987 and 1987-1988, in order to estimate a difference in regression discontinuity model, the RD-DD.[15] By including the non-reform cutoff years we estimate discontinuities between children born in October-December and January-March. Intuitively, the discontinuity at the cutoff in January 1984 will be a combination of the true effect of the reform and month of birth effects: $\alpha_{RD} = \tau_{reform} + \tau_{b_i}$. Under the assumption

[13] Adding controls improves precision and help us reduce any bias due to potential differences in pre-determined characteristics of individuals to the left and right of the cutoff. We add a vector of controls for pre-determined student characteristics $\mathbf{X}_{ic}$, including gender and an indicator variable equal to 1 if the individual obtained a grade of pass with distinction or special distinction in mathematics in lower secondary school. The lower secondary mathematics grade is included as a control for mathematics ability since we hypothesize that this ability is an important determinant of a student's choice of upper secondary courses, in particular whether to substitute the Mathematics C course for another course under the new flexible curriculum introduced as part of the reform. We include a vector of parent characteristics $\mathbf{W}^p_{ic}$, which contains information on whether at least one parent had a low level of education (defined as not having completed three years of upper secondary school), the earnings of the father averaged over age 14-16 of the child, and parents' immigration status (equal to 1 if both parents immigrated to Sweden)

[14] See for example Fredriksson and Öckert (2014) and Black et al. (2011) for good examples of the importance of school starting age.

[15] Stricter entry requirements to upper secondary school programs was introduced in 1998 and hence affecting cohorts born from 1982 onward. The cohort born 1982 was also the first cohort to receive criterion referenced grades when in the 8th grade in elementary school as well as the first cohort that did not receive a course grade in upper secondary school courses if a student's absence was high. Therefore, we are not able to include a wider window of non-reform cutoff years.

that month of birth effects are stable across cutoff years and do not interact with the true reform effect (Carneiro et al., 2015), we can estimate the average discontinuities in outcomes for the five non-reform cutoff years: $\alpha_{RD_{noreform}} = \tau_{b_i}$. By subtracting $\alpha_{RD_{noreform}}$ from α_{RD}, we cancel out the month of birth effect and leave only the true, unconfounded impact of the reform:

$$\alpha_{RD-DD} = \alpha_{RD} - \alpha_{RD_{noreform}} = (\tau_{reform} + \tau_{b_i}) - (\tau_{b_i}) = \tau_{reform}$$

The reform effect is thus the difference between the discontinuity in outcomes for students entering school after the reform and the discontinuity for students entering in nearby non-reform years.

2.3.1 Validity of the RD-DD

We will consistently estimate the impact of reduced mathematics under the crucial assumption that individuals are unable to precisely manipulate the running variable. Use of age-based discontinuities, such as date of birth as the running variable, is common (Lee and Lemieux, 2010), and due to the difference in time between when individuals were born and when they entered upper secondary school, we can be sure that the reform was unknown at the birth date.[16]

Note that we will investigate educational outcomes and annual expected earnings of a restricted part of the full population, namely upper secondary social science students. If the curriculum reform itself, in particular the introduction of the Technology program, caused sorting of students into different upper secondary programs, the comparison between the student samples enrolled in the Social Science program before and after the reform is confounded by selection. As shown in appendix, Table A2, 8 percent of the students born in 1984 enrolled in the new Technology program. To investigate sorting, we estimate the effect of the reform on the probability of enrolling in the Social Science program through estimation of equation 1 for a pooled sample and by gender.[17] The plausibility of the RD-DD estimation assumes that unobserved characteristics are similar across the cutoff for the treatment year as well as the control year. We perform a balancing test to ensure that there is no selection on predetermined observables and show the results in Table 2.4.

[16] We include a histogram of the frequency of birth in the relevant years, see Appendix Figure A5. There is a strong seasonality in timing of birth but it is not systematically different across the relevant years.

[17] Boys are more likely than girls to enter the new Technology program. According to Appendix Table A2, the fraction of males in the Technology program was 90 percent in the first cohort after the program was introduced.

A well-known concern with RD-analysis employing a discrete running variable is the fact that there is no continuity at the cutoff. Cattaneo et al. (2018) recommends supplementing the analysis with estimations from a local randomization (henceforth LR) framework. The idea is that there exist a window around the cutoff where assignment to treatment can be viewed as a local experiment. When the data is discrete and the cutoff known to the researcher, the smallest window is defined by the two observations closest to the cutoff. Though, as we have discussed previously, comparing children in December and January will yield significant differences irrespective of reform exposure due to the age effect.

Nevertheless, we can make use of the LR-framework to further elaborate on the validity and plausibility of continuity assumption with few mass points.[18] We show, in the appendix, Figures A1-A4 containing point estimates from separate RD regressions using a 3 month bandwidth, a 6 month bandwidth and the simple difference in means (the statistics of choice in the Local randomization framework) for women and men respectively. The point estimates are shown for both reform and all non-reform years separately. Regarding the discreteness, the point estimates from the continuity based RD are similar to those adopting the LR framework, simply comparing the difference between the mean of the control (one mass point to the left) and mean of treated (one mass point to the right). This is reassuring since the continuity based approach could mask selection or sorting right at the cutoff but we rule this out due to the similarities between the set of estimators. While the RD and LR estimates are comparable across the years, they are not directly comparable to the main RD-DD results due to the school age starting effect.[19] The latter is the difference in discontinuity between the reform cutoff and the cutoff for the *pooled* control years.

2.4 Data

We use Swedish registry data provided by Statistics Sweden. Statistics Sweden links several administrative registers by personal identification numbers and we obtain information about individuals' birth month and year, educational attainment, school grades and field of study in upper secondary school as well as in

[18] Note that the three mass points to the left and the right cutoff in our main specification contains information from a large number of students, the efficient number of observations at every cutoff is approximately 4500 to the left and 5800 to the right of the cutoff. The difference is due to the fact that the cohort size is larger in January-March relative to October-December.

[19] As the LR simply compares the mean across the cutoff, we have tested the sensitivity of the RD-estimates using a polynomial of degree 0 instead of 1 for a more similar comparison. We have done the same check for the RD-DD estimates. Neither the RD, nor the RD-DD estimates are sensitive to the choice of polynomial.

tertiary education. Individuals are linked to their parents (biological or adoptive) and information on the parents' background characteristics. The data set contains the entire population of individuals born in Sweden between January 1982-December 1988 who have completed elementary school. We restrict our main sample to contain first-time enrollees into the Social Science program in upper secondary school.

2.4.1 Variables

The outcome variables of interest are several measures of tertiary education and annual expected earnings, measured at 27 years as this is the oldest age at which we can observe this information in the dataset. The tertiary education outcomes comprise a set of indicator and discrete variables capturing educational attainment on both the extensive and the intensive margin.

For impacts on the intensive margin, we construct an indicator variable, *MaC-field*, which is equal to 1 if an individual has her or his highest attained education in the field of business, economics, architecture or real estate management. Entry to all of these university programs requires prior completion of Mathematics C in upper secondary school.[20] Inclusion of this outcome variable is motivated by its direct dependence on students' mathematics choices in upper secondary school.

	High Returns	Low Returns
Enroll	(1) +	(2) -
Not enroll	(3) -	(4) +

Table 2.1: Mathematics C choice

Given students' potential returns to mathematics studies, one could roughly define one group of students who *should* (high returns) enroll in the Mathematics C course and one group who *should not* (low returns). A strict, non-flexible, course curriculum ensures that all students with potentially high returns enroll in the course, but it also forces students with low returns to take the course even if they would be better off studying something else; cells (1) and (2) in Table 2.1. Introducing choice under a flexible curriculum may lead to the desirable outcome that low-return students opt out, i.e., cell (4), while high-return students continue to enroll, i.e., cell (1). If this is the case, we expect no impact of the reform on the

[20]Obviously, there are other university fields, for example in the natural sciences, that also require Mathematics C or more. However, graduating from the upper secondary Social Science program does not make individuals eligible for these fields independent of whether they chose to take Mathematics C. Hence, the course choice is not pivotal for eligibility, in contrast to the fields of study included in *MaC-field*.

outcome variable *MaC-field*. However, introducing choice raises the concern that students with low potential returns who ideally should not enroll in the course continue to do so, i.e., cell (2). An even greater concern is that students with potentially high returns may refrain from taking the course under the flexible curriculum, i.e., cell (3), and forego the eligibility to enter mathematics-intensive post-secondary academic fields they would have pursued absent the reform. Under such circumstances we expect to find a negative impact on *MaC-field*.

We also include a discrete variable, $Speed$, measuring the speed to entry into tertiary education. The variable ranges from 0 to 5. It is equal to 0 if an individual started tertiary education in the same year as she or he graduated from upper secondary school and 5 if she or he started tertiary education five years after completing upper secondary school.[21] We expect to find an impact here if students regret their choices induced by the reform and therefore have to take adult education classes to gain the desired eligibility for certain study fields in tertiary education.

For general tertiary education outcomes, we have constructed the indicator variable $AnyTE$, equal to 1 if the individual ever attended any tertiary education, to capture the impact of the reform on the extensive margin. We further include the indicator variable *Degree*, which is equal to 1 if an individual exited tertiary education with an academic or vocational degree. This variable does not distinguish between the different durations of tertiary education programs needed to earn a certain degree.

Given the time span of our data, the students are too young for us to study actual earnings (Bhuller et al., 2017). Students born in 1988 are at most 27 years old in the most recent data – an age at which the differential life cycle trajectories in earnings based on study choice have not yet materialized. For the full sample we estimate the impact of the reform on expected returns to education, through imputing an outcome variable for an individual's annual expected earnings in middle age.[22] Table 2.2 summarizes the mean and standard deviation of the main variables for the sample of upper secondary social science students born in the pre- and post-reform years 1983 and 1984, respectively. Before the reform, 64

[21]We cannot extend the time to more than five years due to data restrictions. However, approximately 50 percent of graduating upper secondary students in Sweden enter university within five years (Holmlund et al., 2007). Note that this is a lower bound since less than 100 percent of students ever enter university.

[22]We take the average earnings for individuals aged 43–45 in 2015, stratified by gender and detailed information on field of tertiary education. Our data includes 116 detailed tertiary education fields. We impute this value to the individuals in the relevant sample as the annual expected mean income, in Swedish kronor (SEK). We further stratify by level of education, in addition to field and gender, to capture the quantity of tertiary education in a separate measure of annual expected earnings.

	Before reform cohort 1983			After reform cohort 1984		
	Mean	Std. dev	Obs.	Mean	Std. dev	Obs.
Tertiary Education Outcomes						
Math C-field	0.16	0.37	20721	0.17	0.37	20704
Speed	2.57	1.27	12777	2.60	1.28	12898
Any tertiary education	0.63	0.48	22487	0.64	0.48	22277
Degree	0.31	0.46	22487	0.33	0.47	22277
Labor Market Outcome						
Annual expected earnings (SEK)	308 330	106 056	22487	309 996	104 877	22277
Upper Secondary School						
GPA	14.20	2.86	19594	14.38	3.10	19834
Mathematics C enrollment	0.64	0.48	22185	0.39	0.49	21973
Mathematics B enrollment	0.93	0.26	22185	0.93	0.26	21973
STEM enrollment	0.03	0.17	23278	0.03	0.18	23079
Non-STEM enrollment	0.08	0.28	23278	0.18	0.38	23079
Background Characteristics						
High math ability	0.43	0.49	23278	0.43	0.50	23079
Male	0.37	0.48	23278	0.36	0.48	23079
Immigrant	0.12	0.32	21943	0.13	0.33	21693
$LowEducation_p$	0.66	0.48	21759	0.64	0.48	21522
$LogAvgWage_f$	10.99	3.80	22140	11.03	3.80	21927

Table 2.2: Summary statistics for the Social Science Program

percent of social science students took Mathematics C. After the reform, the share shrunk to 39 percent. 63 percent of the students enrolled in any tertiary education both before, relative to 64 percent after the reform. 31 percent of the students who started upper secondary school before the reform went on to complete a higher education degree while the corresponding number figure after the reform is 33 percent. The mean of speed to enter tertiary education is approximately 2.6 years for both groups, which implies that the average student enters tertiary education 2–3 years after graduating from upper secondary school. The fraction of males in the sample is approximately 37 percent before and 36 percent after the reform. The low fraction of males is due to the sample restriction to include only students in the Social Science program, which traditionally has a high share of female students. Background characteristics are similar in both groups.

2.5 Results

2.5.1 Sorting

As discussed in Section 2.2, the reform introduced a third higher education preparatory program. To separate the effect of increased course flexibility from the effect

of the introduction of the new program, we must find out whether the sample of students in the Social Science program was similar in terms of background characteristics before and after the reform. We estimate the impact of the reform on the probability of enrolling in the Social Science program, using both the RD and the RD-DD estimator. Recall that the difference between the two is that the RD-DD is augmented with neighboring non-reform years to enable us to subtract a possible month of birth effect from the reform effect. Note that the entire population of potential upper secondary school students is included in this estimation, i.e. all students graduating from 9th grade in high school. We also estimate the regression separately by gender since the new Technology program is strongly male dominated.[23]

The results in Table 2.3 reveal that the reform affected the probability of students choosing the Social Science program, at least for females.[24] The pairwise difference across columns is the inclusion of control variables. In the RD-DD specification, the results are robust with respect to inclusion of different controls.

The impact on the probability of choosing the Social Science program among the entire population of female students is increased, as is evident in column 1 to 4. Regarding the male students, the comparison between RD and RD-DD estimates reveals difference between the RD and RD-DD estimates.[25] After the reform, females were on average 2.5 percentage points more likely to choose the Social Science program. In relative terms, the fraction of female students was 8 percent higher after the reform.

[23] As seen in the extended summary statistics in Appendix Table A2, 90 percent of the students in the Technology program were males.

[24] Regression results of the impact of the introduction of the reform on other upper secondary programs than Social Science are presented in Appendix Table A3.

[25] When comparing with the same estimation on the probability of attending Vocational and Natural Science programs, we find similar differences between the RD and RD-DD estimates, see Table A3. We interpret this as evidence of fluctuating probabilities of choosing specific programs for all cutoff years, not exclusively for the reform cutoff.

Social Science	RD	RD	RD-DD	RD-DD
		All		
Reform	0.011***	0.008***	0.013***	0.012***
Standard Error	0.001	0.001	0.002	0.002
Observations	42,288	42,288	268,835	268,835
R^2	0.026	0.063	0.023	0.054
Pre-reform Mean	0.237	0.237	0.237	0.237
		Females		
Reform	0.041***	0.032***	0.027***	0.025***
Standard Error	0.002	0.003	0.005	0.005
Observations	20,636	20,636	131,286	131,286
R^2	0.028	0.052	0.034	0.046
Pre-reform Mean	0.307	0.307	0.307	0.307
		Males		
Reform	-0.013***	-0.012***	-0.000	-0.000
Standard Error	0.003	0.003	0.006	0.005
Observations	21,652	21,652	137,549	137,549
R^2	0.034	0.043	0.023	0.032
Pre-reform Mean	0.170	0.170	0.170	0.170
Controls		✓		✓

The table reports the impact of the reform on the probability of enrolling in the Social Science program for the full universe graduates from 9th grade. The first two columns show the RD regression results using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The pairwise difference across columns is the inclusion of control variables.

Table 2.3: Probability of enrolling in Social Science

The sorting in to the Social Science program may change the sample before and after the reform. Therefore, it is crucial to address whether the sample selection led to a compositional change among the students enrolled in the Social Science program. For any RD design to be credible, i.e., to separate the treatment effect from any effects of the change in composition, we need to investigate the impact of the reform on pre-determined covariates. To put it differently, even though students cannot manipulate the running variable they can sort themselves into the program. As would manipulation of the running variable, sorting results in predetermined covariates being unbalanced across the threshold. We test the balance of predetermined covariates in Table 2.4. Note that the sample is different in Table 2.3 versus Table 2.4. The latter contains the group of first time enrollees

in Social Science students while the former contained all students that finished elementary school.

	$HighMath_i$	$Male_i$	$Loweduc_p$	$Foregin_p$	$\mathrm{Ln}\overline{Earnings}_f$
	(1)	(2)	(3)	(4)	(5)
			All		
RD	0.101***	-0.022*	0.002	0.009***	-0.056*
Standard Error	0.005	0.010	0.011	0.001	0.024
Observations	10,359	10,359	10,359	10,359	10,359
RD-DD	-0.010	-0.022**	-0.005	0.001	0.004
Standard Error	0.012	0.011	0.008	0.003	0.036
Observations	60,026	60,026	60,026	60,026	60,026
Pre-reform Mean	0.412	0.367	0.647	0.113	11.077
			Females		
RD	0.095***		0.012	0.008**	0.104**
Standard Error	0.008		0.010	0.003	0.034
Observations	6,552		6,552	6,552	6,552
RD-DD	-0.008		-0.009	0.004	0.033
Standard Error	0.010		0.012	0.006	0.070
Observations	37,779		37,779	37,779	37,779
Pre-reform Mean	0.437		0.669	0.110	11.06
			Males		
RD	0.117***		0.006	0.014*	-0.306
Standard Error	0.012		0.028	0.007	0.122
Observations	3,807		3,807	3,807	3,807
RD-DD	-0.012	0.000	0.003	-0.004	-0.021
Standard Error	0.017	0.000	0.013	0.008	0.136
Observations	22,247	22,247	22,247	22,247	22,247
Pre-reform Mean	0.368	1.000	0.608	0.120	11.106

The table reports the impact of the reform on pre-determined characteristics. In the first panel we present the RD regression results using a three months bandwidth on each side of the cut-off and a triangular kernel. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cut-off. In the second panel we present the RD-DD estimates where we augment the regression with students born in October to March in nearby non-reform years 1982-1983, 1984-1985, 1985-1986, 1986-1987 and 1987-1988.

Table 2.4: Balancing test of pre-treatment characteristics Social Science

The results reveals a strong selection on the mathematics grade in lower secondary school.[26] However, from Table 2.4 it is clear that in our preferred specification, the RD-DD, we have no such selection suggesting that the RD was picking

[26]In Appendix Table A4, we present an additional balancing test of pre-determined characteristics for the full population of upper secondary students where similar discontinuities in the RD-estimates are found.

up school starting age effects.[27] We interpret this as evidence of school starting age effects that will confound comparisons of children born in January to children born in December. Henceforth we present only the RD-DD estimates in the main analysis. All corresponding RD-estimates are available upon request.

Besides gender, the results in Table 2.4 suggest no evidence of a compositional change since the covariates balance before and after the reform. The pairwise comparison shows that the estimates are robust to adding control variables. The probability of being male within the Social Science program is slightly lower after the reform. In sum, besides the change in gender composition, we cannot reject that there was no systematic selection to the Technology program, with respect to the other observable characteristics. However, recent developments in the RD-literature recommends additional robustness checks when basing the analysis on the continuity assumption when the running variable is discrete (Cattaneo et al., 2018). Figures A1 and A2 elaborates on the sensitivity to choice of bandwidth and shows the sensitivity to the assumption of continuity. Overall, the point estimates from three different estimations are similar and not significantly different from each other which provides support for using the local linear approach. Importantly for the validity and credibility of the RD design, there is no evidence of a specific jump at the reform cutoff. The figure clearly shows the non-randomness in mathematics grade across the cutoff but is it similar for the reform and control years.

2.5.2 Course-taking Behavior

Did the increase in course selection flexibility significantly alter social science students' course-taking pattern? Table 2.5 presents the regression estimates from the effect of the reform on the probability of taking different courses. The pairwise difference across columns is the inclusion of control variables. Even though the control variables increase R^2, they make little difference to the point estimates.

[27]In particular with respect to controlling for final lower secondary grade in mathematics. For example, McEwan and Shapiro (2008) show that test scores are significantly affected by school starting age.

	MaC (1)	MaC (2)	MaB (3)	MaB (4)	STEM (5)	STEM (6)	non-STEM (7)	non-STEM (8)
				All				
RD-DD	-0.237***	-0.234***	-0.001	-0.001	0.004	0.004	0.082***	0.081***
S.E.	0.004	0.004	0.012	0.011	0.003	0.003	0.005	0.005
Observations	57,668	57,668	57,668	57,668	60,026	60,026	60,026	60,026
R^2	0.095	0.201	0.022	0.056	0.044	0.047	0.093	0.100
Pre-reform $\bar{y}$	0.641	0.641	0.925	0.925	0.035	0.035	0.095	0.095
				Females				
RD-DD	-0.227***	-0.226***	-0.002	-0.001	0.006*	0.006*	0.083***	0.082***
S.E.	0.012	0.012	0.004	0.004	0.003	0.003	0.006	0.006
Observations	36,436	36,436	36,436	36,436	37,779	37,779	37,779	37,779
R^2	0.089	0.198	0.025	0.060	0.044	0.046	0.105	0.111
Pre-reform $\bar{y}$	0.628	0.628	0.925	0.925	0.030	0.030	0.095	0.095
				Males				
RD-DD	-0.250***	-0.245***	-0.002	-0.001	0.002	0.002	0.080***	0.080***
S.E.	0.017	0.013	0.010	0.010	0.005	0.005	0.009	0.010
Observations	21,232	21,232	21,232	21,232	22,247	22,247	22,247	22,247
R^2	0.120	0.220	0.032	0.061	0.066	0.071	0.093	0.097
Pre-reform $\bar{y}$	0.663	0.663	0.924	0.924	0.045	0.045	0.094	0.094
Controls		✓		✓		✓		✓

The table reports the impact of the reform on enrollment in selected courses: Mathematics C, Mathematics B, STEM courses (i.e., courses in science, technology, engineering, and mathematics), and non-STEM courses. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The pairwise difference across columns is the inclusion of control variables.

Table 2.5: Course-taking behavior

As is evident from Table 2.5, the reform did have a significant impact on students' course-taking behavior. In particular, there was considerable substitution of Mathematics C after the reform. That is, we find a highly statistically significant post-reform decrease in the fraction of students enrolled in this course.[28] The estimates suggest a decrease by 23.4 percentage points after the reform, approximately equivalent to a 37 percent decrease given the pre-reform enrollment rate of 64 percent. From a policy point of view, one may worry that the overall decrease in mathematics attainment is driven by a relatively larger course substitution among female subjects after the reform. Researchers have started to search for underlying explanations to observed gender differences in choices of a more mathematics/science intensive curricula. The answer may be attributed

[28]Our register data contains a complete list of grades from each course in upper secondary school and we define attainment as having a grade from Mathematics C. The actual grade does not matter, so students who received a failing grade for the course are still defined as having attained the course.

to a lower taste for competitiveness (Buser et al., 2014) and/or lower valuation of own mathematics abilities among girls (Rapoport and Thibout, 2018). As previously mentioned the mathematics C course is pivotal for a number of study fields in tertiary education with high earning prospects. Therefore, heterogeneous responses in course selection across gender after the reform could potentially induce/exacerbate a gender gap in these mathematics related fields. However our point estimate differs only marginally across genders: compared with the baseline, females were 36 percent less likely to take Mathematics C after the reform; the corresponding decrease for males was 37 percent. Adding controls for predetermined characteristics in column 4 barely affects the magnitude of the coefficient estimates.

To ensure that this drop is not driven by a general decline in mathematics attainment, we also estimate the impact on the preceding math course, Mathematics B. We do not find any changes on the preceding mathematics course. More course choice flexibility did not increase students' probability of enrolling in STEM related courses. Instead, under the flexible curriculum, students chose to enroll in non-STEM elective such as arts and humanities and media. We estimate a 8.1 percentage point increase in the probability of enrolling in non-STEM electives after the reform.[29] From Table 2.5, we conclude that the students experienced a large decrease in mathematics attainment and, importantly, this was not compensated by selecting other STEM-related courses.

2.5.3 Tertiary Education Outcomes and Expected Earnings

Next, we proceed to estimate the impact of the reform on tertiary educational outcomes and annual expected earnings for students in the Social Science program.

[29]Note that STEM and non-STEM are a subset of *electives* offered at the majority of schools. Core content is not included in any of these categories. see Table A1 for a lengthier discussion on what is included in these two categories.

	MaC-field	MaC-field	Speed	Speed	Any TE	Any TE	Degree	Degree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All								
RD-DD	0.001	0.003	0.025	0.028	0.020	0.019*	0.014*	0.012**
Standard Error	0.004	0.004	0.028	0.028	0.013	0.011	0.007	0.006
Observations	53,555	53,555	34,038	34,038	58,126	58,126	58,126	58,126
R^2	0.015	0.040	0.026	0.048	0.020	0.112	0.020	0.079
Pre-reform mean	0.162	0.162	2.591	2.591	0.626	0.626	0.306	0.306
Females								
RD-DD	0.009	0.010*	0.055*	0.052	0.016**	0.017***	0.019***	0.020***
Standard Error	0.006	0.005	0.032	0.032	0.008	0.006	0.006	0.007
Observations	34,422	34,422	23,465	23,465	36,553	36,553	36,553	36,553
R^2	0.016	0.034	0.032	0.048	0.023	0.091	0.026	0.059
Pre-reform mean	0.151	0.151	2.524	2.524	0.683	0.683	0.359	0.359
Males								
RD-DD	-0.013	-0.010	-0.042	-0.036	0.020	0.023	-0.001	0.001
Standard Error	0.009	0.009	0.055	0.057	0.025	0.023	0.013	0.012
Observations	19,133	19,133	10,573	10,573	21,573	21,573	21,573	21,573
R^2	0.039	0.064	0.042	0.059	0.033	0.110	0.025	0.061
Pre-reform mean	0.181	0.181	2.747	2.747	0.527	0.527	0.214	0.214
Controls		✓		✓		✓		✓

The table reports the impact of the reform on tertiary education outcomes. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The pairwise difference across columns is the inclusion of control variables.

Table 2.6: Tertiary education outcomes

Table 2.6 presents the impact of the schooling reform on tertiary education outcomes. The RD-DD approach enables us to disentangle the school starting age effect on y_i from the reform effect under the mild assumption that school starting age effects are constant across the neighboring cutoff years (Carneiro et al., 2015). This assumption cannot be explicitly tested but we complement the analysis with an RD regression identical to equation 2.1 for all of our control years.[30] For women, the point estimates on tertiary education variables for pre- and post-reform control years have similar magnitudes. For men the control cutoff prior to the reform has a slightly higher point estimate for the probability of taking a field requiring Mathematics C, relative to the control cutoffs after the reform. However, this leads to an underestimation of the magnitude of the RD-DD estimate.

We estimate no impact of the reform on the students' probability of choosing a field in tertiary education that requires Mathematics C for eligibility (columns

[30]The RD estimate per control year is plotted in Appendix Figures A3 and A4 for women and men respectively.

1 and 2) for the pooled sample. We find a marginal increase in this probability for women, by 1 percentage points. The result indicates that students opting out of Mathematics C under the flexible curriculum were students who would not have continued their academic career in fields where the course is pivotal under a strict curriculum where the course was mandatory. The positive impact on females showed that some women even continued with higher education in math-related fields which they would not have done prior to the reform. Nor do we find a clear effect on the speed of entering higher education after graduating from upper secondary school (columns 3 and 4). We hypothesized that we would find an impact on these two variables if students opted out of mathematics under the flexible curriculum and then regretted their choice when transferring to the tertiary education cycle. However, the results do not support this hypothesis.

We find a positive and statistically significant effect of the reform on the probability of attending tertiary education (columns 5 and 6). The estimated size of the effect is robust to the inclusion of controls. The control variables are included for two reasons: first, to increase precision, and secondly, they allow us to assess the possible presence of a sorting bias based on observable characteristics. Adding the control variables increases the explained variation in outcomes, R^2, but does not significantly alter the magnitude of the point estimates. If females and males responded differently to a more flexible curriculum and/or to the larger share of female peers in the program after the reform, the estimates are expected to differ. We estimate, on average, a 1.9 percentage point increase in the probability of enrolling in tertiary education, which is equivalent to a 3 percent increase given the pre-reform mean of 63 percent. We lose precision when we split the sample by gender. The estimated coefficient suggest a positive impact of the reform on the probability to enroll in tertiary education for both women and men. However, the effect is only precisely estimated in our female sample. Relative to the baseline, the reform induced a 2.5 percent increase in the probability of enrolling in tertiary education for females.

The reform also led to an increase in the probability of exiting tertiary education with a degree (columns 7 and 8). Here, we can conclude that this increase is entirely driven by women, and after the reform, females were 5.6 percent more likely to exit the tertiary education cycle with a degree.

Taken together, the results regarding the effect of the reform on tertiary education outcomes show a positive impact on students' probability of entering tertiary education. The increase in tertiary education enrollment translates into a higher fraction of students earning a degree. Our results show that the impact is largely driven by a positive impact on females who are significantly more likely to both

enter tertiary education and exit tertiary education with a degree after the reform. With respect to males, we have too low precision to make a conclusion about the effect of the reform on their probability to enroll in tertiary education. However, our results show no impact of the reform on their probability to earn a degree. Hence, our results may signal that females benefit more than males from a flexible curriculum and/or from being in an even more female-dominated group of peers.

In Figures A3 and A4 we show that the results are robust to using estimates with a large bandwidth of 6 months as well as taking the continuity-based assumption into account by adopting the LR framework and show estimates from difference in means comparisons. After the introduction of a new grading system, first applied for the cohort born in 1982, there is a trend of teachers using higher grades in lower secondary school in the first years after introducing the new grading system.[31] This results in a distributional shift both at the 82/83 and 83/84 cutoffs which can be seen in Figure A5. As a robustness check, we re-run the results using only the pre-treatment cutoff 82/83 as a control cohort since they experience a similar shift in distribution but are unexposed to the curriculum reform. The results are shown in Table A5 in the Appendix. The magnitude of the estimated coefficients are in general slightly larger using the restricted control group but the main results remain unchanged. Out of all estimations, only one estimate change from being insignificant using the pooled control years to significant using the pre-treatment control year only.[32] The results are also robust to excluding two municipalities from the analysis, due to a reform change in admission rules to oversubscribed upper secondary schools in a handful of municipalities.[33] In Table A6 we exclude the affected municipalities, Stockholm and Malmö. Since these are two large municipalities, the exclusion leads to lower precision. Overall, the results are qualitatively similar to the main results presented in Table 2.6.

[31]See Holmlund et al. (2014) for a thorough analysis of the Swedish school reforms in 1990s. Major reforms include 1991: Municipalities take over the main responsibility over schools "The municipalication of Swedish schools"; 1992: Charter School reform 1992-1995: Swedish upper secondary school get 16 national programs and one additional program for students who do not get into any of the national programs because of bad grades; 1994: New curriculum (Lpf 94) and criterion referenced grades in upper secondary school; Possibility to choose another school than the one closest to a students home; 1995: New curriculum in elementary school (Lpo 94); 1996: Criterion references grades in elementary school; 1998 Stricter entry requirements to enter upper secondary school.

[32]The probability of having higher education in a mathematics related field becomes significantly lower for boys after the reform.

[33]Söderström and Uusitalo (2010) studies the impact of this reform in the municipality including the capital of Sweden, Stockholm, and find that the new admission rule increased sorting in inner city schools. Molin (2019) expands the analysis and finds that the admission reform changed the socio-demographic composition of students only in two municipalities: Malmö and Stockholm.

We proceed by estimating the impact of the reform on expected earnings in middle age.[34] These are presented in the Appendix. In the two first columns of Table A7 annual expected earnings are based on gender and field of tertiary education studies. We find a modest increase in expected earnings by field for women by approximately 1.2 percent, significant at the 5 percent level.

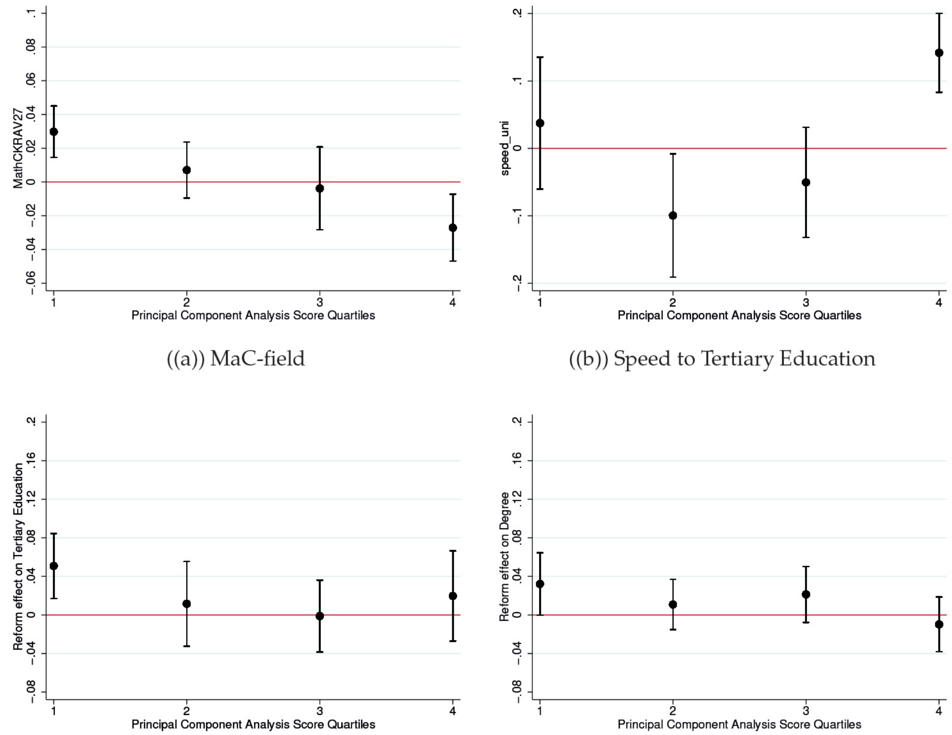
2.5.4 Treatment Heterogeneity

Existing research finds a strong and robust association between an individuals' educational outcomes and parents' SES (Björklund and Salvanes, 2011). According to Björklund and Salvanes (2011), parents' location in the SES distribution may affect a child's educational outcomes through differences in parents' choice of investments in child education and the quantity and quality of information provided to the child about educational prospects. Educational policies and school reforms have the capacity to reduce or reinforce the association between family background and students' educational outcomes and earnings. To investigate the distributional impact of the reform on students educational outcomes along the dimension of SES, we construct an index based on a principal component analysis.[35] The results presented in Figure 2.1 suggest some treatment heterogeneity on outcomes based on parents' SES.

There are no clear patterns in the heterogeneity, the exception being a negative trend with regard to the probability of enrolling in a university program with Mathematics C as an entry requirement. Somewhat surprisingly, our estimates show a negative impact in the higher part of the SES distribution while the impact is positive in the lower quartiles. The most advantaged students have a reduced probability of attending aforementioned programs, combined with a lower speed to tertiary education. This group seem to have opted out of mathematics C and substituted more difficult programs with fast, and easier ones. Somewhat surprisingly, in the first quartile, the magnitude of the effect is sizeable. We estimate a 2.4 percentage point increase in the probability of choosing a post-secondary academic field that requires Mathematics C, equivalent to an increase of 19 percent given the low baseline of 12.4 percent.[36] The lowest SES quartile also drives the increase in the probability of pursuing any tertiary education and the probability of earning a degree. We conclude that low SES students did not fare worse after the reform. Instead, these students are found to benefit the most from the flexibility induced by the reform. On the other hand, the drop

[34] Earnings are measured in 2015 values and the exchange rate per December 31, 2015.
[35] See Appendix Table A8 for details on the construction of the index.
[36] See Appendix, Table A9 for point estimates.

in mathematics requirements led the most advantaged students to substitute the relatively difficult education for fast and easy education.



((a)) MaC-field ((b)) Speed to Tertiary Education

((c)) Any Tertiary Education ((d)) Degree

The figure show the reform coefficient per quartiles from the baseline RD-DD specification with a 3 month bandwidth. The upper and lower bound is calculated at the 90 percent level of significance. The quartiles are based on the distribution of a SES-index constructed based on principal components.

Figure 2.1: Educational Outcomes by Socio-Economic Status Quartile

2.5.5 Possible mediator

One possible pathway to the higher enrollment in tertiary education is through an effect on students' upper secondary GPA. For example, Graetz et al. (2020) employ Swedish register data to estimate students' responses to higher scholastic aptitude test (SAT) and find that higher SAT scores causally increase enrollment and graduation rates, especially among low SES students. Receiving a higher GPA in upper secondary school may also have a positive effect on students' level of aspirations for tertiary education as well as subjective completion beliefs (Kunz and Staub, 2016; DesJardins et al., 2019). Such completion beliefs has further been shown to be predictive of actual completion of tertiary education (Kunz and

Staub, 2016). Hence, students' receiving a higher GPA after the reform may also mediate some of the positive impact of the reform on the estimated probability of earning a degree from tertiary education. Prior to the reform, more than a fifth of all social science students failed the course, and we take this as evidence of the course being particularly difficult for this group of students. A failing grade naturally decreases a student's grade point average (GPA) when applying for tertiary education. Hence, after the reform, students had the option to replace Mathematics C with a course from which they expected to receive a higher grade and thus boost their overall GPA. In general, more overall flexibility in course selection may induce students to act strategically by taking relatively simpler courses or choose courses based on innate ability and preferences. Either of these behavioral responses can be expected to increase students' overall GPA and make them more competitive in the tertiary education application process.

	GPA	GPA	GPA	GPA
	(1)	(2)	(3)	(4)
All				
RD-DD	0.348***	0.343***	0.378***	0.390***
Standard Error	0.110	0.081	0.067	0.038
Observations	54,511	54,511	42,490	42,490
Pre-reform mean	14.134	14.134	14.296	14.296
Females				
RD-DD	0.368***	0.373***	0.344***	0.393***
Standard Error	0.103	0.085	0.109	0.096
Observations	34,843	34,843	27,711	27,711
Pre-reform mean	14.534	14.534	14.648	14.648
Males				
RD-DD	0.270	0.308**	0.401**	0.416***
Standard Error	0.174	0.149	0.174	0.151
Observations	19,668	19,668	14,779	14,779
Pre-reform mean	13.371	13.371	13.584	13.584
Controls		✓		✓

The table reports the impact of the reform on students' grade point average. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The pairwise difference across columns is the inclusion of control variables.

Table 2.7: Grade point average

We estimate the impact of the reform on individual students' final upper secondary school GPA. This is a continuous variable ranging from 0 to 20. During the period of interest, Swedish upper secondary school students received the grade *fail*, *pass*, *pass with distinction* or *pass with special distinction*. If students received a passing grade in all courses, they obtain a final upper secondary GPA of 10. The regression results presented in Table 2.7 confirm a post-reform increase in average GPA.[37] From the RD-DD estimation, we find an approximate 0.34 increase in average GPA. In terms of magnitude, this is approximately equivalent to replacing a grade of pass in Mathematics C with a grade of pass with special distinction in another course. The point estimate is similar for males and females but females come from a slightly higher baseline, resulting in a marginally lower relative increase in GPA. Column 3 and 4 are restricted to students that graduated from the Social Science program (as opposed to first time enrollment). The results are similar when focusing on students that did not change program.

2.6 Conclusion

A rigid and non-flexible curriculum regime provides a tool for policy makers to ensure a desirable level of human capital accumulation and keep the level of acquired knowledge fixed among students, especially within growth enhancing STEM skills (Hanushek and Kimko, 2000). However, it also denies individuals the freedom to take courses they are interested in and that are in line with their personal aspirations. Ultimately, ignoring individual heterogeneity may even cause less able students to shy away from further education.

Our paper contributes to understanding this curriculum trade-off by exploring an upper secondary school reform in Sweden implemented in year 2000. A main feature of the reform involved an increase in the share of elective course work. However, it did so at the cost of reduced mandatory mathematics course load in the most popular upper secondary school program, the Social Science program.

Using detailed register data, we provide evidence that students' course-taking behavior changed after the reform. In particular, mathematics attainment experienced a sharp and robust decrease while enrollment in elective courses in non-STEM fields increased dramatically. Our results show that female and male

[37]One concern is the possibility that GPAs trended upwards due to factors unrelated to the reform, e.g., grade inflation. While we cannot assess such inflation concerns by looking at upper secondary school GPAs, we plot the distribution of lower secondary GPAs in Figure A6 for the 1982–1988 cohorts. Except for the changes in the distribution for early cohorts, discussed in Section 2.5.3, there is no clear evidence of grade inflation and we conclude that the increase in upper secondary school GPA was mainly driven by substitution of courses.

students discard mathematics at a similar rate after the reform. Among neither females nor males was the decrease in mathematics attainment compensated by an increase in STEM-field related coursework.

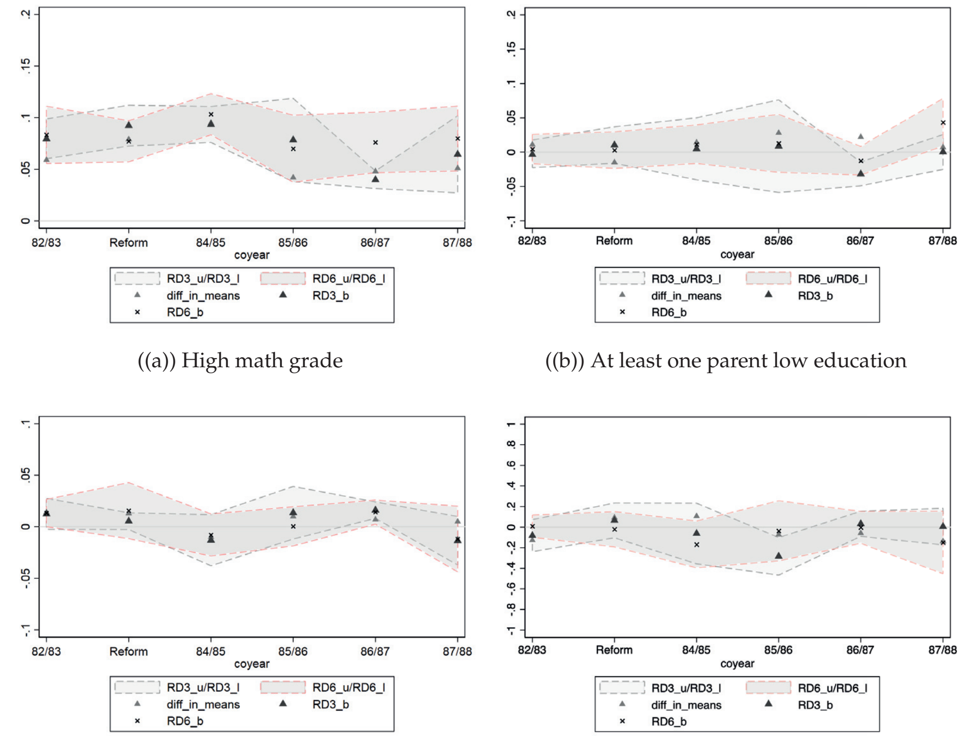
We estimate a positive causal impact of the reform on students' probability of ever enrolling in tertiary education, an increase of 3 percent. The positive impact on Social Science students' enrollment in tertiary education translates into an increase in the probability of students exiting tertiary education with a degree. Estimating the effect by gender shows that the positive impact on the probability of earning a degree was driven by a large and positive impact for females. Interestingly, a marginally significant effect for women and no impact for men is found on the probability of having the highest degree in a relatively mathematics-intensive field. Nor does the reform affect the speed of students entering tertiary education after graduating from upper secondary school, on average. However, the average outcome masks the distributional effects of the reform. Our heterogeneity analysis reveals that relatively disadvantaged students (measured along a socio-economic status index) were not negatively affected by the curriculum reform. Rather, students in the lowest SES quartile seem to have benefited the most from the more flexible curriculum and have a large increase of 19 percent in the probability of entering a mathematics intensive program. On the other hand, the most advantaged students had a reduced probability of attending the same program as well as a lower speed to enter tertiary education. To the extent that majors in Business and Economics give relatively higher earnings, this group were harmed by the reform.

We provide evidence that the decreased required course load in mathematics and the increase in GPA can explain part of the increase in transmission from upper secondary school to tertiary education. The increase in GPA is in line with results found in Yu and Mocan (2018), the paper most closely related to our work, explicitly investigating curriculum flexibility. They, too, find a positive impact on GPAs when students in China were given more course choice flexibility.

Our results are informative for policy makers speculating about the optimal level of flexibility and mathematics content. Increasing flexibility had a positive impact on academic outcomes. The decline in mathematics attainment lead relatively more disadvantaged students in particular to choose more advanced programs than their peers. In particular, the most advantaged students were negatively affected by the reform in terms of chosen programs in higher education. As such, the reform possibly lead to a dismantling of the socio-demographic gradient in educational attainment.

Appendices 2

Figures



((a)) High math grade

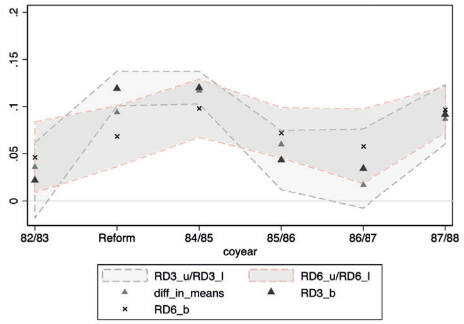
((b)) At least one parent low education

((c)) At least one parent foreign

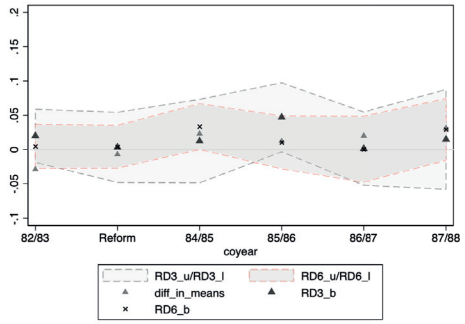
((d)) Log of fathers avg earnings

The figure shows the reform coefficient from the baseline RD estimate from separate local linear regressions using a 3 and 6 month window around the cutoff and its corresponding confidence interval for students on the Social science program. The reform from a simple difference-in-means is, equivalent to having a 1 month bandwidth and a polynomial of degree 0. The year along the x-axis represent the birth cohorts. The upper and lower bounds are calculated at the 95 percent level of significance.

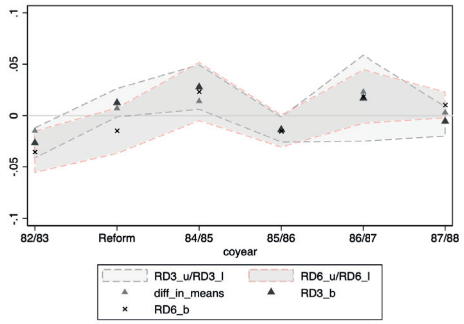
Figure A1: Pre-determined covariates - Women



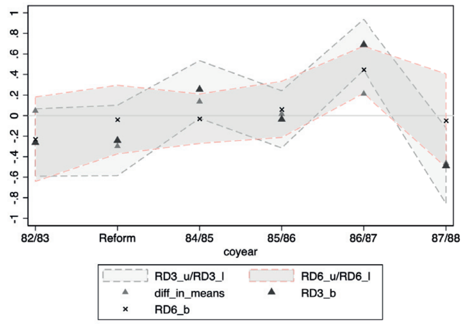
((a)) High math grade



((b)) At least one parent low education



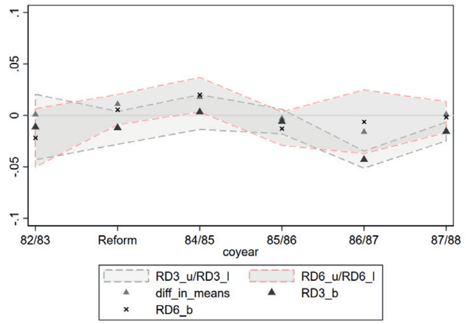
((c)) At least one parent foreign



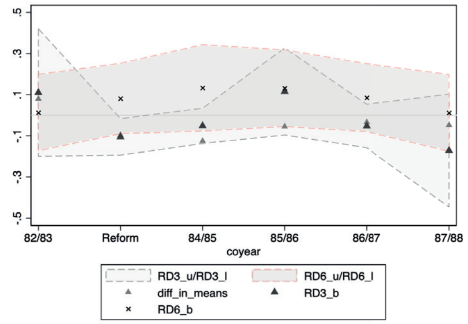
((d)) Log of fathers avg earnings

The figure shows the reform coefficient from the baseline RD estimate from separate local linear regressions using a 3 and 6 month window around the cutoff and its corresponding confidence interval for students on the Social science program. The reform from a simple difference-in-means is, equivalent to having a 1 month bandwidth and a polynomial of degree 0. The year along the x-axis represent the birth cohorts. The upper and lower bounds are calculated at the 95 percent level of significance.

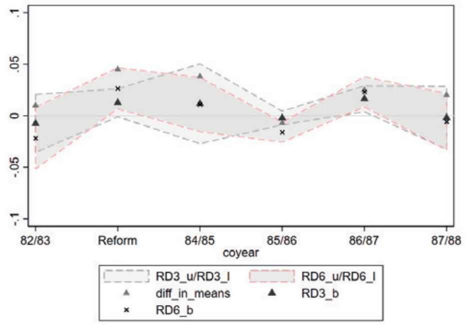
Figure A2: Pre-determined covariates - Men



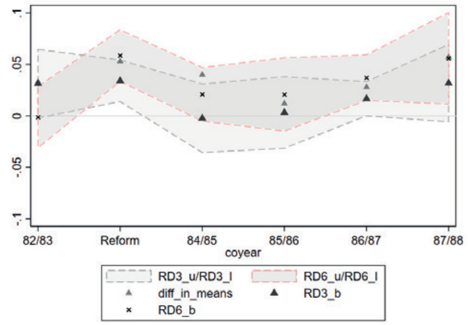
((a)) MaC-field



((b)) Speed to Tertiary Education



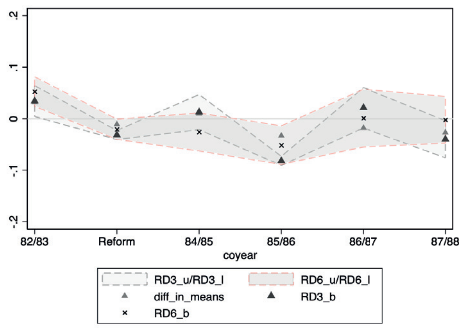
((c)) Any Tertiary Education



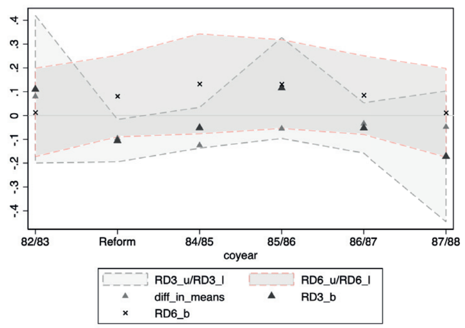
((d)) Degree

The figure shows the reform coefficient from the baseline RD estimate from separate local linear regressions using a 3 and 6 month window around the cutoff and its corresponding confidence interval for students on the Social science program. The reform from a simple difference-in-means is, equivalent to having a 1 month bandwidth and a polynomial of degree 0. The year along the x-axis represent the birth cohorts. The upper and lower bounds are calculated at the 95 percent level of significance.

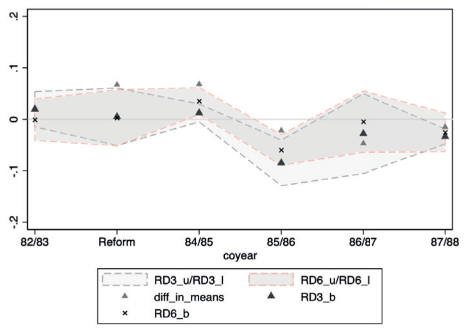
Figure A3: Academic outcomes - Women



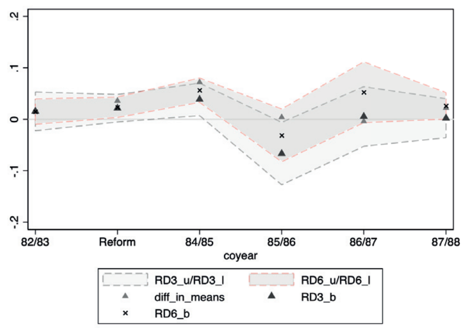
((a)) MaC-field



((b)) Speed to Tertiary Education



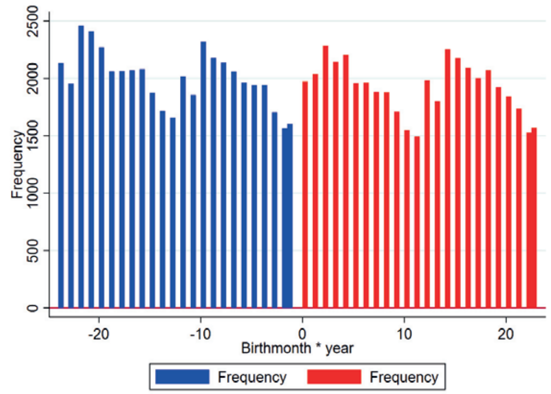
((c)) Any Tertiary Education



((d)) Degree

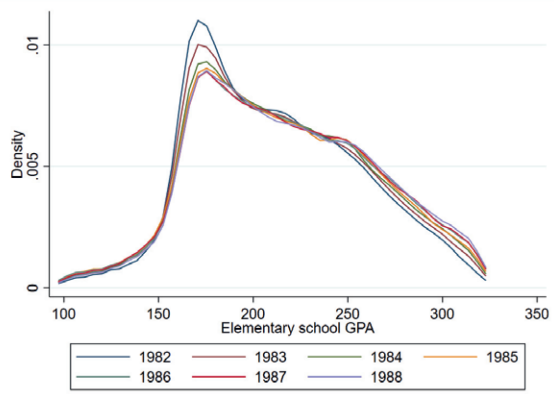
The figure shows the reform coefficient from the baseline RD estimate from separate local linear regressions using a 3 and 6 month window around the cutoff and its corresponding confidence interval for students on the Social science program. The reform from a simple difference-in-means is, equivalent to having a 1 month bandwidth and a polynomial of degree 0. The year along the x-axis represent the birth cohorts. The upper and lower bounds are calculated at the 95 percent level of significance.

Figure A4: Academic outcomes - Men



The histogram shows the frequency of observations among social science students born two years before and after the reform. The sample is restricted to individuals born January 1982–December 1985. Zero (0) denotes the cutoff date, born in January 1984, and include individuals who entered upper secondary school after the reform

Figure A5: Histogram, birthdate in months



The figure shows the distribution of student lower secondary school GPA for the entire population of students born 1982–1988.

Figure A6: Density plot of lower secondary school GPA

Tables

STEM	Credits
Biology A	100
Biology B	50
Physics A	100
Chemistry A	100
Mathematics D	100
Environment	100
non-STEM	**Credits**
Cultural history and contemporary art	50
Leadership	100
Media	100
Multimedia	100
Humanities	100
Text communication A	100

Table A1: STEM and non-STEM courses

Courses included in STEM and non-STEM categories are selected based on the following premises. First, the courses are included on the Swedish National Agency for Education's list of suitable elective courses for the Social Science program (GY2000:16). Second, the courses are not listed as mandatory in any of the specialization tracks within the Social Science program. Third, the courses must have existed both before and after the implementation of GY2000. The remaining courses are coded as STEM if they are mandatory for students enrolled in any of the specialization tracks within the Natural Science program. The rest of the courses are coded as non-STEM. The included STEM courses basically cover all possible STEM course electives that could be offered in the elective course package. However, the included non-STEM courses likely cover only a small fraction of all possible non-STEM courses that upper secondary schools may offer their students.

	Before reform cohort 1983			After reform cohort 1984			
	Mean	Std. Error	Obs.	Mean	Std. Error	Obs.	Diff. in means
Social Science							
High math ability	0.43	0.49	23278	0.43	0.50	23079	0.0083
Lower secondary school grade	221.89	39.68	23278	225.20	42.39	23079	3.3081***
Male	0.37	0.48	23278	0.36	0.48	23079	-0.0116**
Immigrant	0.12	0.32	21943	0.13	0.33	21693	0.0114***
$LowEducation_p$	0.66	0.48	21759	0.64	0.48	21522	-0.0195***
$LogAvgWage_f$	10.99	3.80	22140	11.03	3.80	21927	0.0316
Natural Science							
High math ability	0.81	0.39	18825	0.85	0.36	14789	0.0346***
Lower secondary school grade	248.81	41.05	18818	256.62	41.42	14788	7.8047***
Male	0.60	0.49	18825	0.55	0.50	14789	-0.0478***
Immigrant	0.13	0.34	17879	0.15	0.35	13942	0.0167***
$LowEducation_p$	0.51	0.50	17781	0.48	0.50	13843	-0.0307***
$LogAvgWage_f$	11.16	3.73	18034	11.23	3.69	14091	0.0768
Technical Program							
High math ability				0.54	0.50	6384	
Lower secondary school grade				215.75	37.68	6384	
Male				0.90	0.30	6384	
Immigrant				0.10	0.30	6080	
$LowEducation_p$				0.67	0.47	6052	
$LogAvgWage_f$				11.22	3.52	6143	
Vocational							
High math ability	0.18	0.39	35588	0.18	0.39	36126	0.0015
Lower secondary school grade	183.24	40.95	35587	185.06	42.69	36125	1.8172***
Male	0.53	0.50	35588	0.51	0.50	36126	-0.0234***
Immigrant	0.09	0.29	33403	0.09	0.29	33861	-0.0007
$LowEducation_p$	0.84	0.36	33127	0.84	0.37	33543	-0.0045
$LogAvgWage_f$	10.49	4.09	33793	10.58	4.05	34280	0.0901**
All Students							
High math ability	0.41	0.49	77821	0.41	0.49	80378	-0.0015
Lower secondary school grade	210.68	48.82	77813	212.19	50.07	80376	1.5070***
Male	0.50	0.50	77821	0.50	0.50	80378	0.0042
Immigrant	0.11	0.31	73324	0.11	0.32	75576	0.0034*
$LowEducation_p$	0.71	0.46	72759	0.70	0.46	74960	-0.0044
$LogAvgWage_f$	10.80	3.93	74073	10.88	3.88	76441	0.0792***
Social Science	0.30	0.46	77821	0.29	0.45	80378	-0.0120***
Natural Science	0.24	0.43	77821	0.18	0.39	80378	-0.0579***
Technical				0.08	0.27	80378	
VOC	0.46	0.50	77821	0.45	0.50	80378	-0.0079**

Table A2: Summary statistics of background characteristics

Vocational	RD	RD	RD-DD	RD-DD
	All			
Reform	-0.060***	-0.028***	-0.012***	-0.012***
Standard Error	0.002	0.002	0.004	0.004
Observations	42,288	42,288	268,835	268,835
Pre-reform Mean	0.394	0.394	0.394	0.394
	Girls			
Reform	-0.052***	-0.018***	0.005	0.005
Standard Error	0.002	0.002	0.006	0.005
Observations	20,636	20,636	131,286	131,286
Pre-reform Mean	0.385	0.385	0.385	0.385
	Boys			
Reform	-0.067***	-0.037***	-0.028***	-0.028***
Standard Error	0.003	0.003	0.005	0.004
Observations	21,652	21,652	137,549	137,549
Pre-reform Mean	0.403	0.403	0.403	0.403
Natural Science/Technology				
	All			
Reform	0.068***	0.032***	0.043***	0.043***
Standard Error	0.002	0.003	0.005	0.004
Observations	42,288	42,288	268,835	268,835
Pre-reform Mean	0.188	0.188	0.188	0.188
	Girls			
Reform	0.027***	-0.003***	0.010***	0.010***
Standard Error	0.001	0.001	0.003	0.003
Observations	20,636	20,636	131,286	131,286
Pre-reform Mean	0.153	0.153	0.153	0.153
	Boys			
Reform	0.104***	0.064***	0.075***	0.075***
Standard Error	0.006	0.006	0.008	0.006
Observations	21,652	21,652	137,549	137,549
Pre-reform Mean	0.222	0.222	0.222	0.222
Controls		✓		✓

The table reports the impact of the reform on the probability of enrolling in vocational programs and the Natural Science/Technology programs in upper secondary school. The two first columns show the RD regression results using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The pairwise difference across columns is the inclusion of control variables.

Table A3: Probability of enrolling in other programs

	$HighMath_i$	$Male_i$	$Loweduc_p$	$Foregin_p$	$\text{Ln}\overline{Earnings}_f$
	(1)	(2)	(3)	(4)	(5)
			All		
RD	0.104***	0.017***	0.001	-0.001	0.042
Standard Error	0.002	0.003	0.003	0.001	0.059
Observations	35,368	35,368	35,368	35,368	35,368
RD-DD	-0.003	0.017	-0.009***	0.001	0.063
Standard Error	0.007	0.004	0.003	0.003	0.062
Observations	213,291	213,291	213,291	213,291	213,291
Pre-reform Mean	0.382	0.496	0.717	0.108	10.837
			Females		
RD	0.083***		-0.009**	0.005**	0.233**
Standard Error	0.003		0.003	0.001	0.068
Observations	17,641		17,641	17,641	17,641
RD-DD	-0.014*		-0.016***	0.003	0.115
Standard Error	0.007		0.004	0.003	0.083
Observations	106,196		106,196	106,196	106,196
Pre-reform Mean	0.399		0.728	0.106	10.791
			Males		
RD	0.128***		0.014**	-0.007**	-0.148**
Standard Error	0.002		0.005	0.002	0.049
Observations	17,727		17,727	17,727	17,727
RD-DD	0.008		-0.001	-0.000	0.011
Standard Error	0.010		0.004	0.005	0.070
Observations	107,095		107,095	107,095	107,095
Pre-reform Mean	0.366		0.705	0.110	10.884

The table reports the impact of the reform on pre-determined characteristics. In panel C we present the RD regression results using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff. In panel D We present the RD-DD estimates where we augment the regression with students born in October to March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The pairwise difference across columns is the inclusion of control variables.

Table A4: Balancing test of pre-treatment characteristics: All students

	MaC-field (1)	MaC-field (2)	Speed (3)	Speed (4)	Any TE (5)	Any TE (6)	Degree (7)	Degree (8)
All								
RD-DD	-0.007	-0.006	0.014	0.014	0.025	0.021	0.022***	0.019***
Standard Error	0.004	0.005	0.029	0.030	0.014	0.012	0.004	0.004
Observations	18,832	18,832	11,673	11,673	20,383	20,383	20,383	20,383
Pre-reform mean	0.162	0.162	2.591	2.591	0.626	0.626	0.306	0.306
Females								
RD-DD	0.018**	0.019**	0.035	0.025	0.027**	0.027**	0.032**	0.032**
Standard Error	0.007	0.008	0.024	0.023	0.012	0.010	0.012	0.012
Observations	12,151	12,151	8,133	8,133	12,843	12,843	12,843	12,843
Pre-reform mean	0.151	0.151	2.524	2.524	0.683	0.683	0.359	0.359
Males								
RD-DD	-0.050***	-0.050***	-0.043	-0.036	0.014	0.016	0.002	0.002
Standard Error	0.007	0.006	0.086	0.087	0.024	0.024	0.014	0.012
Observations	6,681	6,681	3,540	3,540	7,540	7,540	7,540	7,540
Pre-reform mean	0.181	0.181	2.747	2.747	0.527	0.527	0.214	0.214
Controls		✓		✓		✓		✓

The table reports the impact of the reform on tertiary education outcomes. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring pre-treatment non-reform year 1982–1983. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The pairwise difference across columns is the inclusion of control variables.

Table A5: Tertiary educational outcomes - using only pre-treatment cutoff

	MaC-field (1)	MaC-field (2)	Speed (3)	Speed (4)	Any TE (5)	Any TE (6)	Degree (7)	Degree (8)
All								
RD-DD	-0.002	0.004	0.028	0.031	0.024*	0.023**	0.013*	0.012*
Standard Error	0.006	0.006	0.029	0.031	0.014	0.011	0.007	0.006
Observations	47,441	47,441	29,967	29,967	51,515	51,515	51,515	51,515
Pre-reform mean	0.158	0.158	2.595	2.595	0.622	0.622	0.308	0.308
Females								
RD-DD	0.005	0.006	0.047	0.044	0.017***	0.019***	0.017**	0.020**
Standard Error	0.008	0.008	0.034	0.034	0.006	0.005	0.008	0.009
Observations	30,604	30,604	20,777	20,777	32,502	32,502	32,502	32,502
Pre-reform mean	0.151	0.151	2.532	2.532	0.679	0.679	0.362	0.362
Males								
RD-DD	-0.004	-0.002	-0.017	-0.009	0.029	0.031	-0.000	0.000
Standard Error	0.010	0.009	0.056	0.057	0.033	0.030	0.015	0.013
Observations	3,298	3,298	1,711	1,711	3,706	3,706	3,706	3,706
Pre-reform mean	0.181	0.181	2.747	2.747	0.527	0.527	0.214	0.214
Controls		✓		✓		✓		✓

The table reports the impact of the reform on tertiary education outcomes. We present the RD-DD estimates where we augment the regression with students born in October–March in the pre-treatment non-reform year 1982–1983. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The pairwise difference across columns is the inclusion of control variables.

Table A6: Tertiary educational outcomes - excluding Stockholm and Malmö

	$E[Earnings]_f$ (1)	$E[Earnings]_f$ (2)	$E[Earnings]_{fl}$ (3)	$E[Earnings]_{fl}$ (4)
		All		
RD-DD	0.006	0.011	0.002	0.008
S.E.	0.007	0.007	0.006	0.005
N	58,126	58,126	58,080	58,080
		Females		
RD-DD	0.012*	0.012**	0.008	0.009*
S.E.	0.006	0.005	0.006	0.005
N	36,553	36,553	36,529	36,529
		Males		
RD-DD	0.007	0.009	0.004	0.006
S.E.	0.012	0.011	0.009	0.009
N	21,573	21,573	21,551	21,551
Controls		✓		✓

The table reports the impact of the reform on students' annual expected earnings, by field and field/level, and actual earnings at age 30. All earning variables are logarithmic. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel. The pairwise difference across columns is the inclusion of control variables.

Table A7: Earning Outcomes

Principal Component Analysis

The SES index is constructed based on a principal component analysis (PCA). The idea is that many socio-demographic characteristics are correlated. For example, having a low educated mother is strongly associated with lower average income. We can exploit the correlation structure to construct one variable, an index, combining the correlated variables. PCA reduces the dimensionality by finding linear combinations of the separate components that explain the most variability. We use the first component since it explains the maximal variation in the original set of predicting variables.

Variable	PC1
$LowEducation_f$	0.4440
$LowEducation_m$	0.4837
$Foreign$	0.3344
$LnAvgEarnings_f$	-0.4702
$LnAvgEarnings_f$	-0.4858
Percent of variation explained	32%

Table A8: Results of the PCA analysis

The principal component has unit length such that:
$PC1 = 0.4440^2 + 0.4837^2 + 0.3344^2 - 0.4702^2 - 0.4858^2 = 1.$
Based on the principal component (PC1) we create a new variable (PCA) that predicts the individual's SES index according to a weighted linear combination of the original set of predicting variables:
$PCA_i = 0.4440loweduc_f + 0.4837loweduc_m + 0.3344foreign - 0.4702lnavgearn_f - 0.4858lnavgearn_m$
Finally, we invert the index so that higher values of the index correspond to higher SES and vice versa.

RD-DD	1st Q(Lowest)	2nd Q	3rd Q	4th Q (Highest)
MaC-field	0.024**	0.003	-0.009	-0.043***
Standard Error	0.011	0.011	0.016	0.015
Observations	11,082	11,072	11,124	10,831
R^2	0.030	0.031	0.031	0.032
Pre-reform mean	0.124	0.151	0.194	0.241
Speed to Uni	0.103	-0.095	-0.084	0.143***
Standard Error	0.070	0.095	0.064	0.033
Observations	6,365	6,861	7,708	8,880
R^2	0.061	0.056	0.053	0.041
Pre-reform mean	2.481	2.655	2.650	2.460
Any TE	0.053***	0.005	0.034	0.020
Standard Error	0.018	0.021	0.023	0.019
Observations	11,725	11,721	11,867	11,679
R^2	0.036	0.039	0.028	0.028
Pre-reform mean	0.551	0.614	0.716	0.811
Degree	0.046**	0.005	0.043*	-0.030
Standard Error	0.019	0.014	0.024	0.023
Observations	11,725	11,721	11,867	11,679
R^2	0.039	0.044	0.036	0.034
Pre-reform mean	0.257	0.303	0.365	0.403

The table reports the impact of the reform on tertiary education outcomes by SES quartile. We present the RD-DD estimates where we augment the regression with students born in October–March in the neighboring non-reform years 1982–1983, 1984–1985, 1985–1986, 1986–1987, and 1987–1988. The discontinuity in outcomes is estimated with a local linear regression with separate trends on each side of the cutoff, using a 3-month bandwidth on each side of the cutoff and a triangular kernel.

Table A9: PCA: Tertiary education outcomes by SES quartile

Appendix B - References in Swedish

Ds 2008:13 (2008). En ny betygsskala. Stockholm: Utbildningsdepartementet.

Grevholm B.(1999). Varför och hur revideras kursplanerna för gymnasieskolan? Nämnaren 26(1), 41-44

GY2000:14 (2000). Naturvetenskapsprogrammet: programmål, kursplaner, betygskriterier och kommentarer. Stockholm: Skolverket.

GY2000:16 (2000). Samhällsvetenskapsprogrammet:programmål, kursplaner, betygskriterier och kommentarer. Stockholm: Skolverket.

GY2000:17 (2000). Teknikprogrammet: programmål, kursplaner, betygskriterier och kommentarer. Stockholm: Skolverket

GyVux 1994/95:14 (1995). Naturvetenskapsprogrammet: Programmål, kursplaner, betygskriterier och kommentarer. Stockholm: Skolverket

GyVux 1994/97:16 (1997). Samhällsvetenskapsprogrammet:Programmål, kursplaner, betygskriterier och kommentarer. Stockholm: Skolverket

GyVux 1994/97:17 (1997). Programhandledning: Programöversikter samt förteckning över ämnen och kurser. Stockholm: Skolverket

Högskoleverket (2010). Swedish universities & university colleges Short version of annual report 2010.
http://english.uka.se/download/18.6b3261a315a296ca0f3e77bd/1487932600878/annual−report−2010.pdf (accessed Aug 31, 2018)

Proposition 1997/98:169 (1998). Gymnasieskola i utveckling – kvalitet och likvärdighet. Stockholm: Utbildningsdepartementet.

SACO (2018). Fastighetsmäklare.
https://www.saco.se/studieval/yrken−a−o/fastighetsmaklare/ (accessed Aug 30, 2018)

SCB (2011). Utbildningssystemet.
https://www.scb.se/sv/Hitta−statistik/Publiceringskalender/Visa−detaljerad−information/?publobjid=1739 (accessed Aug 31, 2018)

Skolverket (1998). Redovisning av regeringens uppdrag 1998- 03-12 angående gymnasieskolans utveckling – U98/1135/S(Dnr. 98:956)

Skolverket (2000).Beskrivande data om barnomsorg och skola 2000 (Dnr:2000:2208)

Skolverket (2001). Beskrivande data om barnomsorg,skola och vuxenutbildning 2001. (Dnr: 71-2001:3248)

Skolverket (2005). Var fjärde 20-åring saknar slutbetyg från gymnasiet. *https : //web.archive.org/web/*20071022062047*/http : //www.skolverket.se/sb/d/*204 %3*Bjsessionid*%3*DFDCB*3425*AAA*8*BFF*25*B*3450*F*92*DC*5C112 (accessed Aug 30, 2018)

Universitets- och högskolerådet (UHR) (2016). Tabell för områdesbehörigheter. Available at: *https : //www.uhr.se/globalassets/*$_u$*hr.se/studier − och − antagning/tilltrade − till − hogskolan/tabell − for − omomradesbehorigheter −* 2016.*pdf* (accessed Jan 31, 2018)

Chapter 3

LOCAL MEDIA INFORMATION AND CHOICE OF PRIMARY HEALTH CARE PROVIDER

Co-authored with Jens Dietrichson and Gustav Kjellsson

Abstract

Patient choice in health care markets requires the patients to be well-informed about quality and act on this information. How patients choose providers, and how that is affected by various sources of information, remain open questions. We study how information from local media coverage influences choices of primary care providers. We use a novel source of information, local newspaper articles, as treatment in a staggered difference-in-difference framework to examine how media coverage affects the number of patients enrolled with a given primary care provider. We compare outcomes between treated and untreated providers before and after a negative or positive publication and perform an event study and a static difference-in-differences estimations. The main analysis does not detect any effect of either positive or negative coverage. The heterogeneity analysis reveals larger but still small and insignificant effects for articles with stronger negative and positive coverage, and in rural compared to urban areas. The small and insignificant effects provide important information for the functioning of patient choice markets and are in line with earlier findings on the topic, using a different source of information.

3.1 Introduction

In the last decades, many governments have introduced choice in health care markets where patients choose from a menu of providers. Patient choice may improve the matching of patients to providers and strengthen providers' incentives to compete on quality compared to a situation with no choice (Besley and Ghatak, 2003; Gaynor et al., 2016). However, to realise the potential improvements, patients need to be sufficiently informed of provider quality and act on this information when they choose providers. While market efficiency in general is dependent on consumers rewarding high quality providers, it may be even more important to have well-informed consumers in health care markets as prices are fixed or non-existent. In regular markets, prices have a coordinating function, which is missing in consumer choice markets (Hayek, 1944; Gode and Sunder, 1993).

The health economic literature points out that patients may have problems with identifying high quality providers (Arrow, 1963; Cutler, 2011). While there is a growing literature on how patients make choices in health care (e.g., Dahlgren et al., 2021; Biørn and Godager, 2010; Iversen and Lurås, 2011; Santos et al., 2017; Chandra et al., 2016; Brekke et al., 2014; Avdic et al., 2019; Varkevisser et al., 2012), and how such choices are affected by initiatives to disseminate quality, through postal mail, publications of report cards, rankings, and online ratings (e.g., Anell et al., 2021; Chen and Lee, 2021; Bensnes and Huitfeldt, 2021; Luca and Vats, 2013; Chartock, 2021), there is little knowledge on how other sources of information affect these markets.

In this study, we focus on media coverage and its effect on choices in health care markets. More specifically, we study how information from local newspapers affects the demand for primary care providers in two Swedish regions (Region Skåne and Region Västra Götaland). Media information differs from other types of quality dissemination, since news articles generally reports specific (negative or positive) events and package the information as a story on a given provider. Stories about, for example, mistreatment of patients or new services may inform and, crucially, engage the local public more than report cards and rankings do. Such engagement may increase the probability that patients act on the information by switching from low-quality to high-quality providers.

Reports in media have been shown to affect health behaviour, such as vaccine uptake (e.g., Anderberg et al., 2011; Smith et al., 2008; Chang, 2018; Suppli et al., 2018; Brilli et al., 2020) as well as mobility and adherence to recommendations during the Covid pandemic (Zhuang et al., 2022; Kim et al., 2020). King

et al. (2017) further demonstrate that newspaper articles, also in small news media outlets, influence individuals' views and initiate discussion in social media. Examining newspaper articles may thus capture broader effects of other media on providers' reputations. There is therefore likely scope for media coverage to affect choices of primary care providers.

It is of specific interest to study the effects of information in primary care markets. First, the choice of primary care provider is of potential importance for the patient. There is considerable variation in primary care quality, both in Sweden and in other contexts (Anell et al., 2021, 2022; Ginja et al., 2022; Chartock, 2021), and the choice may have large effects on patient health outcomes, including on mortality (e.g., Ginja et al., 2022).[1]

Second, previous studies indicate that increasing patient choice in markets with regulated prices have not led to substantial improvements of primary care quality (Dietrichson et al., 2020; Gravelle et al., 2019). A potential explanation for this is that information friction may be substantial, and more significant than in other health care sectors. Since primary health care is more multifaceted and quality is more multidimensional than for secondary care, it may be more challenging to observe quality. Patients may also lack the guidance of a GP or a family doctor, which may be present when choosing a hospital or a specialist.[2] When choosing a primary care provider, patients have to rely on other sources of information and may therefore, to a larger extent, have to search for information online, look for public ratings, or rely on word of mouth among neighbors and friends.[3]

However, newspapers and other media may potentially mitigate information frictions. That is, if we find strong responses to newspaper articles, then patients may be reasonably well informed and the lack of quality improvements should instead be attributed to supply-side issues.[4] Thus, examining the responses to

[1]Primary care is the part of the health care system most people interact with on a regular basis. Primary care providers are often the patients first point of contact with the health care system and are also guiding the patient to other parts of the system. For patients with chronic diseases, primary care providers are responsible for disease management. The institutional details are described in more detail in section 3.2.

[2]However, the finding that the effects of increasing patient choice on quality are often small or mixed on hospital markets may be an indication that patients are often not sufficiently informed of hospital choices (e.g., Cooper et al., 2011; Gaynor et al., 2013; **?**; Moscelli et al., 2016, 2018; Skellern, 2017).

[3]Searching for information is costly, which may explain why few patients (across healthcare settings and countries) actively search for and use comparative information before they make their choice of provider (Victoor et al., 2012).

[4]Examples of such issues include the fact that acquiring new facilities comes with fixed costs, which limits the incentives to expand beyond current capacity constraints (Vengberg et al., 2019); that quality determinants like the human capital of individual physicians and nurses may not be scalable (e.g., recruitment is a challenge according to a majority of primary care managers in Sweden; Petterson and Jaktlund, 2013); that physicians are uninformed of the reasons behind patients' choices

newspaper articles may give us a better understanding of why improvements are lacking, which is needed to design an adequate policy response.

To study how media coverage affects patients' choices of provider, we estimate the effect of newspaper articles on (the logarithm of) the provider list size (i.e., number of registered patients) in an event study and difference-in-difference setting. Treatment is based on the publishing date of articles mentioning primary care providers - either in positive or negative terms - retrieved from a database of all articles from the relevant (local) newspapers from 2013 to 2016. We use individual-level enrollment data and geographical background data to compute the list size in a given day for each of the primary care providers. Using these data sources, we compare monthly changes in list size of providers mentioned in negative or positive terms in newspaper articles with the development of non-mentioned providers. By examining the effects at the level of the providers, we capture both push and pull factors. Push factors occur when patients switch to other providers at a higher (lower) rate than usual when their current provider gets negative (positive) publicity. Pull factors occurr when negative (positive) newspaper coverage on their current provider may make patients less (more) likely to switch to the affected center.

Since there is variation in the timing of the publishing date of the news articles, the treatment is staggered. To address the shortcomings of standard difference-in-differences (DiD) and two-way fixed effects (TWFE) models in such a setting, we use an estimation strategy inspired by Cengiz et al. (2019) and Deshpande and Li (2019).[5] This approach, referred to as stacked regression, generates one dataset per treatment date of treated and control units and allows for both static TWFE DiD and a dynamic event study estimation. The method is also easy to combine with the approach of Bilinski and Hatfield (2018) to address violations of the parallel trend assumption.

A first result of the study is namely that we find differentiated pre-treatment trends for treated and control providers for both positive and negative articles. Since this pattern indicates that at least some patients have and act on quality information *before* the articles come out, it is interesting in relation to our research question. However, although the trend differences are small, they are also problematic for identification. We address the problem by using the 12 months before the publication date to estimate a differentiated pre-trend (separately for each *stack*) and estimate how the treated group deviates from the extrapolated trend after the publication date.

of providers (Vengberg et al., 2019); and that capitation-based payment systems create incentives to limit service provision (Newhouse, 1996; Ellis, 1998; Anell et al., 2022).

[5] See for example Goodman-Bacon (2021a) for an overview of the issues.

The main analysis is not able to detect any significant effect of either positive or negative coverage. While the effect of positive articles is close to zero, the event study (and the DiD) suggests that there is a trend break following the treatment. However, like the trend itself, the effects are small and insignificant. We test heterogeneity between different groups of articles by categorizing both positive and negative articles into those more or less likely to affect patients' enrollment. We find a more pronounced effect among articles classified as strongly negative or positive, but the estimates are still small and insignificant. When splitting the data between providers located in different types of markets, namely in urban and rural towns, we record a stronger, yet insignificant effect among rural providers both with regards to positive and negative news.

We contribute to the previous literature in two ways. First, the lack of substantial effects of media coverage on choice of provider is in line with studies suggesting that information from postal mail or online rating tools has small effects on the choice of providers (Anell et al., 2021; Chen and Lee, 2021; Bensnes and Huitfeldt, 2021; Luca and Vats, 2013; Chartock, 2021). In general, the choice of providers tend to be sticky and individuals are primarily changing provider if they move. For example, experiments conducted in one of our study regions show that while information interventions affect choices, they do so only to a limited degree, and it is not clear that information leads to better choices (Anell et al., 2021, 2022).[6] Our results thus further our understanding of the functioning of patient choice markets. If patients do not punish low-quality providers or reward high-quality providers, then the suggested positive effect of patient choice seems unlikely to appear. In turn, our results increase the likelihood that information frictions is a major explanation of the lack of substantial quality improvements resulting from patient choice in primary care markets.

Second, our also study relates to a growing, but still limited, literature aiming to understand patients' choices of health care provider, and in particular the strand focusing on choices of GPs, primary care practice or family medicine providers (e.g., Dahlgren et al., 2021; Biørn and Godager, 2010; Iversen and Lurås, 2011; Santos et al., 2017). Our results adds to the literature on how quality dissemination affects health care choices by focusing on a new type of information. According to several literature reviews, public reporting has had little impact on patients' choices of provider (e.g., Totten et al., 2012; Mukamel et al., 2014). An exception is Pope (2009), who shows that rankings published in newspapers, which some of our articles also contain, increase demand and revenues for hos-

[6] The related literature on health plan choices also typically find small or no absolute effects of information (e.g., Knutson et al., 1998; Hibbard et al., 2002; Abaluck and Gruber, 2016; Ericson et al., 2017; Domurat et al., 2021), unless the provided information is unidimensional as in Kling et al. (2012).

pitals. However, it is unclear how much the choice of hospital depends on the expert guidance given by the patient's family doctor, and the study does not provide evidence on how more general media reporting influences patients' choice of provider. To the best of our knowledge there is no such previous study related to health care markets.

The remainder of the paper is structured as follows. Section 3.2 describes the institutional features of the Swedish primary health care market, Section 3.3 describes the data in greater detail and Section 3.4 discusses the identification strategy used in this paper. Section 3.5 presents the main results, examines treatment heterogeneity, and the results of several robustness checks. Finally, Section 3.6 contains our concluding remarks.

3.2 Background

The Swedish primary care market

The Swedish health care system is mainly tax-funded, has universal coverage for the citizens and is decentralized. There are 21 independent regions that are responsible for the financing and organisation of health care. This study is set in two of the largest regions in Sweden; Region Skåne (1.3 million residents) and Region Västra Götaland (1.7 million residents). Primary care, which is the first line of care, is provided in multiprofesional group practices called primary care centers (PCC). These are typically larger than single GP practices in many other European or American contexts. Swedish PCCs have on average about 4 GPs employed (Anell, 2015). In addition to GPs, these centers are also generally staffed with nurses and other health care professionals.

Whereas the Swedish health care system has traditionally been classified as a Beveridge type of system with care provided by public providers, the share of private care providers has increased during the last decades. Within the primary care sector the share of private providers increased not least due to a choice reform implemented between 2008 and 2010 (Anell, 2015). While the PCCs are all publicly financed, there are both publicly and privately operated PCCs. During the study period, there were 164 PCCs operating in Skåne and 214 in Västra Götaland. (In our analysis, we will only include the 329 units that were operating during the whole period.)

Following the choice reform all residents are obliged to enroll at a PCC. Residents can freely choose between all PCCs in the region and there are no restrictions on switching since the PCCs are not allowed to refuse new enrollments.[7] Individuals moving into a region are automatically assigned to their closest PCC.[8] To switch PCC individuals may log in to an online platform, or fill out a paper form to be sent to, or handed in, at the practice. There are neither restrictions on the number of switches nor the timing of the switches. Annually, about 8% of the individuals in both regions switch PCC.[9]

The reimbursement of both private and public PCCs are mainly based on capitation (Anell et al., 2018). That is, the PCC receives a fixed amount per enrolled individual (which is risk-adjusted for demographic and socioeconomic factors as well as the expected health care consumption). Thus, the PCCs have incentives to increase (or at least not to decrease) the list size. Public providers may not be profit-maximising, but their activities are restricted by their budget and it is in the interest of the managers to attract funds.[10]

User fees for health care utilization are set by the regional health care authorities, and may therefore vary between regions but not between PCCs within the same region and should therefore not affect the choice of provider. The basic fee for visiting a GP was in SEK 160 for visits at the PCC where the patient was enrolled. In both regions there was a surcharge for visits to other PCCs. This surcharge was 25% in Skåne and about 60% in Västra Götaland. In both regions there was also an annual cap of SEK 1,100 for any health care spending (excluding pharmaceutical spending).

The local newspaper market

During the time of our study, the role of printed media in Sweden was fairly large. In 2015 65 percent of the population aged 9-79 are regular readers of daily newspapers. This figure is mainly concentrated among the elderly, with the share of readers being 83 percent in the group aged 65-79 and declining with decreasing age (Nordicom, 2015). The media markets in this paper are confined to Västra Götaland and Skåne. The largest newspapers in Skåne (in 2015) are Sydsvenskan and Helsingborgs Dagsblad. The majority of the newspapers are local and concentrated to smaller geographical units. With regards to Västra Götaland, the largest newspaper is by far Göteborgs-Posten, followed by local newspapers

[7] All Swedish regions introduced similar choice systems in 2007-2010. For more information on the reforms, see Anell (2011, 2015).

[8] The definition of the closest center varies between the two regions. In Region Skåne the default center is defined based on the straight line distance, in Region Västra Götaland the default center is the closest center by road distance within the same municipality.

[9] Based on own calculations using the annual enrollment data from 2012 to 2017.

[10] Notably, there is also a literature showing that public primary care providers in Sweden respond to financial incentives (e.g., Ellegård et al., 2018; Dackehag and Ellegård, 2019; Ellegård, 2020).

(Presstödsnämnden, 2015). To obtain a more comprehensive view of where the articles are published, and thereby the reach of different articles, we present the articles used in this paper by date, classification and newspaper in Appendices B1-B4.

3.3 Data

3.3.1 Enrolment data

In this paper, we utilize data from the regional health care registers held by Region Skåne and Region Västra Götaland, including information on PCC enrolment and the times of switching during the years 2012-2017. These data include the start and end date of each enrolment spell. The enrolment data are linked to individual-level background information from Statistics Sweden, including information on place of residence from the official population register on the 31 of December each year. Using these data, we generate the list size of each PCC in a given month. To this aggregated data, we link information about the treatment, which is media exposure in either negative or positive terms.

3.3.2 Sample Restrictions

To ensure that the sample is relevant we impose several restrictions. From the universe of individuals being observed in the enrollment data in either of the two regions, we include all individuals that are observed as registered residents in the official population registers of Statistics Sweden in a given region for at least two consecutive years.[11] By this sample restriction we aim to exclude individuals that are not actually visiting the provider they are enrolled with and will therefore be unlikely to be observed to react to coverage in the local newspaper.[12]

We also apply additional restrictions to address issues related to openings and closures of competing providers. The outcome variable, the number of enrolled patients, is very sensitive to such openings and closings since these cause large fluctuations in the enrolment data (not related to treatment). To address these issues, we apply two further sample restrictions. First, we only keep units (both in the treated and in the control group) that remain open throughout the entire panel. This restriction mitigates the concern that units opening up will pull

[11] As we observe the place of residence only at the 31 of December this means that we exclude all individuals that are not registered in the same two consecutive New Years Eve

[12] In Region Skåne, there is a substantial share of individuals that are registered with a provider although no longer registered at an address within the region. We re-run the main results using the full sample of resident individuals in Section 3.5.4.

individuals from existing units. Second, we exclude, from the enrolment data all individuals that have previously been enroled at units that close during the study period. This restriction primarily exclude switches from closing units that mechanically cause large immediate jumps in the number of listed individuals.[13]

3.3.3 The newspaper articles

We searched for newspaper articles using the archive service Mediearkivet Retriever that covers most of Sweden's daily newspapers. The search was limited to newspapers covering Skåne and Västra Götaland respectively, and used a truncated search with the Swedish word for primary care center ("vårdcentral*").[14] Each article was then classified as negative or positive and linked to one or several named PCCs.[15] If an article was classified as neutral, we did not include it in the analysis. Following the initial classification as negative or positive, the classification was further divided into four categories depending on how likely the article was to affect listing behaviour. Negative and positive articles are further divided into two groups as follows: 1 = strongly negative; 2 = weakly negative; 3 = weakly positive; 4 = strongly positive. For each article we record the date and the newspaper or newspapers where it was published. If a unit is treated multiple times, we only use the first date of publication. We present a complete list of articles used in Appendices B1-B4.

The type of information provided in the articles varies, but articles in the strongly negative category often tell a story of the PCC missing signs of illnesses with harmful consequences for a patient, complaints about insufficient staffing, or getting very low quality ratings from patients. Articles classified as weakly negative are more characterized by negative news for the unit, but not necessarily for the patient. Examples include misdiagnosing with no clear impact on patients, or a lack of staff for a short period. With regards to the positive articles, we expect strongly positive articles to have a more positive impact relative

[13] This restriction possibly leads to an underestimation of the potential negative effects, if closing units closed because they were treated by a newsarticle. However, the frequency of treatment among the excluded units is actually lower than the corresponding figure among the included units. It is only a small share of the units that close during the observed period. We excluded 28 units since they close during the sampling period, approximately 7% of the full sample. Among the 28 closing units, three are treated with two negative articles and 1 positive newsarticle, or approximately 10.7%. The corresponding ratio of treated versus control units in the estimating sample is approximately 35%. Thus, it does not seem to be the case that closing units are particularly prone to be treated with negative news.

[14] The newspapers comes from Swedish Press and Broadcasting Authority (*Myndigheten för press, radio och tv* (MPRT)) who annually publishes a list of operating newspaper per region (*Dagstidningsförordningen*)

[15] The classification was done independently by the authors of the paper and two research assistants.

to those in the weakly positive category. Examples in the former group typically informs about good quality ratings from the National Patient Survey (*Nationella Patientenkäten*). Articles in the latter group usually mention the treated PCC in positive terms but do not include information which necessarily be interpreted as a signal of improved quality of care.

Over the years 2013-2016, the total number units treated by negative articles is 71 and the total number treated by positive articles is 48.[16] The number of units treated by negative news are 71 (whereof 25 are in Region Skåne and 46 are in Region Västra Götaland). The number of units treated by positive news are 48 (whereof 34 are in Region Skåne and 14 are in Region Västra Götaland).[17] To reduce spillovers we use only the remaining 202 unexposed units as controls for both type of coverage, i.e. we exclude units treated by negative coverage when estimation the effect of being exposed in positive newsarticles and vice versa.

Table 3.1: Article Summary

	Total	2013	2014	2015	2016
Negative articles					
Number	71	17	25	17	12
Positive articles					
Number	48	15	15	9	9

Due to the fact that we only use the first article that mentions a PCC, there is an overweight of the number of articles from the earlier years during the study period.

3.3.4 Variables

The treatment variable is a binary variable equal to 1 if the PCC unit is exposed to the media and 0 otherwise. Though the timing of treatment can vary, treatment status is constant across units. From the enrolment data, we construct the outcome variable by summarizing the total number of listed individuals per PCC every fourth week.[18] The heterogeneity analysis refines the data into urban and rural health care markets. The definition of rural/urban areas is based on the

[16]We require that the treated units in 2013 are not mentioned in local media during 2012 and that articles published in 2016 have at least 12 months of observations after the treatment date.

[17]While the share of units exposed to negative coverage are similar across regions, the share of units exposed to positive articles is much larger in Region Skåne.

[18]In Region Västra Götaland, these shifts are only registered every fourth week. For Region Skåne, we are able to generate more granular data as changes are recorded on a daily level.

towns where the PCC is located. Urban towns are defined as having > 17000 inhabitants in 2015. These also corresponds to towns having more than 2 primary care centers.[19] An overview of the data is presented in Table 3.2.

Table 3.2: Summary Statistics

	Mean	Std. dev
PCC Characteristics		
List size	7 491	3 343
Log of list size	8.81	0.49
Rural	0.44	0.50
Number of PCC:s		
Before any sample restrictions	353	
Dropping units$<$ full sampling period	329	
Dropping units with list size < 30	325	
Units treated with negative news	71	
Units treated with positive news	48	
Article Characteristics		
Strongly negative	52	
Weakly negative	19	
Strongly positive	21	
Weakly positive	27	

The average list size across all PCC:s is 7 491 individuals but the standard deviation is large. In the analysis we use the logarithmic transformation of monthly list size. The total number of PCC throughout the period is 353 - however, after the sample restrictions described earlier in this section are applied, the effective sample size is reduced to 325 units. 44% of the units are located in towns classified as rural.

3.4 Estimation Strategy

The idea of this paper is to study how media coverage affects patients' choice of provider by estimating the effect of a news-article mentioning a PCC on that PCC's list size (i.e., registered number of patients). This section discusses how we identify this effect.

[19]The exception is the rural town Tomelilla, with about 8 000 inhabitants, that has three primary care centers during the study period (although one of them is a small practice with a single GP that closes during the study period). See Anell et al. (2021) for a similar classification.

We utilize variation that comes from the timing of the publishing date using a difference-in-differences design. We compare treated and control primary care units (first difference) before and after the newspaper publication (second difference). The main specification is an event study using monthly data with staggered adoption of treatment. There is a large and burgeoning literature on potential issues with Two-Way Fixed Effects (TWFE) models (see, for example, Goodman-Bacon (2021a) and Baker et al. (2022) for overviews). In particular, DiDs with staggered treatment adoption as well as dynamic post-treatment effects may generate biased ATTs. One particular concern is the comparison of late adopters using early adopters as the control group. As a consequence, there are many proposed methods to circumvent the issues raised in previously cited papers (e.g., Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021a; Sun and Abraham, 2021).

The estimation strategy in this paper builds on Cengiz et al. (2019) and Deshpande and Li (2019), and is commonly referred to as a the "stacked regression" approach (Baker et al., 2022). The idea is to create one treatment-specific dataset, or a *stack*, and within each stack identify the treatment effect by comparing one or more treated units to never treated units. Each stack corresponds to the timing of treatment - in other words, there will be be fewer stacks than number of articles if multiple articles, covering multiple PCCs are published in the same month and year. Within each stack, adoption of treatment is not staggered. The stacks are then collected into a large data set which allows for both static TWFE DiD and a dynamic event study estimation, including stack-specific unit and time-specific fixed effects (Cengiz et al., 2019).[20]

The choice of the stacked regression approach is motivated by two factors. First, as we discuss in more detail in Section 3.4.1, the simplicity of the model allows us to facilitate more adaptations to the model, such as accounting for pre-trends. Second, we can estimate and present one treatment effect per stack, which is informative when evaluating the full effect to ensure that it is not driven by outliers. The result from this exercise is found in Figure A7.

Using the stacked data set, we first estimate a dynamic model as follows:

$$y_{ims} = \alpha_{is} + \lambda_{ms} + \sum_{l=-K}^{-1} \mu_l D^l_{im} + \sum_{l=1}^{L} \mu_l D^l_{im} + \epsilon_{ims} \tag{3.1}$$

where $D^l_{im} = \mathbb{I}[t - E_i = k]$ is an indicator for a treatment unit i in treatment cohort E_i (equals 1 for the period of treatment) being k periods away from the start

[20]The stacked regression approach performs well in the simulations reported in Baker et al. (2022).

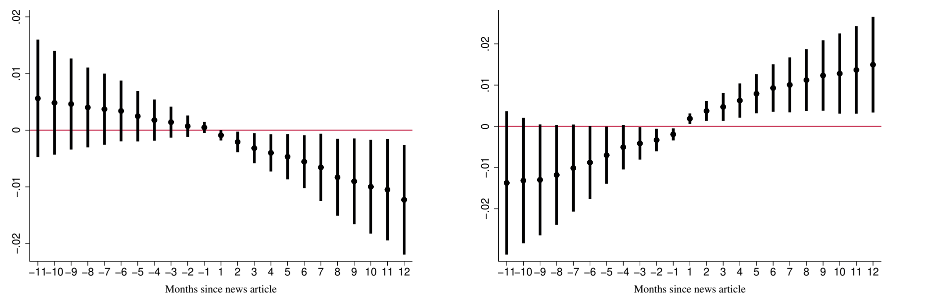
of treatment. The dynamic specification comes into effect from the summation expressions, which includes a set of relative time-indicators. The first summation expression in equation 3.1 captures the time periods leading up to treatment, i.e., the months before treatment, and the second summation includes the months after treatment. In accordance with common practice, we leave out the relative time indicator for the period before the treatment is switched on. The interpretation of μ_l is the difference in list size between treated and and control groups l time periods away from treatment, relative to the outcome difference in the excluded period prior to treatment (period 0). To account for the fact that the estimation data set consists of several event-stacks we interact the time (m) and group (i) fixed effects with each stack (s). Standard errors are clustered at the PCC level (addressing the fact that the same observation may be included in several stacks). In our preferred specifications, we focus on a window of 12 months before and after the event.

3.4.1 Accounting for the trend

The main assumption for causal inference in difference-in-difference and dynamic event study models is the parallel trend assumption, i.e., that the treated and control PCCs would have followed the same trend absent treatment. A visualization of the data in Figure 3.1 reveals clear trends in enrolment behaviour both before and after the time of treatment. The graph shows the coefficients from a dynamic event study graph as specified in equation 3.1 for the 24 months around the month of publication. Although most of the coefficients on the pre-treatment dummies are insignificant at conventional levels and not large in absolute terms, the decreasing/increasing pattern clearly suggests that there are differentiated pre-trends for treated and untreated units. These trends indicate a violation of the parallel trend assumption, and imply different enrolment trajectories for control and treated units before the onset of treatment. These underlying trends generate a difference in difference in the list size between the two groups obscuring any actual treatment effect. These trends are present for both types of articles, but are somewhat more pronounced for the positive subset of articles.[21]

[21]We plot the event studies separately by region in the Appendix in Figures A1 and A2. The results on negative articles are mainly driven by Region Skåne.

Figure 3.1: Dynamic Event Study specification



((a)) Negative articles ((b)) Positive articles

The figure show a subset of monthly regression coefficients from a fully saturated regression of all the months before and after treatment. Enrollment data from the time period 2013-2016 and articles from 2012-2017 classified as negative or positive. The standard errors are clustered at the PCC level. Period 0 is the baseline. The confidence intervals are at the 95% level.

While handling these trends in the TWFE framework is not straightforward, the stacked regression approach provides an appealing solution since it allows for estimating standard dynamic regressions on each of the stacked data sets.[22] To account for the differentiated trend in treated and the control units, we follow Goodman-Bacon (2021b) and Bilinski and Hatfield (2018). We estimate the pre-period trends and extrapolate these trends into the post-treatment periods.[23] To estimate the differential trend in the pre-period, we regress the outcome on a linear time trend interacted with treatment status, and include interactions between the treatment indicator and all post-treatment periods (Bilinski and Hatfield, 2018):

$$y_{ims} = \alpha_{is} + \lambda_{ms} + \gamma T_{ims} + \sum_{l=1}^{L} \mu_l D_{im}^{l} + \epsilon_{ims} \tag{3.2}$$

where T_{ims} is a linear time trend interacted with treatment status. Effectively, this procedure estimates the degree to which the treated group deviates from the extrapolated differentiated trends. In our preferred specifications, we base the estimations of the pre-trend on the last 12 months before treatment by augmenting the model in equation 3.2 with an interaction with a dummy, W, equal to 1 if $m < -12$ and time, $\rho T_{ims} W_m + \zeta W_m$. This approach has the advantage of basing

[22]This makes handling trends more straightforward, and easier to implement relative to, for example, user-written implementations of other estimators such as Sun and Abraham (2021) and Callaway and Sant'Anna (2021).

[23]As Goodman-Bacon (2021b) notes, only adding a linear unit-specific time trend to equation 3.1 using data both before and after the event will absorb dynamic treatment effects and may counfound the treatment effects with pre-existing trends. This is particularly problematic when, as in our case, the response to treatment is expected to be sluggish (see e.g., Wolfers, 2006; Lee and Solon, 2011).

the pre-trend on the same number of observations for all stacks.[24] In practice, we estimate the model using the full data set, but interact the trend with a dummy for the periods before the last 12 months preceding the reference period, allowing for a trend shift both in levels and slope in the year prior to treatment treatment. We believe that this specification will more accurately capture the differential trends in the period leading up to the news article, compared to estimating a linear trend for the full pre-periods of various length.[25]

To visualise the detrended data, we first predict the residuals from equation 3.1, and add back the constant and the coefficient values from the post-treatment dummies μ_l. We then use these residuals as a detrended outcome in the event study model in equation 3.1 and plot the coefficients on the treatment dummies in pre- and post-treatment periods, as depicted in Figure 3.2. While this exercise is useful for illustrating the performance of this detrending approach, the displayed standard errors are not corrected for the two-step procedure of adjusting the trend. To make valid inference (and so as to be able to plot correct standard errors for the post-period), we rely directly on the coefficient estimates and standard errors from equation 3.2. Basically, each of these coefficients tests how a change in the list size in a given month after treatment deviates from the estimated pre-trend.[26]

To obtain an estimate of the treatment effect corresponding to a DiD coefficient (but accounting for the differentiated trend), we compute the average of the coefficient of the first 12 post-period treatment effects.

$$\beta = \sum_{l=1}^{12} \hat{\mu}_l / 12 \tag{3.3}$$

Where $\hat{\mu}$ are the estimated coefficients from the post-periods estimated in equation 3.2. Notably, if we were to exclude the differentiated time trend in equation 3.2, this approach is equivalent to estimate a static DiD comparing the first post-treatment month with the 12-month period leading up to treatment.[27] For comparison, we therefore also estimate such a model. For the full data period,

[24]See Goodman-Bacon (2021b) for a more thorough discussion of this issue.

[25]However, in Section 3.5.4 we elaborate further on the sensitivity of our results to this addition by detrending the data excluding this variable, thereby using the full pre-treatment period to predict the linear trend.

[26]As the coefficients from the detrended approach tell us how each month differs relative to the month before the treatment, when we have washed away the extrapolated pre-trend, these will be very similar.

[27]In contrast to using detrended data to estimate a static DiD (in a two-step approach), this approach correctly estimates the standard errors of the coefficients.

this is equivalent to estimating

$$y_{ims} = \alpha_{is} + \lambda_{ms} + \gamma_1 D^{-3}_{ims} + \gamma_2 D^{-2}_{ims} + \gamma_3 D^{-1}_{ims} + \gamma_4 D^{+1}_{ims} + \gamma_5 D^{+2}_{ims} + \gamma_6 D^{+3}_{ims} + \gamma_7 D^{+4}_{ims} + \epsilon_{ims} \tag{3.4}$$

where, again, α_{is} and λ_{ms} are stack-specific (s), unit-specific (i), and month-specific (m) fixed effects. The monthly pre- and post-treatment dummies are replaced by yearly equivalents, three for the years prior to treatment and four for the years after treatment. The year before treatment (D^0_{ims}) is left out and thus serves as a reference period. In the tables we display only the coefficient for the first year post-treatment, γ_4, which effectively compares the list size between treated and control units during the year right before and the year after the publishing of the article.[28]

3.5 Results

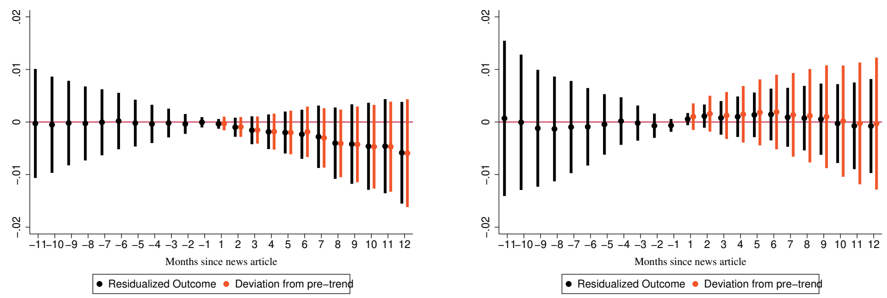
3.5.1 Dynamic event study

The aim of this paper is to evaluate the effect of information about PCCs, by means of articles in local newspapers, on individuals' enrollment. As the event is defined at the PCC level, we analyze the effect of media at the aggregate, PCC, level. For the results below we have used enrollment data from the time period 2013-2016 and articles from 2012-2017 classified as negative or positive.

Figure 3.1 illustrated that treatment and controlgroups of units were on different trends in the pre-treatment period. The list size of PCCs treated by negative articles decreased compared to controls, and the list size of PCCs treated by positive articles increased compared to controls. As this pattern indicates that at least some patients have, and act on, quality information *before* the articles come out, it is interesting in relation to our research question. As mentioned, although the trend differences are small, they are problematic for identification.

[28]This is equivalent to using the 24-month period around the publication date to estimate $y_{im} = \alpha_{is} + \lambda_{ms} + \gamma D_{im} + \epsilon_{im}$ where D_{im} is an interaction term between an indicator variable equal to 1 for the treated unit, D_i, and an indicator for the period after treatment, $Post_m$.

Figure 3.2: Dynamic Event Study specification, adjusted for trend



((a)) Negative articles ((b)) Positive articles

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The detrending procedure is discussed in greater detail in section 3.4.1. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

As described in Section 3.4.1, we predict the residuals from a regression of the outcome and an (extrapolated) linear pre-trend for each stack. Figure 3.2 displays the coefficients from equation 3.1 using the detrended residuals as the outcome variable (shown in black), and the coefficients from equation 3.2 indicating deviations from the extrapolated pre-trend for each month after treatment (shown in red).

In contrast to what is shown in Figure 3.1, there are no longer any significant post-treatment coefficients once the differentiated pre-trends are accounted for in Figure 3.2. The inclusion of an extrapolated pre-trend removes the effect found in Figure 3.2 for both positive and negative articles. While the coefficients are very close to zero for positive articles, there are indications of a negative response, in terms of a trend break from the null effects in the pre-treatment period for units treated with negative articles. However, the estimates are not statistically different from zero at any conventional level. The magnitudes of the coefficients are also small. For example, the effect on the first post-treatment dummy (the month in which the article is published) is a decrease in the number of enrolled individuals by 0.035 percentage points, and by 0.1 percentage points in the post-second month relative to the month before treatment.[29]

[29]Figure A3 reveals that the negative result is mainly driven by units in Region Skåne - in Region Västra Götaland, there is no effect.

3.5.2 Difference-in-difference

To accompany the dynamic event study, Table 3.3 presents results from the static DiD models. Column 1 in Panels A and B uses the unadjusted monthly list size as the outcome, while column 2 in Panels A and B presents estimates from testing the linear combination of the 12 first months after treatment, as specified in equation 3.3. This is the equivalent to the mean of the 12 post-treatment periods using the red coefficients in Figure 3.2.

Table 3.3: Difference-in-Difference

	(1) Unadjusted	(2) Adjusted for trend
	Panel A: Negative articles	
Year_{t+1}	-0.0092*	-0.0029
	(0.0047)	(0.0026)
Observations	519649	519649
Number of clusters	275	275
Number of stacks	32	32
	Panel B: Positive articles	
Year_{t+1}	0.0168**	0.0010
	(0.0069)	(0.0037)
Observations	436870	436870
Number of clusters	252	252
Number of stacks	27	27

Notes: The outcome is the log of monthly listed patients per unit. The coefficient in column 1 in Panel A and B is estimated as the effect of the first year post-treatment relative to the year before. Column 2 presents estimates from testing the linear combination of the 12 first months after treatment since we include a linear trend variable; this is equivalent to estimating deviations from the extrapolated pre-trend. This procedure produces correct standard errors. The standard errors are clustered at the PCC level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

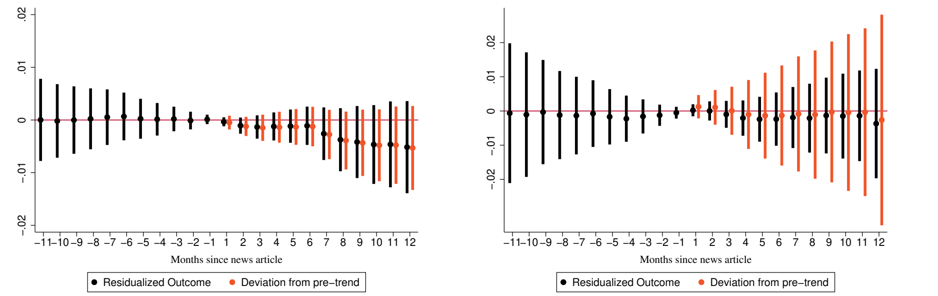
Table 3 confirms the results from the event study. There is a significant effect for both positive and negative news events, which is reduced when adjusting for the trend. The effect of positive articles is completely wiped out, while the effect of negative news is reduced by about two thirds and becomes statistically insignificant. We next turn to a heterogeneity analysis, since the response to media information may be asymmetric in different types of health care markets and to different types of articles.

3.5.3 Treatment heterogeneity

While the overall effect of news coverage is small and insignificant, there is a small trend break after negative reporting, and indications of a small effect immediately after a positive news article. These patterns call for further investigations, since heterogeneous effects due to type of article or market characteristics may blur/hide an actual effect of news coverage. We first examine heterogeneity between different groups of articles, testing if the lack of response according to Figure 3.2 is mitigated by dividing the articles into those with more or less negative/positive news. This classification, which was described in more detail in Section 3.3, divides the articles into subclasses based on whether they are more or less likely to affect patients' enrollment (ex ante). Since the choice of switching PCC is likely to be affected by the availability of options and other market characteristics, we also separately estimate Equation 3.1 for rural and urban PCCs. These are defined by the size of the town that the PCC is located in (which also corresponds to the number of local competitors).

Figures 3.3 and 3.4 display the heterogeneity among the news articles based on the classification of the content being more or less likely to impact patients' perception of the PCC.

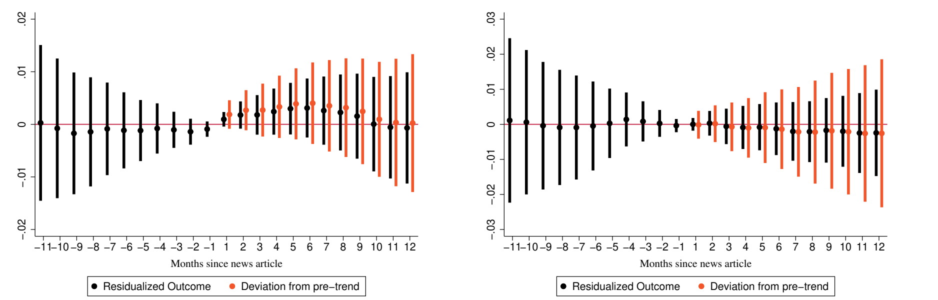
Figure 3.3: Event study specification, negative, by article subgroup



((a)) Strongly negative ((b)) Weakly negative

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The detrending procedure is done separately for different article groups. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

Figure 3.4: Event study specification, positive, by article subgroup



((a)) Strongly positive ((b)) Weakly positive

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The detrending procedure is done separately for different article groups. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

Similarly to the figures illustrating the main results, figures 3.3 and 3.4 display coefficients from a dynamic event study using the trend-adjusted residual as the outcome, and coefficients reflecting the deviations from the extrapolated pre-trend (in the post period). For the negative class of articles, group 1 includes news where we expect a stronger impact on enrollment (strongly negative), relative to group 2, where we perceive the articles to be less informative about poor quality (and thus likely having a limited effect on outcome). As expected the effect is both more precisely estimated and more pronounced for PCCs treated by strongly negative articles relative to the control group. Also for the most positive subset of articles, we find a clearer impact for the group of articles classified as more likely to affect choices. In contrast to the effect for strongly negative articles, where the effect is amplified over time, the effect for strongly positive articles is more immediate and instead fades over time.

Table 3.4: Static DiD by type of article

	(1) Strong impact	(2) Weak impact
	Panel A: Negative articles	
Year_{t+1}	-0.0028	-0.0006
	(0.0021)	(0.0078)
Observations	483105	177151
Number of clusters	256	223
	Panel B: Positive articles	
Year_{t+1}	0.0024	-0.0015
	(0.0039)	(0.0061)
Observations	257264	268600
Number of clusters	231	225

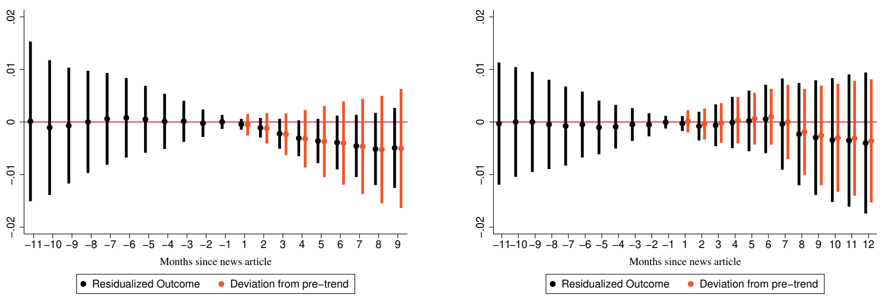
Notes: The coefficients are the yearly average deviation from the pre-trend in the first post-treatment year, relative to the year prior to treatment. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The output in Table 3.4 clearly shows that, as expected, the magnitude of the coefficients in the impact with strongly negative/positive content is larger relative to the treatment effect for those classified as having a weakly impact.

Figures 3.5 and 3.6 show the heterogeneity by market type among the news articles classified as negative and positive, respectively. Even though confidence intervals are large, it seems as there is a difference in how patients respond in different areas. In rural areas, the magnitude of the negative effect is amplified over time, which suggests a sluggish response. With regards to the positive news articles in Figure 3.6, we find no effect in neither the rural nor urban areas. The standard errors when plotting the deviations from the pre-trend for positive articles are much larger than the coefficient using the predicted detrended outcome.

Even though confidence intervals are large it seems as there is a difference in how patients respond in different areas. In rural areas the magnitude of the

Figure 3.5: Event study specification, negative, by market type



((a)) Rural areas ((b)) Urban areas

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The detrending procedure is done separately for urban and rural markets. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

negative effect is amplified over time which suggest an sluggish response. With regards to the positive news articles in Figure 3.6, we find no effect in neither the rural nor urban areas. Thus there seem to only be a reaction among listed patients on PCCs in rural areas. The standard errors when plotting the deviations from the pre-trend for positive are much larger than the coefficient using the predicted detrended outcome.

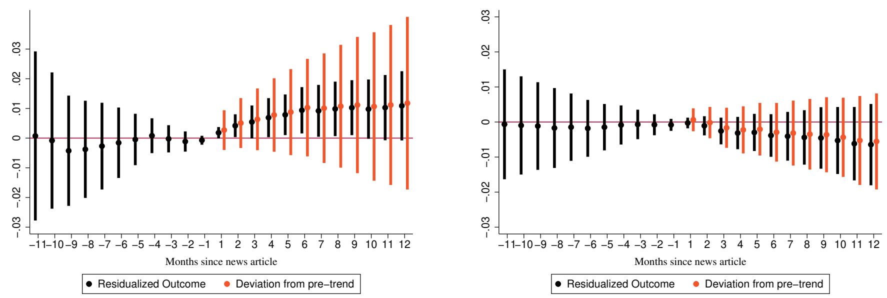
Table 3.5 presents the coefficients from the static DiDs for rural and urban PCCs, separately. There is a negative and insignificant effect on list size for negative news coverage for rural PCCs. As can be seen in the appendix (Table A2), this is purely driven by Skåne. The interpretation is that, on average, a negative news publication decreases the list size by 0.4 percentage points in the year following publication relative to the year before. With regards to the positive articles, the magnitude of the coefficient is 3 times larger in rural areas relative to urban areas.

3.5.4 Robustness

In this section we test the sensitivity of our main results to different robustness checks. The results of the robustness checks are presented in the Appendix, in Figures A7 to A9.

First, to ensure that there are no outliers in any stack, we plot the treatment effect by stack in Figure A7. Note that a stack may contain multiple articles, since a stack is defined by the month of treatment. The majority of treatment effects are similar in effect size.

Figure 3.6: Event study specification, positive, by market type



((a)) Rural areas ((b)) Urban areas

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The detrending procedure is done separately for different markets. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

Second, we relax the sample restriction that only included individuals resident in the relevant regions for two consecutive years. We add individuals who were registered in Region Skåne or Region Västra Götaland at least once throughout the sampling period. The results are presented in Figure A8 and are remarkably similar. For example, with regards to the negative articles, the coefficients for the first and second month after treatment are -0.035 and -0.1 percentage points respectively using the restricted sample, and -0.03 and -0.11 percentage points respectively using the less restricted sample. The effects for positive news displays a similar overall dynamic pattern - there is a small increase in the (log of) number of listed individuals for the first and second month using the sample with more listed individuals. However, this does not lead to qualitatively different conclusions.

Third, we reconsider using all the available pre-data to predict the pre-treatment trend. To this end, we repeat the detrending procedure but exclude the dummy W_{ims} and its interactions, implying that we estimate the pre-trend using not the last year prior to the onset of treatment but the whole pre-period. The result is presented in Figure A9. The evidence is quite different compared to that presented in Figure 3.2. Primarily, the prior similarity between using the deviations from the predicted pre-trend and the regression coefficients using the residualized monthly list size is now gone - the deviations have larger magnitudes and much larger confidence intervals. It should be noted that the data used for the pre-period is long, dating at most 60 months prior to treatment (and varies between stacks). As a consequence, predicting the pre-trend using all available data

Table 3.5: Static DiD by type of market

	(1) Rural	(2) Urban
	Panel A: Negative articles	
Year_{t+1}	-0.0040	0.0010
	(0.0042)	(0.0031)
Observations	200572	319077
	Panel B: Positive articles	
Year_{t+1}	0.0089	0.0028
	(0.0088)	(0.0042)
Observations	168191	268600

Notes: The coefficients are the yearly average deviation from the pre-trend in the first post-treatment year, relative to the year prior to treatment. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

is more vulnerable to changes to population size, and to factors that affect care quality.[30] The pattern shown in Figure A9 provides support for our initial decision to only use only data from 12 months prior to treatment. The results are more similar for positive articles in Figure A9 and the conclusions are unchanged.

3.6 Concluding remarks

Health care markets, and in particular markets characterized by patient choice, are dependent on patients acting on information on quality. The empirical literature shows that the effects of choice and competition on care quality in primary care markets with regulated prices are small or absent (Dietrichson et al., 2020; Gravelle et al., 2019). One potential explanation for this is that individuals are not sufficiently informed about quality, or do not act on such information. Another explanation would be supply-side issues. For example, highly demanded providers may have trouble scaling up their organization, or physicians may be uninformed of the reasons behind patients' choices of providers and therefore do not change their practice even if patients leave. From the earlier empirical literature, we know that dissemination of quality either by information interventions or doctor ratings online may affect choices of provider, but only to a certain degree (Anell et al., 2021; Chen and Lee, 2021; Bensnes and Huitfeldt, 2021; Luca and Vats, 2013; Chartock, 2021).

[30] For example, 60 months is a long enough period that a previously high-quality PCC may have lost most of its staff, and therefore decreased its care quality. A PCC is therefore not necessarily a good comparison to itself over long periods of time.

Our paper adds to the understanding of these markets, and specifically how patients make these choices. Using a stacked regression approach adjusting for pre-trends, we estimate both dynamic event studies and static DiD models. We are not able to detect any significant effect of negative or positive media coverage. There is a pre-trend, which may indicate that patients have and act on information about providers that later become negatively and positively treated by newspaper articles, and there appears to be a trend break at the time of publication of the negative articles. However, the magnitudes of both the trend and the effect estimates are small (all point estimates are below one percentage point). The effect is stronger, yet still small and statistically insignificant, for articles where the content of the article is strongly negative or positive. For PCCs treated by strongly negative articles, there is a negative response which is amplified over time. By contrast, the effect for units treated by clearly positive news is more immediate and rather fades over time (which may be an indication that popular providers have trouble expanding). We also find a more pronounced negative and positive effect in rural areas when splitting the sample by the size of the town in which the practice is located.[31] However, despite some indications of heterogeneity across articles and markets, also the largest of these estimates are small and not statistically significant (the largest estimate is around one percentage point).

This paper has a few limitations that should be acknowledged. First, we only have data on news articles and listing behaviour for two, though very large, regions. Given the null effect in this paper, it is unlikely that adding data from more regions would change our conclusions. Second, we have not considered the type of patients. There could be heterogeneity in how different patients respond to the information used in this paper. However, ex ante, it is difficult to foresee groups that should react more than others. On the one hand, quality may be particularly important for frequent users of primary care. However, due to their many interactions with the unit, information about quality may already be known. Similarly, very infrequent users may have limited information about the quality of the PCC. But it is unclear if they would make an active choice to change their listed unit, since they rarely visit it anyway. Thus, we avoid searching for

[31] That the effects are larger in rural areas may be counterintuitive at first, and contrast the findings in Anell et al. (2021) of larger effects of an information intervention in urban areas (in the same market). However, at least speculatively, media coverage may have larger effects on the general discussions at workplaces and among neighbours in rural towns compared to more urban areas. In line with the "reputation good"-literature, the word of mouth may travel faster and be more accurate in markets where individuals are more likely to have individuals within their network of peers that have been exposed to the various providers (e.g., Pauly and Satterthwaite, 1981). If one already have some information about the competing practice, one may be more likely to react to negative news.

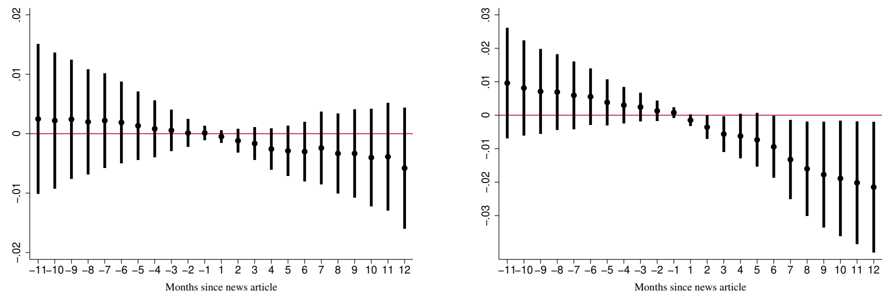
effects when we do not have clear hypotheses and, in addition, when effects in the main analysis are insignificant.

Overall, the small or absent effects of media coverage are of interest when designing these patient choice markets. Unless the information reported in the local newspaper is already known to the public, these results suggest that patients do not turn away from low quality providers - even in case of reports of mistreatment. It may be one thing if patients generally do not reward high-quality providers, but if they do not even punish low-quality providers there is little scope for markets with patient choice to generate more positive outcomes. Thus, newspaper articles do not seem to substantially mitigate information frictions. One major explanation for the lack of quality improvements exercised by patient choice and provider competition may still be that patients either do not have, or do not act on, information on provider quality. These results are of particular interest in the primary care context where patients to a large degree are left alone to make decisions of where to enrol - without guidance by other medical expertise.

Appendices 3

Figures

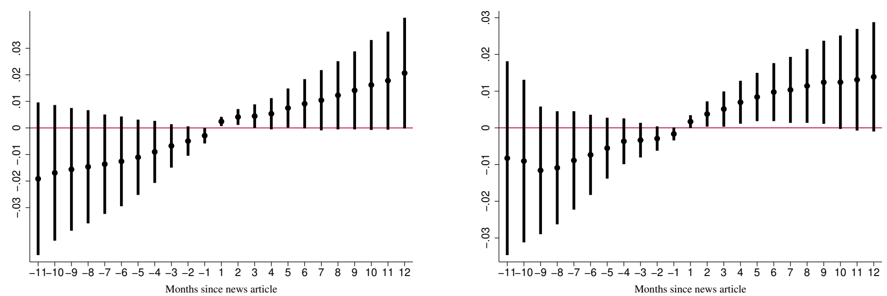
Figure A1: Negative news articles by region



((a)) Västra Götaland ((b)) Skåne

The figure show a subset of monthly regression coefficients from a fully saturated regression of all the months before and after treatment. The standard errors are clustered at the PCC level. Period 0 is the baseline. The confidence intervals are at the 95% level.

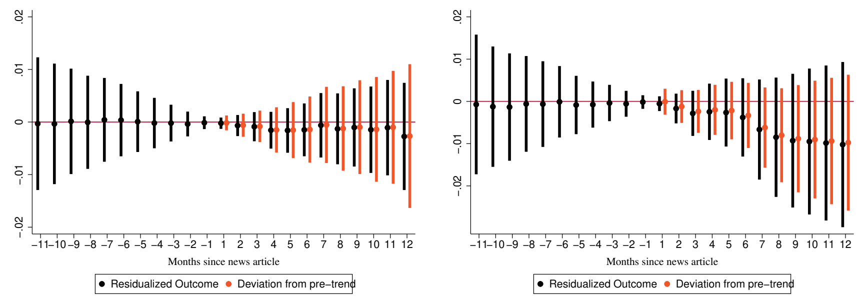
Figure A2: Positive news articles by region



((a)) Västra Götaland ((b)) Skåne

The figure show a subset of monthly regression coefficients from a fully saturated regression of all the months before and after treatment. The standard errors are clustered at the PCC level. Period 0 is the baseline. The confidence intervals are at the 95% level.

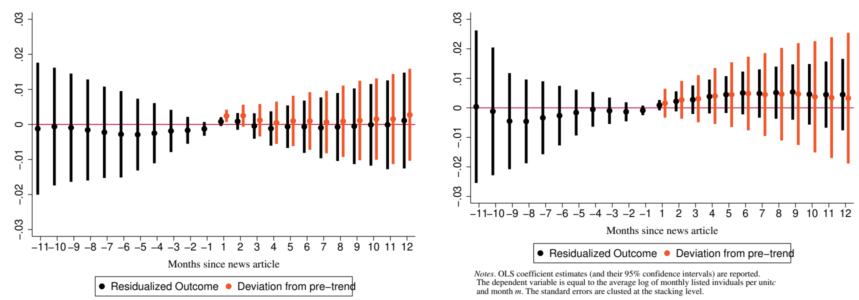
Figure A3: Negative news articles - adjusted for trend



((a)) Västra Götaland ((b)) Skåne

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The trend is predicted separately for each region. The detrending procedure is discussed in greater detail in section 4.2. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

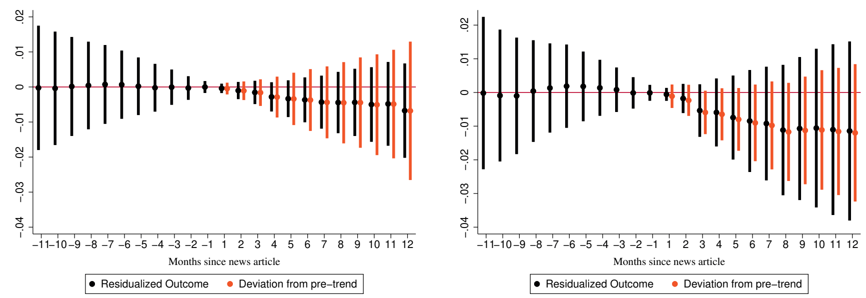
Figure A4: Positive news articles - adjusted for trend



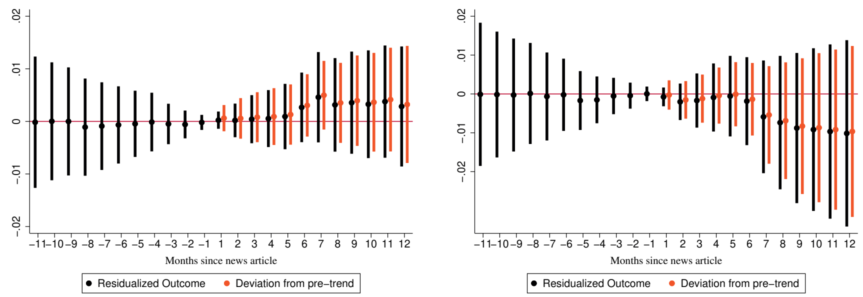
((a)) Västra Götaland ((b)) Skåne

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The trend is predicted separately for each region. The detrending procedure is discussed in greater detail in section 4.2. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

Figure A5: Negative news articles - by region and market



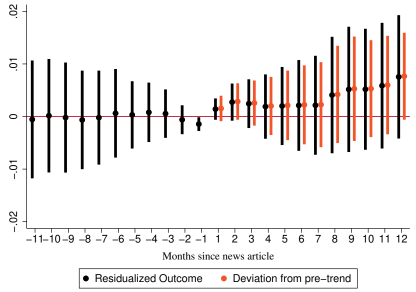
((a)) Rural: Västra Götaland ((b)) Rural: Skåne



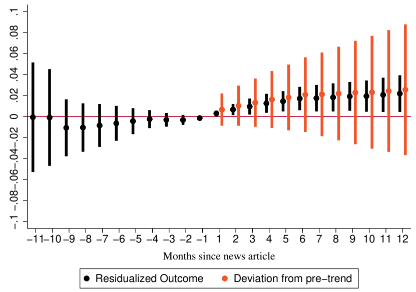
((c)) Urban: Västra Götaland ((d)) Urban: Skåne

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The trend is predicted separately for each region and market. The detrending procedure is discussed in greater detail in section 4.2. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

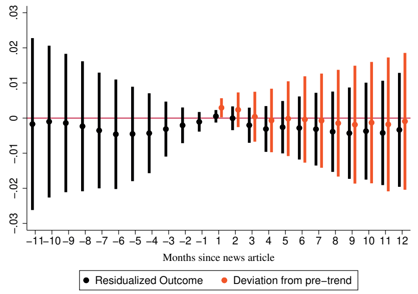
Figure A6: Positive news articles - by region and market



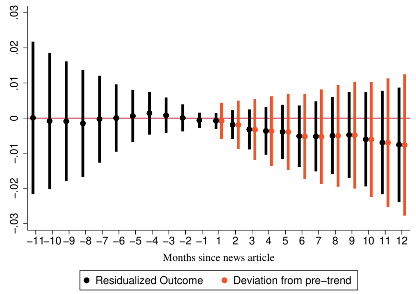
((a)) Rural: Västra Götaland



((b)) Rural: Skåne



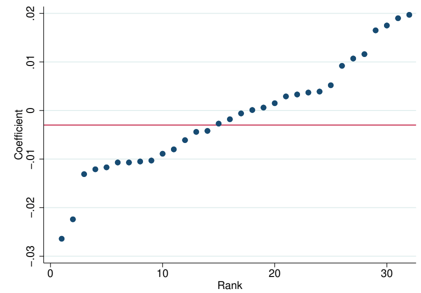
((c)) Urban: Västra Götaland



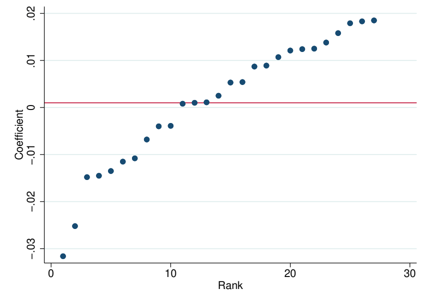
((d)) Urban: Skåne

The figure show monthly regression coefficients of the periods before and after treatment from a regression on a detrended residualized outcome (black coefficients and confidence intervals) and deviations from the extrapolated pre-trend (red coefficients and confidence intervals). The trend is predicted separately for each region and market. The detrending procedure is discussed in greater detail in section 4.2. Period 0 is the baseline. The confidence intervals are at the 95% level. Standard errors are clustered at the PCC level.

Figure A7: Treatment effect per stack



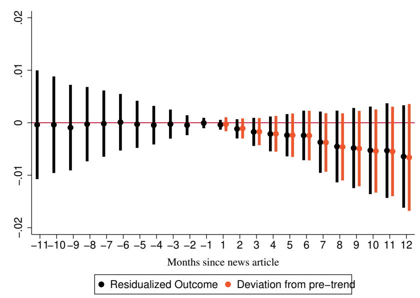
((a)) Negative articles



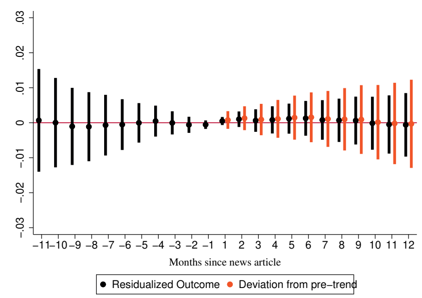
((b)) Positive articles

The figure show the mean of the first 12 months after treatment from the detrending regression, i.e. deviations from the pre-trend for each stack. Each dot represents one stack. The effects are ranked from the least to the largest magnitude. The vertical line corresponds to the effect in the overall, stacked, sample.

Figure A8: Event Study, detrended using a different sample



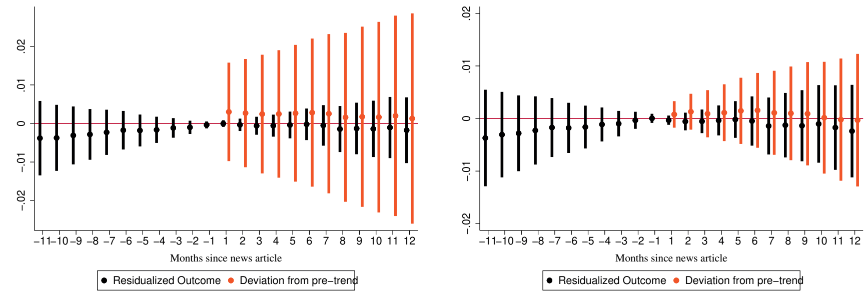
((a)) Negative articles



((b)) Positive articles

The figure show monthly regression coefficients of the leads and lags during the event window from a regression on a detrended residualized outcome. The detrending procedure is discussed in greater detail in section 4.2. Period 0 is the baseline. The confidence intervals are at the 95% level.

Figure A9: Event Study, adjusted for trend using a different detrending procedure



((a)) Negative articles ((b)) Positive articles

The figure show monthly regression coefficients of the leads and lags during the event window from a regression on a detrended residualized outcome. The detrending procedure is discussed in greater detail in section 4.2. Period 0 is the baseline. The confidence intervals are at the 95% level.

Tables

Table A1: Static DiD by region

	(1) All	(2) VGR	(3) RS
	Panel A1: Unadjusted negative articles		
Year_{t+1}	-0.0092*	-0.0042	-0.0163*
	(0.0047)	(0.0053)	(0.0086)
	Panel A2: Trend-adjusted negative		
Year_{t+1}	-0.0029	-0.0011	-0.0052
	(0.0026)	(0.0033)	(0.0043)
Observations	519649	317883	201766
	Panel B1: Unadjusted positive articles		
Year_{t+1}	0.0168**	0.0209*	0.0151*
	(0.0069)	(0.0124)	(0.0090)
	Panel B: Trend-adjusted positive articles		
Year_{t+1}	0.0010	0.0014	0.0037
	(0.0037)	(0.0040)	(0.0066)
Observations	436870	265598	171272

Notes: The outcome is the log of monthly listed patients per unit. The coefficient in Panels A1 and B1 is estimated as the effect of the first year post-treatment relative to the year before while Panels A2 and B2 presents estimates from testing the linear combination of the 12 first months after treatment, since we include a linear trend variable; this is equivalent to estimating deviations from the extrapolated pre-trend. This procedure produces correct standard errors. The standard errors are clustered at the PCC level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2: Static DiD by region and type of market

	(1) All	(2) VGR	(3) RS
	Panel A1: Negative articles rural		
Year$_{t+1}$	-0.0040	0.0036	-0.0084
	(0.0042)	(0.0047)	(0.0057)
Observations	200572	118886	81686
	Panel A2: Negative articles, urban		
Year$_{t+1}$	0.0010	0.0026	-0.0044
	(0.0031)	(0.0032)	(0.0056)
Observations	328732	205176	123556
	Panel B1: Positive articles, rural		
Year$_{t+1}$	0.0089	0.0037	0.0186
	(0.0088)	(0.0033)	(0.0189)
Observations	168191	98513	69678
	Panel B2: Positive articles, urban		
Year$_{t+1}$	-0.0028	-0.0003	-0.0046
	(0.0042)	(0.0061)	(0.0061)
Observations	268600	167085	101515

Notes: The outcome is the log of monthly listed patients per unit. The coefficient in Panels A1 and B1 is estimated as the effect of the first year post-treatment relative to the year before while Panels A2 and B2 presents estimates from testing the linear combination of the 12 first months after treatment, since we include a linear trend variable; this is equivalent to estimating deviations from the extrapolated pre-trend. This procedure produces correct standard errors. The standard errors are clustered at the PCC level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix B - Information about the included articles

Table B1: Group 1: Strongly negative

Newspaper	Date	Information	Name of PCC
Göteborgs-Posten	20130307	Wrong diagnosis	Sätila vårdcentral
Helsingborgs Dagblad	20130412	Bad access	Vårdcentralen Laxen
Borås Tidning	20130510	Wrong diagnosis	Sandareds Vårdcentral
Kristianstadsbladet	20130515	Maltreatment	Vårdcentralen Östermalm
Mölndals-Posten	20130718	Negative coverage	Lindome Vårdcentral
Bohusläningen	20130830	Wrong diagnosis	Vårdcentralen Ljungskile
Sydsvenskan	20130909	Bad access/Maltreatment	Vårdcentralen Delfinen
Borås Tidning	20131106	Wrong diagnosis	Skene Vårdcentral
Bohusläningen	20131204	Negative coverage	Vårdcentral Orust
Alingsås Tidning	20131223	Reduced patient quality	Sörhaga Vårdcentral
Borås Tidning	20140118	Mistreated	Hälsobrunnen i Ulricehamn
Dalslänningen	20140131	Mistreated	Vårdcentralen Dals-Ed
Borås Tidning	20140204	Wrong diagnosis	Vårdcentralen i Bollebygd
Borås Tidning	20140303	Discontinued physiotherapy	Vårdcentralen Fristad
Borås Kuriren	20140312	Low quality NPE	Södra Torget Vårdcentral
Kristianstadbladet	20140322	Low quality NPE	Hälsoringen i Bromölla
Kristianstadbladet	20140322	Low quality NPE	Hälsoringen i Osby
Kristianstadbladet	20140322	Low quality NPE	Läkarmottagningen i Bjärnum
Kristianstadbladet	20140322	Low quality NPE	Osby vårdcentral
Kristianstadbladet	20140322	Low quality NPE	Capio Göinge
Göteborgs-Posten	20140410	Mistreated	Capio Vårdcentral Mölndal
Helsingborgs Dagblad	20140428	Bad access	Rydebäcks vårdcentral
Sydsvenskan	20140507	Low quality	Vårdcentralen Kirseberg
Borås Tidning	20140514	Mistreated	Vårdcentralen Herkules
Falköpings tidning	20140612	Bad access	Oden Vårdcentral
Sydsvenskan	20140624	Missing remit	Vårdcentralen Måsen
Bohusläningen	20141021	Mistreated	Vårdcentralen Bäckefors
Skånska Dagbladet	20141024	Low quality	Vårdcentralen i Skurup
Bohusläningen	20141202	Mistreated	Vårdcentralen Dagson
Bohusläningen	20141211	Mistreated	Achima Care Uddevalla
Partille tidning	20150129	IVO critique	Capio Vårdcentral Sävedalen
Nya Lidköpings-tidningen	20150206	Maltreatment	Vårdcentral Nossebro
Lerums Tidning	20150304	Low availability	Vårdcentralen i Gråbo
Ulricehamns Tidning	20150314	IVO critique	Närhälsan Ulricehamn
Nya Lidköpings-tidningen	20150318	IVO critique	Kinnekullahälsan
Skaraborgs Läns Tidning	20150318	Lex Maria	Vårdcentral Götene
Borås Tidning	20150407	Fee due to insufficient staffing	Närhälsan Horred
Partille tidning	20150409	Bad rating	Partille Vårdcentral
Bohusläningen	20150909	IVO critique	Närhälsan Tanumshede
Skaraborgs Läns Tidning	20150928	Maltreatment	Vilans vårdcentral

Sydsvenskan	20151010	Bad coverage	Multi-Clinic
Helsingborgs Dagblad	20151114	Maltreatment	Ekeby vårdcentral
Kungälvs-Posten	20151204	Water leakage	Vårdcentralen Kusten
Borås Tidning	20160210	Insufficient staffing	Närhälsan Trandared
Trelleborgs Allehanda	20160421	Negative patient story	Valens Läkargrupp
Borås Tidning	20160424	Lex maria	Boda VC
Provinstidningen	20160521	IVO critique	Balderkliniken
Bohusläningen	20160622	Lex maria	Vårdcentralen Silentzvägen
Norra Skåne	20160823	Negative patient story	Vårdcentralen Vilan
Helsingborgs Dagblad	20160824	Bad coverage	Solljungahälsan
Lerums Tidning	20160928	Patient-written article	Floda vårdcentral
Mariestads-Tidning	20161118	Bad physician availability	Vårdcentralen Hova

Table B2: Group 2: Weakly negative

Dalslänningen	20130104	Man was beaten down	Vårdcentralen Ekbacken
Skånska Dagbladet	20130722	Lack of serivces	Vårdcentralen Husie
Helsingborgs Dagblad	20130807	Bad coverage	Vårdcentralen Söderåsen
Helsingborgs Dagblad	20130807	Bad coverage	Vårdcentralen Åstorp
Skaraborgs Allehanda	20131219	Negative coverage	Närhälsan Skövde
Skaraborgs Allehanda	20131219	Negative coverage	Vårdcentralen Hjo
Skaraborgs Allehanda	20131219	Negative coverage	Vårdcentralen i Hentorp
Helsingborgs Dagblad	20140208	Diagnosis without basis	Berga Läkarhus
Helsingborgs Dagblad	20140208	Diagnosis without basis	Roslunds
Helsingborgs Dagblad	20140208	Diagnosis without basis	Båstad Bjäre läkarpraktik
Helsingborgs Dagblad	20140509	Non-legitimized physician	Vårdcentralen Centrum
Kristanstadsbladet	20140527	Insufficient staff	Vårdcentralen Degeberga
Strömstads Tidning	20150623	Patient-written article	Backa Läkarhus
Borås Tidning	20151015	Inreased visits due to refugee inflow	Sjöbo Vådcentral
Dalslänningen	20151023	Insufficient staffing	VC Nygård
Ulricehamns tidning	20151203	Insufficient staffing	Vårdcentralen Kinna
Ulricehamns tidning	20160219	Massflykt av chefer	Vårdcentralen Dalsjöfors
Göteborgs-Posten	20160903	Fee	Hönö vårdcentral
Göteborgs-Posten	20160903	Fee	Närhälsan Öckerö

Table B3: Group 3: Weakly positive

Norra Skåne	20130204	Increased local cooperation	Vårdcentralen Vänhem
Skånska Dagbladet	20130315	Child death wasn't carecenters fault	Vårdcentralen Näsby
Trelleborgs Allehanda	20130426	Positive patient story	Kattens Läkargrupp, Trelleborg
Sydsvenskan	20130722	Good rating	Victoria Vård & Hälsa
Sydsvenskan	20130722	Good rating	Örestadsklinikens vårdcentral
Helsingborgs Dagblad	20130804	Good rating	Capio Citykliniken Helsingborg
Sydsvenskan	20130911	Good rating	NOVA-kliniken, Ystad
Helsingborgs Dagblad	20131008	Positive coverage	Vårdcentralen Närlunda
Helsingborgs Dagblad	20131203	Increased services	Vårdcentralen Granen
Kristianstadsbladet	20140125	Quick administration	Näsums hälsocentral
Kristianstadsbladet	20140322	Good rating, NPE	Vårdcentralen Brösarp
Helsingborgs Dagblad	20140401	New collaboration	Capio Citykliniken Båstad
Helsingborgs Dagblad	20140401	New collaboration	Vårdcentralen Förslöv
Helsingborgs Dagblad	20140401	New collaboration	Vårdcentralen Klippan
Helsingborgs Dagblad	20140401	New collaboration	Capio Citykliniken Klippan
Helsingborgs Dagblad	20140401	New collaboration	Kungsgårdshälsan
Helsingborgs Dagblad	20140401	New collaboration	Läkargruppen Munka Ljungby
Helsingborgs Dagblad	20140401	New collaboration	Vårdcentralen Sjöcrona
Helsingborgs Dagblad	20140401	New collaboration	Vårdkliniken i Ängelholm
Helsingborgs Dagblad	20140401	New collaboration	Vårdcentralen Örkelljunga
Sydsvenskan	20140909	Positive review	Näsets läkargrupp
Partille tidning	20150115	Sufficient staffing	Mössebergs VC
Borås tidning	20150314	Renovation	Dalabergs vårdcentral
Ulricehamns tidning	20150618	New locals	Vårdcentralen Furulund
Skaraborgs Allehanda	20150904	No Ivo critique	Vårdcentralen Stenstorp
Kristianstadsbladet	20160920	Good rating	Brahehälsan Löberöd
Kristianstadsbladet	20160920	Good rating	Hälsomedicinskt Center Landskrona
Kristianstadsbladet	20160920	Good rating	Capio Citykliniken Olympia

Table B4: Group 4: Strongly positive

Helsingborgs Dagblad	20130215	Good rating	Solklart Vård i Bjuv
Helsingborgs Dagblad	20130215	Good rating	Familjehälsan Åstorp
TTELA	20130312	Positive coverage	Vargöns Vårdcentral
Skånska Dagbladet	20130403	Good rating	Vårdcentralen Knislinge
Mölndals-Posten	20130502	Positive coverage	Nötkärnan Kållered
Ystads Allehanda	20130608	New services	Vårdcentralen Tomelilla
Göteborgs-Posten	2014030	Good rating NPE	Vårdcentralen Herrestad
Kristianstadsbladet	20140322	Good rating, NPE	Kristianstadkliniken
Norra Skåne	20140524	Good rating	Vårdcentralen Perstorp
Partille tidning	20150409	High rating	Adinahälsan
Strömstads Tidning	20150410	MR camera	Cityläkarna
Nya Lidköpings-tidningen	20150601	sufficient staffing	Vårdcentralen Guldvingen
Norra Skåne	20150918	Prize winner	Vårdcentralen Råå
Skånska Dagbladet	20151202	Prize winner	Vårdcentralen Tåbelund
Partille tidning	20160201	Sufficient staffing	Bergsjön Vårdcentral
TTELA	20160206	High rating	Maria Alberts Vårdcentral
TTELA	20160810	Positive patient storye	Vårdcentralen Nordstan
Göteborgs-Posten	20160921	High rating	Närhälsan Stora Höga
Helsingborgs Dagblad	20160824	Good rating	Vårdcentralen Ljungbyhed
Skånska Dagbladet	20161220	New facilities	Vårdcentralen Södervärn

Bibliography

Abaluck, J., L. Agha, D. C. C. Jr, D. Singer, and D. Zhu (2020). Fixing Misallocation with Guidelines: Awareness vs. Adherence.

Abaluck, J. and J. Gruber (2016). Improving the quality of choices in health insurance markets. NBER Working Paper 22917.

Adda, J. (2020). Preventing the Spread of Antibiotic Resistance. AEA Papers and Proceedings" 110, 255–59.

Altonji, J. G. (1995). The effects of high school curriculum on education and labor market outcomes. The Journal of Human Resources 30(3), 409–438.

Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. Annual Review of Economics 4(1), 185–223.

Anderberg, D., A. Chevalier, and J. Wadsworth (2011). Anatomy of a health scare: education, income and the MMR controversy in the UK. Journal of Health Economics 30(3), 515–530.

Anell, A. (2011). Choice and privatisation in swedish primary care. Health Economics, Policy and Law 6(4), 549–569.

Anell, A. (2015). The public-private pendulum – patient choice and equity in sweden. New England Journal of Medicine 372(1), 1–4.

Anell, A., M. Dackehag, and J. Dietrichson (2018). Does risk-adjusted payment influence primary care providers' decision on where to set up practices? BMC Health Services Research 18, 179.

Anell, A., M. Dackehag, J. Dietrichson, L. M. Ellegård, G. Kjellsson, et al. (2022). Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform. Technical Report Working Paper 2022:15, Department of Economics, Lund University.

Anell, A., J. Dietrichson, L. M. Ellegård, G. Kjellsson, et al. (2022). Well-informed choices? effects of information interventions in primary care on care quality. Technical report.

Anell, A., J. Dietrichson, L. M. Ellegård, and G. Kjellsson (2021). Information, switching costs, and consumer choice: Evidence from two randomised field experiments in Swedish primary health care. Journal of Public Economics 196, 104390.

Angrist, J. D. and J.-S. Pischke (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. American Economic Review 53(5), 941–973.

Avdic, D., G. Moscelli, A. Pilny, and I. Sriubaite (2019). Subjective and objective quality and choice of hospital: Evidence from maternal care services in Germany. Journal of Health Economics 68, 102229.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? Journal of Financial Economics 144(2), 370–395.

Bakx, P., B. Wouterse, E. van Doorslaer, and A. Wong (2020, sep). Better off at home? Effects of nursing home eligibility on costs, hospitalizations and survival. Journal of Health Economics 73, 102354.

Bejaoui, S., M. Poulsen, and C. East (2020). The impact of early life antibiotic use on atopic and metabolic disorders: Meta-analyses of recent insights. Evolution, Medicine, and Public Health.

Bensnes, S. and I. Huitfeldt (2021). Rumor has it: How do patients respond to patient-generated physician ratings? Journal of Health Economics 76, 102415.

Bertrand, M., M. Mogstad, and J. Mountjoy (2019). Improving educational pathways to social mobility: Evidence from Norway's "Reform 94". NBER Working Paper Series (Working Paper 25679).

Besley, T. and M. Ghatak (2003). Incentives, choice, and accountability in the provision of public services. Oxford Review of Economic Policy 19(2), 235–249.

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020). Incarceration, Recidivism, and Employment. Journal of Political Economy 128(4).

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017). Life-cycle earnings, education premiums, and internal rates of return. Journal of Labor Economics 35(4), 993–1030.

Bilinski, A. and L. A. Hatfield (2018). Nothing to see here? non-inferiority approaches to parallel trends and other model assumptions. arXiv preprint arXiv:1805.03273.

Biørn, E. and G. Godager (2010). Does quality influence choice of general practitioner? An analysis of matched doctor–patient panel data. Economic Modelling 27(4), 842–853.

Björklund, A. and K. G. Salvanes (2011). Chapter 3 - education and family background: Mechanisms and policies. Handbook of the Economics of Education 3, 201–247.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011). Too young to leave the nest? the effects of school starting age. The Review of Economics and Statistics 93(2), 455–467.

Brekke, K. R., H. Gravelle, L. Siciliani, and O. R. Straume (2014). Patient Choice, Mobility and Competition Among Health Care Providers. In R. Levaggi and M. Montefiori (Eds.), Health Care Provision and Patient Mobility, Volume 12, pp. 1–26. Milano: Springer Milan.

Brilli, Y., C. Lucifora, A. Russo, and M. Tonello (2020). Influenza vaccination behavior and media reporting of adverse events. Health Policy 124(12), 1403–1411.

Brunello, G., M. Giannini, and K. Ariga (2007). The Optimal Timing of School Tracking. In Schools and the Equal Opportunity Problem. MIT Press.

Buser, T., M. Niederle, and H. Oosterbeek (2014). Gender, Competitiveness, and Career Choices. The Quarterly Journal of Economics 129(3), 1409–1447.

Cabana, M. D. and S. H. Jee (2004). Does continuity of care improve patient outcomes? The Journal of Family Practice 53.

Cadieux, G., R. Tamblyn, D. Dauphinee, and M. Libman (2007, oct). Predictors of inappropriate antibiotic prescribing among primary care physicians. Canada Medical Association Journal 177(8), 877–883.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. Journal of Econometrics 225(2), 200–230. Themed Issue: Treatment Effect 1.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015). A flying start? maternity leave benefits and long-run outcomes of children. Journal of Political Economy 123(2), 365–412.

Cattaneo, M., N. Idrobo, and R. Titiunik (2018). A Practical Introduction to Regression Discontinuity Designs: Volume II. In Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge University Press.

Célind, J., L. Södermark, and O. Hjalmarson (2014). Adherence to treatment guidelines for acute otitis media in children. The necessity of an effective strategy of guideline implementation. International Journal of Pediatric Otorhinolaryngology 78(7), 1128–1132.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The Effect of Minimum Wages on Low-Wage Jobs*. The Quarterly Journal of Economics 134(3), 1405–1454.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson (2016). Health care exceptionalism? Performance and allocation in the US health care sector. American Economic Review 106(8), 2110–44.

Chandra, A. and D. O. Staiger (2007). Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks. The Journal of Political Economy 115(1), 103.

Chang, L. V. (2018). Information, education, and health behaviors: Evidence from the MMR vaccine autism controversy. Health Economics 27(7), 1043–1062.

Chartock, B. L. (2021). Quality disclosure, demand, and congestion: Evidence from physician ratings. Technical report, Department of Economics, Bentley University.

Chen, Y. and S. Lee (2021). User-Generated Physician Ratings and Their Effects on Patients' Physician Choices: Evidence from Yelp.

Cheng, M.-Y., J. Fan, and J. S. Marron (1997). On automatic boundary corrections. The Annals of Statistics 25(4), 1691–1708.

Cooper, Z., S. Gibbons, S. Jones, and A. McGuire (2011). Does Hospital Competition Save Lives? Evidence from the English NHS Patient Choice Reforms. Economic Journal 121(554), 228–260.

Coxeter, P., C. Mar, and T. Hoffmann (2017). Parents' Expectations and Experiences of Antibiotics for Acute Respiratory Infections in Primary Care. Annals of family medicine 15(2), 149–154.

Cunningham, S. (2021). Causal Inference: The Mixtape. Yale University Press.

Currie, J., W. Lin, and J. Meng (2014). Addressing antibiotic abuse in China: An experimental audit study. Journal of Development Economics 110, 39–51.

Currie, J., W. B. MacLeod, and J. Van Parys (2016). Provider practice style and patient health outcomes: The case of heart attacks. Journal of Health Economics 47, 64–80.

Cutler, D. M. (2011). Where are the health care entrepreneurs? the failure of organizational innovation in health care. Innovation Policy and the Economy 11(1), 1–28.

Dackehag, M. and L. M. Ellegård (2019). Competition, Capitation, and Coding: Do Public Primary Care Providers Respond to Increased Competition? CESifo Economic Studies 65(4), 402–423.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family Welfare Cultures. The Quarterly Journal of Economics 129(4), 1711–1752.

Dahlgren, C., M. Dackehag, P. Wändell, and C. Rehnberg (2021). Simply the best? The impact of quality on choice of primary healthcare provider in Sweden. Health Policy.

Dalsgaard, S., H. S. Nielsen, and M. Simonsen (2014). Consequences of ADHD medication use for children's outcomes. Journal of Health Economics 37(1), 137–151.

Daysal, N. M., H. Ding, M. Rossin-Slater, and H. Schwandt (2021). Germs in the Family: The Long-Term Consequences of Intra-Household Endemic Respiratory Disease Spread. NBER Working Paper number 29524.

de Chaisemartin, C. and X. D'Haultfœuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review 110(9), 2964–96.

Deshpande, M. and Y. Li (2019). Who Is Screened Out? Application Costs and the Targeting of Disability Programs. American Economic Journal: Economic Policy 11(4), 213–48.

DesJardins, S. L., R. K. Toutkoushian, D. Hossler, and J. Chen (2019). Time may change me: Examining how aspirations for college evolve during high school. Review of Higher Education 43(1), 263–294.

Dietrichson, J., L. M. Ellegård, and G. Kjellsson (2020). Patient Choice, Entry, and the Quality of Primary Care: Evidence from Swedish Reforms. Health Economics 29(6), 716–730.

Dobbie, W., J. Goldin, and C. S. Yang (2018). The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges. American Economic Review 108(2), 201–40.

Domurat, R., I. Menashe, and W. Yin (2021). The role of behavioral frictions in health insurance marketplace enrollment and risk: Evidence from a field experiment. American Economic Review 111(5), 1549–74.

Doyle, J. J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. American Economic Review 97(5), 1583–1610.

Ellegård, L., G. Kjellsson, and L. Mattisson (2021). An App Call a Day Keeps the Patient Away? Substitution of Online and In-Person Doctor Consultations Among Young Adults An App Call a Day Keeps the Patient Away?

Ellegård, L. M., J. Dietrichson, and A. Anell (2018). Can pay-for-performance to primary care providers stimulate appropriate use of antibiotics? Health economics 27(1), e39–e54.

Ellegård, L. M. (2020). Effects of pay-for-performance on prescription of hypertension drugs among public and private primary care providers in Sweden. International Journal of Health Economics and Management.

Ellegård, L. M., J. Dietrichson, and A. Anell (2018). Can pay-for-performance to primary care providers stimulate appropriate use of antibiotics? Health Economics 27(1), 39–54.

Ellis, R. P. (1998). Creaming, skimping and dumping: Provider competition on the intensive and extensive margins. Journal of Health Economics 17(5), 537–555.

Ericson, K. M. M., J. Kingsdale, T. Layton, and A. Sacarny (2017). Nudging leads consumers in Colorado to shop but not switch ACA marketplace plans. Health Affairs 36(2), 311–319.

Erikson, B.-M., Å. Melhus, and J. Sjölin (2014). Nya rekommendationer för akut faryngotonsillit kan leda fel. Läkartidningen 01.

Finkelstein, A., P. Persson, M. Polyakova, and J. Shapiro (2021). A Taste of Their Own Medicine: Guideline Adherence and Access to Expertise. IFN Working Paper (1421).

Fogelberg, S. (2013). Effects of Competition between Healthcare Providers on Prescription of Antibiotics. IFN Working Paper (949).

Fredriksson, P. and B. Öckert (2014). Life-cycle effects of age at school start. The Economic Journal 124(579), 977–1004.

Gardner, D. P. (1983). A Nation At Risk: The Imperative For Educational Reform. An Open Letter to the American People. A Report to the Nation and the Secretary of Education.

Gaynor, M., R. Moreno-Serra, and C. Propper (2013). Death by market power: Reform, competition and patient outcomes in the British National Health Service. American Economic Journal: Economic Policy 5(4), 134–166.

Gaynor, M., C. Propper, and S. Seiler (2016). Free to choose? reform, choice, and consideration sets in the english national health service. American Economic Review 106(11), 3521–57.

Gelman, A. and G. Imbens (2018). Why high-order polynomials should not be used in regression discontinuity designs. Journal of Business & Economic Statistics 0(0), 1–10.

Ginja, R., J. Riise, B. Willage, and A. Willén (2022). Does Your Doctor Matter? Doctor Quality and Patient Outcomes. NHH Dept. of Economics Discussion Paper (08).

Glenngård, A. (2015). Primärvården efter vårdvalsreformen: Valfrihet, effektivitet och produktivitet. SNS Förlag.

Gode, D. K. and S. Sunder (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. Journal of Political Economy 101(1), 119–137.

Goodman, J. (2017). The labor of division: Returns to compulsory high school math coursework. NBER Working Paper Series (Working Paper 23063).

Goodman-Bacon, A. (2021a). Difference-in-differences with variation in treatment timing. Journal of Econometrics 225(2), 254–277.

Goodman-Bacon, A. (2021b). The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes. American Economic Review 111(8), 2550–93.

Görlitz, K. and C. Gravert (2018). The effects of a high school curriculum reform on university enrollment and the choice of college major. Education Economics 26(3), 321–336.

Graetz, G., B. Öckert, and O. Nordström Skans (2020). Family Background and the Responses to Higher SAT Scores. IZA Institute of Labor Economics (IZA DP No. 13343).

Gravelle, H., D. Liu, C. Propper, and R. Santos (2019). Spatial competition and quality: Evidence from the english family doctor market. Journal of Health Economics 68, 102249.

Groth, A., F. Enoksson, A. Hermansson, M. Hultcrantz, J. Stalfors, and K. Stenfeldt (2011). Acute mastoiditis in children in Sweden 1993–2007—No increase after new guidelines. International Journal of Pediatric Otorhinolaryngology 75(12), 1496–1501.

Hanushek, E. A. and D. D. Kimko (2000). Schooling, labor-force quality, and the growth of nations. American Economic Review 90, 1184–1208.

Hayek, F. A. (1944). The use of knowledge in society. American Economic Review 35(4), 519–530.

Hedin, K. (2015). Luftvägsinfektioner hos förskolebarn. Technical report, Smittskyddsinstitutet.

Hibbard, J. H., N. Berkman, L. A. McCormack, and E. Jael (2002). The impact of a CAHPS report on employee knowledge, beliefs, and decisions. Medical Care Research and Review 59(1), 104–116.

Holmlund, B., Q. Liu, and O. Nordstrom Skans (2007). Mind the gap? Estimating the effects of postponing higher education. Oxford Economic Papers 60(4), 683–710.

Holmlund, H., J. Häggblom, E. Lindahl, S. Martinson, A. Sjögren, U. Vikman, and B. Öckert (2014). Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola.

Huang, S. and H. Ullrich (2021). Physician Effects in Antibiotic Prescribing Evidence from Physician Exits. DIW Discussion Paper (1958).

Iversen, T. and H. Lurås (2011). Patient Switching in General Practice. Journal of Health Economics 30(5), 894–903.

Jernberg, C., S. Löfmark, C. Edlund, and J. K. Jansson (2010). Long-term impacts of antibiotic exposure on the human intestinal microbiota. Microbiology 156(11), 3216–3223.

Joensen, J. S. and H. S. Nielsen (2016). Mathematics and gender: Heterogeneity in causes and consequences. Economic Journal 126(593), 1129–1163.

Jourdan, A., B. Sangha, E. Kim, S. Nawaz, V. Malik, R. Vij, and S. Sekhsaria (2020, jan). Antibiotic hypersensitivity and adverse reactions: Management and implications in clinical practice. Allergy, Asthma and Clinical Immunology 16(1), 1–7.

Kim, E., M. E. Shepherd, and J. D. Clinton (2020). The effect of big-city news on rural america during the covid-19 pandemic. Proceedings of the National Academy of Sciences 117(36), 22009–22014.

King, G., B. Schneer, and A. White (2017). How the news media activate public expression and influence national agendas. Science 358(6364), 776–780.

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

Kling, J. R. (2006). Incarceration Length, Employment, and Earnings. American Economic Review 96(3), 863–876.

Kling, J. R., S. Mullainathan, E. Shafir, L. C. Vermeulen, and M. V. Wrobel (2012). Comparison friction: Experimental evidence from Medicare drug plans. Quarterly Journal of Economics 127(1), 199–235.

Knutson, D. J., E. A. Kind, J. B. Fowles, and S. Adlis (1998). Impact of report cards on employees: A natural experiment. Health Care Financing Review 20(1), 5–27.

Kunz, J. and K. E. Staub (2016). Subjective Completion Beliefs and the Demand for Post-Secondary Education. SSRN Electronic Journal (10344).

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. Journal of Economic Literature 48(2), 281–355.

Lee, J. Y. and G. Solon (2011). The fragility of estimated effects of unilateral divorce laws on divorce rates. The BE Journal of Economic Analysis & Policy 11(1).

Levine, P. B. and D. J. Zimmerman (1995). The benefit of additional high-school math and science classes for young men and the benefit of additional high-school. Journal of Business & Economic Statistics 13(2), 137–149.

Luca, M. and S. Vats (2013). Digitizing doctor demand: The impact of online reviews on doctor choice. Cambridge, MA: Harvard Business School.

Lundborg Ander, A. and R. Eggertsen (2004). Akut otitis media behandlas inte enligt rekommendationer. Läkartidningen 101(41).

Maestas, N., K. J. Mullen, and A. Strand (2013). Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. American Economic Review 103(5), 1797–1829.

Malamud, O. and C. Pop-Eleches (2010). General education versus vocational training: Evidence from an economy in transition. Review of Economics and Statistics 92(1), 43–60.

Malamud, O. and C. Pop-Eleches (2011). School tracking and access to higher education among disadvantaged groups. Journal of Public Economics 95(11-12), 1538–1549.

McEwan, P. J. and J. S. Shapiro (2008). The benefits of delayed primary school enrollment. Journal of Human Resources 43(1), 1–29.

Milos Nymberg, V., L. M. Ellegård, G. Kjellsson, M. Wolff, B. Borgström Bolmsjö, T. Wallman, and S. Calling (2021). Trends in remote health care consumption in Sweden: A comparison before and during the first wave of the Covid-19 pandemic (Preprint). JMIR Human Factors.

Molin, E. (2019). School choice and student sorting - the impact of changing admission criteria. Unpublished results.

Moscelli, G., H. Gravelle, and L. Siciliani (2016). Market structure, patient choice and hospital quality for elective patients. CHE Research Paper 139.

Moscelli, G., H. Gravelle, L. Siciliani, and R. Santos (2018). Heterogeneous effects of patient choice and hospital competition on mortality. Social Science & Medicine 216, 50–58.

Mueller-Smith, M. (2015). The Criminal and Labor Market Impacts.

Mukamel, D. B., S. F. Haeder, and D. L. Weimer (2014). Top-down and bottom-up approaches to health care quality: The impact of regulation and report cards. Annual Review of Public Health 35, 477–497.

National Board of Health and Welfare (2014). Vägledning för barnhälsovården. Technical report, Socialstyrelsen.

Newhouse, J. P. (1996). Reimbursing health plans and health providers: Efficiency in production versus selection. Journal of Economic Literature 34(3), 1236–1263.

Ning, J. (2014). Do stricter high school math requirements raise college STEM attainment? Unpublished results.

Noddings, N. (2011). Schooling for democracy. Democracy & Education 19(1).

Nord, M., S. Engström, and S. Mölstad (2013). Mycket varierande förskrivning av antibiotika i primärvården. Läkartidningen (110).

Nordicom (2015). Mediebarometern 2015. Technical report, Göteborgs Universitet.

Norris, S., M. Pecenco, and J. Weaver (2021). The effects of parental and sibling incarceration: Evidence from ohio. American Economic Review 111(9), 2926–63.

OECD (2018). Education at a Glance 2018: OECD Indicators. OECD Publishing, Paris., http://dx.doi.org/10.1787/eag–2018–en.

Oliveira, I., C. Rego, G. Semedo, D. Gomes, A. Figueiras, F. Roque, and M. Herdeiro (2020). Systematic review on the impact of guidelines adherence on antibiotic prescription in respiratory infections. Antibiotics (Basel) 9(9), 546.

Pallon, J. (2022). Pharyngotonsillitis in primary health care - Aetiology and clinical findings. Ph. D. thesis, Lund University.

Pauly, M. V. and M. A. Satterthwaite (1981). The pricing of primary care physicians services: A test of the role of consumer information. The Bell Journal of Economics 12(2), 488–506.

Petterson, S. and Å. Jaktlund (2013). Primärvårdens läkarbemanning—öppna jämförelser mellan landsting och driftsformer av primärvårdens försörjning av specialistläkare 2012: Sveriges läkarförbund. Sveriges Läkarförbund.

Pope, D. G. (2009). Reacting to rankings: Evidence from "America's best hospitals". Journal of Health Economics 28(6), 1154–1165.

Presstödsnämnden (2015). Dagstidningsförordningen 2015.

Rapoport, B. and C. Thibout (2018). Why do boys and girls make different educational choices? The influence of expected earnings and test scores. Economics of Education Review 62(1), 205–229.

Riksrevisionen (2014). Primärvårdens styrning - efter behov eller efterfrågan? Technical report.

Rose, H. and J. R. Betts (2004). The effect of high school courses on earnings. Review of Economics and Statistics 86(2), 497–513.

Sabbatini, A. K., K. E. Kocher, A. Basu, and R. Y. Hsia (2016). In-Hospital Outcomes and Costs Among Patients Hospitalized During a Return Visit to the Emergency Department. JAMA 315(7), 663–671.

Santos, R., H. Gravelle, and C. Propper (2017). Does quality affect patient's choice of doctor? Evidence from England. Economic Journal 127(600), 445–494.

Schwartz, D. J., A. E. Langdon, and G. Dantas (2020, sep). Understanding the impact of antibiotic perturbation on the human microbiome. Genome medicine 12(1).

Sievertsen, H., G. van der Berg, and P. Bingley (2021). Antibiotics in Early Childhood and Subsequent Cognitive Skills.

Skellern, M. (2017). The hospital as a multi-product firm: The effect of hospital competition on value-added indicators of clinical quality. CEP Discussion Papers 1484.

Smith, M. J., S. S. Ellenberg, L. M. Bell, and D. M. Rubin (2008). Media coverage of the measles-mumps-rubella vaccine and autism controversy and its relationship to MMR immunization rates in the United States. Pediatrics 121(4), 836–843.

Söderström, M. and R. Uusitalo (2010). School choice and segregation: Evidence from an admission reform. Scandinavian Journal of Economics 112(1), 55–76.

Sosa, A. (2016). Impact of mathematics course taking during high school on earnings : Evidence from shocks to teachers ' supply. Unpublished results.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics 225(2), 175–199. Themed Issue: Treatment Effect 1.

Suppli, C. H., N. D. Hansen, M. Rasmussen, P. Valentiner-Branth, T. G. Krause, and K. Mølbak (2018). Decline in HPV-vaccination uptake in Denmark–the association between HPV-related media coverage and HPV-vaccination. BMC Public Health 18(1), 1–8.

Totten, A., J. Wagner, A. Tiwari, C. O'Haire, J. Griffin, and M. Walker (2012). Public reporting as a quality improvement strategy. Closing the quality gap: Revisiting the state of the science. Evidence Report No. 208.

Vård och Omsorgsanalys (2021a). Fast kontakt i primärvården. Technical report, Report 2021:1.

Vård och Omsorgsanalys (2021b). Vården ur befolkningens perspektiv 2020. Technical report, Report 2021:4.

Varkevisser, M., S. A. van der Geest, and F. T. Schut (2012). Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands. Journal of Health Economics 31(2), 371–378.

Vengberg, S., M. Fredriksson, and U. Winblad (2019). Patient choice and provider competition–quality enhancing drivers in primary care? Social Science & Medicine 226, 217–224.

Victoor, A., D. M. Delnoij, R. D. Friele, and J. J. Rademakers (2012). Determinants of patient choice of healthcare providers: A scoping review. BMC Health Services Research 12(272).

WHO (2022). Antibiotic resistance. accessed on 02.01.2022.

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? a reconciliation and new results. American Economic Review 96(5), 1802–1820.

Youngster, I., J. Avorn, V. Belleudi, A. Cantarutti, J. Díez-Domingo, U. Kirchmayer, B. J. Park, S. Peiró, G. Sanfélix-Gimeno, H. Schröder, K. Schüssel, J. Y. Shin, S. M. Shin, G. S. Simonsen, H. S. Blix, A. Tong, G. Trifirò, T. Ziv-Baran, and S. C. Kim (2017). Antibiotic Use in Children – A Cross-National Analysis of 6 Countries. The Journal of Pediatrics 182, 239–244.

Yu, H. and N. Mocan (2018). The impact of high school curriculum on confidence, academic success, and mental and physical well-being of university students. NBER Working Paper Series (Working Paper 24573).

Zhuang, M., M. Garz, et al. (2022). Media coverage and pandemic behaviour: Evidence from Sweden. Technical report, Stockholm School of Economics, Mistra Center for Sustainable Markets (Misum).

Previous doctoral theses in the Department of Economics, Gothenburg

Avhandlingar publicerade innan serien Ekonomiska Studier startades
(Theses published before the series Ekonomiska Studier was started):

Östman, Hugo (1911), Norrlands ekonomiska utveckling
Moritz, Marcus (1911), Den svenska tobaksindustrien
Sundbom, I. (1933), Prisbildning och ändamålsenlighet
Gerhard, I. (1948), Problem rörande Sveriges utrikeshandel 1936/38
Hegeland, Hugo (1951), The Quantity Theory of Money
Mattsson, Bengt (1970), Cost-Benefit analys
Rosengren, Björn (1975), Valutareglering och nationell ekonomisk politik
Hjalmarsson, Lennart (1975), Studies in a Dynamic Theory of Production and its Applications
Örtendahl, Per-Anders (1975), Substitutionsaspekter på produktionsprocessen vid massaframställning
Anderson, Arne M. (1976), Produktion, kapacitet och kostnader vid ett helautomatiskt emballageglasbruk
Ohlsson, Olle (1976), Substitution och odelbarheter i produktionsprocessen vid massaframställning
Gunnarsson, Jan (1976), Produktionssystem och tätortshierarki – om sambandet mellan rumslig och ekonomisk struktur
Köstner, Evert (1976), Optimal allokering av tid mellan utbildning och arbete
Wigren, Rune (1976), Analys av regionala effektivitetsskillnader inom industribranscher
Wästlund, Jan (1976), Skattning och analys av regionala effektivitetsskillnader inom industribranscher
Flöjstad, Gunnar (1976), Studies in Distortions, Trade and Allocation Problems
Sandelin, Bo (1977), Prisutveckling och kapitalvinster på bostadsfastigheter
Dahlberg, Lars (1977), Empirical Studies in Public Planning
Lönnroth, Johan (1977), Marxism som matematisk ekonomi
Johansson, Börje (1978), Contributions to Sequential Analysis of Oligopolistic Competition

Ekonomiska Studier, utgivna av Nationalekonomiska institutionen vid Göteborgs Universitet. Nr 1 och 4 var inte doktorsavhandlingar. (The contributions to the department series 'Ekonomiska Studier' where no. 1 and 4 were no doctoral theses):

2. **Ambjörn, Erik** (1959), Svenskt importberoende 1926-1956: en ekonomisk-statistisk kartläggning med kommentarer
3. **Landgren, K-G.** (1960), Den "Nya ekonomien" i Sverige: J.M. Keynes, E. Wigfors och utecklingen 1927-39
5. **Bigsten, Arne** (1979), Regional Inequality and Development: A Case Study of Kenya
6. **Andersson, Lars** (1979), Statens styrning av de kommunala budgetarnas struktur (Central Government Influence on the Structure of the Municipal Budget)
7. **Gustafsson, Björn** (1979), Inkomst- och uppväxtförhållanden (Income and Family Background)
8. **Granholm, Arne** (1981), Interregional Planning Models for the Allocation of Private and Public Investments
9. **Lundborg, Per** (1982), Trade Policy and Development: Income Distributional Effects in the Less Developed Countries of the US and EEC Policies for Agricultural

Commodities

10. **Juås, Birgitta** (1982), Värdering av risken för personskador. En jämförande studie av implicita och explicita värden. (Valuation of Personal Injuries. A comparison of Explicit and Implicit Values)
11. **Bergendahl, Per-Anders** (1982), Energi och ekonomi - tillämpningar av input-output analys (Energy and the Economy - Applications of Input-Output Analysis)
12. **Blomström, Magnus** (1983), Foreign Investment, Technical Efficiency and Structural Change - Evidence from the Mexican Manufacturing Industry
13. **Larsson, Lars-Göran** (1983), Comparative Statics on the Basis of Optimization Methods
14. **Persson, Håkan** (1983), Theory and Applications of Multisectoral Growth Models
15. **Sterner, Thomas** (1986), Energy Use in Mexican Industry.
16. **Flood, Lennart** (1986), On the Application of Time Use and Expenditure Allocation Models.
17. **Schuller, Bernd-Joachim** (1986), Ekonomi och kriminalitet - en empirisk undersökning av brottsligheten i Sverige (Economics of crime - an empirical analysis of crime in Sweden)
18. **Walfridson, Bo** (1987), Dynamic Models of Factor Demand. An Application to Swedish Industry.
19. **Stålhammar, Nils-Olov** (1987), Strukturomvandling, företagsbeteende och förväntningsbildning inom den svenska tillverkningsindustrin (Structural Change, Firm Behaviour and Expectation Formation in Swedish Manufactury)
20. **Anxo, Dominique** (1988), Sysselsättningseffekter av en allmän arbetstidsförkortning (Employment effects of a general shortage of the working time)
21. **Mbelle, Ammon** (1988), Foreign Exchange and Industrial Development: A Study of Tanzania.
22. **Ongaro, Wilfred** (1988), Adoption of New Farming Technology: A Case Study of Maize Production in Western Kenya.
23. **Zejan, Mario** (1988), Studies in the Behavior of Swedish Multinationals.
24. **Görling, Anders** (1988), Ekonomisk tillväxt och miljö. Förorenings-struktur och ekonomiska effekter av olika miljövårdsprogram. (Economic Growth and Environment. Pollution Structure and Economic Effects of Some Environmental Programs).
25. **Aguilar, Renato** (1988), Efficiency in Production: Theory and an Application on Kenyan Smallholders.
26. **Kayizzi-Mugerwa, Steve** (1988), External Shocks and Adjustment in Zambia.
27. **Bornmalm-Jardelöw, Gunilla** (1988), Högre utbildning och arbetsmarknad (Higher Education and the Labour Market)
28. **Tansini, Ruben** (1989), Technology Transfer: Dairy Industries in Sweden and Uruguay.
29. **Andersson, Irene** (1989), Familjebeskattning, konsumtion och arbetsutbud - En ekonometrisk analys av löne- och inkomstelasticiteter samt policysimuleringar för svenska hushåll (Family Taxation, Consumption and Labour Supply - An Econometric Analysis of Wage and Income Elasticities and Policy Simulations for Swedish Households)
30. **Henrekson, Magnus** (1990), An Economic Analysis of Swedish Government Expenditure
31. **Sjöö, Boo** (1990), Monetary Policy in a Continuous Time Dynamic Model for Sweden
32. **Rosén, Åsa** (1991), Contributions to the Theory of Labour Contracts.
33. **Loureiro, Joao M. de Matos** (1992), Foreign Exchange Intervention, Sterilization and Credibility in the EMS: An Empirical Study
34. **Irandoust, Manuchehr** (1993), Essays on the Behavior and Performance of the Car Industry
35. **Tasiran, Ali Cevat** (1993), Wage and Income Effects on the Timing and Spacing of Births in Sweden and the United States
36. **Milopoulos, Christos** (1993), Investment Behaviour under Uncertainty: An Econometric Analysis of Swedish Panel Data
37. **Andersson, Per-Åke** (1993), Labour Market Structure in a Controlled Economy: The Case of Zambia
38. **Storrie, Donald W.** (1993), The Anatomy of a Large Swedish Plant Closure
39. **Semboja, Haji Hatibu Haji** (1993), Energy and Development in Kenya
40. **Makonnen, Negatu** (1993), Labor Supply and the Distribution of Economic Well-Being: A Case Study of Lesotho
41. **Julin, Eva** (1993), Structural Change in Rural Kenya
42. **Durevall, Dick** (1993), Essays on Chronic Inflation: The Brazilian Experience
43. **Veiderpass, Ann** (1993), Swedish Retail Electricity Distribution: A Non-Parametric Approach to Efficiency and Productivity Change
44. **Odeck, James** (1993), Measuring Productivity Growth and Efficiency with Data Envelopment Analysis: An Application on the Norwegian Road Sector
45. **Mwenda, Abraham** (1993), Credit Rationing and Investment Behaviour under Market Imperfections: Evidence from Commercial Agriculture in Zambia
46. **Mlambo, Kupukile** (1993), Total Factor Productivity Growth: An Empirical Analysis of Zimbabwe's Manufacturing Sector Based on Factor Demand Modelling
47. **Ndung'u, Njuguna** (1993), Dynamics of the Inflationary Process in Kenya
48. **Modén, Karl-Markus** (1993), Tax Incentives of Corporate Mergers and Foreign Direct Investments
49. **Franzén, Mikael** (1994), Gasoline Demand - A Comparison of Models
50. **Heshmati, Almas** (1994), Estimating Technical Efficiency, Productivity Growth And Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture
51. **Salas, Osvaldo** (1994), Efficiency and Productivity Change: A Micro Data Case Study of the Colombian Cement Industry
52. **Bjurek, Hans** (1994), Essays on Efficiency and Productivity Change with Applications to Public Service Production
53. **Cabezas Vega, Luis** (1994), Factor Substitution, Capacity Utilization and Total Factor Productivity Growth in the Peruvian Manufacturing Industry
54. **Katz, Katarina** (1994), Gender Differentiation and Discrimination. A Study of Soviet Wages
55. **Asal, Maher** (1995), Real Exchange Rate Determination and the Adjustment Process: An Empirical Study in the Cases of Sweden and Egypt
56. **Kjulin, Urban** (1995), Economic Perspectives on Child Care
57. **Andersson, Göran** (1995), Volatility Forecasting and Efficiency of the Swedish Call Options Market
58. **Forteza, Alvaro** (1996), Credibility, Inflation and Incentive Distortions in the Welfare State
59. **Locking, Håkan** (1996), Essays on Swedish Wage Formation
60. **Välilä, Timo** (1996), Essays on the Credibility of Central Bank Independence

61. **Yilma, Mulugeta** (1996), Measuring Smallholder Efficiency: Ugandan Coffee and Food-Crop Production
62. **Mabugu, Ramos E.** (1996), Tax Policy Analysis in Zimbabwe Applying General Equilibrium Models
63. **Johansson, Olof** (1996), Welfare, Externalities, and Taxation; Theory and Some Road Transport Applications.
64. **Chitiga, Margaret** (1996), Computable General Equilibrium Analysis of Income Distribution Policies in Zimbabwe
65. **Leander, Per** (1996), Foreign Exchange Market Behavior Expectations and Chaos
66. **Hansen, Jörgen** (1997), Essays on Earnings and Labor Supply
67. **Cotfas, Mihai** (1997), Essays on Productivity and Efficiency in the Romanian Cement Industry
68. **Horgby, Per-Johan** (1997), Essays on Sharing, Management and Evaluation of Health Risks
69. **Nafar, Nosratollah** (1997), Efficiency and Productivity in Iranian Manufacturing Industries
70. **Zheng, Jinghai** (1997), Essays on Industrial Structure, Technical Change, Employment Adjustment, and Technical Efficiency
71. **Isaksson, Anders** (1997), Essays on Financial Liberalisation in Developing Countries: Capital mobility, price stability, and savings
72. **Gerdin, Anders** (1997), On Productivity and Growth in Kenya, 1964-94
73. **Sharifi, Alimorad** (1998), The Electricity Supply Industry in Iran: Organization, performance and future development
74. **Zamanian, Max** (1997), Methods for Mutual Fund Portfolio Evaluation: An application to the Swedish market
75. **Manda, Damiano Kulundu** (1997), Labour Supply, Returns to Education, and the Effect of Firm Size on Wages: The case of Kenya
76. **Holmén, Martin** (1998), Essays on Corporate Acquisitions and Stock Market Introductions
77. **Pan, Kelvin** (1998), Essays on Enforcement in Money and Banking
78. **Rogat, Jorge** (1998), The Value of Improved Air Quality in Santiago de Chile
79. **Peterson, Stefan** (1998), Essays on Large Shareholders and Corporate Control
80. **Belhaj, Mohammed** (1998), Energy, Transportation and Urban Environment in Africa: The Case of Rabat-Salé, Morocco
81. **Mekonnen, Alemu** (1998), Rural Energy and Afforestation: Case Studies from Ethiopia
82. **Johansson, Anders** (1998), Empirical Essays on Financial and Real Investment Behavior
83. **Köhlin, Gunnar** (1998), The Value of Social Forestry in Orissa, India
84. **Levin, Jörgen** (1998), Structural Adjustment and Poverty: The Case of Kenya
85. **Ncube, Mkhululi** (1998), Analysis of Employment Behaviour in Zimbabwe
86. **Mwansa, Ladslous** (1998), Determinants of Inflation in Zambia
87. **Agnarsson, Sveinn** (1998), Of Men and Machines: Essays in Applied Labour and Production Economics
88. **Kadenge, Phineas** (1998), Essays on Macroeconomic Adjustment in Zimbabwe: Inflation, Money Demand, and the Real Exchange Rate
89. **Nyman, Håkan** (1998), An Economic Analysis of Lone Motherhood in Sweden
90. **Carlsson, Fredrik** (1999), Essays on Externalities and Transport
91. **Johansson, Mats** (1999), Empirical Studies of Income Distribution
92. **Alemu, Tekie** (1999), Land Tenure and Soil Conservation: Evidence from Ethiopia
93. **Lundvall, Karl** (1999), Essays on Manufacturing Production in a Developing Economy: Kenya 1992-94
94. **Zhang, Jianhua** (1999), Essays on Emerging Market Finance
95. **Mlima, Aziz Ponary** (1999), Four Essays on Efficiency and Productivity in Swedish Banking
96. **Davidsen, Björn-Ivar** (2000), Bidrag til den økonomisk-metodologiske tenkningen (Contributions to the Economic Methodological Thinking)
97. **Ericson, Peter** (2000), Essays on Labor Supply
98. **Söderbom, Måns** (2000), Investment in African Manufacturing: A Microeconomic Analysis
99. **Höglund, Lena** (2000), Essays on Environmental Regulation with Applications to Sweden
100. **Olsson, Ola** (2000), Perspectives on Knowledge and Growth
101. **Meuller, Lars** (2000), Essays on Money and Credit
102. **Österberg, Torun** (2000), Economic Perspectives on Immigrants and Intergenerational Transmissions
103. **Kalinda Mkenda, Beatrice** (2001), Essays on Purchasing Power Parity, RealExchange Rate, and Optimum Currency Areas
104. **Nerhagen, Lena** (2001), Travel Demand and Value of Time - Towards an Understanding of Individuals Choice Behavior
105. **Mkenda, Adolf** (2001), Fishery Resources and Welfare in Rural Zanzibar
106. **Eggert, Håkan** (2001), Essays on Fisheries Economics
107. **Andrén, Daniela** (2001), Work, Sickness, Earnings, and Early Exits from the Labor Market. An Empirical Analysis Using Swedish Longitudinal Data
108. **Nivorozhkin, Eugene** (2001), Essays on Capital Structure
109. **Hammar, Henrik** (2001), Essays on Policy Instruments: Applications to Smoking and the Environment
110. **Nannyonjo, Justine** (2002), Financial Sector Reforms in Uganda (1990-2000): Interest Rate Spreads, Market Structure, Bank Performance and Monetary Policy
111. **Wu, Hong** (2002), Essays on Insurance Economics
112. **Linde-Rahr, Martin** (2002), Household Economics of Agriculture and Forestry in Rural Vienam
113. **Maneschiöld, Per-Ola** (2002), Essays on Exchange Rates and Central Bank Credibility
114. **Andrén, Thomas** (2002), Essays on Training, Welfare and Labor Supply
115. **Granér, Mats** (2002), Essays on Trade and Productivity: Case Studies of Manufacturing in Chile and Kenya
116. **Jaldell, Henrik** (2002), Essays on the Performance of Fire and Rescue Services
117. **Alpizar, Francisco, R.** (2002), Essays on Environmental Policy-Making in Developing Countries: Applications to Costa Rica
118. **Wahlberg, Roger** (2002), Essays on Discrimination, Welfare and Labor Supply
119. **Piculescu, Violeta** (2002), Studies on the Post-Communist Transition
120. **Pylkkänen, Elina** (2003), Studies on Household Labor Supply and Home Production
121. **Löfgren, Åsa** (2003), Environmental Taxation – Empirical and Theoretical Applications
122. **Ivaschenko, Oleksiy** (2003), Essays on Poverty, Income Inequality and Health in Transition Economies

123. **Lundström, Susanna** (2003), On Institutions, Economic Growth and the Environment
124. **Wambugu, Anthony** (2003), Essays on Earnings and Human Capital in Kenya
125. **Adler, Johan** (2003), Aspects of Macroeconomic Saving
126. **Erlandsson, Mattias** (2003), On Monetary Integration and Macroeconomic Policy
127. **Brink, Anna** (2003), On the Political Economy of Municipality Break-Ups
128. **Ljungwall, Christer** (2003), Essays on China's Economic Performance During the Reform Period
129. **Chifamba, Ronald** (2003), Analysis of Mining Investments in Zimbabwe
130. **Muchapondwa, Edwin** (2003), The Economics of Community-Based Wildlife Conservation in Zimbabwe
131. **Hammes, Klaus** (2003), Essays on Capital Structure and Trade Financing
132. **Abou-Ali, Hala** (2003), Water and Health in Egypt: An Empirical Analysis
133. **Simatele, Munacinga** (2004), Financial Sector Reforms and Monetary Policy in Zambia
134. **Tezic, Kerem** (2004), Essays on Immigrants' Economic Integration
135. INSTÄLLD
136. **Gjirja, Matilda** (2004), Efficiency and Productivity in Swedish Banking
137. **Andersson, Jessica** (2004), Welfare Environment and Tourism in Developing Countries
138. **Chen, Yinghong** (2004), Essays on Voting Power, Corporate Governance and Capital Structure
139. **Yesuf, Mahmud** (2004), Risk, Time and Land Management under Market Imperfections: Applications to Ethiopia
140. **Kateregga, Eseza** (2005), Essays on the Infestation of Lake Victoria by the Water Hyacinth
141. **Edvardsen, Dag Fjeld** (2004), Four Essays on the Measurement of Productive Efficiency
142. **Lidén, Erik** (2005), Essays on Information and Conflicts of Interest in Stock Recommendations
143. **Dieden, Sten** (2005), Income Generation in the African and Coloured Population – Three Essays on the Origins of Household Incomes in South Africa
144. **Eliasson, Marcus** (2005), Individual and Family Consequences of Involuntary Job Loss
145. **Mahmud, Minhaj** (2005), Measuring Trust and the Value of Statistical Lives: Evidence from Bangladesh
146. **Lokina, Razack Bakari** (2005), Efficiency, Risk and Regulation Compliance: Applications to Lake Victoria Fisheries in Tanzania
147. **Jussila Hammes, Johanna** (2005), Essays on the Political Economy of Land Use Change
148. **Nyangena, Wilfred** (2006), Essays on Soil Conservation, Social Capital and Technology Adoption
149. **Nivorozhkin, Anton** (2006), Essays on Unemployment Duration and Programme Evaluation
150. **Sandén, Klas** (2006), Essays on the Skill Premium
151. **Deng, Daniel** (2006), Three Essays on Electricity Spot and Financial Derivative Prices at the Nordic Power Exchange
152. **Gebreeyesus, Mulu** (2006), Essays on Firm Turnover, Growth, and Investment Behavior in Ethiopian Manufacturing
153. **Islam, Nizamul Md.** (2006), Essays on Labor Supply and Poverty: A Microeconometric Application
154. **Kjaer, Mats** (2006), Pricing of Some Path-Dependent Options on Equities and Commodities
155. **Shimeles, Abebe** (2006), Essays on Poverty, Risk and Consumption Dynamics in Ethiopia
156. **Larsson, Jan** (2006), Four Essays on Technology, Productivity and Environment
157. **Congdon Fors, Heather** (2006), Essays in Institutional and Development Economics
158. **Akpalu, Wisdom** (2006), Essays on Economics of Natural Resource Management and Experiments
159. **Daruvala, Dinky** (2006), Experimental Studies on Risk, Inequality and Relative Standing
160. **García, Jorge** (2007), Essays on Asymmetric Information and Environmental Regulation through Disclosure
161. **Bezabih, Mintewab** (2007), Essays on Land Lease Markets, Productivity, Biodiversity, and Environmental Variability
162. **Visser, Martine** (2007), Fairness, Reciprocity and Inequality: Experimental Evidence from South Africa
163. **Holm, Louise** (2007), A Non-Stationary Perspective on the European and Swedish Business Cycle
164. **Herbertsson, Alexander** (2007), Pricing Portfolio Credit Derivatives
165. **Johansson, Anders C.** (2007), Essays in Empirical Finance: Volatility, Interdependencies, and Risk in Emerging Markets
166. **Ibáñez Díaz, Marcela** (2007), Social Dilemmas: The Role of Incentives, Norms and Institutions
167. **Ekbom, Anders** (2007), Economic Analysis of Soil Capital, Land Use and Agricultural Production in Kenya
168. **Sjöberg, Pål** (2007), Essays on Performance and Growth in Swedish Banking
169. **Palma Aguirre, Grisha Alexis** (2008), Explaining Earnings and Income Inequality in Chile
170. **Akay, Alpaslan** (2008), Essays on Microeconometrics and Immigrant Assimilation
171. **Carlsson, Evert** (2008), After Work – Investing for Retirement
172. **Munshi, Farzana** (2008), Essays on Globalization and Occupational Wages
173. **Tsakas, Elias** (2008), Essays on Epistemology and Evolutionary Game Theory
174. **Erlandzon, Karl** (2008), Retirement Planning: Portfolio Choice for Long-Term Investors
175. **Lampi, Elina** (2008), Individual Preferences, Choices, and Risk Perceptions – Survey Based Evidence
176. **Mitrut, Andreea** (2008), Four Essays on Interhousehold Transfers and Institutions in Post-Communist Romania
177. **Hansson, Gustav** (2008), Essays on Social Distance, Institutions, and Economic Growth
178. **Zikhali, Precious** (2008), Land Reform, Trust and Natural Resource Management in Africa
179. **Tengstam, Sven** (2008), Essays on Smallholder Diversification, Industry Location, Debt Relief, and Disability and Utility
180. **Boman, Anders** (2009), Geographic Labour Mobility – Causes and Consequences
181. **Qin, Ping** (2009), Risk, Relative Standing and Property Rights: Rural Household Decision-Making in China

182. **Wei, Jiegen** (2009), Essays in Climate Change and Forest Management
183. **Belu, Constantin** (2009), Essays on Efficiency Measurement and Corporate Social Responsibility
184. **Ahlerup, Pelle** (2009), Essays on Conflict, Institutions, and Ethnic Diversity
185. **Quiroga, Miguel** (2009), Microeconomic Policy for Development: Essays on Trade and Environment, Poverty and Education
186. **Zerfu, Daniel** (2010), Essays on Institutions and Economic Outcomes
187. **Wollbrant, Conny** (2010), Self-Control and Altruism
188. **Villegas Palacio, Clara** (2010), Formal and Informal Regulations: Enforcement and Compliance
189. **Maican, Florin** (2010), Essays in Industry Dynamics on Imperfectly Competitive Markets
190. **Jakobsson, Niklas** (2010), Laws, Attitudes and Public Policy
191. **Manescu, Cristiana** (2010), Economic Implications of Corporate Social Responsibility and Responsible Investments
192. **He, Haoran** (2010), Environmental and Behavioral Economics – Applications to China
193. **Andersson, Fredrik W.** (2011), Essays on Social Comparison
194. **Isaksson, Ann-Sofie** (2011), Essays on Institutions, Inequality and Development
195. **Pham, Khanh Nam** (2011), Prosocial Behavior, Social Interaction and Development: Experimental Evidence from Vietnam
196. **Lindskog, Annika** (2011), Essays on Economic Behaviour: HIV/AIDS, Schooling, and Inequality
197. **Kotsadam, Andreas** (2011), Gender, Work, and Attitudes
198. **Alem, Yonas** (2011), Essays on Shocks, Welfare, and Poverty Dynamics: Microeconometric Evidence from Ethiopia
199. **Köksal-Ayhan, Miyase Yesim** (2011), Parallel Trade, Reference Pricing and Competition in the Pharmaceutical Market: Theory and Evidence
200. **Vondolia, Godwin Kofi** (2011), Essays on Natural Resource Economics
201. **Widerberg, Anna** (2011), Essays on Energy and Climate Policy – Green Certificates, Emissions Trading and Electricity Prices
202. **Siba, Eyerusalem** (2011), Essays on Industrial Development and Political Economy of Africa
203. **Orth, Matilda** (2012), Entry, Competition and Productivity in Retail
204. **Nerman, Måns** (2012), Essays on Development: Household Income, Education, and Female Participation and Representation
205. **Wicks, Rick** (2012), The Place of Conventional Economics in a World with Communities and Social Goods
206. **Sato, Yoshihiro** (2012), Dynamic Investment Models, Employment Generation and Productivity – Evidence from Swedish Data
207. **Valsecchi, Michele** (2012), Essays in Political Economy of Development
208. **Teklewold Belayneh, Hailemariam** (2012), Essays on the Economics of Sustainable Agricultural Technologies in Ethiopia
209. **Wagura Ndiritu, Simon** (2013), Essays on Gender Issues, Food Security, and Technology Adoption in East Africa
210. **Ruist, Joakim** (2013), Immigration, Work, and Welfare
211. **Nordén, Anna** (2013), Essays on Behavioral Economics and Policies for Provision of Ecosystem Services
212. **Yang, Xiaojun** (2013), Household Decision Making, Time Preferences, and Positional Concern: Experimental Evidence from Rural China
213. **Bonilla Londoño, Jorge Alexander** (2013), Essays on the Economics of Air Quality Control
214. **Mohlin, Kristina** (2013), Essays on Environmental Taxation and Climate Policy
215. **Medhin, Haileselassie** (2013), The Poor and Their Neighbors: Essays on Behavioral and Experimental Economics
216. **Andersson, Lisa** (2013), Essays on Development and Experimental Economics: Migration, Discrimination and Positional Concerns
217. **Weng, Qian** (2014), Essays on Team Cooperation and Firm Performance
218. **Zhang, Xiao-Bing** (2015), Cooperation and Paradoxes in Climate Economics
219. **Jaime Torres, Monica Marcela** (2015) Essays on Behavioral Economics and Policy Design
220. **Bejenariu, Simona** (2015) Determinants of Health Capital at Birth: Evidence from Policy Interventions
221. **Nguyen, Van Diem** (2015) Essays on Takeovers and Executive Compensation
222. **Tolonen, Anja** (2015) Mining Booms in Africa and Local Welfare Effects: Labor Markets, Women's Empowerment and Criminality
223. **Hassen, Sied** (2015) On the Adoption and Dis-adoption of Household Energy and Farm Technologies
224. **Moursli, Mohamed-Reda** (2015) Corporate Governance and the Design of Board of Directors
225. **Borcan, Oana** (2015) Economic Determinants and Consequences of Political Institutions
226. **Ruhinduka, Remidius Denis** (2015) Essays on Field Experiments and Impact Evaluation
227. **Persson, Emil** (2016) Essays on Behavioral and Experimental Economics, Cooperation, Emotions and Health.
228. **Martinangeli, Andrea** (2017) Bitter divisions: inequality, identity and cooperation
229. **Björk, Lisa** (2017) Essays on Behavioral Economics and Fisheries: Coordination and Cooperation
230. **Chegere, Martin Julius** (2017) Post-Harvest Losses, Intimate Partner Violence and Food Security in Tanzania
231. **Villalobos-Fiatt, Laura** (2017) Essays on forest conservation policies, weather and school attendance
232. **Yi, Yuanyuan** (2017) Incentives and Forest Reform: Evidence from China
233. **Kurz, Verena** (2017) Essays on behavioral economics: Nudges, food consumption and procedural fairness
234. **Yashodha, Yashodha** (2017) Contract Choice and Trust in Informal Groundwater Markets
235. **Meles, Tensay Hadus** (2017) Power Outages, Increasing Block Tariffs and Billing Knowledge
236. **Mühlrad, Hanna** (2018) The Impact of Reproductive and Birth Technologies on Health, Fertility and Labor Outcomes
237. **Gakii Gatua, Josephine** (2018) Primary Health Care Interventions and Social Ties in Kenya
238. **Felgendreher, Simon** (2018) Essays on behavioral economics and the effects of the colonial rule on Java
239. **Demeke, Eyoual** (2019) Essays on Environmental and Behavioral Economics
240. **Kiss, Tamas** (2019) Predictability in Equity Markets: Estimation and Inference
241. **Khomenko, Maksym** (2019) Essays on the Design of Public Policies and .

Regulations

242. **Mukanjari, Samson** (2019) Climate Policy and Financial Markets
243. **Schürz, Simon** (2020) Economic and Intergenerational Decision-Making in Families
244. **Lau, Debbie** (2020) Empirical Essays on Education and Health Policy Evaluation
245. **Ho, Hoang-Anh** (2020) Essays on Culture, Institutions, and Economic Development
246. **Rubio Ramos, Melissa** (2020) The Economics of Coercive Institutions, Conflict, and Development
247. **Lindahl, Anna** (2021) Empirical tests of exchange rate and stock return models
248. **Sjöholm, Carolin** (2021) Essays on public policy in the informal sector context
249. **Tian, Ruijie** (2021) Impacts of Climate Policy and Natural Disasters: Evidence from China
250. **Larsson, Sebastian** (2021) Providers and Profiteers: Essays on Profits and Competition in the Provision of Public Services
251. **Norrgren, Lisa** (2022) Carpe Diem or Seize your Health? The Economics of Time Preferences, Health, and Education
252. **Andersson, Lina** (2022) Emotions in Game Theory: Fear, friendliness and hostility
253. **Enlund, Jakob** (2022) Three essays in applied economics
254. **Berggren, Andrea** (2023) Choices among Doctors, Students and Primary Care Providers: Empirical Evidence from Sweden