Social Spillovers: Experimental Evidence on Peer Effects and Social Change

Leonhard Vollmer

Dissertation
Munich, 2021

SOCIAL SPILLOVERS: EXPERIMENTAL EVIDENCE ON PEER EFFECTS AND SOCIAL CHANGE

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2021

vorgelegt von
Leonhard Vollmer

Referent: Prof. Davide Cantoni, PhD
Korreferent: Prof. Dr. Fabian Kosse
Promotionsabschlussberatung: 2. Februar 2022

Tag der mündlichen Prüfung: 21. Januar 2022
Namen der Berichterstatter: Davide Cantoni, Fabian Kosse, Lukas Buchheim

Für Alice und meine Eltern.
(Darum streiten werden sie sich wohl nicht.)

ACKNOWLEDGMENTS

The completion of this thesis would not have been possible without the support of a great number of people. Foremost, I have been immensely fortunate to have Davide Cantoni as my advisor. From day one Davide encouraged me to explore my ideas freely and to be ambitious; he provided guidance whenever needed and revitalized my spirits, when projects foundered. Thank you, Davide, for your unwaveringly generous guidance and encouragement.

Many thanks are also in order to Fabian Kosse and Lukas Buchheim for serving as my second and third examiners, respectively. While our interactions had to take place predominantly online, they always provided a refreshing variation to my pandemic routine, seamlessly transitioning between insightful discussions about economics and entertaining chitchat. Both of you have impressed me with your intellectual thoroughness and deep knowledge in your respective fields. I have learned a lot from my conversations with you, for which I am very grateful. Equally, I am indebted to Mathias Bühler whose creativity and econometric prowess have invaluably contributed to my personal development as an economist. Thank you, Mathias, for carrying our joint project and for the informal supervision on our many runs through *Englischer Garten*. I would also like to express my gratitude to Jérôme Schäfer for opening many doors and for letting me benefit from his positively encyclopedic knowledge of every conceivable literature in economics and political science.

My PhD was funded by the Elite Network of Bavaria through the Evidence-Based Economics program. I am very grateful both for the curriculum, which taught me a great deal about how to conduct applied economics research, and for the generous financial support, which gave us doctoral students considerable leeway with regards to the type of projects we could pursue. This generous level of support can by no means be taken for granted. Thank you to Florian Englmaier and Joachim Winter for running the program and creating an excellent environment for doctoral students. In fact, I owe Joachim a special debt of gratitude for his support throughout the years, not least for handling the extra work caused by last-minute IRB applications and data protection queries. A huge thank you is also due to Britta Pohr for her help in all administrative matters and especially for her optimism as well as humor in the face of the bureaucratic machinery's many entanglements.

This journey would not have been nearly as enjoyable without the company of great peers: thank you, Fabian, Manfei, Pavel, and Vera, for countless moments of fun in *Office 017*. Thank you, Anne and Sebastian, for being constant sources of counsel and companionship, not least during our extensive field studies of the nighttime economy. Finally, a huge thank you to my very supportive and quick-witted colleagues at *Lehrstuhl Cantoni* – I will dearly miss you and our frequent coffee breaks.

Among them, I would like to express my special gratitude to my co-author Emilio. Your eye for detail, paired with an impressive capacity to organize and prioritize, has been invaluable throughout our joint endeavors. Our shared sense of humor and appreciation of life's pleasures have made the entire process even more enjoyable.

One person I cannot thank enough is Johannes, with whom I had the pleasure of collaborating closely over the past four years. Granted, we got off to a rocky start, when on the first day of our doctoral studies you told me that you did not care for beer and hated the Oktoberfest; but we have come a long way since then. I am enormously grateful for the friendship that we formed and I would not want to miss a single of the innumerable hours spent together devising and carrying out research projects. Our papers would have never got to where they are without your incredible energy and enthusiasm for new endeavors. Thank you, Johannes, for everything you have done for me and for always having my back. I sincerely hope the joint papers in our respective dissertations will not have been the last collaboration of the "Cantoni twins".

I am deeply grateful to my friends and family. I am indebted to them in countless ways, and it is fair to say that they are the key source of energy and optimism in my life. Most of all, however, I would like to express my gratitude to Alice and my parents, to whom I dedicate this thesis. Thank you, Alice, for your heroic stamina in discussing economics with me. Your sharp comments have left their mark on all papers in this dissertation. Thank you for being my intellectual sparring partner, for always believing in me and for bringing joy to me, every day. Finally, my deepest gratitude goes to my parents Jutta and Peter for their unconditional love, support and sacrifice. I can never thank you enough for the opportunities you have afforded me. I owe it all to you.

CONTENTS

INTRODUCTION

This dissertation consists of three self-contained chapters. Although they cover diverse fields, all three are concerned with *social spillovers*, i.e., externalities arising in social interactions. Using (quasi-)experimental evidence – from two settings with explicit randomization and one natural experiment in German history – each chapter sheds light on novel aspects about the role of social spillovers.[1]

Such externalities – frequently referred to as *peer effects* – are viewed as particularly relevant in the human capital accumulation process and have received sustained attention by social scientists and policymakers alike. A famous example is the Coleman report that highlighted peer interactions as a major influence on educational equality and has been widely influential for education policy in the United States (Coleman 1966; Hill 2017). Some authors go as far as arguing that even parents influence their children's outcomes predominantly by affecting the peer groups to which they are exposed (Harris 2009). While this is likely an overstatement, a large body of research in economics – summarized by Sacerdote (2011, 2014) – has shown that student characteristics such as academic achievement or gender do indeed impact the academic outcomes of others substantially. However, our understanding as to which peer characteristics are primarily responsible for social spillovers is still incomplete; comprehending the nature of peer effects thus remains an important research agenda.

CHAPTER 1 aims to contribute to this agenda. I study whether individuals' fundamental economic preferences – their *patience, risk taking,* and *prosociality* – produce externalities on the academic performance of their peers. Exploiting randomly matched study groups I document that collaborating with more patient peers raises student performance, while being matched with more prosocial peers, surprisingly, harms achievement. These effects are comparatively large and distinct from the impact of other

[1] I employ the term *social spillovers* as follows: first, I use it to refer to *peer effects*, i.e., cases in which individuals' backgrounds, behavior, outcomes, or views generate externalities for others in their social environment. The limitation to externalities distinguishes peer effects (and social spillovers) from market-based effects or deliberate attempts at influencing others, e.g., via persuasion, authority, or coercion (Sacerdote 2011). Second, I use social spillovers to describe an analogous process playing out at more aggregate levels, i.e., where peer interactions among one group of people produce externalities for other – possibly large – groups of people, or even entire societies.

peer characteristics including peer achievement and gender. Peer risk taking does not significantly influence outcomes in my setting. As for the mechanisms behind these results, I provide suggestive evidence that more patient peers raise students' effort levels and change effort timing, likely contributing to improved performance. Higher levels of peer altruism, a component of peer prosociality, seem to be responsible for students' waning conviction that academic success is a function of effort. This possibly results in behavioral changes reducing study effectiveness, such as free-riding on altruistic team members' efforts. Given the central role of patience, risk taking, and prosociality in human decision making, it is perhaps no surprise that they would also produce important externalities in people's social environment. To my knowledge, this paper is the first to provide causal evidence that they do.

A prominent body of literature on social pressure suggests that concerns about our *social image* are a key mechanism often responsible for the occurrence of peer effects: in such situations, our mere beliefs about the views of our peers, or about the prevailing social norms, are sufficient to substantially alter our behavior. Recent field experimental evidence – summarized by Bursztyn and Jensen (2017) – documents that social image concerns are indeed a powerful motive when our decisions are observable to others. This holds in a wide range of decision making contexts: students alter educational effort when it is "cool to be smart" (Bursztyn, Egorov, and Jensen 2018), households engage in conspicuous consumption to signal their wealth (Bertrand and Morse 2016), and individuals donate – in part – out of social pressure to appear generous (DellaVigna, List, and Malmendier 2012).

CHAPTER 2, co-authored with Emilio Esguerra and Johannes Wimmer, is inspired by this experimental literature and adds a simple point: when behavior is observable to others, people do not just care about how their actions are perceived (*social image*); in some situations, individuals also anticipate that their own behavior may set an example for others and actively incorporate this into their decision making. Such *anticipated peer effects* may be particularly relevant in situations where our actions generate positive externalities if those around us follow suit: examples range from signing up early for a health screening to nudge our more present-biased friends, to wearing an "I voted" button to signal the importance of voting to others.

We examine the role of *anticipated peer effects* in the context of Germany's COVID-19 vaccination campaign in early 2021, studying individuals' willingness to register for a vaccination with their central and sole vaccination supplier. We find that individuals' willingness to register for the vaccination almost doubles when they realize that they can influence a peer's registration decision. Our contribution consists in separating anticipated peer effects – conceptually and empirically – from social image effects, as distinct motivators of prosocial behavior. Moreover, our experimental setting allows us to rule out other potential confounders such as experimenter demand and strategic lying by participants. Our findings thus suggest that individuals are indeed more inclined to act prosocially if they anticipate that they can lead by example. More generally, anticipated peer effects may be a relevant propagation mechanism for social change: a desire to lead by example may motivate individuals to bear the private costs of prosocial behavior, such as participating in a protest for civil liberties.[2]

Finally, CHAPTER 3, which is the product of joint work with Mathias Bühler and Johannes Wimmer, addresses social change directly. Recent contributions in the economics literature on social change have emphasized the important roles of technology (Garcia-Jimeno, Iglesias, and Yildirim 2020; Melander 2020; Zhuravskaya, Petrova, and Enikolopov 2020) and civic leadership (Dippel and Heblich 2021) for the success of existing social movements. We contribute to this literature by investigating how social movements, and their leaders, emerge in the first place.

To do so, we examine the role of education in the context of Germany's women's rights movement. The movement started to form in the second half of the nineteenth century and subsequently contributed substantially to changing women's role in society, e.g., by successfully contending for female suffrage (achieved in 1919). To isolate the effect of education on the emergence and success of the movement, we study the arrival of *finishing schools*, Germany's first institutions offering secondary education and teacher training to women, established from the seventeenth century onwards. Using newly collected historical microdata, we demonstrate the impact of finishing schools at three key milestones of the movement's history: first, the establishment of finishing schools increased women's representation among Germany's political, intellectual, and economic elite ("human capital elite"), from which the activist nucleus of the women's rights movement later emerged. Second, cities that had established finishing schools by 1850 exhibited stronger support for the women's cause in the movements' early phase, which we measure using letters to the editor of *Frauen-Zeitung*, one of Germany's first feminist newspapers. Third, places with a legacy of finishing schools hosted more and larger women's rights organizations in the early twentieth century, which were pivotal in carrying demands for women's empowerment into society at large. We argue that finishing schools facilitated the exchange of critical ideas about women's role in society and the formation of activist networks. These early activists disseminated critical ideas among a wider public and converted their nascent movement into a successful social force. Ultimately, not only those women directly exposed to finishing schools were affected, but society as a whole. This paper thus suggests that substantial social spillovers can arise from educational institutions that foster the exchange of critical ideas and provide the space to form networks.

[2]Similarly, peer effects themselves are likely catalysts for social change, since they are responsible for one person's behavior being transmitted to the next.

Chapter 1

Peer Preferences and Performance

1.1 Introduction

Economic theory suggests that *patience*, *risk taking*, and *prosociality* – individuals' preferences over the timing of payoffs, risk, and the payoffs of others – are fundamental determinants of behavior. A thriving empirical literature confirms the remarkable relevance of these economic preferences for individuals' life outcomes, including educational attainment, income, and health.[1] Given their central role in decision making, it seems likely that these preferences – or their behavioral corollaries – would also produce important externalities in people's social environment. Although a large body of literature has documented the peer effects associated with individual characteristics such as gender and academic achievement, evidence on whether individuals' fundamental economic preferences produce social spillovers is surprisingly scarce. A notable exception is recent research by Golsteyn, Non, and Zölitz (2021), who have documented a novel aspect of peer effects by identifying externalities arising from specific personality traits. However, since psychological concepts of personality and economic preferences are complements rather than substitutes in explaining human behavior (A. Becker et al. 2012), I start from the premise that they likely also complement each other in terms of the social spillovers they generate.[2]

[1]Falk et al. (2021) provide an extensive overview of the empirical literature. The importance of *patience* for individuals' life trajectory is highlighted, among others, by Åkerlund et al. (2016), Epper et al. (2020), Falk, Kosse, and Pinger (2020), Golsteyn, Grönqvist, and Lindahl (2014), and Mischel, Shoda, and Rodriguez (1989); for *risk taking* see Anderson and Mellor (2008), Dohmen et al. (2011), and Kimball, Sahm, and Shapiro (2008); the relevance of *prosociality* is discussed by Deming (2017), Dohmen et al. (2009), Kosse et al. (2020), and Weidmann and Deming (forthcoming). Falk et al. (2018), Hanushek et al. (2021), and Sunde et al. (forthcoming) document the importance of economic preferences at more aggregate levels.

[2]With the exception of risk taking, Golsteyn, Non, and Zölitz (2021) focus on psychological personality measures rather than economic preferences. Due to the closeness of the topics studied and since their paper represents the state of the art in this literature, I adopt many of its insights. This includes, to a large degree, the structure in which I present my evidence as well as the selection of analyses I conduct.

This paper examines whether *peer preferences* influence students' academic performance, by exploiting the randomized composition of study groups during an introductory economics course at a German university. Several institutional features facilitate the occurrence of peer effects within this setting: first, the course is a compulsory core module and an important milestone for students in their first two semesters, where study behavior is malleable and peers are likely to leave their mark. Second, within study groups of five to six individuals, students collaborate on the course's weekly assignments, implying frequent and substantive peer interactions. Moreover, for the cohorts I study, all teaching took place online due to the COVID-19 pandemic. This reduced the scope for endogenous peer group formation, raising the relative importance of the randomly matched study groups.

I measure fundamental economic preferences of students using the *Preference Survey Module* developed by Falk et al. (2018, forthcoming). To estimate the impact of peer preferences on student performance, I combine the preference measures with administrative data on student grades across two cohorts. For my analysis of underlying mechanisms I draw on measures of peer interaction as well as student expectations and effort, elicited in surveys after student interactions have occurred. In addition, the specific distance learning setting allows me to collect a unique revealed preference measure of student effort, by exploiting data on student attendance in the course's weekly online lectures – students' *Zoom hours*.

Results show that students who were randomly matched with more patient group members improve their exam performance substantially: in my preferred specification, a one standard deviation increase in mean peer patience boosts students' exam grade by 0.11 standard deviations (significant at the 5 percent level). Surprisingly, collaborating with more prosocial peers harms student achievement. I find that a one standard deviation increase in peer prosociality reduces exam performance by 0.11 standard deviations (significant at the 10 percent level). Peer risk taking does not affect performance significantly in my setting. These findings suggest that peer patience and peer prosociality causally impact student performance. However, despite the substantial effect sizes, they come with the caveat of only marginal statistical significance and should thus be corroborated in a follow-up study with pre-specified hypotheses.

Are the observed effects truly due to peer preferences or do they merely pick up the impact of other peer characteristics? To rule out this alternative explanation, I show that the inclusion of peer academic achievement, peer gender and peer persistence does not change the estimated effects of peer patience and peer prosociality. Incidentally, this analysis confirms the result by Golsteyn, Non, and Zölitz (2021) that more persistent students positively impact their peers' academic performance: I find very similar point estimates for peer persistence, which are, however, insignificant due to my smaller sample size. Similarly, my coefficient sizes for peer academic achievement and peer gender are in line with those reported in the literature review by Sacerdote (2011). Together, these results indicate that peer patience and peer prosociality have an impact on student performance that is independent of the externalities arising from other peer characteristics.[3]

To understand why peer preferences influence student performance, I start by investigating to what extent each component of peer prosociality – peer altruism, peer trust, and peer positive reciprocity – drives the results. Peer altruism emerges as the dominant factor: a one standard deviation increase lowers student performance by 0.16 standard deviations, a result significant at the 1 percent level. Next, I examine (1) the effect of peer preferences on study group interactions, (2) whether students' grade expectations are affected, and (3) how various measures of effort respond over time. I find that neither interaction quality and intensity nor the number of assignments that groups handed in are mechanisms linking peer preferences with student performance. Students' grade expectations, however, increase sharply in the presence of patient peers. This indicates that – insofar as expectations are accurate – students learn from their patient peers what investments are needed to succeed academically. Moreover, students assigned to groups with more patient peers invest substantially more hours per week on studying for the course and they are more likely to attend online course meetings in the middle of the semester, thus likely avoiding a common mid-semester effort slump. This suggests that peer patience affects student achievement by changing both effort levels and timing. Moving on to the mechanisms for the effect of peer prosociality yields a mixed pattern: students working with more prosocial peers exhibit a substantially weakened belief that grades are a function of effort, yet they slightly increase their attendance of online lectures. However, more study time does not imply higher study effectiveness. Given that higher peer altruism seems to be responsible for students' waning conviction that academic success is a function of effort, behavioral changes resulting in lower study effectiveness, such as free-riding on altruistic team members' efforts, could play a role.

I further examine mechanism plausibility in hetereogeneity analyses asking which type of students are affected by peer preferences. I document that students in the lower half of the patience distribution profit particularly from collaborating with more patient peers, suggestive of an adoption of more effective study behavior. Impatient and impersistent students are also the ones (negatively) affected by prosocial peers, which is, again, consistent with the interpretation that free-riding on altruistic team members' efforts – which impatient and impersistent students are perhaps more likely

[3]Additional analyses document that coefficient estimates remain virtually unchanged after winsorizing peer patience and peer prosociality, suggesting that the observed relationships are not artifacts of influential outliers.

to engage in – ultimately harms performance.

The key threat to identification in this paper arises from the fact that in one of the two cohorts studied, preferences are elicited after student interactions have taken place. This might lead to simultaneity and selection bias if peers influence each other's survey participation, preferences, or the measurement thereof. I provide several analyses alleviating these concerns: first, in tests of random assignment I document that individuals' preference measures and other predetermined characteristics such as high school grades or gender are not correlated with the preferences of their study group peers. Second, peer preferences do not affect whether I observe exam grades nor whether students participate in surveys, mitigating concerns of selective attrition. Third, using data from the other cohort, where I elicited patience before and after study group interactions, I show that peer patience measured at baseline does not impact students' patience measure elicited at endline. In sum, the evidence suggests that peers do not influence each other's preferences or the measurement thereof and that selection into the estimation sample is of limited concern.

This study contributes to a large literature on peer effects in education. Existing work typically studies how peer characteristics related to previous academic achievement, gender, or race affect student outcomes. Important contributions on peer achievement effects include Booij, Leuven, and Oosterbeek (2016), Carrell, Fullerton, and West (2009), Carrell, Sacerdote, and West (2013), De Giorgi and Pellizzari (2014), Duflo, Dupas, and Kremer (2011), Feld and Zölitz (2017), Whitmore (2005), and Zimmerman (2003). The impacts of peer gender and race have been studied, among others, by Angrist and Lang (2004), Hoxby (2000), Hoxby and Weingarth (2005), Lavy and Schlosser (2011), Oosterbeek and Van Ewijk (2014), and Shan (2021). Much of the accumulated evidence is summarized in the extensive surveys by Sacerdote (2011, 2014) and Winston and Zimmerman (2004).

The paper most closely related, and which has inspired the present one, is Golsteyn, Non, and Zölitz (2021), who were the first to document social spillovers arising from specific personality traits. Their main focus is on *persistence*, a personality trait measuring students' perseverance in attempting to solve challenging problems. Exploiting a similar setting, in which university students are randomly assigned to teaching sections, they show that students reap enduring benefits from being exposed to more persistent peers.

What sets my paper apart from existing studies is that it analyzes the peer effects arising from individuals' fundamental economic preferences, rather than personality, achievement, gender, or race. This paper provides the first causal evidence that economic preferences – or their behavioral corollaries – produce substantial externalities in people's social environment. I show that the effects of peer preferences are complementary to those of peer persistence (and other peer characteristics). This indicates that economic preferences and psychological personality measures are not only complements in explaining peoples' behavior (A. Becker et al. 2012) but also with regards to the externalities they generate.

The article proceeds as follows. Section 1.2 details the institutional setting including study group assignment. Section 1.3 describes the data and estimation sample. Section 1.4 discusses the empirical strategy and assesses the random assignment of study groups. In Section 1.5, I present the paper's main results, followed by a discussion of mechanisms in Section 1.6. Finally, Section 1.7 concludes.

1.2 Setting

I study peer interactions in small study groups of an introductory economics course at a large German university. The course is offered every semester and is compulsory for students majoring in economics, who are expected to take the course in their first semester, as well as for students pursuing an economics minor, who should do so in their second semester. I collect data on two consecutive cohorts: one consisting of economics major (*Semester 1*) and one of economics minor students (*Semester 2*).[4]

In the semesters I study, all teaching took place online due to the COVID-19 pandemic. In each week, students were expected to engage with pre-recorded lectures, solve assignments and participate in online lectures taking place via *Zoom*. The latter had the aim of enriching students' understanding of key course contents through interactive discussions, allowing them to ask questions accumulated over the week, and discussing the solutions to the weekly assignments. All course materials could be accessed via the online learning platform *Moodle*, which also served as the central communication tool between academic staff and students. At the end of each semester, students took an exam which determined their course grade.

To facilitate student interaction and collaboration despite the unusual distance learning conditions, students were offered the opportunity to participate in a study group. These groups consisted of five to six randomly matched students, who were invited to collaborate, online or in person – if official COVID regulations allowed –, on the weekly assignments, which contained exam-style questions. Participation in these groups was voluntary, but strongly encouraged by the teaching team and incentivized via prizes awarded to the best groups.

Figure 1.1 shows a timeline of events and illustrates several differences in study group allocation and data collection between the two semesters: first, in Semester 1, study

[4]The academic year at German universities is divided into a winter semester and a summer semester, each comprising ca. 13 weeks of teaching.

groups were re-shuffled every four weeks, implying that students could be part of up to three groups, while in Semester 2, students remained in a single group throughout. When analysing Semester 1 data, I focus on peer preference variation from group phase 1, i.e. from students' interactions with their peers during the – arguably very formative – first four weeks of their student life.[5]

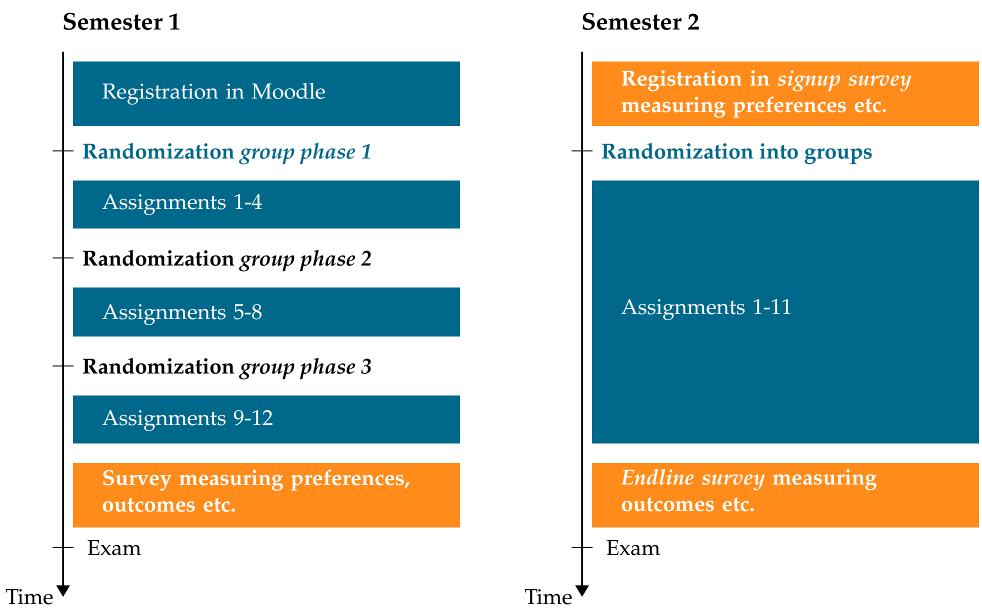


Figure 1.1: Semester timelines

Second, the determination of study group participation differed between the semesters: in Semester 1, students had to actively opt out of study group participation – all students who had registered for the course in *Moodle* were assigned to a group, unless they asked the teaching team to be excluded. In contrast, in Semester 2, participation in a study group was dependent on actively registering in a *signup survey*.

Third, in Semester 2, the *signup survey* measured student's economic preferences and baseline characteristics at the beginning of the semester, i.e. before any student interaction took place, while an *endline survey* administered before students learned their grade collected outcomes and data on potential mechanisms. In Semester 1, on the other hand, the same set of economic preferences, outcomes, and potential mechanisms was measured in a single survey at the end of the semester.

The measurement of student preferences after student interactions have taken place (as in Semester 1) is cause for concern if it gives rise to simultaneity or selection bias. I address these concerns in various sections throughout the paper: in section 1.4, I test for random assignment and document that individuals' preference measures and

[5] I opt against using peer preferences averages from all three group phases since the implied mean reversion reduces much of the variation in peer preferences key for answering my research questions.

other predetermined characteristics are not correlated with the preferences of their peers in the study groups. Section 1.C in the Appendix provides further evidence of the stability of economic preferences in young adults.

In sum, the present setting exhibits several features that facilitate the occurrence of peer effects within the randomly formed study groups: for instance, the lack of in-person teaching due to the pandemic left less room for endogenous peer group formation than usual. Moreover, the sample consists of students at the very start of their university career, where study behavior in the new higher education context is still malleable and peers are more likely to leave their mark.

1.3 Data

1.3.1 Sample and data overview

In Table 1.1, I summarize how the paper's estimation sample is derived, separately for each of the two cohorts. On first look, the figures seem to suggest very large dropout rates; however, these are largely an artifact of several university regulations, which make it difficult to ascertain at the beginning of each semester how many students are truly determined to participate in the course: there is no official course registration list and joining the course in Moodle is possible for every student enrolled at the university. Since students can retake the exam several times, many potential re-takers join the Moodle course at the beginning of the semester with the intention to improve their grade but become inactive after a few weeks.

Table 1.1: Data overview

	N individuals	N groups
Semester 1		
Randomized in group phase 1	400	82
Registered for exam	447	82
Exam taken	313	82
Completed survey	157	70
Exam taken, completed survey, and peer preferences available	116	49
Semester 2		
Randomized after taking signup survey	122	24
Registered for exam	207	24
Exam taken	135	24
Completed endline survey	56	20
Exam taken, completed endline survey, and peer preferences available	36	20

Given these circumstances, all students registered in Moodle at the beginning of Semester 1 were informed that they could opt out of being assigned to a study group. The remaining 400 students, who did not opt out, were randomly allocated to 82 study groups.[6] Semester 1 saw a total of 447 exam registrations – which have to be logged relatively early during the semester and equally include many students later turning inactive – and 313 students took the exam in the end. A total of 157 students – i.e., ca. 50% of those who took the exam – completed the online survey administered at the end of the semester (but before the release of grades), in which economic preferences, outcomes, and potential mechanisms were elicited. Ultimately, the Semester 1 cohort contributes 116 observations to my main estimation sample; it contains those students, which (i) took the exam, (ii) completed the survey at the end of the semester, and (iii) for whom I can construct measures of peer preferences. The latter condition is dependent on at least one of their study group members from group phase 1 also completing the survey.

In Semester 2, 122 students were randomly allocated to one of 24 study groups. In the study group *signup* survey almost all students also completed the elicitation of economic preferences, allowing me to construct measures of peer preferences for all 122 students. 135 students took the exam. For 36 of these, I could also collect data on potential mechanisms in the *endline* survey. These 36 observations add to the 116 from Semester 1, resulting in a final estimation sample of 152 students.[7]

Table 1.2 reports summary statistics for this estimation sample. Panel A contains summary statistics of *individual characteristics* including students' own economic preferences, previous educational achievements, and demographics. Panel B, on the other hand, summarizes the *peer characteristics* encounted by students in their study groups, including those capturing their peers' economic preferences *peer patience*, *peer risk taking*, and *peer prosociality*, which are the main focus of this paper. Finally, Panel C describes outcomes and potential mechanisms: student performance, effort, study group interaction, and beliefs.

1.3.2 Measurement of economic preferences and personality

To measure fundamental economic preferences at the individual level, I draw on the *Preference Survey Module* developed by Falk et al. (2018, forthcoming). The module offers an elicitation of economic preferences based on twelve survey items, which are validated against behavior in standard incentivized choice experiments and predict real-world behavior.

[6]Despite the call to opt out to avoid inactive group members, there was a non-negligible amount of inactive students among the 400 assigned to a study group.

[7]I use larger samples to check for selective attrition with respect to taking the exam and participating in the surveys in Section1.4.

Table 1.2: Summary statistics

	N	Mean	SD	5^{th} pctl.	95^{th} pctl.
A. Individual characteristics					
Preferences/personality					
Persistence	152	5.40	1.01	3.50	6.86
Patience	146	0.00	1.00	-1.79	1.22
Risk taking	136	0.00	1.00	-1.76	1.58
Prosociality	149	0.00	1.00	-1.54	1.56
Altruism	152	0.00	1.00	-1.66	1.60
Trust	152	0.00	1.00	-1.54	1.53
Positive reciprocity	149	0.00	1.00	-1.71	1.39
Negative reciprocity	152	0.00	1.00	-1.30	1.75
Education					
High school grade	152	2.06	0.59	1.00	2.94
High school math grade	152	10.07	3.06	5.00	15.00
Exam re-taken	152	0.05	0.22	0.00	0.45
Demographics					
Female	152	0.38	0.49	0.00	1.00
Year of birth	152	1999.73	3.68	1996.00	2002.00
Born in DE	152	0.87	0.34	0.00	1.00
Migration background	152	0.41	0.49	0.00	1.00
B. Group characteristics					
Peer preferences/personality					
Peer persistence	152	5.33	0.76	4.07	6.50
Peer patience	152	0.00	1.00	-1.86	1.34
Peer risk taking	152	0.00	1.00	-1.82	1.50
Peer prosociality	152	0.00	1.00	-1.53	1.51
Peer altruism	152	0.00	1.00	-1.58	1.56
Peer trust	152	0.00	1.00	-1.50	1.60
Peer positive reciprocity	152	0.00	1.00	-1.70	1.40
Peer negative reciprocity	152	0.00	1.00	-1.51	1.54
Peer education					
Peer high school grade	152	2.10	0.44	1.30	2.80
Peer high school math grade	152	9.99	2.38	6.00	14.00
Peer share exam re-taken	152	0.13	0.19	0.00	0.50
Peer demographics					
Peer share female	152	0.39	0.25	0.00	0.75
Peer birthyear	152	1999.71	2.17	1997.00	2002.00
Peer share born in DE	152	0.84	0.25	0.33	1.00
Peer share migration background	145	0.43	0.39	0.00	1.00
C. Student performance, effort, and further survey measures					
Student performance and effort					
Exam taken	152	1.00	0.00	1.00	1.00
Exam performance (grade)	152	3.10	1.26	1.30	5.00
Exam points	152	60.12	23.36	20.27	97.95
Assignments handed in	152	8.64	3.38	0.55	12.00
Zoom hours	152	13.91	8.97	0.36	26.15
Weekly study hours (self-reported)	152	8.97	4.83	3.55	16.45
Evaluation of peer interaction					
"My study group has functioned well."	148	3.91	1.26	1.00	5.00
"Collaborating in study groups aided my understanding of the subject matter."	147	3.78	1.09	2.00	5.00
"I have prepared/I plan to prepare for the exam with my group members."	152	0.16	0.37	0.00	1.00
"I have also exchanged privately with my group members."	152	0.57	0.50	0.00	1.00
Expectations and beliefs					
Grade expectation	152	2.77	0.87	1.70	4.45
Grades depend on effort	152	6.30	0.81	5.00	7.00

Table 1.3 gives an overview of how the fundamental economic preferences *patience*, *risk taking*, *altruism*, *trust*, and *positive reciprocity* are elicited. In my main analyses below, I follow Falk et al. (2018) and combine *altruism*, *trust*, and *positive reciprocity* into a single measure of *prosociality*.[8]

To benchmark my results against the recent findings of positive externalities arising from *peer persistence* in a similar setting by Golsteyn, Non, and Zölitz (2021), I also elicit their measure of *persistence*, developed by Martin (2009), which is also described in Table 1.3.

Table 1.3: Measurement of economic preferences and personality

Preference	Item description
Patience	Intertemporal choice sequence using staircase method Self-assessment: willingness to wait
Risk taking	Lottery choice sequence using staircase method Self-assessment: willingness to take risks in general
Altruism	Donation decision Self-assessment: willingness to give to good causes
Trust	Self-assessment: people have only the best intentions
Positive reciprocity	Gift in exchange for help Self-assessment: willingness to return a favor
Persistence	Self-assessment: willingness to keep trying to solve a challenging problem (4 items)

Based on Falk et al. (2018, p. 1653) and Golsteyn, Non, and Zölitz (2021, p. 1063).

1.3.3 Student performance and potential mechanisms

The main outcome of interest in this paper is students' *exam performance*, i.e. the grade achieved in the final exam at the end of the semester.[9] Panel C of Table 1.2 shows an average exam grade of 3.1 with a standard deviation of 1.26.[10] For ease of interpretation in the analyses below, I standardize the exam grade and refer to it as *exam performance*, where a higher value implies a better performance.

The present setting allows me to collect a unique revealed-preference measure of student effort: I aggregate each student's time spent in all Zoom lectures of the semester in the variable *Zoom hours*. Since there was no in-person teaching due to the pandemic in both cohorts studied, all synchronous teaching took place via Zoom. Moreover, these Zoom live sessions covered contents particularly relevant for students' exam

[8]Specifically, prosociality is computed as the first principal component of altruism, trust and positive reciprocity.

[9]I have access to the full administrative dataset of exam grades in both cohorts and thus observe this measure for all students who took the exam, not just those who completed the surveys.

[10]The German grading scale ranges from 1 (excellent) to 5 (fail). The lowest passing grade is 4.

performance, not least because they provided the only opportunity for students to have their questions answered in real time. Additionally, the online sessions were used to present solutions for the weekly assignments containing exam-style questions.

Further measures used in the analysis of potential mechanisms are derived from the surveys administered among students at the end of each semester, but before students learn their grade.[11] First, I also collect a self-reported effort measure, namely average weekly *study hours*. Second, following Golsteyn, Non, and Zölitz (2021), I create a *peer interaction index* based on the survey items (i) "My study group has functioned well.", (ii) "Collaborating in study groups aided my understanding of the subject matter.", (iii) "I have prepared/I plan to prepare for the exam with my group members.", and (iv) "I have also exchanged privately with my group members." I construct this index by standardizing the responses to each of the four questions, computing their average, and standardizing again. Finally, I also elicit students' *grade expectation* using the German grading scale described in the footnote above as well as students' belief to what extent they think that *grades depend on effort*.

1.4 Empirical strategy and balancing

1.4.1 Empirical strategy

The main goal of this paper is to identify whether students' academic performance is influenced by their peers' fundamental economic preferences. The peer interactions I exploit arise in randomly matched study groups in an introductory economics course. I estimate the impact of peer preferences on performance using the following regression model:[12]

$$y_{igs} = \overline{\textbf{peer preferences}}_{g-i}\, \boldsymbol{\beta}' + \boldsymbol{X_i}\, \boldsymbol{\gamma}' + \rho_s + \varepsilon_{igs} \tag{1.1}$$

y_{igs} is the exam performance of student i in group g and semester s. The vector $\overline{\textbf{peer preferences}}_{g-i}$ contains the *leave-out mean* economic preferences of students in group g, i.e. the mean preferences in each study group calculated *excluding* student i. $\boldsymbol{X_i}$ is a vector of predetermined individual controls, employed to reduce the standard errors of the main estimates of interest, while ρ_s represents semester fixed effects.[13]

[11] For the technical implementation of the surveys I use the open-source software oTree (Chen, Schonger, and Wickens 2016).

[12] This operationalization of the peer group setting closely follows Golsteyn, Non, and Zölitz (2021).

[13] The individual-level controls are: high school grade, high school math grade, female, year of birth, born in DE, and migration background. I opt against controlling for student's own economic preferences in the main specifications as this would reduce the effective sample size further; Table 1.B.2 in the Appendix shows that results do not change substantially when individual-level economic preferences are included. Additional specifications also control for *group*-level characteristics to test to what extent

Following Abadie et al. (2017), I cluster standard errors at the level of treatment assignment, i.e., at the group level.

To investigate potential mechanisms underlying the effects of peer preferences on performance I estimate versions of Equation 1.1, varying the independent variable y_{igs} (see Section 1.6). I shed further light on mechanisms and test which types of individuals are affected by peer preferences in heterogeneity analyses, by dividing the estimation sample according to individuals' predetermined characteristics such as their own preferences, previous educational achievement and gender (see Section 1.6.5). To ease interpretation and comparison of estimates, I standardize variables, where appropriate, to have mean zero and standard deviation one.

1.4.2 Balancing

Identifying the peer effects of interest requires that the key peer preferences – peer patience, peer risk taking, and peer prosociality – are not systematically associated with other potential determinants of student performance. Table 1.4 documents that peer preferences are indeed unrelated to student characteristics: only two out of 66 estimates reported are significant at the 5 or 10 percent level, suggesting that assignment to study groups was indeed random and that potential selection into the identifying sample is of limited concern.[14] Importantly, neither of the three peer preference variables are systematically related with students' own patience, previous educational achievement or gender – which are the individual characteristics most strongly associated with exam performance in the present setting.[15]

This test of random assignment follows Guryan, Kroft, and Notowidigdo (2009), by additionally controlling for the *semester*-level leave-out mean of the respective peer preference to correct for the mechanical relationship between own and peer preferences, which is a consequence of the sampling of peers without replacement from a finite set.

The present balancing results also help allay concerns regarding simultaneity and selection bias. As section 1.2 explained, in Semester 1, students' economic preferences were elicited in a survey at the end of the semester, i.e. after student interactions had already taken place. This ex-post measurement of preferences poses a threat to identification, e.g., if higher ability students somehow influenced how their group members

the peer economic preferences of interest pick up underlying effects from other peer characteristics (see Sections 1.4.3 and 1.5.2).

[14] To further minimize the risk that pre-existing individual differences drive my results, I control for the two variables that show significant differences: I control for year of birth as part of the *individual controls* in my main specification and I test for the impact of negative reciprocity in Table 1.B.2 in the Appendix.

[15] Table 1.B.1 in the Appendix reports which of the students' own characteristics predict exam performance best.

responded to the survey items eliciting economic preferences or which type of group members completed the surveys in the first place. In the presence of such a simultaneity bias, we would expect individuals' preference measures and predetermined characteristics to be systematically associated with the preferences of their peers in the study groups; yet, Table 1.4 documents the opposite. In Appendix Section 1.C, I provide further evidence of the stability of economic preferences in young adults, using the Semester 2 sample, where I elicited preferences both before and after study group interaction.

Table 1.4: Balancing of individual characteristics

Variable				Peer patience		Peer risk-taking		Peer prosociality	
	N	Mean	SD	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value
Preferences/personality									
Persistence	152	0.00	1.00	0.01	0.88	0.09	0.22	-0.06	0.43
Patience	146	0.00	1.00	0.10	0.28	-0.06	0.49	0.03	0.73
Risk taking	136	0.00	1.00	0.00	0.96	-0.07	0.50	-0.04	0.68
Prosociality	149	0.00	1.00	-0.05	0.60	-0.10	0.24	-0.04	0.66
Altruism	152	0.00	1.00	-0.02	0.84	-0.05	0.51	0.05	0.58
Trust	152	0.00	1.00	-0.04	0.62	-0.12	0.14	-0.04	0.67
Positive reciprocity	149	0.00	1.00	0.00	0.98	0.07	0.46	-0.12	0.17
Negative reciprocity	152	0.00	1.00	0.00	0.96	0.14	0.16	0.17^{**}	0.02
Education									
High school grade	152	0.00	1.00	0.00	0.96	0.00	0.98	0.02	0.83
High school math grade	152	0.00	1.00	-0.04	0.59	0.00	0.97	-0.01	0.89
Exam re-taken	152	0.05	0.22	0.02	0.31	0.02	0.15	0.00	0.92
Demographics									
Female	152	0.38	0.49	-0.01	0.88	0.00	0.90	0.03	0.46
Year of birth	152	1999.73	3.68	-0.28^{*}	0.05	0.20	0.22	0.01	0.92
Born in DE	152	0.87	0.34	-0.04	0.12	-0.01	0.87	0.03	0.24
Migration background	152	0.41	0.49	-0.02	0.65	-0.06	0.14	0.02	0.57
Family background									
Father schooling	147	0.00	1.00	0.00	0.99	0.02	0.83	-0.04	0.70
Mother schooling	147	0.00	1.00	0.00	0.97	-0.03	0.70	0.08	0.30
Family support (EUR/month)	146	497.01	489.20	-5.55	0.88	-43.93	0.25	-21.92	0.57
Federal grant (EUR/month)	151	35.21	137.20	6.30	0.53	-7.87	0.50	-3.20	0.76
Selective attrition									
Exam taken	406	0.64	0.48	-0.01	0.76	-0.01	0.75	-0.01	0.79
Completed survey (Semester 1)	284	0.35	0.48	-0.08	0.15	0.02	0.78	-0.05	0.12
Completed endline survey (Semester 2)	122	0.22	0.42	-0.03	0.76	0.04	0.68	-0.06	0.38

The table reports β-coefficients from a regression of the form $\text{Variable}_{igs} = \alpha + \beta \cdot \overline{\text{peer preference}}_{g-i} + \gamma \cdot \overline{\text{peer preference}}_{s-i} + \varepsilon_{igs}$, where $\overline{\text{peer preference}}_{g-i}$ is the *group*-level leave-out mean for patience, risk taking, or prosociality. In the spirit of Guryan, Kroft, and Notowidigdo (2009) these randomization checks control for the *semester*-level leave-out mean of the respective peer preference ($\overline{\text{peer preference}}_{s-i}$) to correct for the mechanical relationship between own and peer preferences. P-values are robust and clustered at the group-level. All peer preference variables are standardized.

A further requirement for identification is the absence of selective attrition. Selective attrition would lead to biased estimates in my setting, e.g., if peer preferences affected whether high or low ability students took the exam or completed the surveys in the first place. The last three lines of Table 1.4 document that peer preferences do not seem to affect whether I observe exam grades, or whether students complete the surveys eliciting the measures used in my analysis of mechanisms. These findings indicate that selective attrition is unlikely to confound the paper's results.

1.4.3 Peer preferences and other peer characteristics

Since economic preferences are correlated with other individual characteristics, a natural worry is whether regressions based on Equation 1.1 indeed identify the effects of peer preferences on student performance or whether the former just proxy for other peer characteristics such as peer math ability. Table 1.5 reports balancing results of group characteristics, summarizing how they are related with the key peer preferences studied in this paper: groups of more patient peers also exhibit significantly higher average levels of math achievement, which is unsurprising given that patience and math ability are strongly correlated at the individual level. Risk-loving peer groups are characterized by higher levels of negative reciprocity and have lower shares of women as well as of students with a migration background. Finally, more prosocial peer groups have higher average levels of persistence and higher shares of women.[16]

Table 1.5: Balancing of group characteristics

Variable				Peer patience		Peer risk-taking		Peer prosociality	
	N	Mean	SD	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value
Peer preferences/personality									
Peer persistence	152	0.00	1.00	0.12	0.16	0.05	0.57	0.25***	0.00
Peer patience	152	0.00	1.00			-0.09	0.34	0.10	0.25
Peer risk taking	152	0.00	1.00	-0.09	0.34			-0.03	0.77
Peer prosociality	152	0.00	1.00	0.10	0.24	-0.03	0.76		
Peer altruism	152	0.00	1.00	-0.02	0.88	-0.10	0.24	0.62***	0.00
Peer trust	152	0.00	1.00	0.12	0.17	0.02	0.84	0.90***	0.00
Peer positive reciprocity	152	0.00	1.00	0.16	0.15	0.02	0.89	0.44***	0.00
Peer negative reciprocity	152	0.00	1.00	-0.09	0.42	0.22***	0.00	0.16	0.12
Peer education									
Peer high school grade	152	0.00	1.00	0.14	0.15	-0.02	0.81	0.08	0.42
Peer high school math grade	152	0.00	1.00	0.20**	0.02	0.03	0.66	0.06	0.48
Peer share exam re-taken	152	0.00	1.00	0.04	0.59	-0.04	0.52	0.02	0.77
Peer demographics									
Peer share female	152	0.00	1.00	-0.01	0.91	-0.29***	0.00	0.18*	0.05
Peer birthyear	152	1999.71	2.17	-0.34*	0.05	-0.01	0.93	-0.15	0.34
Peer share born in DE	152	0.00	1.00	0.08	0.45	-0.06	0.52	0.14	0.22
Peer share migration background	145	0.00	1.00	-0.10	0.24	-0.19**	0.02	0.00	1.00

See notes of Table 1.4.

[16] Since prosociality is computed as the first principal component of altruism, trust, and positive reciprocity, the significant relationships between peer prosociality and the three respective peer preferences in Table 1.5 are purely mechanical.

Given these patterns, I assess the extent to which peer preferences merely pick up the effects of other peer characteristics by estimating versions of Equation 1.1, in which I control for several candidates of the latter, including previous academic achievement (high school grades), gender, and persistence (Golsteyn, Non, and Zölitz 2021). The results of this exercise are discussed in Section 1.5.2 and suggest that peer patience and peer prosociality do indeed have an independent and sizeable impact on student performance. Moreover, given that the economic preferences are likely measured with a considerable amount of error – as previously discussed by Falk et al. (2021) – effect sizes reported in the following sections likely suffer from attenuation bias and represent lower bounds.

1.5 Results

1.5.1 The impact of peer preferences on performance

Table 1.6 presents the main results, which indicate that peer preferences influence student performance: a one standard deviation increase in peer patience improves exam performance by 9-11% of a standard deviation.

Table 1.6: Peer preferences and exam performance

	Exam performance							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Peer patience	0.09	0.10^{*}					0.10	0.11^{**}
	(0.07)	(0.06)					(0.07)	(0.06)
Peer risk taking			-0.02	-0.02			-0.01	-0.01
			(0.08)	(0.06)			(0.08)	(0.06)
Peer prosociality					-0.10^{*}	-0.10^{*}	-0.11^{*}	-0.11^{*}
					(0.06)	(0.06)	(0.06)	(0.06)
Mean outcome	0	0	0	0	0	0	0	0
SD outcome	1	1	1	1	1	1	1	1
No. clusters	69	69	69	69	69	69	69	69
Joint sign. (p-value)							0.31	0.14
Clustered	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes
Observations	152	152	152	152	152	152	152	152
R^2	0.10	0.42	0.09	0.41	0.10	0.42	0.11	0.43

Joint sign. (p-value) is from an F-test for joint significance of peer patience, peer risk taking and peer prosociality. *Individual controls:* high school grade, high school math grade, gender, year of birth, born in DE, and migration background. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Being randomly matched with more prosocial peers, on the other hand, has a negative impact: a one standard deviation increase in peer prosociality is associated with a reduction in performance by 10-11% of a standard deviation. Coefficient estimates on peer risk taking are an order of magnitude smaller than those on peer patience and peer prosociality and are not statistically significant.

The impact of peer preferences is unchanged when I include individual-level controls and when I estimate the impact of all three peer preferences jointly, both attesting to coefficient stability: the inclusion of individual controls in columns 2, 4, 6, and 8 increases precision but does not alter coefficient estimates, as expected under random assignment. The point estimates do not change in columns 7 and 8 either, where all three peer preferences measures enter the model simultaneously, indicating mutually distinct impacts on student performance. Section 1.5.2 provides further evidence of coefficient stability, by documenting that coefficient estimates remain similar when I additionally control for other peer characteristics.

While the estimated effect sizes are substantial, an important caveat is that the impacts of both peer patience and peer prosociality are only marginally significant: the coefficient on peer persistence, e.g., is significant at the 5 percent level in my preferred specification in column 8 of Table 1.6, but only significant at the 10 percent level in others. Correspondingly, an F-test for the joint significance of peer patience, peer risk taking and peer prosociality in this specification has an implied p-value of 0.14.[17] Further, it is important to note that the standard errors reported here do not correct for multiple comparisons. Given the small sample of this study and the fundamental conceptual uncertainties of how to adequately correct for multiple testing, I refrain from further investigating the multiple testing problem here and argue – in the spirit of Althouse (2016) – that a follow-up study with pre-specified hypotheses would be the most fruitful avenue for corroborating the present results.

Figure 1.2 illustrates the main findings and provides an informal assessment to what extent outliers might be driving them. The three subfigures plot residualized exam performance against the respective peer preference measure and thus depict the regression results reported in column 8 of Table 1.6. To more formally account for potentially influential outliers, I repeat the analysis with winsorized preference measures, where I replace all values below the 5^{th} percentile and above the 95^{th} percentile with the level of the 5^{th} percentile and the 95^{th} percentile, respectively. Table 1.B.3 and Figure 1.A.1 in the Appendix document that coefficient estimates remain virtually unchanged after winsorization, suggesting that the observed relationships are not

[17] In other specifications, the p-value of this type of F-test is as low as 0.05, e.g. in column 2 of Table 1.B.2 in the Appendix, where I also control for students' own patience, which is a strong predictor of exam performance and thus increases precision considerably. The p-value for the joint significance of peer patience and peer prosociality (excluding peer risk taking) in column 8 of Table 1.6 is 0.07.

artifacts of influential outliers.

In sum, the estimates presented in Table 1.6 and Figure 1.2 suggest that peer patience and peer prosociality causally impact student performance, and that the effects are distinct from one another. However, the abovementioned caveat of marginal statistical significance applies.

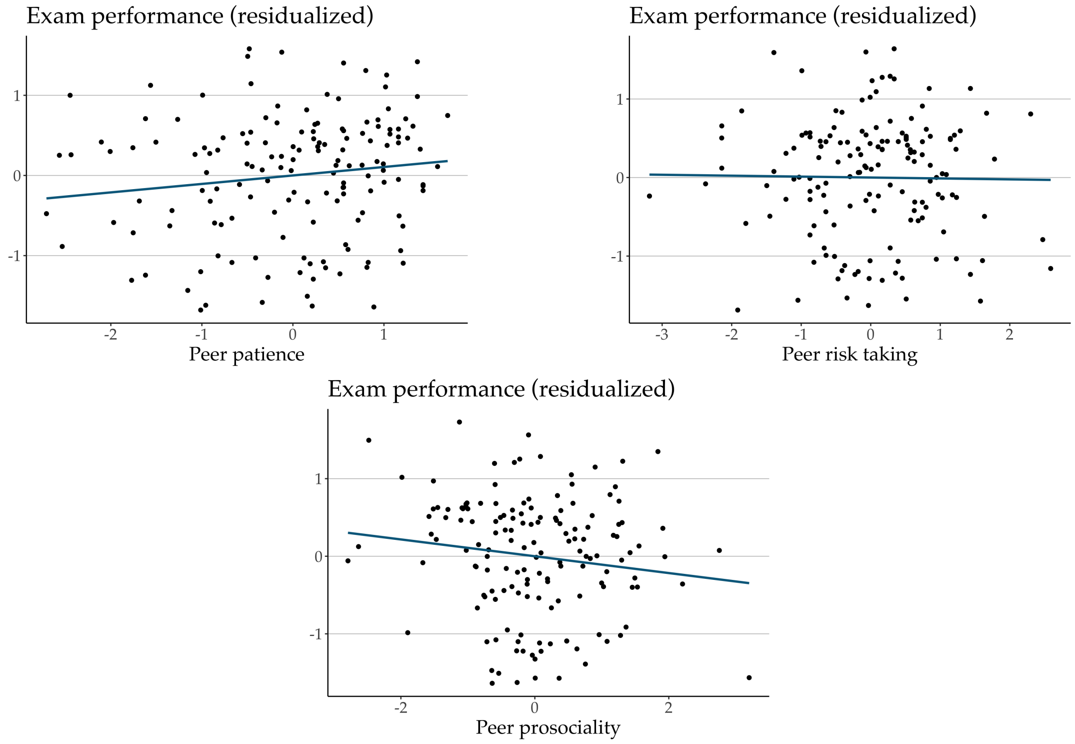


Figure 1.2: Peer preferences and exam performance

Scatter plots corresponding to regression results in column 8 of Table 1.6. All 152 data points shown.

1.5.2 Comparison to other peer effects

Next, in Table 1.7, I investigate whether the observed effects are a product of peer preferences or other peer characteristics. I focus on characteristics which have been identified as relevant in the existing literature on peer effects in education, among them academic achievement, gender, and persistence.

Columns 2 and 4 report coefficient estimates that are very similar to those found by Golsteyn, Non, and Zölitz (2021), whose study encompasses peer effects of risk attitudes as well as persistence, and who were the first to document peer effects arising from specific personality traits. They find that a one standard deviation increase in peer persistence improves performance by 1.9% of a standard deviation (here: perfor-

Table 1.7: Comparison to other peer effects

	Exam performance							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Peer patience	0.10*							0.11*
	(0.06)							(0.06)
Peer risk taking		−0.02						−0.01
		(0.06)						(0.06)
Peer prosociality			−0.10*					−0.13**
			(0.06)					(0.06)
Peer persistence				0.03				0.06
				(0.07)				(0.07)
Peer high school grade					−0.01			−0.07
					(0.05)			(0.07)
Peer high school math grade						0.04		0.06
						(0.06)		(0.09)
Peer share female							0.03	0.05
							(0.06)	(0.06)
Mean outcome	0	0	0	0	0	0	0	0
SD outcome	1	1	1	1	1	1	1	1
No. clusters	69	69	69	69	69	69	69	69
Joint sign. (p-value)								0.11
Clustered	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	152	152	152	152	152	152	152	152
R^2	0.42	0.41	0.42	0.41	0.41	0.41	0.41	0.44

All variables are standardized to have mean zero and standard deviation one. *Joint sign. (p-value)* is from an F-test for joint significance of peer patience, peer risk taking and peer prosociality. *Individual controls:* high school grade, high school math grade, gender, year of birth, born in DE, and migration background. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

mance improvement by 3-6%), while a one standard deviation increase in peer risk tolerance lowers performance by 1.1% of a standard deviation (here: performance reduction by 1-2%). Similarly, coefficient sizes on peers' previous academic achievement and the female share are in line with those reported in the review article by Sacerdote (2011).

Compared to these previously studied peer effects, coefficient sizes on peer patience and peer prosociality in Table 1.7 are substantially larger. Notably, the estimates for peer patience and peer prosociality change only marginally when I include all peer characteristics simultaneously in one model (column 8); and they are the only ones that are significantly different from zero, both in the separate estimations (columns 1-7) and in the joint estimation (column 8).[18]

Together, these results suggest that peer patience and peer prosociality have an impact on student performance that is independent of the externalities arising from other peer characteristics. A further interesting implication is that peer patience and peer persistence seem to capture distinct peer characteristics. This is mirrored by regressions of

[18] Figure 1.A.2 in the Appendix provides a visualization of the results presented in Table 1.7.

exam performance on students' own characteristics, where patience and persistence also exhibit separate effects (see Table 1.B.1 in the Appendix).

1.6 Mechanisms

This section investigates why peer preferences influence student performance. I begin by decomposing the effect of peer prosociality, zooming in on the constituent parts of the prosociality measure: altruism, trust, and positive reciprocity. Next, I investigate study group interactions, student expectations, and student effort to find potential behavioral changes on the part of the affected students, that link variation in peer preferences to diverging exam performance. A final subsection studies effect heterogeneity to further assess mechanism plausibility.

1.6.1 Components of peer prosociality

Perhaps the most surprising result of the previous section is the negative impact of peer prosociality on students' exam performance. To understand the underlying mechanisms, I start by investigating to what extent each component of peer prosociality – peer altruism, peer trust, and peer positive reciprocity – contributes to the relationship. Separate and joint estimations of the components' impact in Table 1.8 reveal peer altruism as the dominant factor: a one standard deviation increase in peer altruism decreases exam performance by 0.16 standard deviations, a result significant at the 1 percent level. In contrast, peer trust and peer positive reciprocity have statistically insignificant and substantially smaller implications for student performance. These results give us some clues as to the nature of potential mechanisms linking peer prosociality and exam performance: they must be plausibly related to peer altruism – or its behavioral correlates – as well as manifest in changes of students' own behavior.

1.6.2 Peer interaction

Peer preferences might influence performance by shaping the interactions within study groups. However, column 1 of Table 1.9 shows that neither of the three peer preferences substantially impacts perceptions of peer interaction quality and intensity, as measured by the *peer interaction index*. This finding suggests that interaction quality and intensity do not drive the effect of peer preferences on exam performance. In addition, column 2 of Table 1.9 documents that a potential implication of peer interaction quality and intensity, namely the number of assignments handed in, is not responsible for the observed effects.

Table 1.8: Components of peer prosociality

	Exam performance				
	(1)	(2)	(3)	(4)	(5)
Peer patience	0.11^{**}	0.10^{*}	0.11^{*}	0.10^{*}	0.09
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Peer risk taking	−0.01	−0.02	−0.01	−0.01	−0.02
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Peer prosociality	-0.11^{*}				
	(0.06)				
Peer altruism		-0.16^{***}			-0.16^{***}
		(0.05)			(0.05)
Peer trust			−0.07		−0.03
			(0.07)		(0.07)
Peer positive reciprocity				−0.01	0.05
				(0.06)	(0.06)
Mean outcome	0	0	0	0	0
SD outcome	1	1	1	1	1
No. clusters	69	69	69	69	69
Clustered	Yes	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Observations	152	152	152	152	152
R^2	0.43	0.44	0.43	0.42	0.45

All variables are standardized to have mean zero and standard deviation one. *Individual controls:* high school grade, high school math grade, gender, year of birth, born in DE, and migration background. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

1.6.3 Expectations

Column 3 of Table 1.9 documents that being matched with more patient peers strongly raises students' expectations about their exam grade: a one standard deviation improvement in peer patience increases students' grade expectations by 20% of a standard deviation. Changes in grade expectations add to our understanding of mechanisms since they are informative of students' confidence in their ability to perform well. To the extent that this confidence is warranted, this finding suggests that peer patience may have raised students' understanding which investments are needed to reap the reward of a good exam grade. Peer prosociality, in contrast, has no impact on expectations.

1.6.4 Effort

Peer preferences might influence student outcomes if students adapt their level of effort in response to being exposed to more patient or prosocial peers. This section examines how peer patience and peer prosociality impact various measures of student effort. Column 5 of Table 1.9 documents that students who collaborated with more patient peers reported significantly higher average weekly study hours: a one standard deviation increase in peer patience is associated with approximately one additional hour of studying for the course per week – corresponding to 0.26 standard deviations. Peer patience also increases the aggregate time spent in Zoom lectures throughout the semester (column 6), although the effect is not significant. These findings suggests that increased effort may be one channel linking peer patience and student performance.

Table 1.9: Mechanisms

	(1) Peer interaction index	(2) Assignments handed in	(3) Grade expectation	(4) Grades depend on effort	(5) Study hours	(6) Zoom hours
Peer patience	−0.03	0.09	0.20**	−0.04	1.09***	0.74
	(0.09)	(0.30)	(0.09)	(0.07)	(0.39)	(0.68)
Peer risk taking	0.13	0.06	0.03	−0.01	0.14	1.01
	(0.08)	(0.23)	(0.07)	(0.09)	(0.31)	(0.62)
Peer prosociality	0.02	−0.01	0.00	−0.16**	−0.18	1.25*
	(0.09)	(0.29)	(0.09)	(0.07)	(0.35)	(0.65)
Mean outcome	0	8.64	0	0	8.97	13.91
SD outcome	1	3.38	1	1	4.83	8.97
No. clusters	69	69	69	69	69	69
Semester FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustered	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	146	152	152	152	152	152
R^2	0.05	0.13	0.17	0.12	0.17	0.19

Specifications as in column 8 of Table 1.6. *Peer interaction index* is based on four survey items capturing interaction quality and intensity and constructed as described in Section 1.3.3. *Assignments handed in* are the number of assignments handed in throughout the entire semester. *Grade expectation* is students' expected grade elicited on the prevalent grade scale, then standardized. *Grades depend on effort* is a belief elicited on a 1-7 Likert scale, then standardized. *Study hours* are average weekly study hours, self-reported; *Zoom hours* is time spent in all Zoom lectures of the semester. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1.3 explores the temporal dimension of student participation in Zoom lectures in more detail. It plots coefficient estimates and 95-percent confidence intervals for each of the three peer preference measures from regressions of the form specified in Equation 1.1. The independent variables in these models are the minutes spent in each of the 13 Zoom lectures of the respective semester.

The top panel of Figure 1.3 depicts the resulting pattern for peer patience which indicates that working with more patient peers seems to have a positive effect on Zoom lecture participation in the middle of the semester, but not at the semester's beginning and end, where student motivation is generally high. Peer patience thus might contribute to avoiding a mid-semester dip in effort, for which students are often unable to compensate when only raising their effort levels right before the exam.[19]

[19] More consistent levels of effort have been identified as superior to "cramming" in several studies including Geller et al. (2018) and McIntyre and Munson (2008). Being matched with more risk loving peers seems to be associated with the higher risk strategy of placing heavy emphasis on attendance at the end of the semester (see middle panel of Figure 1.3).

Having analyzed the impact of peer patience on effort, I now turn to peer prosociality: column 4 of Table 1.9 shows that peer prosociality substantially lowers students' belief that grades are a function of effort. This belief change might, in turn, lead students to reduce effort or it might impact their study effectiveness. Peer prosociality does not, however, reduce students' self-reported study time or the time spent in Zoom lectures; in fact, the latter significantly increases in the aggregate (column 6 of Table 1.9) and during the middle of the semester (bottom panel of Figure 1.3).

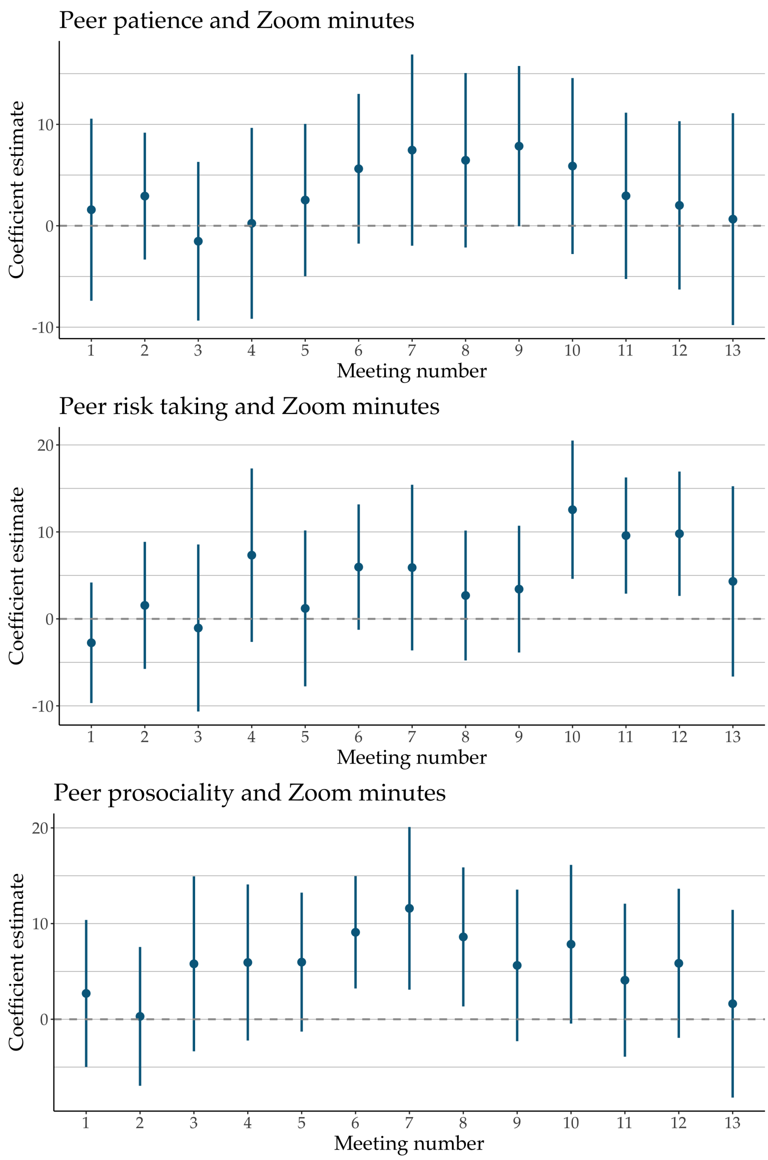


Figure 1.3: Peer preferences and minutes spent in Zoom lectures

Results from regressions of minutes spent in each of the 13 Zoom lectures per semester on peer preferences and controls as laid out in Equation 1.1. Coefficient estimates and 95-percent confidence intervals reported.

To make sense of these findings, it must be noted that neither self-reported study hours nor Zoom hours are perfect proxies for effort; and that they cannot capture study effectiveness fully: e.g., just being present in a Zoom lecture neither requires as much effort nor is it as effective as actively engaging in it and extracting strategies how to best approach exam problems. The pattern of a weakened belief in the importance of study effort despite increased presence in Zoom lectures resulting in lower student performance, could suggest that peer prosociality led students to re-adjust their time investments in a way that reduced study effectiveness. Given that peer altruism is the driving component of prosociality in these results (see Table 1.8), the above findings are consistent with an interpretation whereby students free-riding on their altruistic peers' efforts fall behind in the exam because they acquire the knowledge and skills necessary to succeed in the exam only to a lesser extent.

1.6.5 Effect heterogeneity

Who benefits or loses out from being randomly matched with more patient or prosocial peers? To investigate effect heterogeneity, I split my estimation sample based on individuals' own characteristics and then run regressions of the form specified in Equation 1.1 separately in the resulting subsamples: e.g., I create a subsample of students with above median patience and one of students exhibiting below median patience and then compare the estimated impacts of peer preferences between the two subsamples. Figure 1.4 illustrates the results from this exercise.

Students in the lower half of the patience distribution profit particularly from collaborating with more patient peers, suggesting that comparatively impatient students adopt beneficial behaviors from their patient peers. Furthermore, the results show that more persistent students and those with better high school grades benefit strongly from being matched with patient peers. These findings are consistent with the idea that patient study group members relay strategies enhancing study effectiveness – perhaps related to effort timing – also to those peers which are more likely to succeed in academic contexts to begin with.

Prosocial peers, on the other hand, especially affect impatient and impersistent students negatively, while e.g., persistent students seem to not be affected at all. This finding is, again, consistent with the hypothesis that free-riding on altruistic team members' efforts – which impatient and impersistent students are perhaps more likely to engage in – ultimately harms student performance.

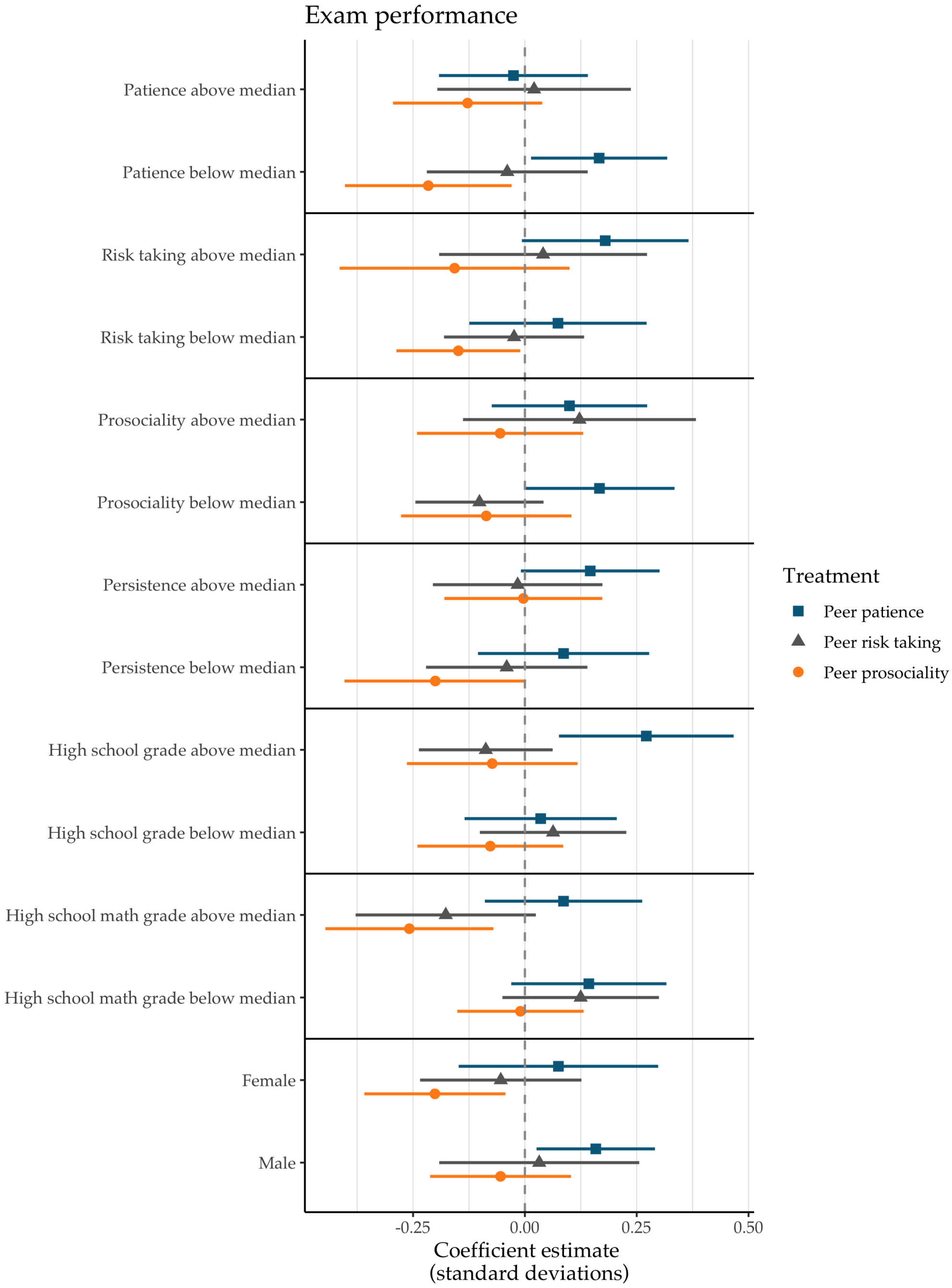


Figure 1.4: Effect heterogeneity

Results from regressions of the form specified in Equation 1.1, estimated in various subsamples created by splitting the full estimation sample based on predetermined student characteristics. Coefficient estimates and 95-percent confidence intervals reported.

1.7 Conclusion

This paper exploits interactions in randomly matched student groups to document that individuals' fundamental economic preferences generate social spillovers among their peers. I find that collaboration with more patient peers raises student performance, while being matched with prosocial peers harms achievement. These effects are comparatively large and distinct from the impact of other peer characteristics including peer achievement, gender and persistence. Peer risk taking does not significantly influence outcomes in my setting. Given the central role of these preferences in human decision making, it is plausible that they would also produce important externalities in people's social environment. To my knowledge, this paper is the first to provide causal evidence that they do.

I present suggestive evidence of the following mechanisms: more patient peers seem to raise students' effort levels and change effort timing, likely contributing to improved performance. Higher levels of peer altruism seem to be responsible for students' waning conviction that academic success is a function of effort. This possibly results in behavioral changes reducing study effectiveness, such as free-riding on altruistic team members' efforts.

These findings are of potential policy relevance if the spillovers documented here generalize: first, the overall benefits of early childhood interventions aiming to raise patience, or those of parental investments therein (Falk et al. 2021), risk being undervalued if spillovers are not considered. Second, there may be unintended consequences of altruistic behavior in the human capital accumulation process when group collaboration is involved. Possibly, these are easily remedied, e.g., via information provision.

Finally, this study has clear limitations: although the effect sizes estimated are substantial, they come from a comparatively small sample and are, depending on specification, only marginally significant. Future research could stress-test these findings in different settings and further investigate underlying mechanisms, critically assessing the ones suggested here.

APPENDIX 1

1.A Additional figures

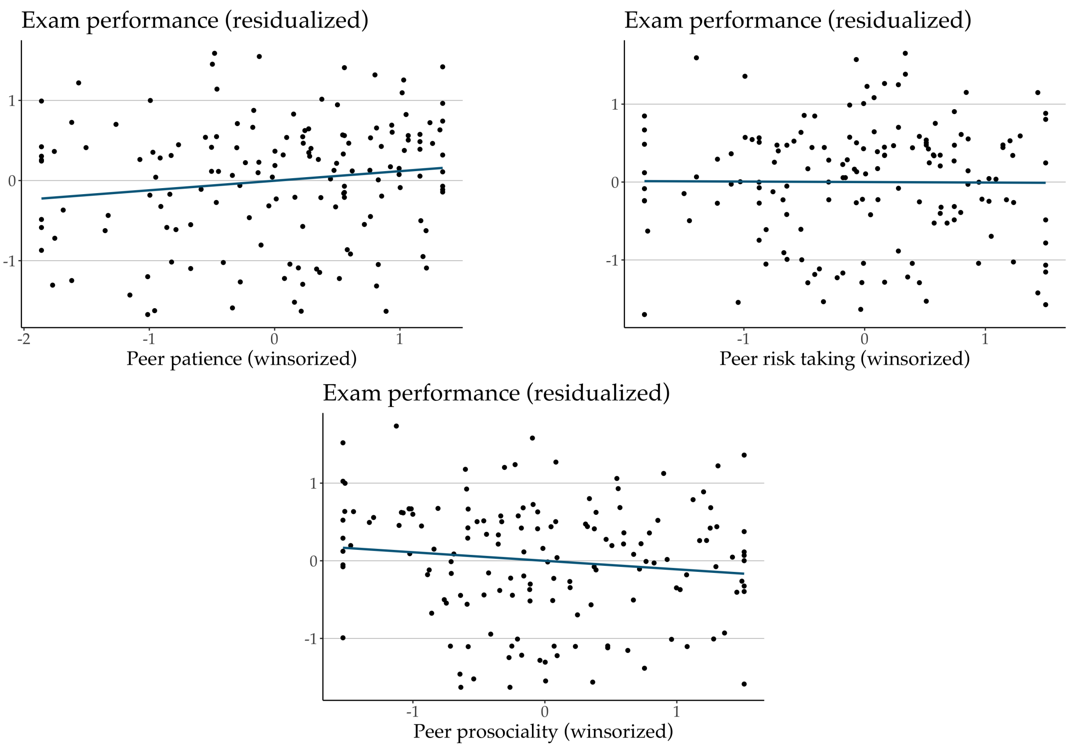


Figure 1.A.1: Winsorized peer preferences and exam performance

Scatter plots corresponding to regression results in column 8 of Table 1.B.3. All 152 data points shown.

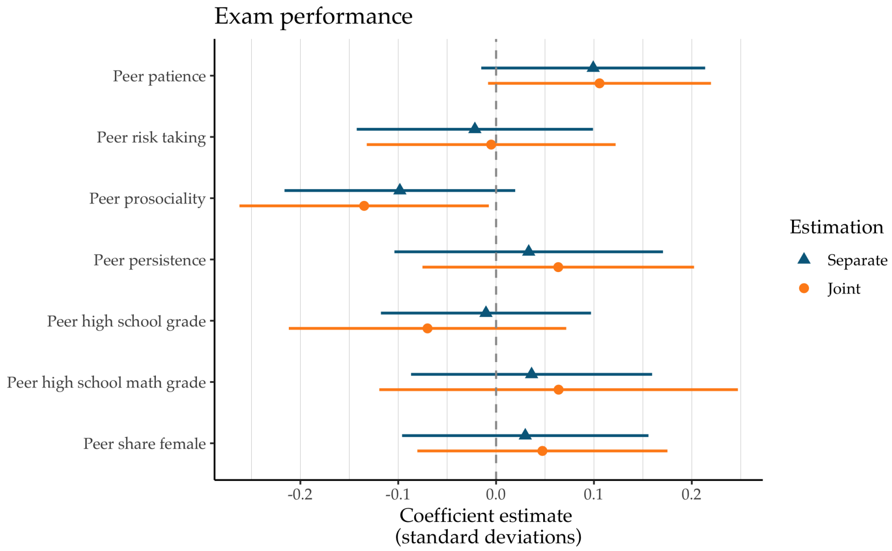


Figure 1.A.2: Comparison to other peer effects

Coefficient estimates and 95-percent confidence intervals corresponding to results in Table 1.7. *Separate* estimations refer to columns 1-7, while *joint* estimation refers to column 8, where all peer characteristics enter the model simultaneously. All variables are standardized to have mean zero and standard deviation one.

1.B Additional tables

Table 1.B.1: Students' own characteristics and performance

	Exam performance							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience	0.26***							0.20**
	(0.08)							(0.08)
Risk taking		−0.03						−0.05
		(0.10)						(0.07)
Prosociality			0.07					0.07
			(0.08)					(0.08)
Persistence				0.17**				0.11
				(0.08)				(0.08)
High school grade					0.43***			0.27**
					(0.06)			(0.12)
High school math grade						0.41***		0.12
						(0.06)		(0.12)
Female							−0.56***	−0.52***
							(0.17)	(0.15)
Mean outcome	0	0	0	0	0	0	0	0
SD outcome	1	1	1	1	1	1	1	1
No. clusters	69	69	69	69	69	69	69	69
Clustered	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	146	136	149	152	152	152	152	130
R^2	0.17	0.08	0.09	0.12	0.28	0.26	0.16	0.41

All variables (except *Female*) are standardized to have mean zero and standard deviation one. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.B.2: Peer preferences and exam performance controlling for own preferences

	Exam performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Peer patience	0.11**	0.11**	0.12**	0.10*	0.11*	0.09
	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Peer risk taking	−0.01	−0.01	−0.03	0.00	−0.01	−0.01
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Peer prosociality	−0.11*	−0.15**	−0.13**	−0.11*	−0.10	−0.12*
	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)
Patience		0.11*				0.13*
		(0.07)				(0.07)
Risk taking			−0.04			−0.03
			(0.08)			(0.07)
Prosociality				0.05		0.04
				(0.06)		(0.06)
Negative reciprocity					−0.05	−0.06
					(0.08)	(0.09)
Mean outcome	0	0.03	0	-0.01	0	0.01
SD outcome	1	0.99	0.99	1	1	0.99
No. clusters	69	67	65	69	69	63
Joint sign. (p-value)	0.14	0.05	0.08	0.17	0.21	0.22
Clustered	Yes	Yes	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	152	146	136	149	152	130
R^2	0.43	0.47	0.43	0.44	0.43	0.47

Joint sign. (p-value) is from an F-test for joint significance of peer patience, peer risk taking and peer prosociality. *Individual controls:* high school grade, high school math grade, gender, year of birth, born in DE, and migration background. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.B.3: Winsorized peer preferences and exam performance

	Exam performance							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Peer patience (winsorized)	0.11	0.11*					0.12	0.13**
	(0.08)	(0.06)					(0.08)	(0.06)
Peer risk taking (winsorized)			−0.01	−0.01			−0.01	−0.01
			(0.09)	(0.07)			(0.09)	(0.07)
Peer prosociality (winsorized)					−0.08	−0.10	−0.10	−0.11*
					(0.07)	(0.07)	(0.07)	(0.07)
Mean outcome	0	0	0	0	0	0	0	0
SD outcome	1	1	1	1	1	1	1	1
No. clusters	69	69	69	69	69	69	69	69
Joint sign. (p-value)							0.31	0.16
Clustered	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes
Observations	152	152	152	152	152	152	152	152
R^2	0.10	0.42	0.09	0.41	0.10	0.42	0.11	0.43

Peer preferences values $< 5^{th}$ and $> 95^{th}$ pctl. winsorized. *Joint sign. (p-value)* is from an F-test for joint significance of peer patience, peer risk taking and peer prosociality. *Individual controls:* high school grade, high school math grade, gender, year of birth, born in DE, and migration background. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

1.C Stability of preferences

The key threat to identification in my setting stems from the measurement of preferences after student interactions have taken place in the Semester 1 subsample. This ex-post measurement of preferences in principle opens the door to simultaneity bias, which arises if peers influence each other's preferences or the measurement thereof. This section presents additional evidence to alleviate this concern.

I exploit data from the Semester 2 subsample, where I elicited individuals' measures of patience and persistence both in the *signup* survey and the *endline* survey, i.e. before *and* after peer interactions took place.[20] Within this subsample, I can directly test the stability of the two measures.

I conduct two types of analysis: in the first, I calculate test-retest correlations between the measures elicited at baseline and at endline, which are shown in Table 1.C.1. The test-retest Spearman correlation of 0.57 for patience is very similar to the one in the online appendix of Falk et al. (2021): they report a test-retest Spearman correlation of 0.67 for two measurements of patience in a sample of young adults taken only one week apart. As Falk et al. (2021) argue, it is unlikely that preferences changed within a one-week interval in that sample, implying that the observed instability is likely due to measurement error rather than changing preferences. The comparable results in Table 1.C.1 are thus reassuring evidence of the stability of patience in my sample of young adults, where the interval between the two measurements was three months, in between which the study group interactions took place.

Table 1.C.1: Test-retest correlations for patience and persistence

	Patience	Persistence
Pearson correlation	0.499	0.569
Spearman rank correlation	0.572	0.539
N	29	36

Based on a three-month interval between test (in the *signup* survey) and retest (in the *endline* survey).

The second analysis directly tests whether peer patience (based on the patience measure elicited at *baseline*) had an impact on the patience measure elicited at *endline*. I thus regress endline patience on peer patience, controlling for individuals' baseline patience, to maximize precision. The results in column 2 of Table 1.C.2 document that there is no significant effect of peer patience on the patience measure elicited at endline. The coefficient is not exactly zero, but the very large standard error suggests

[20] I did not re-elicit all preferences from the GPS in the endline survey as this would have likely impacted the completion rate negatively.

random residual variation in a small sample rather than a systematic relationship.

Table 1.C.2: Baseline and endline preferences

	Patience (endline)		Persistence (endline)	
	(1)	(2)	(3)	(4)
Patience	0.52**	0.51**		
	(0.21)	(0.21)		
Peer patience		0.08		
		(0.22)		
Persistence			0.57***	0.61***
			(0.12)	(0.15)
Peer persistence				−0.15
				(0.13)
Mean outcome	0	0	0	0
SD outcome	1	1	1	1
No. clusters	19	19	20	20
Clustered	Yes	Yes	Yes	Yes
Observations	29	29	36	36
R^2	0.25	0.25	0.32	0.34

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In sum, the present evidence suggests that peers do not influence each other's preferences or the measurement thereof. It thus complements the balancing results in Section 1.4 and further mitigates the concern that simultaneity bias might confound the main findings of this paper.

Chapter 2

Anticipated Peer Effects

2.1 Introduction

What motivates us to act prosocially? A prominent literature in economics has documented that individuals are more likely to act prosocially if their behavior is observable to others and that *social image effects* are an important motive explaining these behavioral responses.[1] In this paper, we propose a complementary explanation as to why individuals are more inclined to act prosocially under observability: *anticipated peer effects*. While social image effects imply that individuals care about how their behavior is perceived by others, anticipated peer effects capture the idea that individuals are also motivated by an anticipation that their own (prosocial) behavior may exhibit a peer effect on others.

We argue that anticipated peer effects are a relevant motive for decision making in many situations where our actions generate positive externalities if those around us follow suit: examples range from signing up early for a health screening to nudge our more present-biased friends, to ordering vegetarian food during a group meal or wearing an "I voted" button to signal the importance of eco-friendly behavior or voting, respectively. In all these situations other motives such as social and self image concerns clearly play important roles, but anticipated peer effects frequently push the benefits of a prosocial action above its costs, such as when we stop at a red traffic light to be a role model to younger observers. Despite their potential importance, empirical evidence on the existence of anticipated peer effects in prosocial settings is scarce, largely due to the difficulty of disentangling them from social image effects.

In this paper, we causally identify anticipated peer effects in the decision whether

[1] Bénabou and Tirole (2006) provide the seminal theoretical exposition of social signaling in the context of prosocial behavior. Bursztyn and Jensen (2017) offer an extensive review of experimental evidence on how observability shapes behavior in various domains including voting, donations to charity, financial decision making, or schooling decisions.

to register for a COVID-19 vaccination, separating them conceptually and empirically from social image effects. Using a survey-based online field experiment, we document that individuals' willingness to register for the vaccination increases sharply when realizing that they can influence a peer's decision. Our results highlight that individuals anticipate and internalize their potential to lead by example in consequential decision environments.

Our experimental design groups survey participants into pairs, where one participant takes on the role of "Sender" (she) and the other acts as "Receiver" (he). To isolate anticipated peer effects, we experimentally vary (1) the *observability* of the Sender's decision to her Receiver and (2) the *timing* when the Receiver is informed. In the baseline condition *"not informing partner"*, we told Senders that their decision whether to register for a vaccination would *not* be reported to their partner. In the *"informing partner after"* condition, Senders learned that their decision would be shared with their Receiver, but only *after* the Receiver himself had already decided whether to register. Finally, in the *"informing partner before"* condition, we told Senders that their Receiver would be informed *before* his own registration decision. We expect anticipated peer effects to influence the behavior of Senders in the *"informing partner before"* condition, while the *"informing partner after"* condition serves as a control group which holds other behavioral factors – in particular social image effects – constant.

Anticipated peer effects almost double Senders' likelihood of registering for a COVID-19 vaccination: 9 percent of Senders in the *"informing partner before"* condition completed the official registration process during our online experiment compared to 5 percent in the *"informing partner after"* condition. We exploit the official state-wide online registration and appointment allocation system for COVID-19 vaccinations in the German state of Bavaria, which constituted the only pathway to receiving a COVID-19 vaccination for Bavarian citizens at the time of the experiment in early April 2021. This allowed us to elicit a verifiable revealed preference measure of individuals' willingness to register for a vaccination. We also document that Senders in the *"informing partner before"* condition were substantially more likely to believe that their registration decision could influence their partner's decision. These Senders arguably internalized that their registration decision might generate externalities – with respect to their partner's personal health as well as herd immunity – if their partners followed their lead.[2]

The interaction between Senders and Receivers in our experimental design is explicitly designed to be anonymous and to rule out future interactions between partners

[2]An alternative two-stage least squares (2SLS) estimation supports this interpretation: if we use our key experimental manipulation as an instrument for Senders' beliefs about the likelihood of exerting a peer effect on their partner, we find statistically significant effects of this belief on Senders' likelihood to sign up for a COVID-19 vaccination.

to further minimize the influence of social image effects. We recruited participants for our online survey using a professional online panel provider sampling from the adult population of Bavaria. After an elicitation of baseline beliefs and attitudes (e.g., regarding vaccination safety and efficacy) participants learn that they will be grouped into pairs to collaborate on a brief joint problem solving task, where they can interact via chat but merely learn each other's nickname, age, gender and state of residence. As intended, the joint task is successful in inducing social proximity, i.e., in establishing the Receiver as a relevant – if temporary – peer to the Sender. However, the interaction between Senders and Receivers is limited to the brief exchange in the joint task and the anonymous online survey setting forestalls future interactions outside of the experiment, which precludes Senders from influencing their partner after the experiment.[3] After the joint task and an elicitation of social proximity, participants proceed to the treatment stage where we vary the *observability* of the Sender's decision and the *timing* of when the Receiver is informed about it. Subsequently, we begin by eliciting *stated* preference outcomes, e.g., asking participants whether they would like to register for a COVID-19 vaccination. We elicit our main *revealed* preference outcome by giving participants the option of actually registering for a COVID-19 vaccination using Bavaria's official online registration platform and subsequently return to our survey, where we verify their registration.[4]

Our experimental setting allows us to rule out various alternative explanations including (i) social image effects, (ii) experimenter demand, (iii) strategic lying, and (iv) cheating. First, we can hold social image effects constant since we identify anticipated peer effects by comparing the share of participants who registered for a COVID-19 vaccination between the *"informing partner before"* and the *"informing partner after"* conditions. Senders in both conditions are subject to being judged by their partner, but only Senders in the *"informing partner before"* condition should infer that they can influence their partner during the experiment. Moreover, a comparison of the *"informing partner after"* condition with the baseline condition *"not informing partner"* reveals no

[3]In a related lab experiment, Karlan and McConnell (2014) use a similar set of experimental manipulations to tease out the impact of leading by example on prosocial behavior. Yet, in their experiment neither the interaction between Senders and Receivers nor the decision taken by Senders is one-shot in nature. This limits the ex-ante potential for identifying anticipated peer effects since Senders in the "*informing partner after*" condition may also anticipate that they will influence their partner's behavior after the experiment.

[4]We verify registrations by asking participants to provide us with specific information from the official confirmation email they would have received from the Bavarian health authorities. To motivate participants to take on these additional time costs, we incentivized the verification via a lottery of Amazon vouchers. An analysis of timing patterns reveals that participants whose registration we verify in this way did not complete the registration process only after they learned that they could win a voucher: on average they were inactive in our survey for six minutes at the stage where we expected them to register (before learning about the incentive) – a realistic duration for completing the official online registration – while they were much faster in providing the verifiable information. These timing patterns did not differ between the relevant treatment conditions as documented in Figure 2.6 and the corresponding Table 2.B.5 in the Appendix.

difference in registration shares between those two groups, suggesting that social image effects only play a minor role in our setting, as expected due to our experiment's highly anonymous and one-shot peer interaction. Second, experimenter demand is unlikely to play a major role in our setting: we show that *stated* registration willingness does not differ between the *"informing partner before"* and the *"informing partner after"* conditions, while *actual* (verified) registration willingness clearly does. This renders it unlikely that an *"informing partner before"*-specific experimenter demand effect biases our main result on the revealed-preference outcome; instead, moving from *stated* to *actual* willingness seems to weed out a general experimenter demand effect regarding registration willingness present in all experimental conditions. Third, this comparison between *stated* and *actual* registration willingness shows that strategic lying does not drive our results: Senders in the *"informing partner before"* condition do not merely pretend to register; instead, they actually follow through, arguably to signal the benefits of taking the vaccine to their partner. Finally, we show that differential incentives to cheat are unlikely to explain our findings, addressing two variants of this concern: either Senders in the *"informing partner before"* condition only signed up after they had learned about the monetary incentives for verifying their registration status; or, Senders in this condition were more likely to tap other sources, e.g., the internet, to pass our verification procedure. Using data on how much time Senders spent on each survey page, we demonstrate that neither of these alternative explanations challenge our findings.

Heterogeneity analyses further support our interpretation that Senders in the *"informing partner before"* condition internalize an anticipated peer effect and that this constitutes a prosocial motive: our main treatment effect of interest – the difference in registration shares between Senders in the *"informing partner before"* and the *"informing partner after"* conditions – is positive and significant for Senders who have a strong pre-treatment belief in the safety and efficacy of COVID-19 vaccines, while it is negative among those who do not believe in the vaccines' safety and efficacy. This result suggests that Senders who can influence their partner indeed choose to lead by example, signalling what they believe is best for the other.[5] This interpretation is further reinforced by the finding that the strength of the treatment effect seems to increase with the level of social proximity between Sender and Receiver: Senders who can influence their partner are slightly more likely to sign up for a vaccination when they "feel closer" to their partner.

In a final set of results, we document that Senders are not successful in influencing their partners – contrary to their own anticipation. Receivers who are informed

[5] In our specific vaccination registration setting, Senders' beliefs about the vaccination certainly influence their view as to which decision – to register or not – would entail a positive versus a negative externality if their partner followed suit.

about their Sender's registration decision *before* they can decide themselves are no more likely to make the same choice as their partner than those who learn about their Sender's choice only *after* their own decision. This finding can be explained by our specific setting in which Senders do not observe Receivers' decisions, implying that Senders cannot verify whether they indeed exert a peer effect. Moreover, social image effects – a potential channel for conformity – play no role on the Receivers' side, as Senders do not observe their decision. We can learn from this result, however, that there are important decision environments in which individuals overestimate their anticipated peer effect; hence, the materialization of an 'actual' peer effect is not always necessary to activate people's desire to lead by example.[6]

In sum, our findings indicate that anticipated peer effects can play a substantial role in decision settings with a prosocial component, i.e., where our actions generate positive externalities if those around us follow suit. This has implications for policy: where anticipated peer effects are a meaningful behavioral motive, policymakers wishing to encourage prosocial behavior may have another "social carrot" at their disposal – leveraging our desire to lead by example – and need not revert to "social sticks". For example, as an alternative to imposing a fine on citizens who miss their vaccination appointment, governments may also target people's willingness to set a good example in order to increase vaccination rates.[7] At a more general level, our findings suggest that anticipated peer effects may constitute one potential channel through which social change propagates: if a desire to lead by example – perhaps based on an overestimation of our impact on others – motivates us to bear the private costs of prosocial behavior, such as participating in a protest for civil liberties in an autocracy, anticipated peer effects may add to our understanding of why people, despite the low chances of being pivotal, are willing to accept these private costs.

The idea that individuals incorporate into their decision making an anticipation of how their behavior might influence others has been an implicit theme in a wide variety of papers in economics. To our knowledge, however, the present paper is the first to offer field experimental evidence that people act upon a desire to lead by example in a prosocial setting. The paper most closely related to ours is a lab experiment by Karlan and McConnell (2014), who hypothesize – in a similar vein as we do – that a desire to influence others might be one reason why donations are higher under observability. Based on a comparable design they find that a desire to influence others does *not*

[6]That anticipated peer effects influence Senders' behavior despite their partners not following suit implies that Senders have out-of-equilibrium beliefs. Due to reasons of statistical power we have to limit ourselves to documenting these misperceptions in this paper. However, we believe that the endogenous formation of such misperceptions constitutes an interesting avenue for future research.

[7]Recently, policy makers in Germany discussed whether to introduce fines for citizens who missed their COVID-19 vaccination appointments (see, e.g., media coverage by *Süddeutsche Zeitung*: https://bit.ly/3zuoQO9, last accessed 2021-09-10).

seem to increase donations. As the authors point out themselves, however, their study does not offer "dispositive evidence" due to a lack of precise null effects as well as concerns regarding the external validity and particular features of their lab setting, e.g., the difficulty of ruling out future interactions between subjects drawn from the same peer group. We benefit from these insights and explicitly design our experiment as a one-shot interaction between two anonymous partners in a consequential decision environment, thereby limiting Senders' ability to influence their partner to one specific moment.

Our research also relates to a large body of literature that studies whether and under which conditions leading by example is successful in increasing contributions in public goods games played in the lab.[8] For example, Potters, Sefton, and Vesterlund (2007) show that leading by example increases public goods contribution under asymmetric information, i.e., when an informed leader can signal information about the value of contributing to an uninformed follower, a result consistent with earlier theoretical work (e.g., Hermalin 1998 and Vesterlund 2003). Leading by example 'works' since followers are very likely to copy the decisions of leaders and leaders tend to correctly anticipate followers' responses, contributing more themselves. Our paper shares with this literature the idea that leaders anticipate that their behavior might influence the behavior of followers. A key difference in these public goods games is, however, that leaders always have a first-order monetary incentive to lead by example: their own monetary payoff is higher if they convince others to follow. We complement this literature by abstracting from such first-order incentives and adding an explicit focus on studying the desire to lead by example in a prosocial field setting.

Finally, we speak to a prominent literature on social signaling in the context of prosocial behavior. At least since the seminal contribution by Bénabou and Tirole (2006), a large body of literature in economics has highlighted that individuals' prosocial behavior depends on the visibility of their actions to others and that social image effects are an important motive explaining these behavioral responses.[9] Early theoretical predictions have been confirmed by a series of field experiments investigating social image effects (Bursztyn and Jensen 2017; Bursztyn, Fujiwara, and Pallais 2017; DellaVigna, List, and Malmendier 2012; DellaVigna et al. 2017; Perez-Truglia and Cruces 2017). We add to this literature by highlighting a social signaling motive that is distinct from social image concerns, insofar as it depends on the anticipation that our behavior can have a peer effect on others. Our paper is thus also related to the study of social signaling in the context of childhood immunization by Karing

[8]Important contributions include Arbak and Villeval (2013), Cappelen et al. (2016), Dannenberg (2015), Drouvelis and Nosenzo (2013), Gächter et al. (2012), Gächter and Renner (2018), Güth et al. (2007), Haigner and Wakolbinger (2010), and Potters, Sefton, and Vesterlund (2007).

[9]In a related earlier model, Bernheim (1994) argues that people's status concerns can generate conformity of behavior.

(2021), in which she highlights the distinction between social signals as transmitters of information about others' actions on the one hand and as a means to signal one's type on the other. Our results have a similar potential for informing policymakers aiming to promote prosocial behavior, e.g., by increasing timely vaccination take-up. The results from our anonymous setting indicate that such policies need not conflict with privacy concerns: revealing anonymous information may suffice to facilitate prosocial behavior via anticipated peer effects.

This paper is organized as follows: in Section 2.2, we discuss our experimental design employed to identify anticipated peer effects in a prosocial setting. Then, in Section 2.3, we present our main results from the experiment and address potential concerns about our findings. Section 2.4 concludes the paper.

2.2 Experimental setup

The objective of our experimental design is to separate two complementary motives why people are more inclined to act prosocially if they are observed by others: anticipated peer effects and social image effects. In this section, we illustrate our experiment's setting and the sample employed and discuss the main features of our experimental design. In the final subsection, we show that Senders' predetermined characteristics are balanced across treatment conditions.[10]

2.2.1 Setting and sample

We conduct a survey-based online field experiment studying decision making in the context of COVID-19 vaccinations in the German state of Bavaria. We examine individuals' willingness to register for a COVID-19 vaccination via the state-wide central appointment allocation system BayIMCO, which at the time of the experiment in April 2021 constituted the only pathway for obtaining a vaccination in Bavaria.[11] Owing to vaccine supply shortages, which prevailed until ca. July 2021, the official vaccination regulations categorized individuals into several priority groups depending on their age and pre-existing health conditions. However, all Bavarian residents had the possibility to register online[12] from January 2021 onwards, regardless of their prioritization

[10]We pre-registered all features of our experimental design at the AEA RCT registry under ID AEARCTR-0007437 before the experiment commenced. The experiment described here was approved by the Ethics Committee of the Department of Economics at LMU Munich, protocol 2021-01. For the technical implementation of our online experiment, we used the open-source software oTree (Chen, Schonger, and Wickens 2016).

[11]Later in the vaccination campaign, the central system was complemented by a decentralized system relying on local doctors' offices. However, as of July 2021, the central system still accounts for roughly 60 percent of all vaccinations in Germany (Bundesministerium für Gesundheit 2021).

[12]Only in exceptional cases was registration via phone also possible.

status. Once vaccine supply and their prioritization status allowed, registered residents received a vaccination appointment through the central system.

We recruited the participants for our survey from the online panel provider CINT. During our experiment's field time, approximately 15 percent of the Bavarian population had already received at least one vaccination and a further 30 percent had registered in the central system. We exclude both of these groups from our experiment by screening them out at the start of our survey. In total, 1,857 participants completed our experiment. We report summary statistics on participant characteristics in Table 2.1: 51 percent of our participants were willing to get vaccinated (elicited pre-treatment), which is – due to our exclusion of already vaccinated and registered individuals – somewhat lower than the vaccination willingness of 65 percent elicited in a nationally representative study at the same point in time (COSMO – COVID-19 Snapshot Monitoring 2021; Betsch, Wieler, and Habersaat 2020). With respect to other key characteristics such as gender, age, and income, our study participants are suitably representative of the Bavarian population as a whole.[13]

Table 2.1: Summary statistics for full sample (Senders and Receivers)

Statistic	Mean	St. Dev.	Min	Max	N
Demographics					
Age	40.90	14.35	18.00	79.00	1,857
Female	0.55	0.50	0.00	1.00	1,857
Monthly household income (net)	2,907.78	1,597.37	1,100.00	7,500.00	1,857
Upper secondary degree	0.39	0.49	0.00	1.00	1,857
Local characteristics*					
Mean incidence rate (second wave)	138.73	40.67	65.64	301.07	1,857
Population in zip	14.81	9.85	0.60	48.05	1,857
Lives in urban area ($\geq$100k)	0.29	0.46	0.00	1.00	1,857
Turnout in 2017	77.52	4.31	59.90	90.20	1,857
AfD vote share in 2017	12.23	3.07	5.49	26.42	1,857
Unemployment rate	2.37	0.93	0.05	5.50	1,857
Beliefs about vaccine					
Safety	3.41	1.96	1.00	7.00	1,857
Efficacy	3.77	1.96	1.00	7.00	1,857
Social desirability	3.62	2.25	1.00	7.00	1,857
Severity of freerider problem	3.26	2.07	1.00	7.00	1,857
Willingness to take vaccine in state (%)	59.11	20.16	0.00	100.00	1,857
Preferences					
Own willingness to take vaccine (%)	51.31	37.09	0.00	100.00	1,857
Altruism	0.01	0.83	−1.96	2.25	1,857
Desire to influence	0.08	0.98	−2.90	1.72	1,857
Social image concern	0.03	1.00	−1.81	2.34	1,857
Social proximity					
Social proximity	0.02	1.00	−1.04	2.27	1,526

Variables marked with * vary on the zip code, county ("Landkreis"), or town ("Gemeinde") of residence level and not on the individual level.

[13]Roughly half of our sample is female; mean age and monthly net income are 40.9 years and 2,907EUR, respectively, compared to the official state averages of 43.7 years in 2017 (Bayerisches Landesamt für Statistik 2019) and 2,549 EUR in 2018 (GESIS – Leibniz-Institut für Sozialwissenschaften 2019).

2.2.2 Experimental design

Our experiment revolves around the interaction within teams consisting of one *Sender* and one *Receiver* and aims to isolate anticipated peer effects from social image effects. It evolves over seven stages, which we detail below.[14]

1. Introduction. We begin by screening out all subjects who had already been vaccinated or registered for a COVID-19 vaccination. From all remaining participants we collect basic demographic information as well as a rich set of attitudes, beliefs, and preferences related to the vaccination (e.g., beliefs about vaccine safety and efficacy).

2. Joint problem solving task. Subsequently, we build teams consisting of two randomly paired participants. Within teams, subjects are randomly assigned either to the role of Sender (she) or Receiver (he). Before teams enter the main stage of the experiment, they work on a joint problem solving task adopted from Goette and Tripodi (2020), which we use to induce social proximity between the partners, i.e., to establish the Receiver as a relevant peer to the Sender. The task consists of four consecutive questions, in which teams are presented with historical paintings and are asked to select the corresponding artist from a list. Each correct answer increases participants' probability of winning an Amazon voucher, but only if their partner selects the correct artist as well. To allow for coordination between partners, we provide them with the option to exchange text messages.[15] Participants are informed as to whether they won any of the vouchers on the final page of the survey.

3. Social proximity. After the joint task, we elicit a measure of social proximity between partners using the "oneness" scale (Cialdini et al. 1997; Gächter, Starmer, and Tufano 2015), again following Goette and Tripodi (2020).[16] We find that the joint problem solving task performs well in establishing social proximity between partners: according to this scale, at least half of the participants perceive their partner in the experiment as an "acquaintance" and 25 percent even think of their partner as a "non-close friend" (for details on the scale, see Gächter, Starmer, and Tufano 2015).

4. Treatment. Next, teams enter the experiment's treatment stage, where we use two experimental manipulations to isolate the impact of anticipated peer effects on Senders' decisions to sign up for a vaccination: we vary (1) the observability of the

[14]We provide the English translation of the complete survey instrument in Section 2.D in the Appendix.

[15]We provide a screenshot of the joint problem solving task showing the chat window in Figure 2.C.1 in the Appendix.

[16]The oneness scale is computed as the unweighted mean of the "Inclusion of Other in the Self" (IOS) scale (A. Aron, E. N. Aron, and Smollan 1992) and the "WE" scale (Cialdini et al. 1997). We provide screenshots of how we elicited the oneness scale in Figure 2.C.2 in the Appendix.

Sender's decision to her partner and (2) the timing of when the partner is informed about the Sender's decision.

The main intuition of our design is illustrated in Figure 2.1. For each treatment condition, we report the key treatment message shown to the Sender and the corresponding decision sequence as implemented in the experiment. Irrespective of the condition to which we assigned teams, Senders were always offered the opportunity to sign up for the vaccination before the Receiver and no Sender learned about the decision of the Receiver.

In the *"not informing partner"* condition, we inform Senders that their decision on whether to register for a vaccination will not be reported to their partner. As a result, neither anticipated peer effects nor social image effects affect Senders' decisions.

In the *"informing partner after"* condition, Senders learn that their decision will be shared with their partner. However, we highlight to Senders that their partner will only be informed about their decision once he (the partner) has himself already decided whether to register. Therefore, while social image effects might arise, Senders cannot influence their partner's decision within the experiment and, consequently, anticipated peer effects should play no role in this condition.

In the third and final condition, *"informing partner before"*, we inform Senders that their partner will learn about their decision *before* he is given the opportunity to register for a vaccination. As above, Senders in this condition are subject to social image effects. In addition, however, they should infer that they can now influence their partner's decision. More formally, in the present condition Senders' beliefs about the likelihood of exerting a peer effect on their partner should, in expectation, be higher than in the *"informing partner after"* condition. Hence, by comparing Senders' willingness to sign up for the vaccination between Senders who can and those who cannot influence their partner, we can isolate anticipated peer effects from social image effects.

5. First stage. As laid out above, the strength of anticipated peer effects is governed by Senders' beliefs about how likely it is that they can influence their partner's decision. As such, changes in this particular belief constitute the "first stage" of our experiment. To measure whether our experimental manipulations indeed induce an upward shift in this first stage belief, we ask Senders how likely they think it is that they can influence their partner's decision of whether to sign up for the vaccination. To elicit this belief, we use a slider ranging from 0 to 100. We pose this question to Senders after the treatment module and before eliciting the main outcome.

6. Main outcome. Next, we elicit our main outcome by asking participants whether they wished to sign up for a COVID-19 vaccination right away. If participants an-

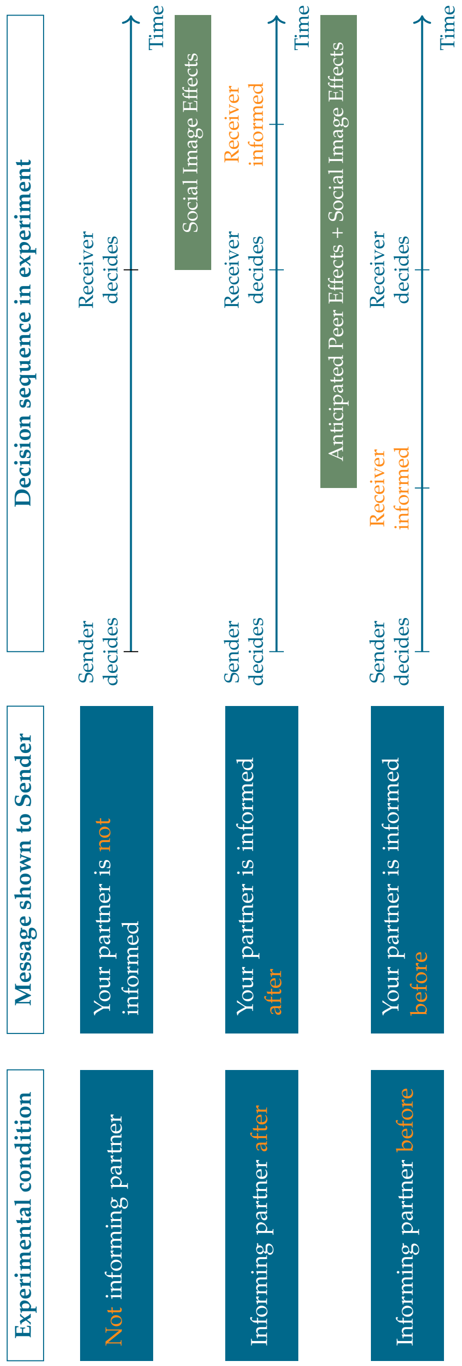


Figure 2.1: Treatment messages and corresponding implementation in survey

swered "yes", they were forwarded to the BayIMCO website outside of our survey.[17] Participants who responded "no" were forwarded to the next page of our survey. On average, it took participants in our experiment five to six minutes to complete the online registration form. Once participants completed the form, they obtained an email from BayIMCO officially confirming their registration. We use this confirmation email to verify whether participants indeed registered for a vaccination by asking them to enter the sending address and the subject line of the confirmation email in a survey form.[18]

The timing of the steps we used to elicit whether participants actually signed up for the vaccination is crucial in this context: when we offered participants the opportunity to sign up for the vaccination, participants did not know that we would ask them to provide proof of their registration. We only informed participants about the confirmation and the corresponding remuneration after they had reported to us that they successfully completed the registration. Hence, participants did not have an incentive to misreport their registration in order to qualify for one of the vouchers. Still, one may worry that participants misreporting their registration status tried to find out the address and the subject line of the confirmation email to nevertheless qualify for one of the vouchers. It is, however, very unlikely that participants successfully managed to cheat, since the address from which the confirmation email was sent changed over time. Thus, even if participants found a screenshot of the confirmation email by searching the internet, the screenshot had to be fairly recent to keep up with the changes of the confirmation email over time.[19]

7. Further outcomes. Finally, we collect post-treatment attitudes and beliefs related to the COVID-19 vaccination, including participants' stated willingness to take the vaccine alongside with their beliefs regarding the safety and efficacy of the vaccine, its social desirability, and associated free riding problems.[20] In addition, we collected further demographic information including income and education as well as participants' county and zip code of residence. On the final page of the survey, we revealed

[17]The official registration website (BayIMCO) provided by the Bavarian Ministry of Health can be accessed at https://impfzentren.bayern/citizen/ (last accessed 2021-09-11). We provide screenshots illustrating how we elicit and verify the registration decision in Section 2.C.3 in the Appendix.

[18]We incentivized participants by informing them that by reporting both pieces of information correctly they would qualify for one of 30 additional 20 EUR Amazon vouchers. Once participants had entered their information, their responses were checked by our system. If both answers were correct, a lottery determined whether participants obtained one of the Amazon vouchers. Participants only learned whether they had won any of the Amazon vouchers after they had answered all questions, i.e., on the final page of the survey.

[19]As we detail in Section 2.3.4, we find no differential indication of Senders successfully bypassing our verification process or completing their registration only after they had learned about the vouchers between experimental conditions.

[20]We also collect the same set of beliefs before the treatment to analyze within-individual changes arising from the treatment.

payoffs to participants and provided them with the opportunity to comment on the survey.

2.2.3 Additional steps taken to identify anticipated peer effects

In order to identify anticipated peer effects, our design aims to maximize the difference in Senders' beliefs about their ability to influence their partner between the *"informing partner before"* and *"informing partner after"* conditions. To achieve this, we designed both the decision Senders take as well as the interaction with their partner to be "one-shot". To ensure that the interaction is one-shot in nature, we paired individuals who had never met before and upheld anonymity throughout the experiment. Anonymity facilitates identifying anticipated peer effects as it limits Senders' chances of influencing their partner to that particular encounter: Senders in the *"informing partner before"* condition should realize that their opportunity to influence their partner's decision is either now, by sending a signal in the experiment, or never. Of course, Senders' decisions within the experiment may influence Receivers' behavior after the experiment has ended, as Receivers may contemplate their partner's decision in the experiment for a while and register for a vaccination at some later point in time. In principle, Senders in the *"informing partner after"* condition may realize as well that their actions during the experiment might influence Receivers' behavior *after* the experiment. If that was the case, anticipated peer effects would also motivate Senders in this condition, potentially attenuating behavioral differences relative to the *"informing partner before"* condition.[21]

As a further step towards identifying anticipated peer effects, we deliberately limited the scope for social image effects from the onset. To this end, Senders in our experiment interacted with individuals they had never met before and not with their neighbors (as in Bursztyn, González, and Yanagizawa-Drott 2020) or classmates (as in Bursztyn and Jensen 2015). Our design further limits the potential impact of social image effects by informing only one individual about the Sender's decision. In existing paradigms, the number of Receivers is usually much larger (e.g., in Perez-Truglia and Cruces 2017). Ultimately, by upholding anonymity throughout the experiment, we rule out future interactions between partners and thereby shut down most instru-

[21]Moreover, the fact that the decision itself – and thus its potential externality on the Receiver – is one-shot, may render it more salient from the perspective of the Sender. Combined, the one-shot decision and the one-shot interaction help us identify anticipated peer effects. The role of these design features also suggests a reason why Karlan and McConnell (2014) – who used a similar set of experimental manipulations – did not find evidence for anticipated peer effects: to conduct their experiment, they recruited participants from the same peer group (college students from the same university). As a result, Senders might have already known Receivers and might have anticipated to meet them again in the future, reducing the relative importance of the signal sent within the experiment. A similar logic applies to the decision they studied: they asked Senders to decide about a donation to a university institution, a decision which Senders could take multiple times in the future.

mental motives underlying social image effects.[22] Taken together, we expect only a weak impact of social image effects on Senders' behavior in this particular context.

2.2.4 Experimental assignment and sample balancing

We used a two-stage random procedure to assign participants into treatment conditions: first, we assigned teams to one of the three treatment conditions *"not informing partner"*, *"informing partner after"*, or *"informing partner before"*. Second, within the teams, we further randomized who was assigned the role of Sender and Receiver, respectively. We report the resulting assignment into experimental conditions in Table 2.2.[23]

Table 2.2: Number of Senders and Receivers assigned to each group

Condition	Treatments	Senders	Receivers
(1) **Not** informing partner	Observability = 0	328	–
(2) Informing partner **after**	Observability = 1 Informed before = 0	554	236
(3) Informing partner **before**	Observability = 1 Informed before = 1	519	220

Since we are primarily interested in Senders' decisions, we opted for an implementation using fewer Receivers than Senders in each group: in some teams a Sender's partner was another Sender and not a Receiver. To avoid deception, the experimental instructions thus involved a degree of uncertainty regarding whether a participant's decision would be shared with their partner. Therefore, we could use the same experimental instructions for all Senders in the same condition irrespective of whether a Sender's partner was another Sender or an "actual" Receiver, while still only employing factually true information.[24] To further reduce the number of Receivers in our experiment, we paired Senders in the *"not informing partner"* condition always with other Senders. Since Senders' decisions in this condition were not shared with their partner from the joint problem solving task anyways, these Senders' partners could also be other Senders without introducing deception.

[22]See Bursztyn and Jensen (2017) for a discussion of the distinction between instrumental and hedonic motives underlying social image effects.

[23]The discrepancy between the number of participants in the *"informing partner after"* and *"informing partner before"* conditions is an artifact of the specific randomization procedure used. We used "on the fly" randomization to assign participants into experimental conditions as they entered the survey. Due to the random nature of the assignment process, the effective share in each condition slightly deviates from the target shares we specified in our pre-analysis plan.

[24]To be precise, we informed Senders that their partner would learn about their decision only "with high probability".

To assess whether Senders' predetermined characteristics are balanced across experimental conditions, we conducted pairwise comparisons of 21 predetermined characteristics across all three conditions using bivariate regressions.[25] In Table 2.3, we report the group means separately for each condition alongside the p-values obtained from these regressions.[26] Out of the 63 estimates reported in Table 2.3, only one is significant at the 5 percent level, suggesting that Senders' predetermined characteristics are well balanced across treatment conditions.[27] These results thus minimize the risk of wrongly attributing potential treatment effects to our experimental manipulations.

Table 2.3: Senders' predetermined characteristics compared across treatment conditions

	Group means			Test for equal means: p-values			
	Before	After	Not	Before vs. After	Before vs. Not	After vs. Not	N
Attrition							
Completed survey	0.73	0.74	0.76	0.87	0.24	0.30	1892
Demographics							
Age	40.67	41.36	40.43	0.43	0.82	0.36	1401
Female	0.56	0.54	0.52	0.42	0.23	0.61	1401
Monthly household income (net)	2846.82	2850.90	2990.55	0.97	0.21	0.21	1401
Upper secondary degree	0.37	0.40	0.40	0.43	0.41	0.88	1401
Local characteristics							
Avg. incidence rate (during second wave)	138.37	140.61	136.96	0.37	0.61	0.20	1401
Population in zip	14.21	15.17	14.91	0.11	0.29	0.71	1401
Lives in urban area ($\geq$100k)	0.29	0.32	0.31	0.31	0.44	0.91	1401
Turnout (%)	77.42	77.60	77.52	0.50	0.74	0.79	1401
AfD vote share (%)	12.29	12.17	12.18	0.55	0.64	0.96	1401
Unemployment rate (%)	2.36	2.41	2.37	0.40	0.91	0.53	1401
Beliefs							
Safety	3.37	3.40	3.42	0.80	0.69	0.86	1401
Efficacy	3.76	3.73	3.76	0.84	0.99	0.87	1401
Social desirability	3.58	3.64	3.62	0.69	0.81	0.92	1401
Severity of freerider problem	3.29	3.10	3.40	0.12	0.46	0.03**	1401
Willingness to take vaccine in state (%)	58.37	58.41	59.83	0.97	0.30	0.30	1401
Preferences							
Own willingness to take vaccine (%)	50.78	51.40	49.57	0.78	0.65	0.48	1401
Altruism	-0.01	0.04	-0.01	0.34	0.97	0.39	1401
Desire to influence	0.10	0.03	0.10	0.24	0.99	0.30	1401
Social image concerns	0.04	0.02	0.02	0.77	0.80	0.99	1401
Social proximity							
Oneness	-0.03	0.07	-0.04	0.13	0.92	0.14	1140
Test for joint significance				0.59	0.93	0.44	

Group means of Senders' predetermined characteristics, reported alongside p-values of tests for equal means. Results are based on the following model: $characteristic_i = \alpha + \beta \cdot treat_i + \epsilon_i$, where $treat_i$ is a dummy variable corresponding to either the *"informing partner after"* or the *"informing partner before"* condition, and where we omit one condition from our sample for every pair-wise comparison. *Not* refers to the *not informing partner* condition, *Before* and *After* are defined analogously. All variables classified as "local characteristics" do not vary on the individual level but on the zip code or town ("Gemeinde") of residence level.

[25]We use regressions of the following form to compare predetermined characteristics between pairs of conditions: $characteristic_i = \alpha + \beta \cdot treat_i + \epsilon_i$, where $treat_i$ is a dummy variable corresponding to either the *"informing partner after"* or the *"informing partner before"* condition, and where we omit one condition from our sample for every pair-wise comparison.

[26]We report the corresponding balancing results for Receivers in Table 2.B.6 in the Appendix.

[27]This is supported by the p-values obtained from tests for joint significance of all predetermined characteristics reported at the bottom of the table.

2.3 Empirical analysis

In this section, we first discuss our empirical strategy to separate anticipated peer effects from social image and other behavioral motives. Subsequently, we document that our experimental variation generated the desired first-stage effect, i.e., it manipulated Senders' beliefs about the likelihood of exerting a peer effect. After discussing our main result – that anticipated peer effects more than doubled Senders' likelihood of signing up for a COVID-19 vaccination – we move on to addressing several potential concerns including experimenter demand, strategic lying and cheating. We also corroborate our interpretation that Senders acted upon a prosocial anticipated peer effect, mediated by the possibility of influencing their partner, in a heterogeneity analysis and a 2SLS estimation. We end this section by documenting that Senders' anticipated peer effects did not translate into "actual" peer effects, i.e., behavioral changes among Receivers.

2.3.1 Regression specification

To identify experimental treatment effects, we estimate regression models of the following form:

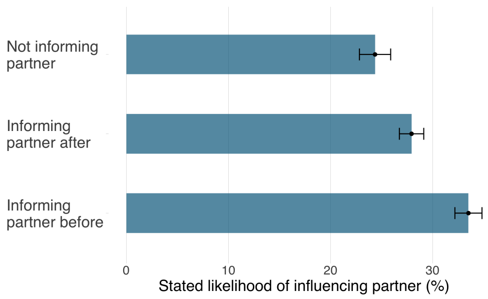
$$y_i = \beta_0 + \beta_1 \cdot \text{informing partner}_i + \beta_2 \cdot \text{informing partner before}_i + \boldsymbol{X_i}\ \gamma' + \epsilon_i \quad (2.1)$$

y_i corresponds to the relevant outcome of interest for Sender i. In our main specifications, y_i is a dummy variable indicating whether Sender i registered for a COVID-19 vaccination and could provide proof of her registration. When testing for the first-stage effect of our experiment, we instead use Sender i's belief about the likelihood of her being able to influence her partner as the outcome variable y_i. In alternative specifications, we also consider Sender i's self-reported registration status and willingness to take the vaccine as well as further measures of her decision to sign up for the vaccination, such as whether Sender i clicked on the link forwarding participants to the BayIMCO website, as dependent variables.

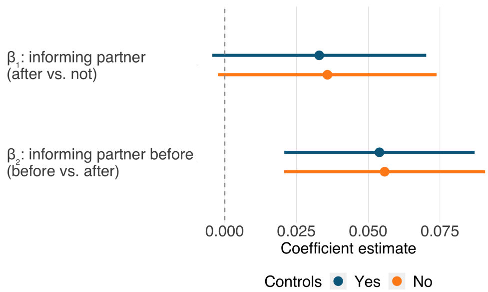
The variables *informing partner*$_i$ and *informing partner before*$_i$ capture the impact of our two experimental manipulations. First, *informing partner*$_i$ is an indicator variable taking value 1 if Sender i learned that we would report her registration decision to her partner in the experiment. Second, *informing partner before*$_i$ takes value 1 if Sender i learned that her partner would be informed about her registration decision *before* her partner himself would have the opportunity to sign up for the vaccination. When using both indicators, *informing partner*$_i$ and *informing partner before*$_i$, simultaneously as specified in Equation 2.1, β_1 captures the social image effect and β_2 the additional anticipated peer effect that only occurs if a Sender's partner was informed *before* rather than *after* his own registration decision. Finally, in some specifications we include control variables: $\boldsymbol{X}_i$ is a vector that includes Senders' predetermined characteristics.[28]

2.3.2 First stage

We first test for the presence of the intended first-stage effect: are Senders' beliefs about the likelihood of exerting a peer effect on their partner shifted upwards if they learn that we will inform their partner *before* rather than *after* the partner has the opportunity to register for the COVID-19 vaccination? The left-hand panel of Figure 2.2 reports the mean belief in each of the three experimental conditions. When we compare the upper (*"not informing partner"*) to the middle bar (*"informing partner after"*), we find that Senders in the latter group are slightly more likely to think that they can influence their partner. This is consistent with the idea that these Senders anticipate that their decision in the experiment might influence Receivers' registration behavior *after* the survey, even though their signal arrives too late for Receivers' decisions *within* the experiment.[29] More importantly, however, when contrasting the middle with the lower bar (*"informing partner before"*), we discover that Senders who learned that their partner was informed *before* rather than *after* are significantly more likely to believe that they can exert a peer effect on their partner.



(a) Conditional means



(b) Coefficient estimates

Figure 2.2: Treatment effects on first stage beliefs

Panel (a) plots Senders' mean stated likelihood of influencing their partner, in percent, for each treatment condition. Error bars indicate 95 percent confidence intervals of the mean. Panel (b) plots coefficient estimates and 95 percent confidence intervals from regressions as laid out in Equation 2.1.

[28]We use all Sender characteristics reported in Table 2.1 with the exception of our measure of social proximity as control variables. We exclude social proximity due to a number of missing observations for this measure from participants who have skipped the corresponding survey items. Our results do not change when including it as an additional control variable.

[29]Arguably, however, if all Senders in the *"not informing partner"* condition had fully understood the experimental instructions, they should have reported that they cannot influence their partner at all. That this belief is not zero is likely explained by some degree of inattention among participants which is not uncommon for this type of online experiment.

This finding is confirmed by our regression results depicted in the right-hand panel of Figure 2.2. In this figure, we report coefficient estimates and the corresponding 95 percent confidence intervals which we obtained from regressions following the specification depicted in Equation 2.1. Here, we employ Senders' beliefs about the likelihood of exerting a peer effect on their partners as the dependent variable. We present results from regressions both with and without controls.[30] The upper pair of coefficients in Figure 2.2 depicts our estimates for β_1 which corresponds to the difference between the upper and the middle bar in the left-hand panel of the figure. Our estimates for β_1 range from 3.29 (with controls, se: 1.90) to 3.58 (without controls, se: 1.94) percentage points and are both significant at the 10 percent level, corresponding to a 13 percent increase over the likelihood stated by Senders whose partners were not informed.

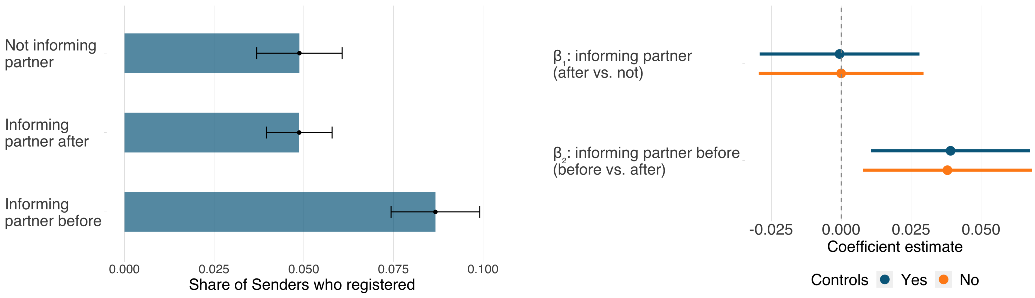
We estimate even larger treatment effects for β_2 (the lower pair of coefficients) which correspond to the difference in first-stage beliefs between the middle and the lower bar. Our estimates for β_2 range from 5.39 (se: 1.69) to 5.57 (se: 1.79) percentage points and are significant at the 1 percent level, irrespective of whether we include controls or not, corresponding to a 20 percent increase over the *"informing partner after"* condition. Taken together, these findings suggest that our experimental manipulations successfully induced the desired shift in beliefs: making Senders' decisions observable to their partners increased Senders' mean beliefs about the likelihood of being able to influence their partners. Crucially, however, informing Senders' partners before rather than after we offer them the opportunity to register induced a significant wedge in Senders' first-stage beliefs, which we leverage to isolate anticipated peer effects from social image effects.

2.3.3 Separating anticipated peer effects from social image effects

Next, we present treatment effects on Senders' likelihood to sign up for a COVID-19 vaccination. The left-hand panel of Figure 2.3 displays the share of Senders who verifiably registered across our three experimental conditions. We find that among Senders in the *"not informing partner"* and *"informing partner after"* conditions, 5 percent decided to sign up for a vaccination during our experiment. When contrasting this with Senders in the *"informing partner before"* condition, we find that Senders in this condition are roughly 80 percent more likely to register (9 vs. 5 percent).

We assess whether the differences in the share of Senders who signed up are statistically significant by running regressions of the form specified in Equation 2.1. We use a dummy variable taking value 1 for Senders who verifiably registered for a COVID-19 vaccination as the dependent variable. The upper pair of coefficients reported in Figure 2.3, right panel, corresponds to β_1 and as such captures the difference between

[30] Full regression results are reported in columns (3) and (4) of Table 2.B.1 in the Appendix.



(a) Conditional means **(b) Coefficient estimates**

Figure 2.3: Treatment effects on Senders' likelihood to register for a COVID-19 vaccination

Panel (a) plots the share of Senders who registered for a vaccination in each treatment condition. Error bars indicate 95 percent confidence intervals of the mean. Panel (b) plots coefficient estimates and 95 percent confidence intervals from regressions as laid out in Equation 2.1.

the upper and the middle bar in the left-hand panel. Irrespective of whether we use controls or not, we obtain precisely estimated zero effects for β_1 (se: 0.01-0.02).[31] This implies that Senders who knew that their partner was informed *after* he had obtained the opportunity to sign up for a vaccination (*"informing partner after"*) are not more likely to register than Senders who knew that their partner was not informed at all (*"not informing partner"*). In other words, Senders' decisions on whether to sign up for a COVID-19 vaccination seem not to be affected by social image concerns due to being observed by their partner in the experiment. In stark contrast to this, we obtain point estimates for β_2 of approximately 4 percentage points (se: 0.01-0.02), both with and without controls, which are highly significant ($p < 0.02$), implying that anticipated peer effects significantly affect Senders' likelihood to sign up.[32] These finding are confirmed by results from Fisher permutation tests summarized in Figure 2.A.1 in the Appendix, which yield a p-value of 0.97 for β_1 and of 0.02 for β_2.[33]

Combined, our findings depicted in Figure 2.3 thus suggest that observability per se does not induce a change in Senders' behavior. However, once Senders can influence their partners' decision whether to sign up for a COVID-19 vaccination within the experiment, they are almost twice as likely to register themselves. In other words, in this particular setting, anticipated peer effects seem to explain why Senders are more

[31] We report full regression results in columns (3) and (4) of Table 2.B.2 in the Appendix.

[32] Given the absence of a difference in the share of Senders who signed up between the *"informing partner after"* and the *"not informing partner"* condition, our estimates for β_2 also correspond to an 80 percent increase over the mean in the *"not informing partner"* condition.

[33] We pre-specified both conventional t-statistics as well as permutation tests for statistical inference in our pre-analysis plan. To derive Fisher p-values, we randomly assign "placebo treatment" status to Senders in our experimental conditions in 5,000 iterations and calculate a distribution of "placebo estimates" for both β_1 and β_2. We then compare the size of the treatment effects we find using the actual treatment assignment (the "true" estimate) to the distribution of "placebo estimates".

likely to act prosocially if their behavior can be observed by others. The absence of social image effects and the relative strength of anticipated peer effects in this context results from the fact that the scope for social image effects was limited by design: instead of leveraging a considerable number of individuals from the Senders' peer group as Receivers as in comparable studies interested in identifying social image effects, we matched Senders with only one stranger and let them interact in a quasi-anonymous online setting without the chance of future interactions.

2.3.4 Robustness

Experimenter demand

Taking a COVID-19 vaccine, and by extension also signing up for a vaccination, is generally perceived as a socially desirable action. Thus, we expect a certain baseline level of experimenter demand effects to be present in all experimental conditions. This type of experimenter demand does, however, not constitute a potential threat to our interpretation of the findings as long as the extent of experimenter demand is uniform across conditions. Yet, if Senders in the *"informing partner before"* condition inferred with a higher probability from our instructions that our experiment's main hypothesis was that a higher share of them would sign up for a COVID-19 vaccination, our estimates could, at least partially, reflect stronger experimenter demand in the *"informing partner before"* condition.

Previous research by de Quidt, Haushofer, and Roth (2018) and Haaland, Roth, and Wohlfart (forthcoming) found that self-reported outcomes are more prone to suffer from experimenter demand effects than revealed preference outcomes since the latter impose an actual economic cost on experimental subjects. We thus compare our estimates for β_1 and β_2 between regressions where we employ our revealed preference outcome (verified registrations) as the dependent variable and those where we use one of the following self-reported outcomes: first, a dummy taking value 1 if a Sender reported that she signed up, which we elicit after participants were offered the opportunity to register for a COVID-19 vaccination via BayIMCO; second, a dummy taking value 1 if a Sender clicked on the link forwarding her to the BayIMCO website; third, a dummy taking value 1 if a Sender replied that she is planning to sign up, which we elicit after Senders saw the treatment messages, yet before we offered them the opportunity to sign up; fourth, the change in a Sender's self-reported willingness to take the vaccine from before to after the treatment.

In Figure 2.4, we plot coefficient estimates and corresponding confidence intervals for both β_1 and β_2 obtained from regressions as specified in Equation 2.1 using our full

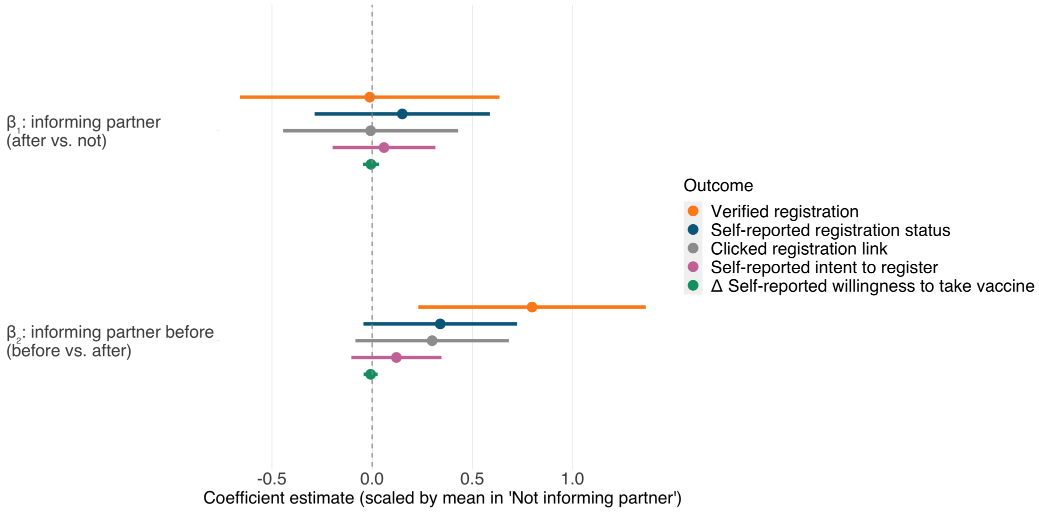


Figure 2.4: Comparing treatment effects across outcomes

Coefficient estimates from regressions as laid out in Equation 2.1. We use the following dependent variables: (i) a Sender's verified registration status; (ii) a dummy variable taking value 1 if a Sender reports that she has registered (elicited before verification); (iii) a dummy variable taking value 1 if a Sender clicked on the registration link forwarding her to BayIMCO; (iv) a dummy variable taking value 1 if a Sender reported to be willing to register (elicited before verification); and (v) the change in a Sender's self-reported willingness to take the vaccine (pre vs. post treatment). All outcomes are scaled using the corresponding mean in the *"not informing partner"* condition. 95 percent confidence intervals reported.

set of controls.[34] All outcomes are scaled using the corresponding mean in the *"not informing partner"* condition to facilitate the interpretation of coefficient sizes. Irrespective of whether we look at self-reported or verified outcomes, we find fairly precisely estimated zero effects for β_1. In contrast, while we find that our estimates for β_2 are positive and highly statistically significant if we employ our revealed preference outcome (verified registrations), we obtain insignificant and considerably smaller estimates when using self-reported outcomes. For example, when employing Senders' *self-reported* registration status as an outcome, we estimate that β_2 only corresponds to about a 30 percent increase over the mean (compared to about an 80 percent increase over the mean when using verified registrations as the dependent variable). We obtain even smaller estimates for β_2 for any of the other self-reported outcomes.[35] The pattern we observe in Figure 2.4 thus suggests that our experimental manipulations did not generate additional experimenter demand in the *"informing partner before"* condition that goes beyond any baseline experimenter demand present in all conditions.

[34]Coefficients and confidence intervals are barely affected by using control variables. Yet, estimates obtained from regressions using our full set of control variables are slightly more precise. Since smaller confidence intervals would, in this particular exercise, work against us when looking at self-reported outcomes, we decided to present results obtained from regressions with controls.

[35]We report full regression results in Table 2.B.3 in the Appendix.

Strategic lying

A related alternative explanation is strategic lying, given that the decision of whether to take the COVID-19 vaccine, and by extension whether to sign up for a vaccination, represents a collective action problem. While the vaccine entails several important benefits, including for society, it also comes with private costs for individuals, e.g., in terms of potential side effects or opportunity costs. Therefore, Senders have an incentive to state that they are willing to register – to nudge their partner to take the vaccine – without actually following through with the registration themselves. Strategic lying poses a threat to our interpretation if the extent of such behavior is more pronounced among Senders in the *"informing partner before"* condition.

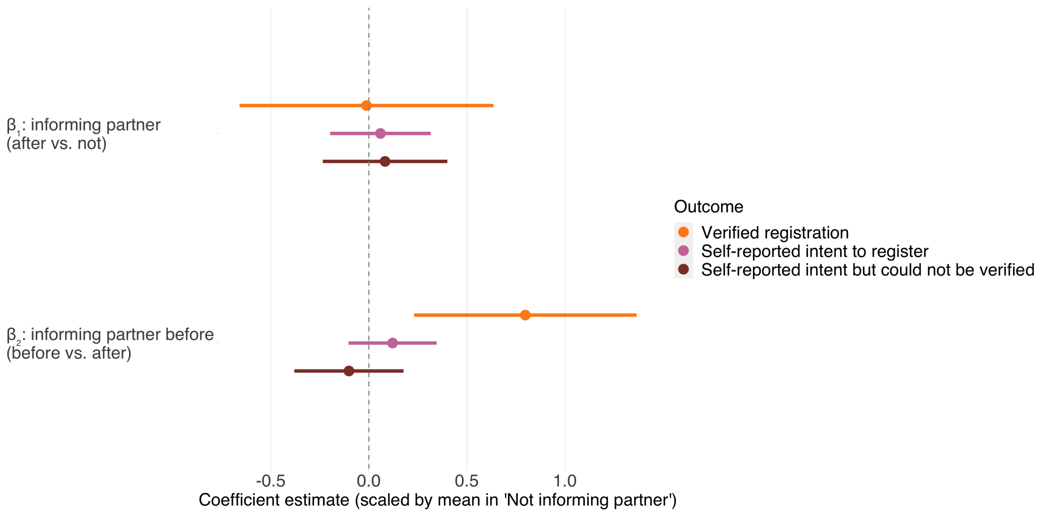


Figure 2.5: Treatment effects on strategic lying

Results from regressions of verified registration status, self-reported intent to register, and a measure of strategic lying on treatment indicators as laid out in Equation 2.1. Strategic lying is measured as a dummy variable taking value 1 for Senders who report that they have signed up but who failed to provide proof of their registration. We plot coefficient estimates and 95 percent confidence intervals. All outcomes are scaled using the corresponding mean in the *"not informing partner"* condition.

In Figure 2.5, we construct a measure of strategic lying – a dummy variable taking value 1 for Senders who reported that they had signed up but failed to provide proof of their registration – and use it as dependent variable in regressions following the general setup specified in Equation 2.1. We compare the coefficient estimates for β_1 and β_2 obtained for this measure of strategic lying with the corresponding estimates for verified registrations and Senders' self-reported intent to register.[36] To facilitate the comparison of effect sizes across outcomes, we scale all three outcomes with the respective mean in the *"not informing partner"* condition. We find fairly precisely estimated zero effects for β_1, irrespective of whether we use controls or not. In contrast,

[36]For full regression results, please consult the corresponding Table 2.B.4 in the Appendix.

we obtain positive and highly significant estimates for β_2 when employing verified registrations as the dependent variable, whereas we find no significant coefficients when using our measure of strategic lying as the dependent variable. The estimate for β_2 is small, statistically indistinguishable from zero, and if anything, negative. It is thus not the case that Senders in the *"informing partner before"* condition are merely more likely to state that they would like to register, while failing to follow through with their registration, than Senders in other conditions. Taken together, anticipated peer effects seem to be a sufficiently strong behavioral motive which reflects a true preference for prosociality rather than a strategic, and thus selfish, concern.

Cheating

A third potential threat concerns differential incentives to "cheat" which may arise from differences in treatment instructions or survey items employed across experimental conditions.[37] There are two related variants of this concern: first, there may exist differential incentives to sign up for a vaccination to qualify for the additional remuneration offered to participants passing our verification process. Second, Senders may face differential incentives to search the internet for the information required to qualify for the remuneration.

To address this type of concern, we exploit the fact that either variant should manifest in similar patterns with respect to how much time Senders devote to each of the survey pages post treatment. We would expect Senders who cheat to spend only a short time on the survey page where the registration should have taken place and considerably more time on the survey page where they are asked to provide proof of their registration: either because they need to register for the vaccination online ex post, i.e., after we had offered them to start their registration from within the survey on the previous page, or because they need to retrieve the address and the subject line of the confirmation email from the internet.

We compare time spent on each survey page after receiving the treatment message in Figure 2.6. We hereby focus only on Senders who could provide proof of their registration to assess the potential severity of (successful) cheating. We provide Senders with the opportunity to sign up for a COVID-19 vaccination on the "Registration" page and ask them to provide verifiable information on the next page ("Confirmation"), where we also informed them about the additional remuneration.[38] Therefore, any increase in time spent due to cheating would manifest on the "Confirmation" page. Yet, contrary to the notion of differential incentives to cheat explaining our findings, we find

[37] In our particular setting, survey items were identical across the *"informing partner before"* and the *"informing partner after"* conditions. Thus, differential incentives to cheat must arise from slight differences in the wording of experimental instructions.

[38] Note that there was no possibility to "return" to a previous page throughout the entire survey.

that participants in both relevant groups spent ca. six minutes inactive on our "Registration" page, a realistic duration to switch to the websites of the Bavarian health authorities and conduct the official registration there. Furthermore, as we document in Figure 2.6 and Table 2.B.5 in the Appendix, Senders in the *"informing partner before"* condition did not spend significantly more (or less) time on any of the pages post treatment than Senders in the *"informing partner after"* condition. Together, these results thus speak against the idea that differential incentives to cheat explain our main finding.

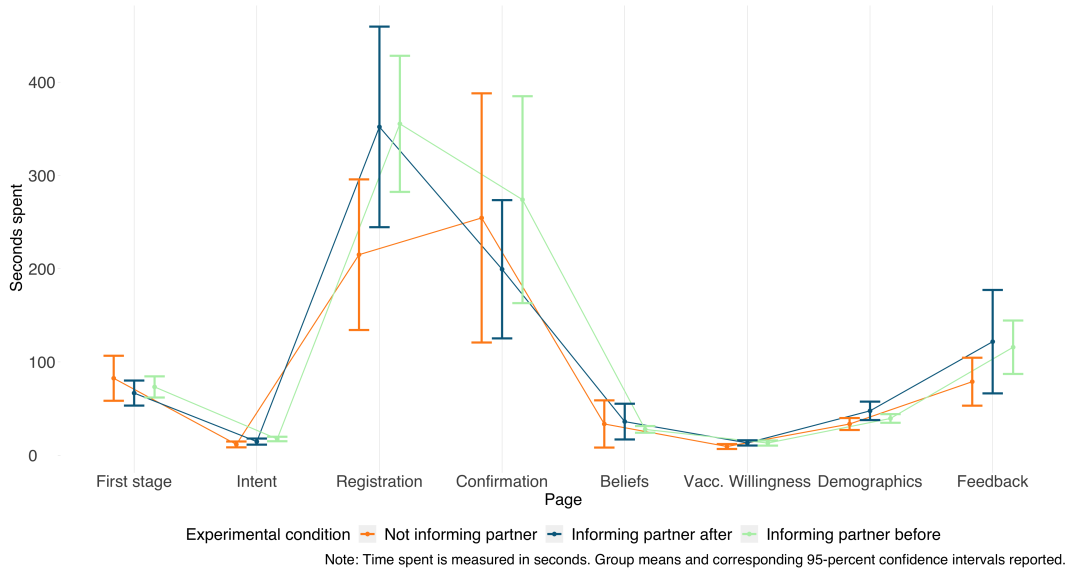


Figure 2.6: Time spent on each page post treatment by experimental condition
Senders' mean time spent on all survey pages after the treatment module, reported by treatment condition and alongside 95 percent confidence intervals. Time spent on each page is measured in seconds. The sample of Senders is limited to those who could provide proof of their registration.

2.3.5 Internal consistency

Heterogeneities

We now investigate potential explanations for why Senders are more willing to incur the cost of signing up during the experiment when they know that they can influence their partner. One such explanation is that Senders wish to send a signal about the quality of the vaccine to their partner, who they may suspect to be less informed. To assess this explanation, we analyze whether treatment effects depend on Senders' own beliefs about the quality of the vaccine. Specifically, we now limit our estimation sample to Senders in the *"informing partner before"* and *"informing partner after"* conditions and estimate the following regression model:

$$y_i = \gamma_0 + \gamma_1 \cdot \text{beliefs quality}_i + \gamma_2 \cdot \text{informing partner before}_i + \gamma_3 \cdot \text{informing partner before} \times \text{beliefs quality}_i + \boldsymbol{X_i}\,\boldsymbol{\theta'} + \epsilon_i \tag{2.2}$$

y_i and *informing partner before*$_i$ are defined as in Equation 2.1.[39] To proxy Sender i's beliefs about the quality of the vaccine, we employ the average of her pre-treatment beliefs about the safety and the efficacy of the vaccine.[40] The interaction term between our treatment indicator and Senders' beliefs about the quality of the vaccine (γ_3) captures whether the likelihood of registering due to anticipated peer effects becomes stronger if Senders are more convinced about the quality of the vaccine.

Table 2.4: Treatment effects conditional on beliefs about vaccine quality

	Verified registration	
	(1)	(2)
Beliefs quality	0.05***	0.01
	(0.01)	(0.01)
Informing partner before	0.04***	0.04***
	(0.01)	(0.01)
Informing partner before x Beliefs quality	0.05***	0.05***
	(0.02)	(0.02)
Controls		Yes
Mean 'Verified registration' (control)	0.05	0.05
Mean 'Beliefs quality' (control)	0	0
SD 'Beliefs quality' (control)	1	1
Observations	1,073	1,073
R^2	0.10	0.12

Results from regressions as laid out in Equation 2.2. Estimation sample limited to Senders in the *"informing partner before"* and *"informing partner after"* conditions. Thus, *"control"* refers to the *"informing partner after"* condition. *Beliefs quality* is standardized using the mean and standard deviation of the control group. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We report the main results from this exercise in Table 2.4 and Figure 2.A.2 in the Appendix where we standardize Senders' beliefs about the quality of the vaccine using the mean and standard deviation in the control group. Our estimates for γ_2 and γ_3 are positive and highly significant. Together, these estimates imply that Senders are more likely to sign up for a vaccination once they can influence their partner if they indeed

[39] See Section 2.3.1 for a detailed description of these variables.

[40] For both survey items we employed a 1–7 scale, where higher numbers represent stronger beliefs in the safety or efficacy of the vaccine.

believe that the vaccine is safe and effective. Conversely, Senders with the lowest levels of trust in the vaccine deliberately decided not to sign up during the experiment if they knew that they could influence their partner. This finding underlines our interpretation that anticipated peer effects constitute a prosocial motive in our setting: Senders who can influence their partners indeed choose to lead by example, signalling what they believe is best for their partner. Anticipated peer effects thus seem to arise if people think that leading by example sends an informative signal about the desirability of a certain action to observing individuals.

Table 2.5: Treatment effects conditional on social proximity between partners

	Verified registration	
	(1)	(2)
Social proximity	0.01	0.01
	(0.01)	(0.01)
Informing partner before	0.05***	0.05***
	(0.02)	(0.02)
Informing partner before x Social proximity	0.02	0.01
	(0.02)	(0.02)
Controls		Yes
Mean 'Verified registration' (control)	0.05	0.05
Mean 'Social proximity' (control)	0	0
SD 'Social proximity' (control)	1	1
Observations	877	877
R^2	0.02	0.13

Results from regressions as laid out in Equation 2.2. Estimation sample limited to Senders in the *"informing partner before"* and *"informing partner after"* conditions. Thus, *'control'* refers to the *"informing partner after"* condition. *Social proximity* is standardized using the mean and standard deviation of the control group. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Naturally, one may expect that our desire to lead by good example is more pronounced if observing individuals include friends and family or other individuals we care about. To test whether perceived social proximity affects the strength of anticipated peer effects in a prosocial setting, we use a measure of Sender i's perception of social proximity between herself and her partner (instead of her beliefs about the vaccine's quality) as the conditioning variable. Results from this exercise are summarized in Table 2.5 and in the corresponding Figure 2.A.3 in the Appendix, where we again standardize the conditioning variable using its mean and standard deviation in the control group. Although we find tentative evidence that anticipated peer effects increase in perceived social proximity, the interaction term between the dummy variable indicating that Sender i was assigned to the *"informing partner before"* condition and the perceived social proximity to her partner is not significant. However, in combination with the fact that more than 50 percent of Senders perceive their partner at least as an "acquaintance"[41], this finding is nevertheless consistent with the notion that anticipated peer effects are more likely to matter for prosocial behavior if people care about observing individuals.

2SLS estimates

The estimated impact of our experimental manipulations on Senders' likelihood to sign up for a COVID-19 vaccination we discussed so far constitutes a "reduced form" effect. Yet, as argued in previous sections, Senders' beliefs about the potential impact of their own decision on their partners – their first-stage belief – may be particularly important for the strength of anticipated peer effects. Therefore, we now investigate the relationship between our experimental manipulations, Senders' first-stage belief, and their willingness to sign up for a COVID-19 vaccination more systematically: to this end, we limit our estimation sample to Senders in the the *"informing partner before"* and *"informing partner after"* conditions and leverage our experimental manipulation as an instrument for Senders' beliefs about the likelihood of being able to exert a peer effect on their partner in a 2SLS framework. We thus estimate a local average treatment effect on those Senders whose beliefs were actually shifted by the experimental intervention.

We report the results from this exercise in Table 2.6. Columns (1) to (4) summarize our findings discussed in previous sections: Senders in the *"informing partner before"* condition exhibit a significantly higher probability to believe that they can influence their partner (columns (1) and (2)) and are significantly more likely to sign up for a COVID-19 vaccination (columns (3) and (4)) than Senders in the *"informing partner after"* condition. Then, in columns (5) and (6) we leverage our experimentally induced variation in Senders' beliefs about the likelihood of influencing their partner to compute the 2SLS estimate of this belief on Senders' likelihood to sign up for a COVID-19 vaccination. Irrespective of whether we use controls or not, we obtain positive and significant estimates for Senders' beliefs about the likelihood of exerting a peer effect on their partner. This confirms our view that anticipated peer effects are governed by Senders' beliefs about their chances of influencing their partner.

[41] For an explanation of the social proximity scale, see Section 2.2.

Table 2.6: Treatment effects estimated using 2SLS

	First Stage		Reduced Form		Second Stage	
	Likelihood of influencing partner		Verified registration			
	(1)	(2)	(3)	(4)	(5)	(6)
Informing partner before	5.57***	5.42***	0.05***	0.05***		
	(1.78)	(1.69)	(0.02)	(0.02)		
Likelihood of influencing partner					0.01**	0.01**
					(0.00)	(0.00)
Controls		Yes		Yes		Yes
Mean 'Likelihood of influencing partner' (control)	27.93	27.93	27.93	27.93	27.93	27.93
Mean 'Verified registration' (control)	0.05	0.05	0.05	0.05	0.05	0.05
F-Statistic for 1st stage			9.74	10.29		
Observations	911	911	911	911	911	911

Results from first-stage (columns (1) and (2)), reduced form (columns (3) and (4)), and 2SLS (columns (5) and (6)) estimations reported. The estimation sample is limited to Senders in the *"informing partner before"* and *"informing partner after"* conditions. Thus, *'control'* refers to the *"informing partner after"* condition. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.3.6 "Actual" peer effects on Receivers

As we have documented in the previous subsection, Senders' beliefs about the likelihood of exerting a peer effect on their partner play a central role in the strength of anticipated peer effects. Yet, it remains an open question as to whether Senders correctly anticipate such a peer effect on their partner. To investigate this question, we analyze whether Senders' decision to sign up for a COVID-19 vaccination actually influenced Receivers' behavior by running the following regression model with our full sample of Receivers:[42]

$$y_i = \phi_0 + \phi_1 \cdot \text{informed before}_i + \boldsymbol{X_i}\ \boldsymbol{\zeta'} + \epsilon_i, \tag{2.3}$$

y_i is a dummy variable which equals 1 if the Receiver decides in the same way as his partner and 0 otherwise. We use Receivers' decision to sign up for a COVID-19 as our main outcome and complement it, among others, with Receivers' self-reported intention to register. Our main explanatory variable informed before$_i$ is a dummy which equals one if Receiver i was informed about his partner's decision before we offered him the opportunity to sign up for a vaccination. Finally, in some specifications we include control variables which we denote by $\mathbf{X}_i$.[43]

[42]As we document in Table 2.B.6 in the Appendix, Receivers' predetermined characteristics are also well balanced. The test for joint significance of all predetermined characteristics yields a p-value of 0.6. Thus, we can be fairly confident that the treatment effects we estimate can actually be attributed to our experimental manipulations and did not arise to differences in predetermined characteristics.

[43]We use the same set of control variables as for our analysis of anticipated peer effects, i.e., all characteristics reported in Table 2.1 with the exception of our measure of social proximity.

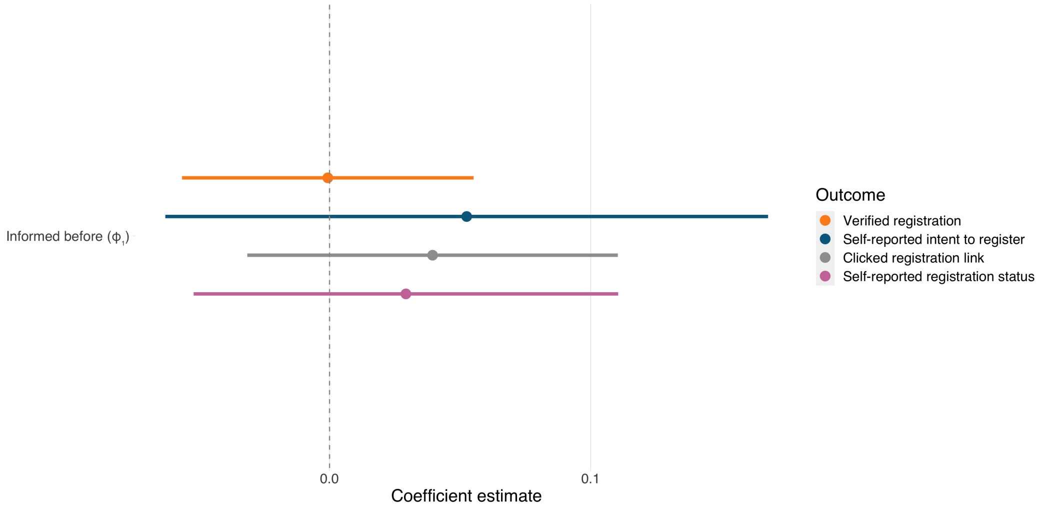


Figure 2.7: Comparing treatment effects on Receivers across experimental conditions

Results from regressions as laid out in Equation 2.3. Outcomes are indicator variables taking value 1 if Receivers decide the same way as their Senders with respect to the measures indicated in the legend. We plot coefficient estimates and 95 percent confidence intervals.

We present regression results in Figure 2.7.[44] The dependent variable is a dummy taking value 1 if the Receiver decided the same way as the Sender. Irrespective of whether we consider revealed-preference or self-reported outcomes, we find small and insignificant estimates throughout.[45]

Contrary to their own anticipation, Senders are thus not successful in influencing Receivers' behavior and attitudes. The absence of actual peer effects can be explained by the specific setting we study: Receivers learned about Senders' decisions, yet not vice versa. As a result, Senders could not verify whether they indeed exerted a peer effect on their partner and Receivers were thus not subject to social image concerns which could explain the missing conformity of Receivers' behavior. At a more general level, this suggests that in certain decision environments people might perceive themselves as more pivotal than they actually are, such that anticipated peer effects can arise even in absence of "actual" peer effects. Hence, leveraging people's desire to lead by example as a measure to promote prosocial behavior can work even if observing individuals do not follow suit.

[44]Full regression results underlying Figure 2.7 can be found in Table 2.B.7 in the Appendix.

[45]The same pattern emerges when we look at changes in attitudes or beliefs, which we report in Table 2.B.8 in the Appendix: three out of four coefficient estimates are statistically indistinguishable from zero. Only for the change in Receivers' beliefs about the severity of the freeriding problem in the roll-out of the mass immunization program do we obtain a negative and statistically significant coefficient (p-value < 0.1).

2.4 Conclusion

We provide evidence that anticipated peer effects constitute a relevant motive for prosocial behavior in a consequential decision environment. Leveraging a survey-based online experiment in the context of the COVID-19 vaccination campaign in Germany, we find that individuals' willingness to register for the vaccination almost doubles when informed that they can influence a peer's decision. We further document a strong first-stage effect of our treatment on subjects' beliefs about the chances of influencing their peer's decision, implying that individuals anticipate and internalize their potential to lead by example. Anticipated peer effects constitute a complementary behavioral mechanism explaining why people are more inclined to act prosocially if they can be observed by others, which operates independently of social image effects. Our findings further highlight that individuals are willing to incur considerable costs to send an encouraging signal to observing peers to follow their lead if they are convinced that an action can generate positive externalities. In line with this interpretation, we find that anticipated peer effects only increase individuals' willingness to register for a vaccination if they are sufficiently convinced about the quality of the vaccine.

The behavioral relevance of anticipated peer effects can hold relevant implications for policy makers seeking to promote prosocial behavior: instead of having to resort to "social sticks" in the form of enforcement or punishment (e.g., fines for missing vaccination appointments or maintenance of personal restrictions for unvaccinated people), they might leverage people's desire to lead by example as a "social carrot", e.g., by encouraging people to publicly signal their decision to get vaccinated. Increasing the benefits of behaving prosocially – rather than raising the costs of failing to do so – is also likely also beneficial from a welfare perspective.

While in this paper we provided evidence for the existence and empirical relevance of anticipated peer effects in a prosocial setting, future work could focus on exploring the underlying mechanisms in more detail. We can think of at least two potential drivers behind people's willingness to lead by good example: first, individuals might simply feel good about shaping the behavior of others and receive a hedonic payoff from leading by example. Second, in the spirit of theories of pure altruism (Andreoni 1989; Bénabou and Tirole 2006), individuals might care about the total provision of a public good (e.g., contributing to ending the pandemic). In that case, Senders could be motivated by an observability-dependent form of altruism, pushing them to set an example of prosocial behavior if they expect that others might follow suit and contribute to the public good as well.

Finally, our results have highlighted that anticipated peer effects can arise even without translating into peer effects, i.e., without actually affecting the decision of observing individuals. This has two interesting implications: first, it indicates that in some settings the mere potential of being able to influence others can be sufficient to promote prosocial behavior. Second, it implies that individuals might hold out-of-equilibrium beliefs about their impact on others. Understanding the sources of such misperceptions as well as their potential importance in motivating people to assume the responsibility of being social leaders – bearing the private costs of prosocial behavior without knowing that others might follow suit – constitutes another interesting avenue for future research.

Appendix 2

2.A Additional figures

2.A.1 Randomization inference

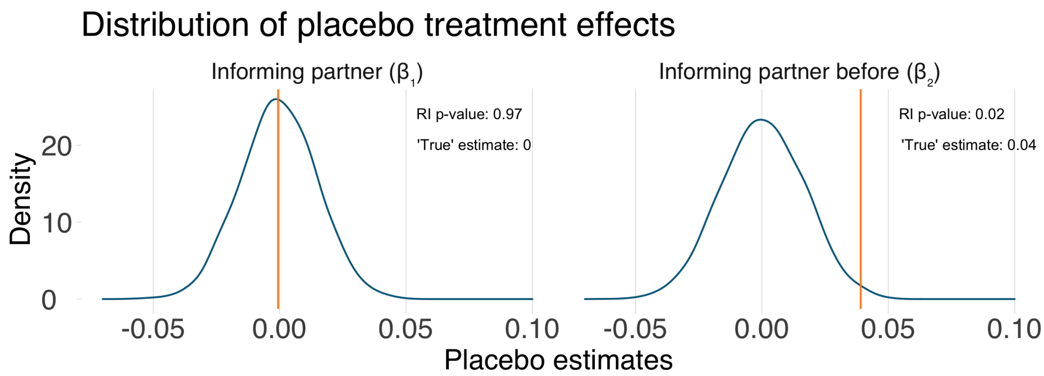


Figure 2.A.1: Results from randomization inference

Distribution of placebo estimates derived from randomly assigning "placebo treatment" status to Senders in 5,000 iterations and calculating the share of "placebo treatment effects" that exceed the "true treatment effect" in (absolute) magnitude. Panel (a) reports the resulting distribution and Fisher exact p-value for β_1 and panel (b) for β_2 as detailed in Equation 2.1. The outcome in both panels is Senders' verified registration status.

2.A.2 Heterogeneities

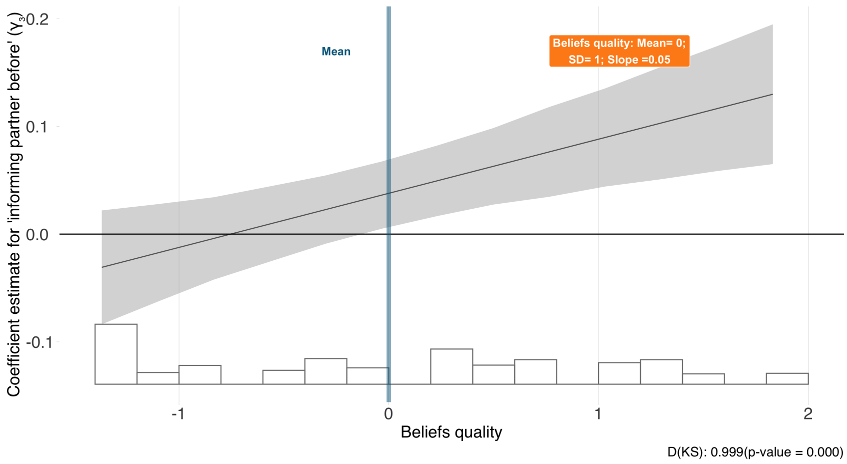


Figure 2.A.2: Treatment effect for *informing partner before* conditional on beliefs about vaccine quality

Treatment effect heterogeneity based on Equation 2.2. Beliefs about the quality of the vaccine are employed as the conditioning variable and are measured as the standardized average of Senders' beliefs about (i) the safety and (ii) the efficacy of the vaccine. The horizontal axis depicts the distribution of this belief among Senders.

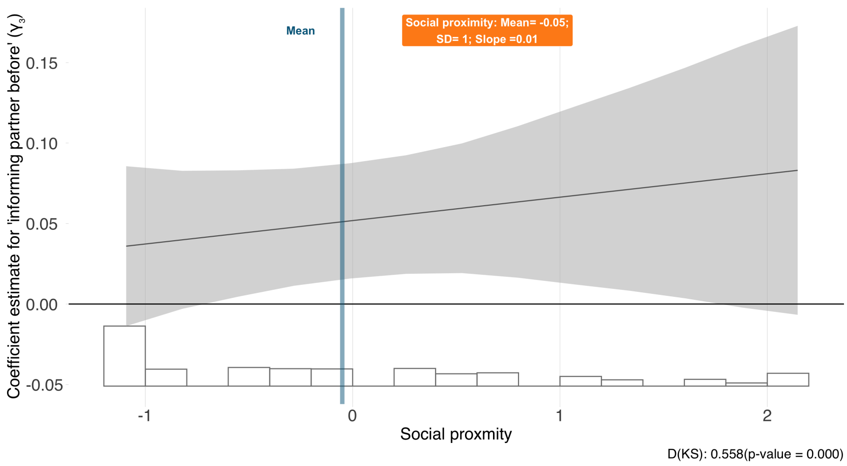


Figure 2.A.3: Treatment effect for *informing partner before* conditional on perceived social proximity to partner

Treatment effect heterogeneity based on Equation 2.2. Social proximity to the partner is employed as the conditioning variable and is measured using the "oneness" scale (see Section 2.2). The horizontal axis depicts the distribution of social proximity among Senders.

2.B Additional tables

2.B.1 First stage

Table 2.B.1: Treatment effects on first stage beliefs

	Likelihood that partner can be influenced (%)			
	(1)	(2)	(3)	(4)
Informing partner	6.26***	5.90***	3.58*	3.29*
	(1.77)	(1.73)	(1.94)	(1.90)
Informing partner before			5.57***	5.39***
			(1.79)	(1.69)
Controls		Yes		Yes
Mean, 'Not informing partner'	24.36	24.36	24.36	24.36
Mean, 'Informing partner after'	27.93	27.93	27.93	27.93
Observations	1,194	1,194	1,194	1,194
R^2	0.01	0.11	0.02	0.12

Results derived from regressions as laid out in Equation 2.1 with Senders' beliefs about the likelihood that their partner can be influenced as dependent variable. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.B.2 Separating anticipated peer effects from social image effects

Table 2.B.2: Treatment effects on signing up for a COVID-19 vaccination

	Verified registration			
	(1)	(2)	(3)	(4)
Informing partner	0.02	0.02	−0.00	−0.00
	(0.01)	(0.01)	(0.02)	(0.01)
Informing partner before			0.04**	0.04***
			(0.02)	(0.01)
Controls		Yes		Yes
Mean, 'Not informing partner'	0.05	0.05	0.05	0.05
Mean, 'Informing partner after'	0.05	0.05	0.05	0.05
Observations	1,401	1,401	1,401	1,401
R^2	0.00	0.10	0.01	0.11

Results derived from regressions as laid out in Equation 2.1 with Senders' verified registration status as dependent variable. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.B.3 Addressing potential concerns

Table 2.B.3: Comparing treatment effects across outcomes

	Verified registration	Δ Self-reported willingness to take vaccine	Self-reported intent to register	Clicked reg. link	Self-reported registration status
	(1)	(2)	(3)	(4)	(5)
Informing partner	−0.00	−0.00	0.01	−0.00	0.01
	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)
Informing partner before	0.04***	−0.00	0.02	0.03	0.03*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Controls	Yes	Yes	Yes	Yes	Yes
Mean, 'Not informing partner'	0.049	0.494	0.198	0.088	0.095
Mean, 'Informing partner after'	0.049	0.505	0.217	0.088	0.11
Observations	1,401	1,401	1,401	1,401	1,401
R^2	0.11	0.85	0.21	0.12	0.13

Results derived from regressions as laid out in Equation 2.1. We employ the following dependent variables: a dummy variable taking value 1 if a Sender reported that she registered for the vaccination and could provide proof of her registration (column 1); the change in a Sender's self-reported willingness to take the vaccine, pre vs. post treatment (column 2); a dummy variable taking value 1 if a Sender reported to be willing to register, elicited before verification (column 3); a dummy variable taking value 1 if a Sender clicked on the registration link forwarding her to BayIMCO (column 4); and a dummy variable taking value 1 if a Sender reported that she had registered, elicited before verification (column 5). Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.4: Strategic lying

	Self-reported intent to register	Verified registration	Self-reported intent not verified
	(1)	(2)	(3)
Informing partner	0.01	−0.001	0.01
	(0.03)	(0.01)	(0.02)
Informing partner before	0.02	0.04***	−0.02
	(0.02)	(0.01)	(0.02)
Controls	Yes	Yes	Yes
Mean, 'Not informing partner'	0.198	0.049	0.149
Mean, 'Informing partner after	0.194	0.044	0.149
Observations	1,401	1,401	1,401
R^2	0.21	0.11	0.11

Results derived from regressions as laid out in Equation 2.1. We employ the following dependent variables: a dummy variable taking value 1 if a Sender reported to be willing to register, elicited before verification (column 1); a dummy variable taking value 1 if a Sender reported that she registered for a vaccination and could provide proof of her registration (column 2); a dummy variable taking value 1 if a Sender reported that she had signed up but failed to provide proof of her registration (column 3). Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.5: Cheating (time spent on survey pages in seconds)

	Before vs. After		Before vs. Not		After vs. Not	
	Diff.	p-value	Diff.	p-value	Diff.	p-value
First Stage						
First stage	6.57	0.448	-9.3	0.460	-15.87	0.229
Main Outcomes						
Intent	2.785	0.165	5.815***	0.003	3.03	0.169
Waitpage	-0.104	0.177	0.044	0.160	0.148**	0.040
Registration	3.274	0.959	140.311***	0.009	137.037**	0.040
Confirmation	82	0.215	19.563	0.814	-62.438	0.388
Additional Outcomes						
Beliefs	-8.37	0.378	-5.833	0.624	2.537	0.867
Vacc. Willingness	-0.141	0.942	3.732**	0.047	3.873**	0.042
Demographics						
Demographics	-8.156	0.130	5.962	0.118	14.118**	0.017
Feedback						
Feedback	-8.541	0.779	34.387*	0.068	42.928	0.155

Differences in Senders' mean time spent (in seconds) on all pages after the treatment module, reported alongside p-values of tests for equal means. Results are based on the following regression model: $time\ spent_i = \alpha + \beta \cdot treat_i + \epsilon_i$, where $treat_i$ is a dummy variable corresponding to either the *informing partner after* ("After") or the *informing partner before* ("Before") condition, omitting one condition for every pair-wise comparison. "Not" refers to the *not informing partner* condition. Significance levels computed based on robust standard errors. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.B.4 Balancing receivers

Table 2.B.6: Receivers' predetermined characteristics compared across treatment conditions

	Group means		p-value	
	Before	After	Before vs. After	N
Attrition				
Completed survey	0.73	0.71	0.50	635
Demographics				
Age	39.76	42.03	0.10	456
Female	0.56	0.56	0.86	456
Income	2861.59	3103.39	0.12	456
Upper secondary degree	0.37	0.37	0.92	456
Local characteristics				
Avg. incidence rate (during second wave)	139.21	137.11	0.58	456
Population in zip	15.30	14.70	0.52	456
Lives in urban area ($\geq$100k)	0.29	0.23	0.20	456
Turnout (%)	77.45	77.57	0.76	456
AfD vote share (%)	11.99	12.53	0.06*	456
Unemployment rate (%)	2.39	2.25	0.10	456
Beliefs				
Safety	3.50	3.42	0.68	456
Efficacy	3.92	3.79	0.48	456
Social desirability	3.74	3.59	0.48	456
Severity of freerider problem	3.37	3.26	0.59	456
Willingness to take vaccine in state (%)	60.74	59.84	0.64	456
Preferences				
Own willingness to take vaccine (%)	53.22	52.91	0.93	456
Altruism	-0.01	-0.01	0.93	456
Desire to influence	0.06	0.15	0.31	456
Social image concerns	0.00	0.05	0.60	456
Social proximity				
Oneness	0.16	-0.05	0.04**	386
Test for joint significance				
			0.6	456

Group means of Receivers' predetermined characteristics, reported alongside p-values of tests for equal means. Results are based on the following regression model: $characteristic_i = \alpha + \beta \cdot informed\ before_i + \epsilon_i$, where *informed before*$_i$ is a dummy taking value 1 for all Receivers in the *informed before* condition. All variables classified as "local characteristic" do not vary on the individual but on the zip code or town ("Gemeinde") of residence level.

2.B.5 'Actual' peer effects on Receivers

Table 2.B.7: Treatment effects on Receivers: Registration outcomes

	$\mathbb{I}$[Receiver decides like Sender]			
	Verified registration	Self-reported intent to register	Clicked reg. link	Self-reported registration status
	(1)	(2)	(3)	(4)
Informed before	−0.00	0.04	0.03	0.02
	(0.02)	(0.04)	(0.03)	(0.03)
Controls	Yes	Yes	Yes	Yes
Mean dependent variable, 'informed after'	0.919	0.712	0.86	0.835
Observations	456	456	456	456
R^2	0.03	0.07	0.06	0.03

Results derived from regressions as laid out in Equation 2.3. Outcomes are indicator variables taking value 1 if Receivers decide the same way as their Senders with respect to the following measures: verified registration status (column 1); self-reported intent to register (column 2); click on link forwarding participant to BayIMCO (column 3); and self-reported registration status (column 4). Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.8: Treatment effects on Receivers: Changes in attitudes and beliefs

	Change in attitudes/beliefs			
	Δ Self-reported willingness to take vaccine	Δ Beliefs safety & efficacy	Δ Beliefs freeriding	Δ Beliefs image
	(1)	(2)	(3)	(4)
Informed before	−0.00	−0.08	−0.19*	−0.23
	(0.01)	(0.08)	(0.10)	(0.16)
Controls	Yes	Yes	Yes	Yes
Mean dependent variable, 'informed after'	0.002	0.307	0.089	1.924
Observations	456	456	456	456
R^2	0.02	0.06	0.04	0.39

Results derived from regressions as laid out in Equation 2.3. We consider the following dependent variables which are all defined as changes from before to after the treatment: willingness to take the vaccine (column 1); beliefs about safety and efficacy of the vaccine (column 2); beliefs about the severity of free-riding in the context of the vaccination program (column 3); and beliefs about the social desirability of the vaccine (column 4). Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.C Screenshots

2.C.1 Joint problem solving task

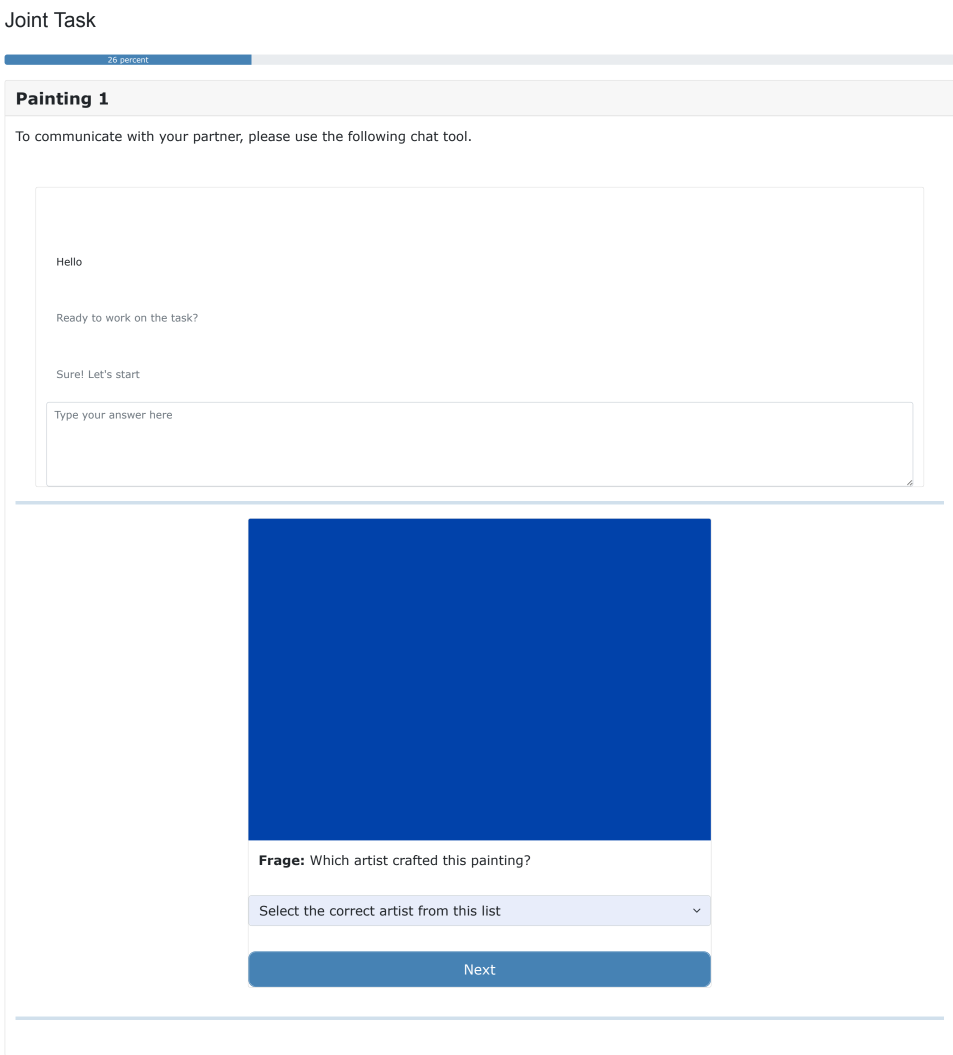
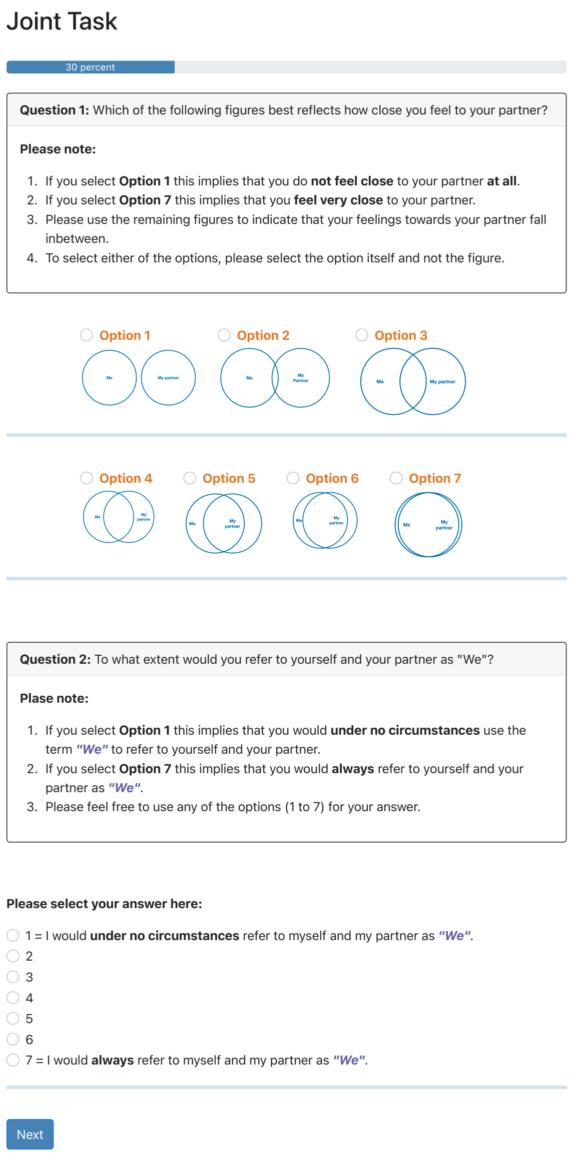


Figure 2.C.1: Survey page showing chat window and historical painting (placeholder)

2.C.2 Oneness elicitation


Joint Task

30 percent

Question 1: Which of the following figures best reflects how close you feel to your partner?

Please note:

1. If you select **Option 1** this implies that you do **not feel close** to your partner **at all**.
2. If you select **Option 7** this implies that you **feel very close** to your partner.
3. Please use the remaining figures to indicate that your feelings towards your partner fall inbetween.
4. To select either of the options, please select the option itself and not the figure.

Option 1 Option 2 Option 3

Option 4 Option 5 Option 6 Option 7

Question 2: To what extent would you refer to yourself and your partner as "We"?

Plase note:

1. If you select **Option 1** this implies that you would **under no circumstances** use the term ***"We"*** to refer to yourself and your partner.
2. If you select **Option 7** this implies that you would **always** refer to yourself and your partner as ***"We"***.
3. Please feel free to use any of the options (1 to 7) for your answer.

Please select your answer here:

1 = I would **under no circumstances** refer to myself and my partner as ***"We"***.
2
3
4
5
6
7 = I would **always** refer to myself and my partner as ***"We"***.

Next

Figure 2.C.2: Survey page showing oneness elicitation

2.C.3 Registration for COVID-19 vaccination

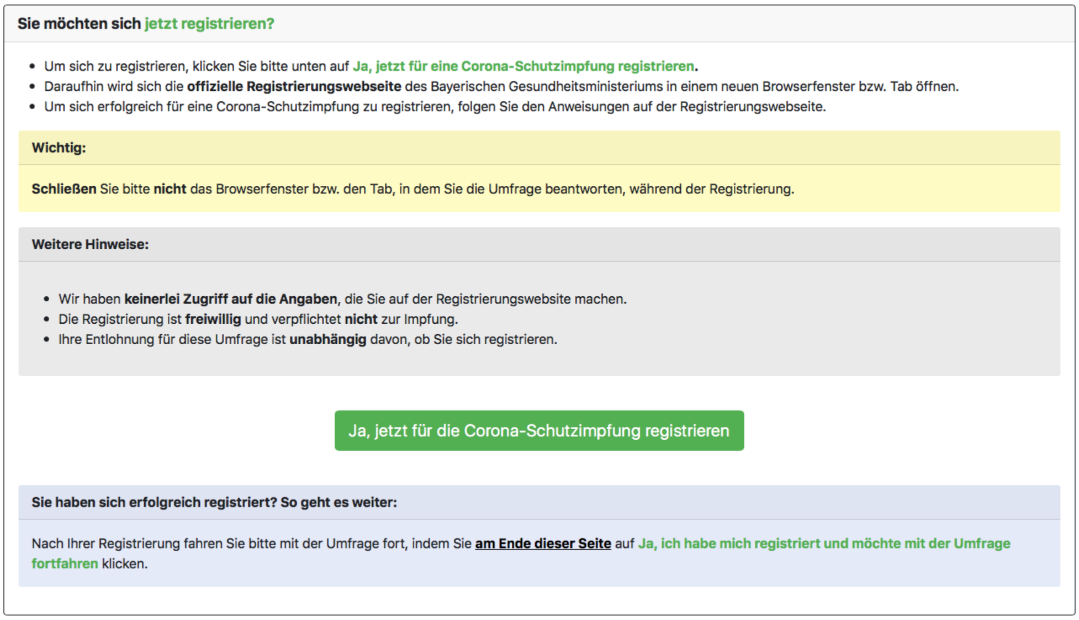


Figure 2.C.3: Survey page measuring registration intent and providing link to official registration website (BayIMCO)

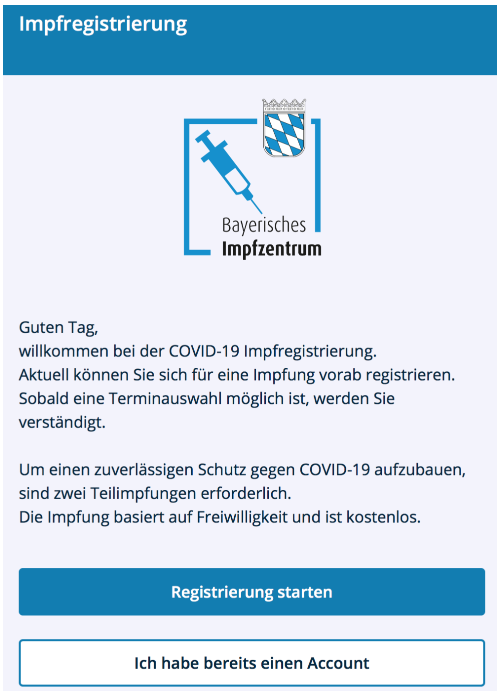
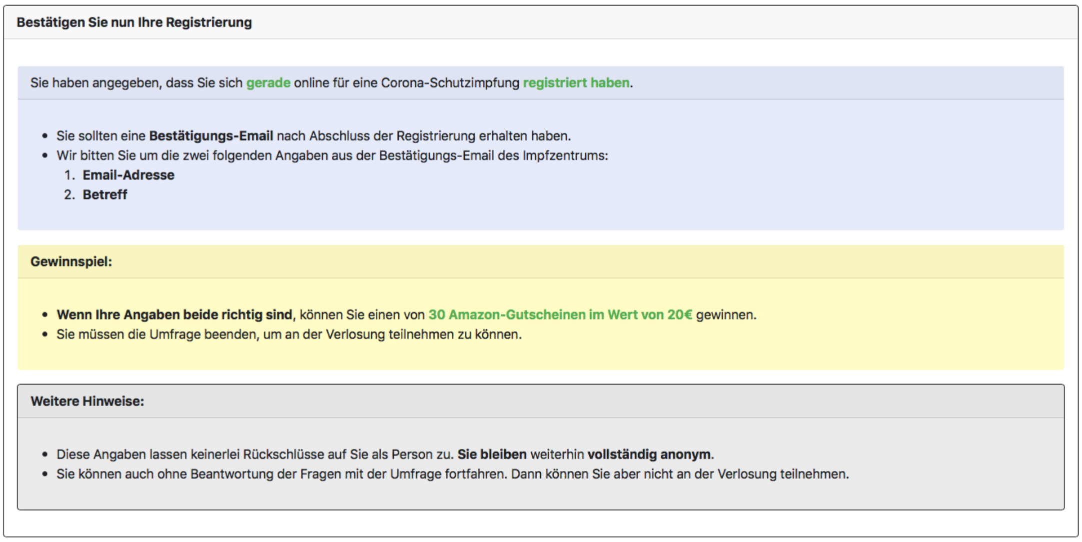


Figure 2.C.4: Starting page of the official registration process (BayIMCO)



Bestätigen Sie nun Ihre Registrierung

Sie haben angegeben, dass Sie sich **gerade** online für eine Corona-Schutzimpfung **registriert haben**.

- Sie sollten eine **Bestätigungs-Email** nach Abschluss der Registrierung erhalten haben.
- Wir bitten Sie um die zwei folgenden Angaben aus der Bestätigungs-Email des Impfzentrums:
  1. **Email-Adresse**
  2. **Betreff**

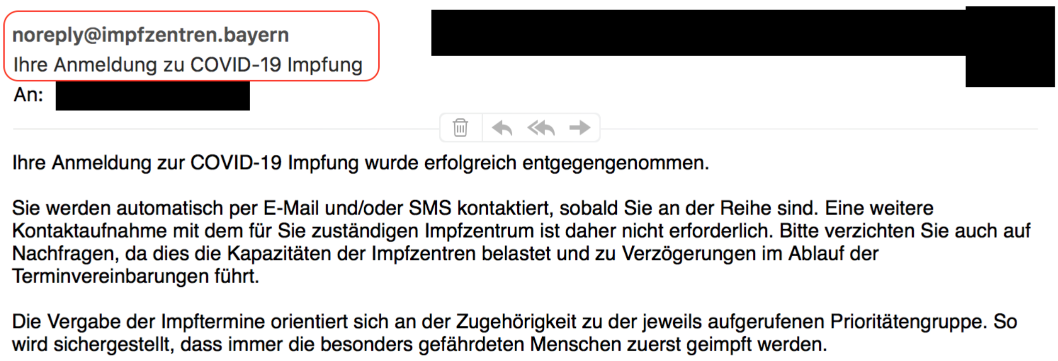
Gewinnspiel:

- **Wenn Ihre Angaben beide richtig sind**, können Sie einen von **30 Amazon-Gutscheinen im Wert von 20€** gewinnen.
- Sie müssen die Umfrage beenden, um an der Verlosung teilnehmen zu können.

Weitere Hinweise:

- Diese Angaben lassen keinerlei Rückschlüsse auf Sie als Person zu. **Sie bleiben** weiterhin **vollständig anonym**.
- Sie können auch ohne Beantwortung der Fragen mit der Umfrage fortfahren. Dann können Sie aber nicht an der Verlosung teilnehmen.

Figure 2.C.5: Survey page explaining verification of registration



noreply@impfzentren.bayern [redacted]
Ihre Anmeldung zu COVID-19 Impfung
An: [redacted]

Ihre Anmeldung zur COVID-19 Impfung wurde erfolgreich entgegengenommen.

Sie werden automatisch per E-Mail und/oder SMS kontaktiert, sobald Sie an der Reihe sind. Eine weitere Kontaktaufnahme mit dem für Sie zuständigen Impfzentrum ist daher nicht erforderlich. Bitte verzichten Sie auch auf Nachfragen, da dies die Kapazitäten der Impfzentren belastet und zu Verzögerungen im Ablauf der Terminvereinbarungen führt.

Die Vergabe der Impftermine orientiert sich an der Zugehörigkeit zu der jeweils aufgerufenen Prioritätengruppe. So wird sichergestellt, dass immer die besonders gefährdeten Menschen zuerst geimpft werden.

Figure 2.C.6: Confirmation email highlighting sending address and subject line

2.D Survey instrument

I Basic demographic information

Question 1: Are you male or female?
Question 2: How old are you?
Question 3: In which federal state do you live?

new page

Since the end of last year (December 2020), *vaccinations against the coronavirus* (COVID-19 vaccinations) have been administered in Germany.

Question: Have you already received a COVID-19 vaccination? Reply options: *Yes or No*

new page

Did you know that?
In Bavaria, it is possible to register for a COVID-19 vaccination already, even though the actual vaccination may not take place for a few months. Registration takes place either online or by telephone at the Bavarian vaccination centres.

Question: Have you already registered for a COVID-19 vaccination? Reply options: *Yes or No*

new page

II Attitudes towards the COVID-19 vaccination

We would like to start by asking you a few basic questions regarding how you feel about the COVID-19 vaccination.
There are now several vaccines against the coronavirus on the German market. Vaccination is officially recommended for adults of all ages (exception: not during pregnancy and breast-feeding for the time being, as no data on safety and efficacy are yet available).

To what extent do you agree with the following statements?

- **Statement 1:** I have full confidence that vaccination against COVID-19 is safe.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*
- **Statement 2:** I have full confidence that vaccination against COVID-19 is effective.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*
- **Statement 3:** I see vaccination as a collective effort against the spread of COVID-19.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*
- **Statement 4:** If everyone is vaccinated against COVID-19, I don't need to get vaccinated too.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*

Question 1: How likely is it that you will get vaccinated against COVID-19?
Instruction: Please use the bar/slider for your answer. Click on the bar at the bottom to reveal the slider. Then move the slider to give your answer. 0 percent means "definitely not willing to get vaccinated". 100 percent means "definitely willing to get vaccinated".

Question 2: What do you think? What proportion of people in Bavaria are willing to get vaccinated against COVID-19?
Instruction: Please use the bar/slider for your answer. Click on the bar at the bottom to reveal the slider. Then move the slider to give your answer. 0 percent means "no one is willing to get vaccinated". 100 percent means "everybody is willing to get vaccinated".

new page

III Broader set of attitudes

How well do the following statements apply to you as a person?

- **Statement 1:** I like it when people accept my suggestions.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely*.
- **Statement 2:** I like it when my ideas and opinions have an impact on other people.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*
- **Statement 3:** I would like the feeling of having influenced other people's lives.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*

new page

How well do the following statements apply to you as a person?

- **Statement 1:** It is important to me to impress others.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely*.
- **Statement 2:** I think a lot about whether I am good enough compared to others.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*
- **Statement 3:** It is important to me how I am perceived by others.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*

new page

We now ask you about your behavior in certain situations.

Question: How much would you be willing to give to a good cause without expecting anything in return?
Reply options: *0: Not at all willing, 10: Extremely willing*

Imagine the following situation: Today you unexpectedly received 1,000 EUR.
Question: How much of the money would you donate to a good cause? *Note: You can enter whole numerical values from 0 to 1,000 here.*

new page

IV Joint task

Please read the following instructions carefully before proceeding with the survey.

- In the next section of our survey, we ask you to solve a short task together with another participant of this survey.
- Your task is to match famous pieces of art to the respective artist together with your partner.
- In this task, you can win one of 30 Amazon vouchers worth 10 EUR.
- You can communicate with your fellow player by means of a chat.
- To facilitate communication, please enter your first name or a nickname below.

Question: What is your first name or nickname?

Hint:

- In order to remain anonymous, please make sure to enter only your first name.
- You can also choose another name here. However, the name should correspond to your gender.

——————————— *new page* ———————————

We ask you to solve the upcoming task together with your partner.
Your partner is: [*name*]

[He/she] is [*xx*] years old. [He/she] lives in Bavaria.

Task: Together with your partner, match the following four pieces of art with the correct artist.

Hints:

1. You and your partner have 60 seconds for each piece of art.
2. If you and your partner correctly match at least three pieces of art, you can win one of 30 Amazon vouchers worth 10 EUR.
3. You must complete the full survey to qualify for one of the vouchers.
4. To increase your chances of winning, it is important that you and your partner work together.
5. You will receive points only if you both give the correct answer.
6. Use the chat window to communicate with your partner via text messages and coordinate your answers. The chat window is available for the entire task.
7. Its a good idea to introduce yourself to your partner with a short message right away.

[***Chat window***]

Final hints before the tasks begins: You may have to wait for a moment until your partner [*name*] has read the instructions and responds to you.
Reminder: You can win one of 30 Amazon vouchers worth 10 EUR.

——————————— *new page* ———————————

[*Painting is shown for 1 Minute.*]

Question: Which artist painted this piece of art?
Reply options: *Participants can choose one artist from a drop-down menu.*

[*This process is repeated four times. During this time the participants have the option to use the chat window to communicate.*]

new page

Question: Which of the following figures best reflects how connected you feel with your partner [*name*]?

Hints:

1. Option 1 means that you do not feel connected to your partner [*name*] at all.
2. Option 7 means that you feel very close to your partner [*name*].
3. Use the remaining options (2-6) to grade your answer.
4. To select one, click on the option in the header and not on the image.

new page

Please still think of your partner [*name*].
Question: To what extent would you refer to yourself and your partner [*name*] as "we".

Hints:

1. Option 1 means that you would definitely not refer to the two of you as "we".
2. Option 7 means that you would definitely speak refer to the two of you as "we".
3. Use the remaining options (2-6) to grade your answer.

new page

V Explanations on the survey

Instructions: In the following, we would like to ask you about your willingness to get vaccinated against COVID-19. Specifically, we would like to know whether you are willing to register for a COVID-19 vaccination right away. With that we are referring to the official registration process required for residents of Bavaria to be able to obtain an appointment at a vaccination center. In this survey, we will provide you with the opportunity to switch to the official registration website of the Bavarian Ministry of Health to complete the registration. Of course, the registration is voluntary and you can also complete the survey without registering.

Task: Confirm that you have understood these instructions by selecting the correct answer below.

Question: During this survey, will you be able to switch to the official registration website of the Bavarian Ministry of Health to complete the registration for a COVID-19 vaccination?
Reply options: *Yes or No*

new page

V.A Instructions Senders *"not informing partner"*

Instructions:
The survey proceeds as follows:

Step 1: You decide whether you want to register for a COVID-19 vaccination right away.
Step 2: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.

Important: We do not tell your partner [*name*] whether you want to register for a vaccination.

You do not find out about the decision of your partner [*name*].

Task: Confirm that you have understood the instructions by selecting the correct answer below.
Question: Will your partner [*name*] find out whether you want to register?
Reply options: *Yes/No*

V.B Instructions Senders *"informing partner after"*

Instructions:
We will tell your partner [*name*] with a high probability whether you want to register for a vaccination. This proceeds as follows:

Step 1: You decide whether you want to register for a COVID-19 vaccination right away.
Step 2: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.
Step 3: We tell your partner [*name*] whether you want to register for a vaccination.
Important: Your partner [*name*] will find out about your registration decision **only after** [he/she] has already decided whether [he/she] wants to register.

You do not find out about the decision of your partner [*name*].

Task: Confirm that you have understood the instructions by selecting the correct answers below.

Question 1: Will your partner [*name*] find out with a high probability whether you want to register?
Reply options: *Yes/No*
Question 2: When will your partner [*name*] find out about your registration decision? **Directly before** or **only after** [he/she] can register for a COVID-19 vaccination?
Reply options: *Directly before/Only after*

V.C Instructions Senders *"informing partner before"*

Instructions:
We will tell your partner [*name*] with a high probability whether you want to register for a vaccination. This proceeds as follows:

Step 1: You decide whether you want to register for a COVID-19 vaccination right away.
Step 2: We tell your partner [*name*] whether you want to register for a vaccination.

Important: Your partner [*name*] will find out about your registration decision **directly before** [he/she] can decide whether [he/she] wants to register.
Step 3: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.

You do not find out about the decision of your partner [*name*].

Task: Confirm that you have understood the instructions by selecting the correct answers below.

Question 1: Will your partner [name] find out with a high probability whether you want to register?
Reply options: *Yes/No*
Question 2: When will your partner [*name*] find out about your registration decision? **Directly before** or **only after** [he/she] can register for a COVID-19 vaccination?
Reply options: *Directly before/Only after*

V.D Instructions Receivers *"informed before"* and *"informed after"*

Instructions: The survey proceeds as follows:

Step 1: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.
Step 2: You decide whether you want to register for a vaccination now. Since you are the second to decide you may have to wait for a moment.

We do not tell your partner [*name*] whether you want to register for a vaccination.

Task: Please confirm that you have understood these instructions by selecting the correct answer below.
Question: Will your partner find out about your decision?
Reply options: *Yes/No*

new page

VI Vaccination willingness

VI.1.A First stage Senders *"not informing partner"*

Reminder: Below we will provide you and your partner [*name*] with the opportunity to go to the official registration website of the Bavarian Ministry of Health to complete the registration process.

Your partner [*name*] will not know whether you wish to register for a COVID-19 vaccination.

Remember: Your partner [*name*] **will not** learn about your registration decision.
Question 1: What do you think? How likely is it that your decision to register or not to register will influence your partner's decision?

Hints:

- Click on the bar at the bottom to reveal the slider.

- Then move the slider to give your answer.
- 0 percent means "there is no way I can influence my partner with my decision".
- 100 percent means "I can definitely influence my partner with my decision".

Remember: Your partner [*name*] **will not** learn about your registration decision.
Question 2: What do you think? How likely is it that your partner will make the same decision as you?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means “my partner will definitely not decide the same way I do”.
- 100 percent means “my partner will definitely decide like me” .

VI.1.B First stage Senders *“informing partner after”*

Reminder: Below we will provide you and your partner [*name*] with the opportunity to go to the official registration website of the Bavarian Ministry of Health to complete the registration process.

Your partner [*name*] will learn with a high probability whether you wish to register for a COVID-19 vaccination.

Remember: Your partner [*name*] will learn about your registration decision only after [he/she] has already decided whether to register for COVID-19 vaccination now.
Question 1: What do you think? How likely is it that your decision to register or not to register will influence your partner′s decision?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means “there is no way I can influence my partner with my decision”.
- 100 percent means “I can definitely influence my partner with my decision”.

Remember: Your partner [*name*] will learn about your registration decision only after [he/she] has already decided whether to register for COVID-19 vaccination now.
Question 2: What do you think? How likely is it that your partner will make the same decision as you?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means “my partner will definitely not decide the same way I do”.
- 100 percent means “my partner will definitely decide like me” .

VI.1.C First stage Senders *'informing partner before'*

Reminder: Below we will provide you and your partner [*name*] with the opportunity to go to the official registration website of the Bavarian Ministry of Health to complete the registration process.

Your partner [*name*] will learn with a high probability whether you wish to register for a COVID-19 vaccination.

Remember: Your partner [*name*] will learn about your registration decision right before [he/she] decides whether to register for a COVID-19 vaccination.
Question 1: What do you think? How likely is it that your decision to register or not to register will influence your partner's decision?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "there is no way I can influence my partner with my decision".
- 100 percent means "I can definitely influence my partner with my decision".

Remember: Your partner [*name*] will learn about your registration decision right before [he/she] decides whether to register for a COVID-19 vaccination.
Question 2: What do you think? How likely is it that your partner will make the same decision as you?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "my partner will definitely not decide the same way I do".
- 100 percent means "my partner will definitely decide like me" .

new page

VI.2.A Registration intent Senders *"not informing partner"*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner [*name*] **will not** learn if you want to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.B Registration intent Senders *'informing partner after'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.

Your partner [*name*] will learn with a high probability if you wish to register for a COVID-19 vaccination.

Important: Your partner [*name*] will learn about your registration decision only after [he/she] has already decided whether to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.C Registration intent Senders *'informing partner before'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner [*name*] will learn with a high probability if you wish to register for a COVID-19 vaccination.

Important: Your partner [*name*] will learn about your registration decision directly before [he/she] decides whether to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.D Registration intent Receivers *'informed after'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner will not know if you want to register.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.E Registration intent Receivers *'informed before'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner will not know if you wish to register.

Important: Your partner [*name*] [would like/would not like] to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

new page

VI.3 Registration for COVID-19 vaccine

Would you like to register now?
To register, please click on **Yes, register now for a COVID-19 vaccination** below.
This will open the official registration website of the Bavarian Ministry of Health in a new browser window or tab. To successfully register for a COVID-19 vaccination, follow the instructions on the registration website.

Important: Please do not close the browser window or tab in which you are answering the survey during registration.

Additional Notes: We do not have any access to the information you provide on the registration website. Registration is voluntary and it does not entail an obligation to get vaccinated. Your reward for this survey is independent of whether you register.

Button: *Yes, register for the COVID-19 vaccination right away.*

[*Opens the link to the official registration website.*]

Have you successfully registered?

Here's how to proceed: once you have registered, please continue with the survey by clicking **Yes, I have registered and would like to continue with the survey** at the bottom of this page.

Don't want to register now?
If you do not wish to register now, you will not be penalized in any way, for example by being paid less for this survey. To continue with the survey, please click **No, I have not registered and would like to continue with the survey** at the bottom of this page.

To continue with the survey, please answer the following question:
Question: have you just register for the COVID-19 vaccination?

Reply options:

- *No, I have not registered and would like to continue with the survey*
- *Yes, I have registered and would like to continue with the survey*

——— *new page* ———

VI.4 Confirmation of registration for COVID-19 vaccination

Now confirm your registration: You have indicated that you have just registered online for a Corona vaccination.
You should have received a confirmation email after completing your registration.
Please provide the following two pieces of information from the confirmation email sent out by the vaccination center:

1. Email Address
2. Subject

Lottery: If both of your answers are correct, you can win one of 30 Amazon vouchers worth 20 EUR.

You must complete the survey to qualify for the lottery.

Further notes: Providing this information does not allow us to infer anything about you as a person. You remain completely anonymous. You can also continue with the survey without answering the questions. However, you will then not be able to participate in the lottery draw.

Question 1: What is the email address from which you received the confirmation email?
Question 2: What is the subject of the confirmation email you received from the vaccination center?

new page

VI.5 What do you think about the COVID-19 vaccine?

Question 1: What do you think? How safe is the COVID-19 vaccination?
Reply option: *Likert scale (1-7) with 1: not at all safe, 7: extremely safe.*

Question 2: What do you think? How effective is the COVID-19 vaccination?
Reply option: *Likert scale (1-7) with 1: not at all effective, 7: extremely effective.*

Question 3: What do you think? To what extent is it socially desirable to get vaccinated against COVID-19?
Reply option: *Likert scale (1-7) with 1: not at all socially desirable, 7: extremely socially desirable*

Question 4: To what extent do you agree with the following statement? Statement: if everyone is vaccinated against COVID-19, I don't need to get vaccinated too.
Reply option: *1: do not agree at all, 7: agree completely*

new page

Question: How likely are you to get vaccinated against COVID-19?
Please use the bar/slider for your answer.

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to make your selection.
- 0 percent means "definitely not willing to get vaccinated."
- 100 percent means "definitely willing to get vaccinated."

new page

VII Further demographic information

To conclude this survey, please provide some general information.

Question 1: What county do you live in?
Question 2: What is your zip code?
Question 3: What was your household's monthly net income last year?
Note: We mean the sum that results from wages, salaries, income from self-employment, pensions, income from public aid, income from letting, housing allowances, child benefits and all

other incomes, after the deduction of taxes and social security contributions.
Reply options:

- Less than 1,100 EUR
- 1.100 - 1.500 EUR
- 1,501 - 2,000 EUR
- 2,001 - 2,600 EUR
- 2,601 - 4,000 EUR
- 4,001 - 7,500 EUR
- More than 7,500 EUR

Question 4: What is your highest educational degree (general or vocational)?

new page

VIII End of survey

Thank you for participating in our survey!

In the following, we list your performance in the task in which you had to assign artworks to artists together with your partner and inform you whether you have won one of the Amazon vouchers. Afterwards, we ask you to answer two more questions about this survey yourself and give you the opportunity to give us feedback on the survey.

- Unfortunately, you have not won one of the raffled Amazon vouchers./Congratulations, you have won one of the raffled Amazon vouchers.
- If you would like to know how you and your partner did on your shared task, please click here. [*Upon clicking the button, participants' answers and the corresponding solutions open in the same window.*]
- *For Receivers 'informed after':* Finally, we would like to inform you that your partner [*name*] [registered/did not register] for a COVID-19 vaccination.
- Thank you again for participating in our survey.

Please answer the following questions to complete the survey:

Question 1: What do you think? What was the purpose of this survey?
Question 2: Where on the political spectrum would you place this survey?
Hints: Please use the slider to tell us the extent to which you felt this survey was leaning more toward the political right or toward the political left.
Click on the bar below to reveal the slider. Then move the slider to make your selection.

Feedback If you would like to give us any feedback on the survey, please feel free to do so here.

Would you like to close the survey now?

Click on ***Close survey***

Chapter 3

Education and the Women's Rights Movement

3.1 Introduction

What determines the emergence and success of social movements? Historically, successful movements often passed three key milestones in their development (Della Porta and Mattoni 2016; Markoff 2015; Tilly, Castañeda, and Wood 2020): (i) a small number of dedicated activists develop critical ideas that challenge the status quo and begin forming networks; (ii) these activists then spread these ideas using available mass media; and (iii) institutionalize their movement. From Dr. Martin Luther King Jr. to Susan B. Anthony, from Nelson Mandela to V. I. Lenin, such leaders are often considerably more educated than their peers. While their education is arguably crucial in a movement's emergence, the arrival of educational opportunities often coevolves with economic development and culture (Duflo 2012; Goldin 2006; Morris and Staggenborg 2004). Thus, it remains unclear whether increasing educational attainment can bring about societal change by facilitating the emergence and success of social activists and movements.

In this paper, we isolate the role of education in the emergence of social movements by studying the women's rights movement in Germany, and its relation to the expansion of educational opportunities for women. By 1919, German women achieved suffrage largely due to the growing influence of women's rights associations (Schraut 2019). By 1909, these associations were present in more than 320 cities, with the women teachers' association alone organizing more than 23,000 female teachers. Much like women's rights movements in other countries at the time, early members utilized female-led newspapers (e.g., *Frauen-Zeitung*, 1849–1852) to expand public support for their cause beyond their own demographic of educated teachers, writers, and artists.

In many cases, these early leaders obtained their education at Germany's first institutions providing secondary education and teacher training to women: so-called finishing schools (*Höhere Töchterschulen*). Finishing schools only admitted women and were present in more than 170 cities by 1850. The first finishing schools in Germany were opened by foreign Catholic orders dedicated to female education: Ursuline nuns (Aachen, 1626) and the Congregation of Jesus (Munich, 1627). Despite focusing on religious teachings and manners, these nuns also critically engaged with the ecclesiastical and social discrimination against women and supported the educational and sociopolitical principles of the Enlightenment in the early nineteenth century (Conrad 1996). Religious finishing schools complemented their curriculum with instructions in foreign languages and arithmetic. In short, they represented the only possibility for women to obtain secondary education or the necessary qualifications to work and live independently as teachers.[1]

Against this background, we leverage the timing of finishing school establishment as a positive shock to the availability of education for women. Using finishing schools, we highlight the role of education at three stages in the history of the women's rights movement. First, in a panel of cities and notable individuals, women started to represent a larger share of the political, intellectual, and economic elite ("human capital elite") after cities established finishing schools. Second, women from these cities also sent a disproportionate share of editorial letters to the first feminist newspaper in the mid-nineteenth century. Third, cities with historical finishing schools had more, and larger, women's rights organizations by the beginning of the twentieth century. We argue that finishing schools facilitated the exchange of critical ideas about women's role in society and the formation of networks; thus contributing to the rise of a female human capital elite from which the nucleus of the women's rights movement emerged. Crucially, these pioneering women disseminated critical ideas among a wider public and founded local chapters to convert their movement into a successful societal force.

We combine three sets of novel historical microdata, each representing one milestone of the women's rights movement in Germany, with data on the availability of education for women across cities and time. Variation in the availability of education comes from the opening date and city of 225 finishing schools constructed between 1626 and 1850 (Zymek, Neghabian, and Ziob 2005). Our measure of human capital in every city and period is derived from the *Neue Deutsche Biographie*. This biographical collection reports the places of birth and occupation for more than 150,000 individuals born between 800 CE and today. Its editors ignored local and time-bound personalities and only included individuals in a high position of responsibility who impacted the

[1] As Albisetti (1988, p. xiv) hypothesizes, "*the formal and informal curricula of these schools, when compared to those of the classical Gymnasien attended by boys from the same social groups, could stimulate in young girls an early awareness of, and a protest against, their 'second-class citizenship' rather than a submissive conformity to the 'German ideal of womanhood'.*"

general societal course. Thus, these data provide the most comprehensive historical account of Germany's political, intellectual, and economic elite. We measure the dissemination of critical ideas by digitizing all letters to the editor in one of the first, and quickly banned, feminist newspapers in German history (*Frauen-Zeitung*, 1849–1852), which contain the sender's city of origin and first name. Finally, we obtain variation in the institutionalization of the women's rights movement in Germany from a comprehensive survey on more than 1,200 local chapters conducted by Germany's Imperial Statistical Office in 1909.

For the first milestone, an increased representation of women among the human capital elite, we merge the timing of finishing school opening and the birthplaces of notable individuals to a balanced panel of German cities. In an event-study design with city and period fixed effects, we find that the share of women among the human capital elite rose from 1.8% prior to the opening of schools, to 4% within 50 years. Notably, the share of unmarried women also increased from 2.2% to 3.6%, indicating that finishing schools improved women's opportunities to live independently and be recognized for their achievements.

Cities that establish finishing schools may be selected on a wide range of characteristics. Such a selection process would be of concern to our interpretation if it correlates with women's status in society or a city's economic potential; then cities would exhibit different trends prior to school establishment. However, we find no evidence for differential pre-trends in women entering the human capital elite. Our findings are robust to including city and period fixed effects, linear time trends, and flexibly controlling for a rich set of predetermined educational, economic, and religious covariates separately in each period.

Differential population growth between cities might affect our interpretation if larger cities disproportionately attracted individuals from the human capital elite. We address this potential concern by dividing our main outcome variable by the total number of notable individuals born, thus controlling for the size of the elite in every city and period.[2] In addition, we use women from the nobility, a demographic educated by private tutors, as a placebo to capture potentially different population growth rates. We find no evidence that population trends confound our estimates.

If cities establish finishing schools in response to changes in (local) attitudes towards women, we would wrongly attribute the effect of social change to the expansion of education. Thus, to distinguish the impact of education from other social changes, we test whether other important economic and cultural events predict a similar increase in the representation of women among the human capital elite. To this end, we em-

[2]Our estimated effect is then identified within a city's elite, net of population growth, if the share of the elite relative to a city's population remains constant over time.

ploy a series of placebo exercises and test whether nonlinear changes in (i) economic activity, (ii) the returns to education, and (iii) gender-specific changes in culture predict a similar increase in the emergence of notable women. First, using construction data from Cantoni, Dittmar, and Yuchtman (2018), we find that the establishment of finishing schools did not coincide with a surge in economic activity. Second, we document that the staggered introduction of male schools does not predict women entering the human capital elite; similarly, finishing schools have no impact on men entering the human capital elite. Third, to alleviate concerns about nonlinear gender-specific changes, we employ four markers of gender-specific cultural change as placebo treatments and find that none coincide with a rise in the female human capital elite. Finally, we show that our results are not driven by the Protestant Reformation arriving in cities.

In a final step, we take a different approach to deal with the potentially endogenous adoption of finishing schools. We first show point estimates from a classical difference-in-differences design, adopting recent advances in Baker, Larcker, and Wang (2021), and second, report estimates from an instrumental variables strategy: first, we define sets of cities based on whether they established a finishing school by 1850 (treatment group) or not (control group), and compare the shares of women entering the human capital elite after the opening of the first finishing school in 1626 (post period). Second, we instrument our treatment group using monasteries constructed before 1300 coupled with religious competition near the religious divide. Throughout all specifications, we find no differential pre-trends, but a significant increase in women entering the human capital elite after the first finishing school was constructed. These findings carry over when analyzing every treatment period separately: even finishing schools established in the nineteenth century, when women were already more common among the human capital elite, significantly increase women's representation among the human capital elite.

We thus argue that finishing schools had an independent impact on women's representation among the human capital elite. The increased representation is driven by the very demographic that represented the core of the women's rights movement: the share of female teachers and writers increased from 1.9 to 3.6% post finishing school opening (compared to male teachers and writers). Further, using their biographies to identify activists campaigning for equal rights and women's suffrage, we show that the likelihood of an activist being born in a city increased from 1.6% to 6.9%.

This activist nucleus started to form networks early on. We find that after the opening of finishing schools, the probability that a notable woman is mentioned in another woman's biography from the same city increased threefold.[3] To show that these net-

[3]These connections are only recorded if they were substantial: for example, if women collaborated on

works, and not finishing schools per se, matter for increased human capital representation, we identify 500 women who migrated during their lifetime. While cities do not differentially attract women before the establishment of finishing schools, women start migrating to cities in which a native notable woman has already established a network.[4]

After leaders of successful social movements have developed critical ideas and formed an early network, they begin spreading their ideas using available mass media and institutionalize their movement. We document this second historical milestone of the German women's rights movement by linking the presence of finishing schools to letters to the editor of an early feminist newspaper (*Frauen-Zeitung*, 1849–1852) in a cross-sectional analysis. Compared to cities without finishing schools, cities with finishing schools are three times as likely to send a letter to *Frauen-Zeitung* in support of the women's cause, indicating a more successful propagation of critical ideas.

The third historical milestone of the women's rights movement we study is its institutionalization. Local chapters of the German women's rights movement sprung up from 1848, with the first organization specifically targeting female education being founded in the 1880s. Yet by 1909, only 37% of cities without finishing schools had established a women's rights organization, compared to 78% of cities with finishing schools founded by 1850. This difference is even more pronounced for educational organizations, at 5% and 29% respectively; these organizations also have an order of magnitude more members when located in a city with finishing schools.

In these cross-sectional results, unobserved differences between cities, previously captured by fixed effects, might reemerge and bias our estimates. We thus control for economic, religious, and educational covariates throughout to mitigate the threat from differential attitudes towards women. In addition, bias-adjusted point estimates (Oster 2019), estimates from an instrumental variables strategy using monasteries in 1300 coupled with religious competition as an instrument, as well as estimates using propensity score matching show a robust and stable impact of finishing schools on all cross-sectional outcomes.

In sum, our findings indicate that educational institutions, which foster the exchange of critical ideas and provide the space to form networks, can function as important catalysts for the emergence of a group of leading activists. Using newspapers to dis-

the foundation of a local chapter of a women's rights association. An example is the connection between Helene Lange and Gertrud Bäumer who jointly published the feminist newspaper *Die Frau* from 1893 onwards.

[4]These migrating women are a subset that – in our main results – are assigned to their cities of birth. We only assign them to their city of death to identify whether finishing schools were a pull factor in their migration decision. Our results are not the result of a violation of the stable unit treatment value assumption (SUTVA), and are robust to excluding these women, excluding neighboring cities and choosing a larger unit of observation (Appendix 3.D).

seminate critical ideas and founding local chapters to institutionalize their movement, these leading activists turned an initially upper-class movement into a broad societal force. Their legacy is still felt today, as cities with finishing schools in 1850 have brought forth a higher number of female members of parliament in any democratically elected parliament since 1919.

Our paper expands upon a thriving literature in economics studying the increasing representation of women starting in the late nineteenth century (Bertocchi and Bozzano 2016; Fernández 2013; Goldin 1990, 2006; Nekoei and Sinn 2021). First, by disentangling the availability of secondary education from other cultural and societal changes, we show that education was a key driver behind the women's rights movement and the increasing status of women in society. Second, and at a more general level, our results indicate that the positive effects of education are not limited to students themselves. In the case at hand, women from various backgrounds benefited from extending education to an initially limited number of women. Thus, our paper also informs a large body of literature in development economics studying the effects of interventions targeted at reducing gender inequality in education (Beaman et al. 2009; Chattopadhyay and Duflo 2004). By providing evidence on the effects of secondary education for women from the historical case of Germany, our paper highlights the potential long-run benefits of such interventions for society at large.

This paper also complements a recent literature in economics, which has highlighted the importance of civic leadership (Dippel and Heblich 2021) and technology (Enikolopov, Makarin, and Petrova 2020; Garcia-Jimeno, Iglesias, and Yildirim 2020; Melander 2020) in promoting the success of existing social movements. We extend this literature by studying how social movements, and their leaders, emerge in the first place. A prominent theory in sociology holds that educational capital is the key resource for leaders, even when leaders arise from poorer segments of society (Morris and Staggenborg 2004). By leveraging data spanning several centuries, we can study the emergence of the women's rights movement from before its very beginning until it reached key milestones, such as women's suffrage in 1919. Our findings support the notion that educational institutions that foster the exchange of critical ideas and network formation can serve as important catalysts of the emergence and success of social movements.

Finally, our findings also speak to the literature studying the role of an emerging human capital elite in early-modern Europe and beyond. Here, the human capital elite constituted a herald of economic change in the lead-up to the Industrial Revolution (Diebolt and Perrin 2013; Mokyr, Vickers, and Ziebarth 2015; Squicciarini and Voigtländer 2015). The dispersion of this upper-tail human capital over space and time was shaped by the institutional environment including welfare and educational policies (Dittmar and Meisenzahl 2019; Squicciarini 2020; Tabellini and Serafinelli 2020). Countries with highly educated leaders showed higher rates of economic growth (Besley, Montalvo, and Reynal-Querol 2011) and democratic participation (Glaeser, Ponzetto, and Shleifer 2007). We extend these existing studies in two dimensions: first, we explicitly focus on the female human capital elite; second, we show that in the context of the emergence of the German women's rights movement, this female human capital elite – through early activists' efforts to disseminate critical ideas and institutionalize the movement – constituted an important determinant of social change in and of itself.

The paper is structured as follows: In Section 3.2, we discuss the historical link between finishing schools and the women's rights movement. We discuss our data sources and dataset construction in Section 3.3, before discussing the identification assumptions of our empirical strategy in Section 3.4. In Section 3.5 we present our main findings on finishing schools' impact on female representation among the human capital elite. In Section 3.6, we conduct several placebo exercises to rule out confounding economic and cultural changes. In Section 3.7, we show that finishing schools facilitated network formation and immigration of women. We discuss the long-run results on the dissemination of critical ideas, the organization of the women's rights movement, and modern-day representation in parliaments in Section 3.8. Finally, Section 3.9 concludes.

3.2 Historical background

We begin by illustrating the links between the women's rights movement in the late nineteenth century and the emergence of religious finishing schools. In the aftermath of the Protestant Reformation, foreign Catholic women's orders began establishing finishing schools that focused on religious teachings but also included limited aspects of secular secondary education. At these finishing schools, students and teachers alike found access to critical ideas and a network of like-minded women. Several graduates eventually disseminated critical ideas in feminist newspapers and founded the women's rights movement. Religious finishing schools thus contributed to the formation of a group of pioneering women among the human capital elite, who acted as catalysts for social change.

3.2.1 Finishing schools

For the largest part of German history, only daughters from privileged families could obtain secondary education in the form of private tutoring. Access to secondary education for women improved when the orders of the Ursulines and the Congregation of

Jesus, founded in Italy 1535 and Flanders 1609 respectively, expanded into Germany. In the aftermath of the Protestant Reformation, these orders aimed to strengthen women's adherence to Catholicism in religiously competitive areas of Germany: Ursulines founded one of the first finishing schools in Cologne with the explicit goal of creating a "bulwark against emerging Protestantism" (Lewejohann 2014, p. 57), while the Congregation of Jesus established their school near Munich to educate young women in "good Christian manners, virtues and other studies [Wissenschaften]" (Riedl-Valder 2020, p. 2). In response, Pietists opened the first school in 1698, to combine biblical doctrine with a similar focus on Christian life and piety. Some ruling families took pride in sponsoring finishing schools in their territory, but compared to Catholic rulers of Bavaria and Württemberg, "Prussian monarchs did not move as vigorously as others to support secondary schools for girls" (Albisetti 1988, p. 29). By and large, city governments and Prussian rulers only became active in the field of female secondary education in the nineteenth century.[5]

Finishing schools' primary goal was to strengthen women's adherence to the respective faith, while parents sent their girls to finishing schools to improve marriage opportunities. This focus on religious teachings and marketable housekeeping skills emphasizes that religious finishing schools were not established with the explicit aim of empowering women. However, these finishing schools also included limited instruction in German, foreign languages, and arithmetic, and were among the first in German history to provide education at the secondary level to women. In contrast to the rollout of secondary education in the United States (Goldin and Katz 2003), women generally received lower-quality education than men as female teachers were denied the same quality of education as male teachers. By 1850, more than 200 finishing schools provided secondary education to thousands of young women.

3.2.2 The German women's rights movement

Starting in 1848, early women's rights activists around Louise Otto-Peters publicly demanded equal access to education, equal occupational opportunities and the right to vote (Berndt 2019; Gerhard 1990; Nagelschmidt and Ludwig 1996). Similar in spirit to the agenda of contemporary women's rights movements in the United States or Great Britain, they particularly emphasized the necessity of obtaining equal access to education as a key enabling factor for securing the other two central demands, the right to vote and equal occupational opportunities (Schötz 2019).

[5]The establishment of finishing schools in Protestant areas only gained momentum after 1750, by which time 40 finishing schools had already been established in Catholic regions. When including covariates, we always control for religion and ruler fixed effects to capture these different tendencies. In addition, we provide a specification separating schools into 'Early' and 'Late' schools, to assess the severity of this potentially demand-driven bias.

Initially, only women from the upper class formed the nucleus of the German women's rights movement. To gain broader support and turn the movement into a societal force, early women's rights activists pursued two complementary strategies: the dissemination of critical ideas about women's role in society and an institutionalization of the movement (Berndt 2019; Gerhard 1990; Nagelschmidt and Ludwig 1996). First, the movement started to publish a newspaper in 1849, *Frauen-Zeitung*, to disseminate critical ideas about the role of women in society among interested women and the general public alike; *Frauen-Zeitung* remained the main relay of the German women's rights movement until World War I (Schötz 2019).[6] Second, to coordinate its members, the movement started to establish associations with an increasing number of local chapters throughout Germany.

The first of these women's rights associations, *Allgemeiner Deutscher Frauenverein* (German Association of Female Citizens), was founded in Leipzig in 1865 and soon organized more than 20,000 women in 48 local chapters (Kaiserliches Statistisches Amt 1909). An important part of the local chapters' activity was to file petitions to (state) governments: they demanded the equality of women and men in the civil code (1876), the admission of women to universities (1876), and the improvement of the quality of teacher training for women (1887) (Schraut 2019). Reflecting the central importance of teachers, the *Allgemeiner Deutscher Lehrerinnenverein* (German Association of Female Teachers), founded in 1890 to advocate for equal access to education for women and adequate training for female teachers, quickly grew to a membership of more than 23,000 teachers spread across 108 local chapters by 1909.

In total, more than one million women joined women's rights associations by 1909 (Kaiserliches Statistisches Amt 1909, p. 17); many also joining political parties when the ban on female entry was lifted in 1908 (Evans 1980). In the first democratically elected parliament of the Weimar Republic (1919), at least 40% of female members of parliament had attended a finishing school and more than 50% had actively fought for women's rights in one of more than 1,200 women's rights associations in Germany.

3.2.3 Finishing schools and the women's rights movement

Several accounts by historians and the biographies of leading women's rights activists, such as the teacher Helene Lange, indicate the importance of finishing schools for the emergence of the women's rights movement in Germany (Albisetti 1988; Ringer 1987; Schaser 2000; Schötz 2019). Based on these accounts, we discuss two mechanisms that link the establishment of finishing schools to the emergence of the women's

[6] *Frauen-Zeitung* (translated: Women's Newspaper) was renamed *Neue Bahnen* (translated: New Ways) after it was banned by the Prussian government. However, the editorial staff and the ideological orientation remained.

rights movement: access to critical ideas about women's role in society, and reduced cost to form and access networks of like-minded peers. In this way, finishing schools provided the "foundations upon which the whole breadth and force of the women's movement were to depend" (Strachey (1928), p. 124, as quoted in Albisetti (1988), p. xiii).

First, despite their general focus on religious piety, Ursuline nuns and Mary Ward sisters also critically engaged with the ecclesiastical and social discrimination against women and demanded the 'spiritual' recognition of the equality of the sexes. They also actively supported the educational and socio-political principles of the Enlightenment in the early nineteenth century and augmented their religious teachings with secular subjects such as arithmetic and foreign languages.[7] Knowledge of English and French allowed women to access the critical writings of early feminist thinkers (e.g. Olympe de Gouge), which influenced the formation of the women's rights movement in Germany (Hauch 2019). Their ideas likely stimulated a critical questioning of women's role in society among the young women and teachers at finishing schools, especially when contrasting their opportunities with those afforded to their male counterparts (Albisetti 1988).

Second, finishing schools reduced the costs to form and access networks of like-minded women. In contrast to life outside schools, students at finishing schools lived together without the supervision of their families, being taught by female teachers who pursued an independent lifestyle unthinkable outside the teaching profession. This provided young women at a formative stage in life with access to a network of students and teachers which could strengthen opposition to their status as second-class citizens (Albisetti 1988; Ringer 1987). Finishing schools thus facilitated the exchange of ideas between teachers and fueled the rapid spread of local women's rights associations across Germany, as illustrated by the more than 23,000 teachers active in the *Allgemeiner Deutscher Lehrerinnenverein* (German Association of Female Teachers) in 1909.

More than any other profession, female teachers at finishing schools shaped the direction and force of the women's rights movement in Germany by influencing the lives of generations of women. This does not stand in contrast to the achievements of the working-class women's movement (Evans 1980), but complements the views of Albisetti (1988) and Wolff (2018, p. 19), who emphasize the importance of the "association and print media structures built since the 1860s" in carrying the demand for women's suffrage into society at large.[8]

[7]Authors' translation, adapted from Conrad (1996) p. 256 and p. 262.

[8]Our findings are consistent with the idea that both the bourgeois and the working-class women's movement made important contributions to improving women's opportunities in general and to gaining suffrage in particular. Both, female leaders of the SPD such as Clara Zetkin and leaders of the "radical

Without finishing schools, neither teachers nor students would have had comparable access to critical ideas and a network of like-minded women. Thus, they contributed to the formation of a group of pioneering women among the human capital elite, united by their opposition against women's status as second-class citizens. Crucially, these pioneering women disseminated their ideas to the broader public and institutionalized their movement, thus acting as catalysts for societal change.

3.3 Data

We assemble a novel dataset to study the role of secondary education in promoting the emergence of a female human capital elite. Our main outcome variable is derived from the biographies of all notable individuals born between 800 and 1950 CE within modern-day boundaries of Germany. Our explanatory variable finishing schools captures the availability of secondary education for women between 1626 and 1850 in all German cities. We combine these data to a balanced panel of cities in half-century periods, indicating the birth of notable women and the availability of secondary education at the nearest city.

Biographies of notable women. We obtain detailed microdata on the universe of notable German women and men for the period 800 to 1950 CE from the *Neue Deutsche Biographie* (NDB) to construct measures of women's representation among the human capital elite. The NDB is "considered the single most relevant biographic encyclopedia of the German language" and includes biographies detailing the professions and nobility of historically relevant men and women (Bayerische Akademie der Wissenschaften, Historische Kommission 1953).[9] It incorporates its direct predecessor, the *Allgemeine Deutsche Biographie* (ADB) (Königliche Academie der Wissenschaften, Historische Commission 1875), and in scope is comparable to the *Dictionary of National Biography* for British notable men and women.[10] We link 2,363 non-noble secular

wing" of the bourgeois women's movement such as Anita Augspurg, Minna Cauer, Lida Gustava Heymann, Gertrud Bäumer, either studied, received teacher training or taught at a finishing school at one point in their life.

[9] "Those personalities are to be included whose deeds and works reflect the development of German history in science, art, trade, and commerce; in short in every branch of political, intellectual and economic life." (Bayerische Akademie der Wissenschaften, Historische Kommission 1953, pp. VII-VIII). There is no evidence that editors or experts are selected based on the existence of finishing schools: "[t]he editors don't just rely on their own judgment; [the collection] bases its decisions on the advice of experts, on the advice of scientific institutes, and professional organizations. Essentially, it is assumed that the local and time-bound personalities have to be eliminated. In the areas of intellectual culture, it is primarily the independent, forward-looking achievement that decides, in the case of persons in a high position of responsibility, the impact on the general social course." (Bayerische Akademie der Wissenschaften, Historische Kommission 1953, p. IX, authors' own translation).

[10] The contents of NDB and ADB are freely available online (Bayerische Akademie der Wissenschaften, Historische Kommission 2019).

women to cities of birth within in the modern-day boundaries of Germany after 800 CE, as well as 261 women from the nobility, who we use as a placebo to ensure our estimates are not affected by differential population growth between cities. Thus, for each city and period, our data records the number of women born who later became recognized for their achievements. Of all 2,624 women, 32% became notable for being an artist, 21% for being a writer, 10% for being born into nobility, and 6% each for being an academic or a politician (Table 3.1). We use the place and date of birth of notable women alongside the reported biographical information to trace women's representation among the human capital elite across cities and periods. Our main dependent variables are (i) an indicator for whether at least one woman was born in a given city and period who became notable later in life, (ii) the log number of notable women, (iii) and the share of notable women among all notable individuals. These variables measure the extensive and intensive margin of women's representation among the human capital elite.

Finishing schools. We link the birthplaces of all notable women to the historical emergence of finishing schools providing secondary education obtained from the *Data Handbook of German Education History*. This handbook covers traditional female finishing schools constructed between 1626–1850 and their location as shown in Figure 3.1 (Zymek, Neghabian, and Ziob 2005).[11] We match finishing schools to our data on notable women based on their location and opening date. The first finishing schools were established by female orders of the Catholic church who, following invitations by ruling houses, often settled near existing monasteries to educate and "protect the women's mind from the falsities of their time".[12] Protestant or city schools started to emerge in significant numbers only after 1750. In total we record 209 school openings in 129 cities between 1626 and 1850, without a clear spatial pattern in location or timing (Figure 3.1).[13]

[11]We focus on these schools with continuous operation selected by Zymek, Neghabian, and Ziob (2005) as the most comprehensive data on finishing schools (*Höhere Töchterschulen*) in Germany before the emergence of the women's rights movement. Other schools existed, especially in later years, but Zymek, Neghabian, and Ziob (2005) do not include these schools for two main reasons: first, these schools often operated only for a few years and closed down quickly for unknown reasons. Second, it is often unclear whether these schools provided curricula that extended beyond primary education. Since such other schools are more likely to appear in the later years of our dataset, we divide the data into 'Early Schools' prior to 1750, and 'Late Schools' post 1750 in Table 3.C.5. We find no differential impact, and thus no evidence for a bias arising from the omission of these temporary existing schools.

[12]"*...vor allem den unteren Volksschichten das religiöse Leben (zu) heben und den Frauen Ansichten und Grundsätze (zu) vermitteln, durch die sie gegen Irrtümer ihrer Zeit gesichert und für eine gesunde Erweiterung ihres Lebensinhaltes befähigen würden*". Source: https://bit.ly/2WGKe4I, cited from Festschrift der Ursulinenschule, Köln 2014, S. 261, last accessed 2021-02-09.

[13]Some later schools might have been a response to local demand of the population. We report the same results when distinguishing between schools constructed in the period 1650–1750 and those constructed 1750–1850 in Table 3.C.5. We also report no differential pre-trends and similar-sized point estimates for every treatment period in Figure 3.F.2 and Table 3.F.2. Schools are not spatially correlated

Table 3.1: Summary statistics: Finishing schools and notable women

	Cities		
	Without finishing schools (N=259)	With finishing schools (N=129)	Percent of sample
Data: Female finishing schools in Germany			
Finishing schools	0	1.620	
Data: Neue Deutsche Biographie			
Non-Noble Secular (NNS)			
Academic	33	131	0.063
Artists	139	712	0.324
Founders	2	9	0.004
Medicine	17	56	0.028
Not assigned	45	146	0.073
Occupations	39	136	0.067
Politics	43	122	0.063
Sports	0	5	0.002
Teachers and Writers (also NNS)			
Teacher	27	59	0.033
Writers, Publishers	146	416	0.214
Activists (also NNS)			
Activists	36	94	0.050
Unmarried women			
Unmarried	492	1666	0.822
Nobility			
Royals, Wifes, Relatives	91	170	0.099
Nuns			
Religion	25	55	0.030
Population (Bairoch, Batou and Chèvre 1988)			
Population in 1600	5.3	10.4	

Notes: The first row reports the average number of schools in cities without historical finishing schools (259) and with historical finishing schools (129). The average number of finishing schools in cities with schools is 1.62, with 85 cities having one school, 29 cities having two schools, and 15 cities having three or more schools. The subsequent rows detail the absolute number of notable women in each sub-group and their share of the total. Activists and unmarried women are separately coded and could belong to all other groups as well. The last row indicates the average city size in thousands. Cities that have a finishing school by 1850 are nearly twice the size in 1600. While this ratio is very similar for women from the Nobility (factor 1.9) and Nuns (factor 2.2), Non-Noble Secular (unmarried) women are 3.6 (3.3) times more likely to appear in cities with finishing schools. We control for the difference in population by interacting 'Population in 1600' with period fixed effects in all regression with control variables.

Cities. Since birthplaces of notable women and the location of finishing schools do not overlap perfectly, we utilize data from Voigtländer and Voth (2012) to construct a balanced panel of 388 German cities that existed in 1300.[14] For each city, we cre-

(Moran's I: 0.002, p-value 0.156), yet we follow two additional strategies to deal with any remaining spatial autocorrelation. First, we report standard errors corrected for spatial correlation in Table 3.D.1. Second, we randomly distribute the actual number of schools build in every period across Germany and show the distribution of point estimates in Figure 3.D.1.

[14]The "extended sample" of Voigtländer and Voth (2012) includes 1,428 "towns and cities", 739 of which were mentioned before 1300. Many of these "towns and cities" are close to a major city. For example, Voigtländer and Voth (2012) link three suburban towns to Aachen: AACHEN L, town_ids 1,3,4, mentioned in 930, 1118, and 870 CE which are close to the original city of Aachen (AACHEN S, town_id 5, mentioned in 400 CE). We use the latter as our reference city if it lies in present-day borders of Germany to control for spillovers from suburban towns to cities. Results are robust to changing the set of cities to those that existed already in 800 (Table 3.C.1), changing to 25 year periods (Table 3.C.2), and including city×period fixed effects in a panel setting with gender×city×period as the unit of observation

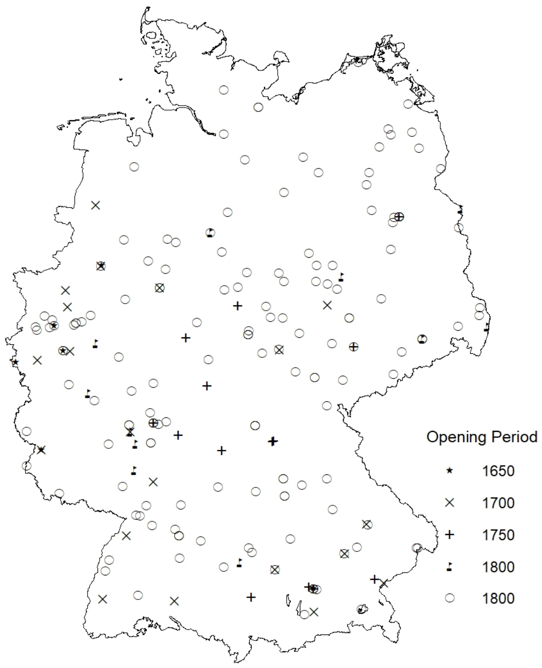


Figure 3.1: Spatial distribution of finishing schools in Germany by opening year

Notes: This figure shows the location of finishing schools by their opening period. In Figure 3.C.1, we additionally illustrate the variation of notable women across cities. We depict finishing schools by opening period and religious denomination in Figure 3.J.1.

ate 50 year periods from 800 until 1950 CE to ensure a sufficient overlap between the opening of a finishing school and its effect on women becoming recognized for their achievements in our biographical database. We then merge the biographies of women and the emergence of finishing schools to the nearest city and period in our sample, thus covering all of modern Germany. This procedure has two advantages: First, it does not rely on any political or geographical boundary as the matching procedure is solely based on distance.[15] Second, we can use the rich set of covariates from Voigtländer and Voth (2012) to flexibly capture economic, religious, and educational factors, as measured in 1300, in every period.

3.4 Empirical strategy

We study the role of secondary education in promoting the emergence of a female human capital elite that later formed the nucleus of the German women's rights move-

(Table 3.3).

[15]In an alternative approach explored in Appendix 3.C.2, we instead use administrative boundaries of territories in 1618 and merge all data based on whether city 'y' was in territory 'x'. As our results remain qualitatively unchanged, we argue that sample selection does not introduce a bias in our setting.

ment. Our empirical strategy combines the staggered introduction of religious finishing schools and unique biographical microdata on the universe of notable women in German history to a balanced panel of 388 cities between 800 and 1950 CE. The key empirical challenge is then to isolate the impact of finishing schools from potential confounders that are correlated with both finishing school opening and the increase in women's representation among the human capital elite.

Cities that establish finishing schools may be selected on a wide range of characteristics. Even if these schools were established for reasons that are arguably uncorrelated with local economic conditions or the demand for education, a causal interpretation of the impact of finishing schools requires that all unobservable factors that influence women's representation among the human capital elite must be orthogonal to finishing school opening. However, as production technologies change, increased returns to education induce a rise in the demand for education. Similarly, wars or natural catastrophes that disproportionately affect the male population increase the demand for female labor and thus the demand for educated women. These local, often unobservable, factors can increase the adoption of educational policies and thus change the relative wages between cities. Under such conditions, cross-sectional evidence or failing to control for local factors risks overstating the true effect of finishing schools on women's representation among the human capital elite.

3.4.1 Specification

We address local differences between cities by including city and period fixed effects in a two-way-fixed-effects (TWFE) setup, capturing all observable and unobservable time-invariant factors that vary between cities and periods in our sample.

$$\begin{aligned} Y_{c,t} &= \beta\, \mathit{Finishing\ school}_{c,t} + \alpha_c + \alpha_t + \alpha_c \times t + && \text{(Baseline)} \\ &+ \sum_{\tau=800}^{T=1950} \left[\boldsymbol{X}_{e,c}\, \boldsymbol{\alpha}'_{e,\tau} + \boldsymbol{X}_{r,c}\, \boldsymbol{\alpha}'_{r,\tau} + \boldsymbol{X}_{s,c}\, \boldsymbol{\alpha}'_{s,\tau} \right] + \varepsilon_{c,t} && \text{(Additional Controls)} \end{aligned}$$

In our baseline specification, we regress a binary outcome of whether a woman who became notable later in life was born in city c and period t, on an indicator of the presence of a finishing school. We use two definitions of this indicator $\mathit{Finishing\ school}_{c,t}$: In our main specification, this indicates whether a finishing school is present in city c at time t. In Appendix 3.F, we abstract from the variation in timing and define this variable as the classical difference-in-differences estimator, comparing 129 cities with finishing schools to 259 cities without after 1650: $\mathit{Finishing\ school}_c \times \mathbf{1}(t \geq 1650)$.[16]

[16] Using this classical difference-in-differences design we find no evidence for pre-trends (Figure 3.F.1) and similar point estimates (Table 3.F.1). Further, we find no evidence of differential pre-trends or heterogeneous treatment effects across treatment periods (Figure 3.F.2 and Table 3.F.2).

We include city α_c and period α_t fixed effects as well as city-specific linear time trends $\alpha_c \times t$. This baseline set of fixed effects captures all unobservable city-specific trends that evolve linearly over time. We cluster our standard errors at the city level c and report standard errors corrected for spatial correlation in Appendix 3.D, Table 3.D.1.

To identify the impact of finishing schools on women's representation among the human capital elite, we must argue that conditional on our set of fixed effects, either school assignment is as good as random or that observed increases in women's representation among the human capital elite can only be attributed to finishing schools. Since the former is unlikely, the latter requires us to relate the increase in the number of notable women being born after the opening of the first finishing school to the long-term trends that determine women's representation among the human capital elite and finishing schools. Then, to identify the impact of finishing schools, cities need not exhibit different trends prior to the establishing of the first finishing school. In addition, since our baseline specification already captures differences between cities that grow linearly over time (e.g. population growth), our identifying assumption necessitates sufficiently capturing all remaining nonlinear, city-specific, confounding factors.

With our additional controls we capture three sets of potential confounders that might nonlinearly predict women's representation among the human capital elite and the opening of finishing schools: economic, religious, and educational characteristics. The first set of covariates capture the potential direct effects of economic characteristics that influence the decision to open finishing schools ($\boldsymbol{X_{e,c}}$). We proxy for the economic and financial development using membership in the Hanseatic League, Jewish settlements and pogroms against Jews (Voigtländer and Voth 2012). We complement these covariates with population data in 1600 from Bairoch, Batou, and Chèvre (1988), female specific labor demand as proxied by religious battles during the Thirty Years' War affecting sex-ratios and local weather conditions affecting agricultural production from Leeson and Russ (2017). Combined, these covariates, measured before the opening of the first school, capture demand factors of productivity and relative wages that may impact the decision to establish a finishing school.

The second set of covariates capture the potential influence of religion on school opening and women's representation among the human capital elite. Since almost all early finishing schools were established by religious orders, this set of covariates captures any direct effects of religious differences across cities ($\boldsymbol{X_{r,c}}$). We include whether the city was a bishopric seat (Voigtländer and Voth 2012) and distance to Wittenberg to proxy for the diffusion of Protestantism (S. O. Becker and Woessmann 2009; Cantoni 2015). We determine which cities were Protestant or Catholic in 1618 by digitizing cartographic material in Engel and Zeeden (1995), and include the distance to the inner-German denominational boundary to capture religious competition between the major religious denominations. In combination, our religious controls thus address two major concerns regarding the comparison between Protestant and Catholic cities: first, early finishing schools were built by Catholic orders and Protestant cities did not establish secondary educational institutions in significant numbers until 1750. Second, as highlighted in S. O. Becker and Woessmann (2009), since Protestantism is generally associated with a greater proportion of women receiving (limited) primary education, we might wrongly attribute an effect of Protestantism to finishing schools.

Finally, we address the direct effects of differential returns to education across cities ($\boldsymbol{X}_{s,c}$) by determining whether a city had a university or provided higher male education in 1650.[17] In addition, we control for different educational preferences of different heads of state by controlling for the ruling house of each city as of 1618 using Engel and Zeeden (1995).[18] Combined, male schools, universities and the educational preferences of ruling houses capture local returns to education across all genders at the time the first finishing schools were established in Germany.[19]

We interact all covariates with period fixed effects to isolate the effects of finishing schools from these confounding factors.[20] Our identifying variation is thus limited to within-city, off the linear time trend of any unobservable confounding factor and the nonlinear evolution of observable economic, religious, and educational differences across time. Hence, all remaining violations of the main identifying assumption must arise from unobservable nonlinear confounding factors which explain both the opening of a finishing school as well as the subsequent increase in women's representation among the human capital elite.

3.4.2 Evaluating pre-trends

We evaluate the validity of our empirical design by testing for differential pre-trends in the event-study graph of Figure 3.2.[21] Here, we limit our sample to all cities in which a finishing school has ever been established and estimate the impact of the

[17] Obtained from https://bit.ly/2OHH4tp and https://bit.ly/3mG9mRr, last accessed 2021-02-09.

[18] An example is Prince Bishop Ferdinand of Bavaria who, in response to the religious competition, pushed for female education to win over the minds of women.

[19] In the spirit of Galor and Weil (1996) we assume that local returns to education are not impacted by directed technical change that would increase the returns to education for one specific gender. However, estimating a panel with city $\times$ year fixed effects and gender $\times$ year fixed effects in Table 3.3 captures this variation and the point estimates are not statistically different from our baseline.

[20] We explore heterogeneity along all covariates and find no heterogeneous impacts nor changes to our main coefficient.

[21] We estimate the event-study equivalent of our baseline equation with and without covariates:

$$Y_{c,t} = \alpha_c + \alpha_t + \sum_s \beta_s \mathbf{1}\{t - E_c = s\} + \varepsilon_{c,t}$$

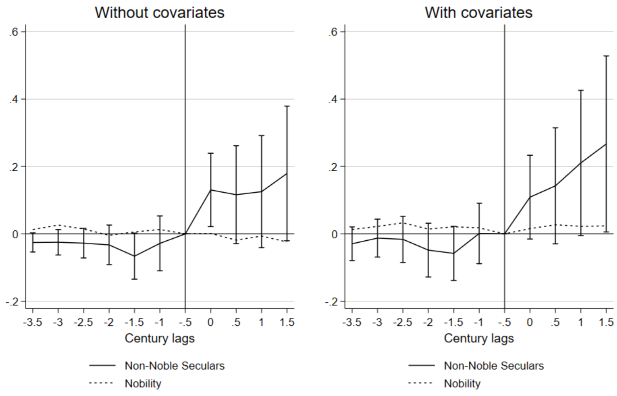
$\{t - E_c = s\}$ denote relative time periods to opening of the finishing schools. Cities enter this sample 400 years prior to the establishing the first school and leave it 150 years after.

first finishing school four centuries before and two centuries after its opening. In Figure 3.2, we provide evidence in favor of our identification assumption as finishing schools have a precisely estimated zero impact in all periods prior to opening. We estimate the impact of finishing schools on two subgroups of women: non-noble secular women (solid line) and the nobility (dashed line). We use women from the nobility as a placebo group and separate them from the remaining notable women, since they likely had access to private tutoring and thus should not be affected by the opening of finishing schools.[22] If the establishment of finishing schools is correlated with an unobserved change in the overall likelihood of being recorded as notable (e.g. population growth or local political change), the point estimate on nobility would be significant in post periods. However, while we find no impact of finishing schools on women from the nobility, the probability of a non-noble secular woman being born in the city and becoming notable later in life increases immediately after the first school opened. This relationship remains robust when including all control variables nonlinearly in the right panel of Figure 3.2a.

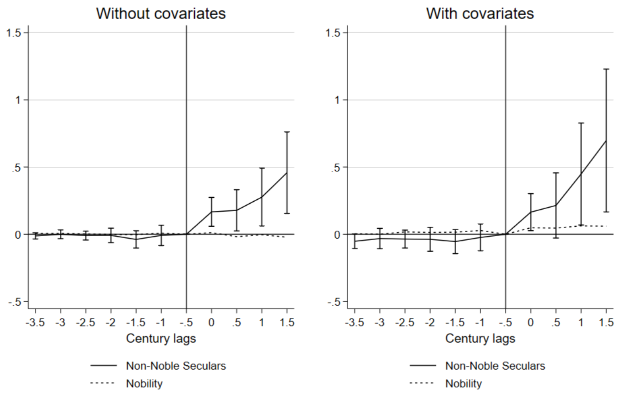
In the remaining panels of Figure 3.2, we document the absence of pre-trends when using the number of women born (Figure 3.2b) and the share of women among all notable individuals born in the same city and period (Figure 3.2c). We observe a significant treatment effect in the first period after opening that is slightly increasing in the right panels when controlling for covariates.

If this slight increase is driven by cohort-specific treatment effects, our TWFE estimator might produce biased estimates. This problem is most pressing in settings without a never-treated control group: Here, later-treated cohorts function as the control-group for earlier-treated cohorts, potentially creating negative treatment weights biasing the estimate (Goodman-Bacon forthcoming). Using the suggested decomposition, we find non-negative weights and point estimates that result from the difference between never-treated cities and cities with finishing schools. We thus leverage cities that never establish a finishing schools as a pure control group in our setting and follow Baker, Larcker, and Wang (2021) in providing three sets of evidence against heterogeneous treatment effects biasing our estimates: first, we provide the main event-study graph with and without controls (Figure 3.2). Second, we provide an assessment of pre-trends by treatment cohort (Figure 3.F.2) and provide estimates for each treatment-cohort (Table 3.F.2). Third, in Appendix 3.E we implement the aggregation methods suggested by Chaisemartin and D'Haultfœuille (2020) and Callaway and Sant'Anna (forthcoming), as well as include never-treated cities in the event-study design. We find no evidence of treatment-effect heterogeneity or differential pre-trends and report similar point estimates for all treatment groups and methods.

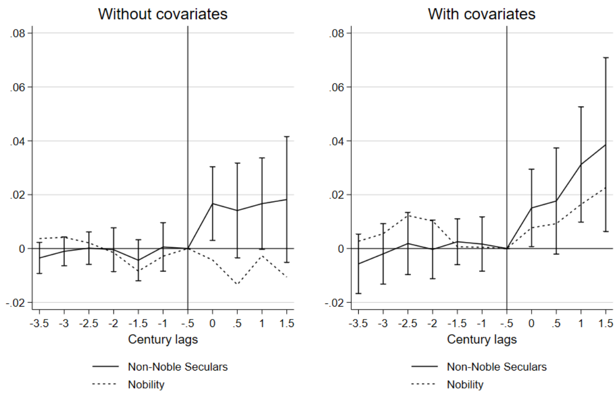
[22]We separate this group not to discredit the efforts and successes of many noble women advocating women's rights, but merely to reflect historical differences in the provision of secondary education.



(a) Indicator: Notable woman born in city



(b) Log. number of notable women born in city



(c) Female share of notable individuals born in city

Figure 3.2: Event-study:
Impact of finishing school establishment on notable women

Notes: Zero is the normalized opening period of the first finishing school in a city. The vertical line marks the reference period, which we choose to be 50 years prior to establishment of the first finishing school. City and period fixed effects included in the left figure and full economic, religious, and educational controls added in the right; 95%-confidence intervals shown only for non-noble secular women, the impact on notable women from the nobility is indistinguishable from zero in all periods and specifications. Alternative approaches are discussed in Section 3.E.

Lastly, choices when compiling our dataset might affect the observability of pre-trends. In our dataset, we merge women and finishing schools to a balanced panel of 388 cities, including never-treated cities, and 50-year periods. This, however, does not fully utilize the exact treatment period of each school. In Appendix 3.C.3, we instead construct 10, 20, 25, and 50 year intervals around each exact opening year of finishing schools and show the resulting event-study graphs. Again, we find no evidence for a pre-trend in any specification, a significant uptick after opening, and point estimates that are not statistically different from our baseline. Thus, we use our balanced panel of cities, allowing us to include never-treated cities and control variables in a TWFE estimation, and take this result as additional evidence against pre-trends or heterogeneous effects biasing our estimates.

3.5 Finishing schools and the human capital elite

Our hypothesis is that the opening of finishing schools increased women's representation among the human capital elite. Women belong to the human capital elite of their city of birth if their names were recorded in the *Neue Deutsche Biographie*. Using data on notable women from 800 to 1950 CE, we document a sustained impact of the opening of finishing schools on an indicator of whether a notable woman was born, the number of notable women, and the share of notable women relative to their male counterparts. Using detailed occupational and biographical data, we provide additional evidence that finishing schools contributed to women entering the human capital elite as teachers and activists. These women later formed the core demographic of the women's rights movement, spreading their ideas in outlets such as *Frauen-Zeitung*, and organizing in women's rights associations throughout the country.

We present our main results in Table 3.2, using our baseline empirical specification including all cities and periods. We report estimates from three different specifications of our dependent variable to address the sparsity in our outcome variable. In columns (1) and (2), we regress an indicator variable of whether a notable woman was born in city c at period t on our indicator variable for finishing schools that turns on after the opening of the first finishing school in city c and period t. Our baseline estimate is reported in column (1) of panel A and suggests a 23-percentage point increase (s.e. 0.029) in the propensity to observe a woman being born and becoming notable later after the establishment of the finishing school. To capture the impact of city-specific differences on the establishment of finishing schools and notable women, we interact economic, religious, and educational covariates with period fixed effects in column (2). The point estimate of 0.164 (s.e. 0.033) suggests a stable impact of finishing schools on women's representation among the human capital elite, with finishing schools doubling the likelihood of observing a notable woman in periods after their establishment.[23]

Table 3.2: Fixed-effects results on the importance of finishing schools

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.230***	0.164***	0.355***	0.204***	0.019***	0.021***
	(0.029)	(0.033)	(0.053)	(0.045)	(0.004)	(0.005)
Mean, untreated	0.150	0.149	0.272	0.272	0.018	0.018
Panel B: Unmarried women						
Finishing school$_{it}$	0.194***	0.147***	0.302***	0.173***	0.011**	0.014**
	(0.030)	(0.034)	(0.049)	(0.043)	(0.005)	(0.006)
Mean, untreated	0.155	0.153	0.275	0.274	0.022	0.022
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.151***	0.104***	0.174***	0.103***	0.019***	0.017***
	(0.027)	(0.026)	(0.034)	(0.029)	(0.006)	(0.006)
Mean, untreated	0.076	0.075	0.096	0.096	0.019	0.019
Panel D: Activists						
Finishing school$_{it}$	0.076***	0.053***	0.064***	0.043***	0.013***	0.011**
	(0.018)	(0.018)	(0.017)	(0.015)	(0.004)	(0.005)
Mean, untreated	0.016	0.016	0.018	0.018	0.005	0.005
Panel E: Nobility						
Finishing school$_{it}$	-0.018	-0.013	-0.009	-0.007	-0.002	-0.002
	(0.016)	(0.017)	(0.015)	(0.018)	(0.008)	(0.009)
Mean, untreated	0.039	0.038	0.050	0.050	0.018	0.018
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	9,312	9,240	9,312	9,240	9,312	9,240

Notes: Main results using a TWFE estimation based on all cities in all periods. We consider three types of dependent variables to capture the extensive and intensive margin of the birth of notable women. $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city had at least one notable woman born in this period. 'log Women' constitutes the natural logarithm of the number of women born plus one. 'Share Women' divides the number of women by the number of women *and* men in the respective category, except for Activists, where we use the number of male politicians instead. We regress the number of non-noble secular women, teachers and writers, and women from the nobility born in a city, as defined in the top row, on our finishing school variable. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city specific linear trends. In columns (2), (4), and (6) we additionally interact city controls with period fixed effects to capture variation from economic, religious, and educational differences. We include the following controls measured in the 13th century: Hanseatic League and bishopric indicators as well as indicators for having a Jewish presence and for pogroms. Additionally, we include the following controls from 1600: distance to Wittenberg, an indicator for confessional battles in the vicinity, distance to the denominational divide, and a Catholicism indicator (as of 1618) to capture religious differences. In addition, we control for the average temperature in 1650 to capture differential agricultural productivity, and hence income. City-level population in 1600 is included to capture different population effects; pre-existing male schools, universities in 1650, and a ruling house indicator are included to capture differential educational preferences. Standard errors clustered by city reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

[23]If our sample of finishing schools suffered from survival bias, and assuming a positive impact of schools, our estimates would be downward biased as control observations would be treated, too. In addition, we report reduced form estimates, unaffected by survival bias, using monasteries in 1300 as an instrument around 10 km of the denominational divide in Figure 3.F.4.

In the remaining columns (3)–(6) we explore the intensive margin of the effect of finishing schools on women's representation among the human capital elite. Using the log number of women born in city c at period t, we find that the number of notable women increases by 20%, even when extensively controlling for economic, religious, and educational factors.[24]

Population in 1650 interacted with period fixed effects might not adequately capture the heterogeneous growth paths of German cities.[25] By using the number of notable men born in each city and period, we are able to capture differential growth in population, prosperity, and creativity, that might lead to the adoption of finishing schools and an increased representation of women among the human capital elite. In columns (5) and (6), we thus divide the number of notable women born by the total number of notable men and women in the same category and period. If the number of notable women in our sample only increased due to a discontinuous change in population, prosperity, or creativity, happening at the same time, this would increase in the number of notable men in the same category, too.[26] Relative to cities without finishing schools in which 1.8% of all notable individuals are women, the share of women among the human capital elite increased to 4% after the establishment of finishing schools.[27] The robust estimates suggest that finishing schools increased women's representation among the human capital elite and did not affect a city's population or its elite's size in particular.

Similarly to other countries (Goldin 2006), the majority of notable women were unmarried and independent. The share of unmarried women, relative to all unmarried men and women, increases from 2.2% to 3.6% after the opening of finishing schools (Table 3.2, panel B, Column 6). While it is possible that measurement error in the data biases this point estimate, the measurement error would have to be correlated to finishing school opening to bias the point estimate upwards. Our results thus support the notion that finishing schools facilitated the emergence of a greater number of women pursuing a more independent lifestyle, free from the constraints of marriage in a patriarchal society.

---

[24]Using the logarithm of a variable with a large amount of zeros is problematic as the $\log(y+1)$ transformation might introduce a bias. We are aware of this and thus refer to columns (1) and (2) as our preferred specification and report all figures using the binary definition (columns 1 and 2) as the outcome variable.

[25]While Aachen and Trier were some of the most important cities at the beginning of our sample period, they have been outpaced by Munich and Berlin at the end. This pattern is not predicted by initial population size or ruling houses in the 17th century.

[26]The number of notable men is equally obtained from the *Neue Deutsche Biographie*.

[27]We address the possibility that people move to neighboring towns with schools, and thus spillovers are impacting our interpretation, in two tables: We increase the catchment area of each city by only using 101 cities that already existed in 800 and show the same effect sizes (Table 3.C.1); In Table 3.D.2 we restrict our sample to 129 cities with schools and 27 non-neighboring cities in 1300. All results are robust and indistinguishable from the baseline empirical specification.

In the remaining panels (C)–(E) of Table 3.2 we explore the effects of finishing schools on different subcategories of notable women based on their professions and the placebo group, women from the nobility. First, we confirm historical accounts arguing that many students went on to become teachers and writers by showing that the likelihood of a female teacher or writer being born and recorded in our data is substantially higher after the opening of a finishing school. Second, we analyze the biographies of all notable women and use keywords to identify women's rights activism.[28] While we record markedly fewer women than in other categories, the relationship is robust and stable in all specifications and suggests a threefold increase in the likelihood of observing an activist after the opening of a finishing school (panel D, columns (2) and (4)).

Finally, we estimate the impact on the subgroup of noble women in panel E. Again, we treat the nobility as a placebo group since the likelihood of being recorded in the *Neue Deutsche Biographie* should not benefit from the establishment of a finishing school. This subgroup captures overall trends in population growth that should equally affect all notable individuals of either category. In line with our argument that the relationship between finishing schools and women's representation among the human capital elite is not mechanically driven by population growth, we find robustly estimated insignificant null effects of finishing schools on the nobility throughout all specifications.[29]

We take the strong and robust results on non-noble secular women, and the non-existent impact on women from the nobility, as evidence that finishing schools indeed increased women's representation among the human capital elite in Germany. We conduct numerous further robustness tests in the Appendix. In Appendix 3.B, we show that our results remain qualitatively unaffected when omitting the linear time-trend, using different covariates (Table 3.B.1), or omitting outliers (Figure 3.B.1). In Appendix 3.C, we gather additional evidence against dataset construction choices biasing our estimates: our results remain unchanged when using alternative sets of cities (Table 3.C.1) or alternative lengths of periods (Table 3.C.2). The estimated effect does not vary greatly by occupation (Table 3.C.3) or the timing of school opening (Table 3.C.5). We dedicate Appendix 3.D to showing that the results are unlikely to be the result of systematic SUTVA violations. To assess whether spillovers affect our interpretation, we create 200 placebo datasets using the true spatial correlation and temporal assignment and find p-values of 0.000 for all outcomes except activists (p-

[28]The top five keywords are (in order): "Frauenrecht" (women's rights), "Frauenbewegung" (women's movement), "Frauenverein" (women's clubs), "Emanzipation" (emancipation), and "Feministin" (feminist). The share of activists is constructed using the number of male politicians as a proxy for the politically active male population.

[29]Controlling for construction activity does not impact our results (Table 3.B.3) and construction activity is not predicted by school establishment (Figure 3.4).

value: 0.020). In Appendix 3.E, we show that our point estimates are also robust to varying weighting techniques from the recent literature on the validity of event-study designs. In Appendix 3.F, we report similar estimates from a classical difference-in-differences setting, dividing cities into those that had established a finishing school by 1850 and those that had not (Table 3.F.1). There is no discernible pre-trend when using all treatment periods jointly (Figure 3.F.1) or when separately identifying pre-trends by school opening period (Figure 3.F.2). We regard the robustness of our results as evidence against a mechanical relationship between finishing schools and notable women which could arise simply due to finishing schools improving record keeping of influential women or increasing the demand for teachers.

3.6 Placebo exercises

To rightfully attribute the increase in women's representation among the human capital elite to the emergence of finishing schools, we discuss whether changes in the returns to education, culture, or economic activity predict a similar increase. To identify such potential confouding factors, we exploit the following city- and time-specific placebo events: in Section 3.6.1, we use the opening of secondary schools for men to capture an increase in the overall returns to education. In Section 3.6.2 we use construction activity as a proxy for economic activity; and in Section 3.6.3, we exploit the end of witch trials, the opening of female monasteries, the consecration of churches to a female saint, and the arrival of the Reformation, to capture gender-specific cultural changes at the local level. No placebo event predicts a subsequent increase in the number of notable women.[30] Unobservable nonlinear and city-specific factors are thus unlikely to confound our finding that finishing schools increase women's representation among the human capital elite.

3.6.1 Returns to education

In our first placebo exercise, we assess whether finishing schools merely capture local changes to the returns to education. We exploit cross-gender variation and show that the number of notable men and women is only affected by the opening of male and female schools, respectively. We thus argue that finishing schools are unlikely to reflect local changes in the returns to education.

To assess the importance of changes in the returns to education, we correlate the occurrence of non-noble secular men, unmarried men, and male teachers and writers, with the opening of male schools. Following Galor and Weil (1996), we interpret schools

[30] These changes are however, correlated to the establishing of finishing schools, suggesting that they are relevant cultural and educational proxies to consider.

for men as an endogenous response to increased returns to education following an increased demand for skilled labor. As such, the estimated effect of male schools on the occurrence of notable men is a combination of (i) increased returns to education and (ii) education itself. By the same token, if female finishing schools were also a result of increased returns to education common to both genders, we would expect to see an increase in the number of notable men in response to the establishment of finishing schools.[31]

Table 3.3: Placebo estimates on the importance of finishing schools: Differential returns to education

	Non-Noble Secular			*Unmarried*			*Teachers & Writers*		
	(1) Female	(2) Male	(3) Panel	(4) Female	(5) Male	(6) Panel	(7) Female	(8) Male	(9) Panel
Panel A: Impact of Finishing Schools									
Finishing school$_{it}$	0.096*	-0.002		0.087*	0.003		0.115**	-0.081	
	(0.054)	(0.039)		(0.052)	(0.037)		(0.049)	(0.061)	
Finishing school$_{it}$ $\times$ women			0.145**			0.100*			0.123*
			(0.059)			(0.058)			(0.066)
Panel B: Impact of Male Schools									
Male school$_{it}$	0.005	0.066		0.015	0.012		0.000	0.075**	
	(0.012)	(0.040)		(0.021)	(0.041)		(0.005)	(0.034)	
Male school$_{it}$ $\times$ men			0.088**			0.072**			0.110***
			(0.038)			(0.036)			(0.030)
City covariates $\times$ period FE	Yes	Yes		Yes	Yes		Yes	Yes	
Religious covariates $\times$ period FE	Yes	Yes		Yes	Yes		Yes	Yes	
City $\times$ period FE			Yes			Yes			Yes
Gender $\times$ period FE			Yes			Yes			Yes

Notes: Testing a panel specification in a window of four centuries before and two centuries after the establishment of finishing schools (N=1,421) or male schools (N=2,161). The outcomes are indicators for the birth of notable women or men. In columns (3), (6), and (9) we construct a panel in which every city $\times$ period cell has two observations: one for women and one for men. This allows us to control for city $\times$ period fixed effects and period fixed effects of the other gender. We include full economic and religious covariates as defined in Table 3.2 in all regressions. Due to collinearity with the 'Male school'-treatment variable, we exclude the educational controls. Standard errors clustered by city reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In panel A of Table 3.3, we limit our sample to 129 cities that ever constructed a finishing school, in a window of four centuries before and two after establishing the first school. In columns (1), (4), and (7) we estimate the impact of finishing schools on notable women, unmarried women and teachers and writers. Despite the reduction in sample size and the omission of educational covariates, the estimated coefficients in this event-study design are close to those of the fixed-effects estimation reported in Table 3.2. Finishing schools do, however, have no impact on the likelihood of observing notable men in our data (columns (2), (5), and (8)). In columns (3), (6), and (9), we construct a panel in which every city-period cell has two observations: one for women and one for men. In this setup, we are able to control for city-by-period fixed

[31] In support of this argument we find that in cities that had both finishing and male schools, the male school was always constructed before the finishing school.

effects and gender-by-period fixed effects to estimate the impact of finishing schools on women, while nonlinearly controlling for the trends in men and city characteristics at any point in time. Our results confirm the pattern observed previously as finishing schools increase the likelihood of a notable woman being born in the city.

In the second panel of Table 3.3, we turn to the impact of male schools on notable women and men. The opening of a male school in a city increases the likelihood of observing a notable man (Columns (2), (5) and (8)), but the impact on women in the same city is a precisely estimated zero (Columns (1), (4), and (7)). Repeating the panel exercise and nonlinearly controlling for city characteristics confirms this pattern and suggests that male schools only had an impact on notable men in the city.

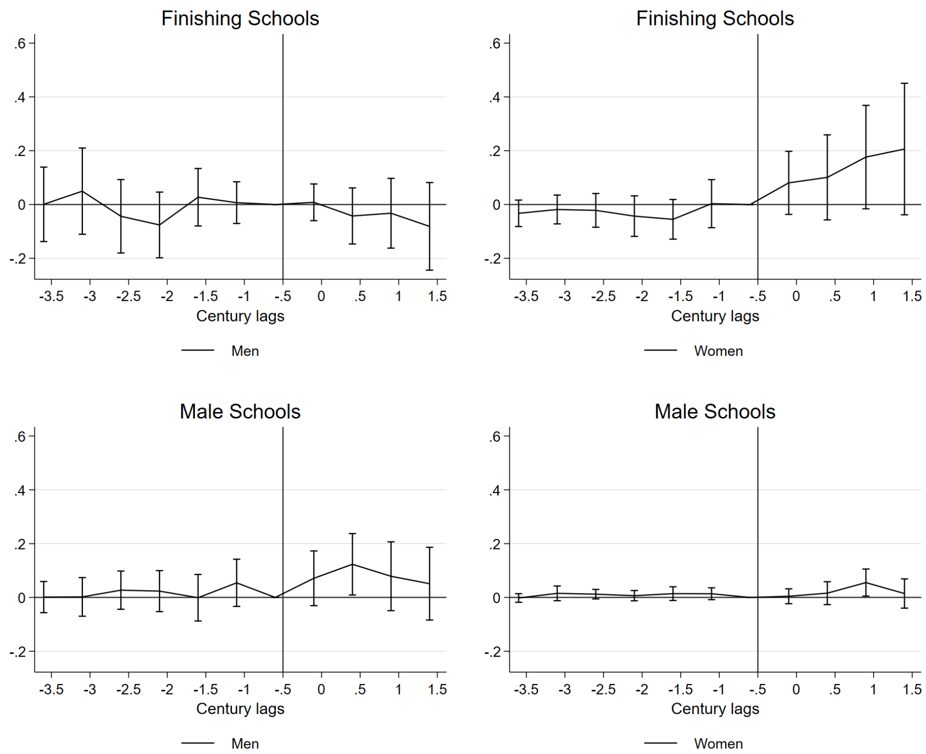


Figure 3.3: Cross-gender impact of male and female schools

Notes: The outcome in the two panels on the left (right) is an indicator equal to one if a notable man (woman) was born in a given city and period. Zero is the normalized opening period of the first finishing school (top panels) or of the first male school (bottom panels) in a city. The vertical line marks the reference period, which we choose to be 50 years prior to establishment of the first school for the respective gender. All figures include full economic and religious controls; educational controls are omitted. 95%-confidence intervals reported.

This evidence is summarized graphically in Figure 3.3. The validity of our point estimates is supported by the absence of pre-trends and the increase in notable women and men after the opening of finishing and male schools, respectively (top right and bottom left). If finishing schools captured local returns to education, in the same way male schools likely do, we would observe a significant increase in the number of men as well (top left). Similarly, if we observed more notable women purely because the returns to education increased, we should observe a similar increase in women when using male schools as the source of variation (bottom right). Since we observe neither, we conclude that differential returns to education are unlikely to explain the increase in the number of notable women after the opening of a finishing school.

3.6.2 Economic growth

In the second placebo exercise, we test whether cities with a steeper growth trajectory established finishing schools earlier. Then, finishing schools might merely reflect the underlying growth potential that attracted the human capital elite.

Under this alternative hypothesis, the increase in notable women born is not a response to the emergence of finishing schools, but a response to increasing income. We identify local economic activity in our panel using city-level construction data by Cantoni, Dittmar, and Yuchtman (2018). If finishing schools are merely a manifestation of increased economic growth, the establishment of finishing schools should be a good predictor of future construction activity. However, this is not borne out in our data: even when defining a subset of growth-specific construction that excludes religious, military, and palace buildings, we find no impact of finishing schools on economic activity in Table 3.4, nor in any period around the opening of finishing schools (Figure 3.4).

Table 3.4: Placebo estimates on the importance of finishing schools: Construction Activity

	$\mathbb{I}[>0]$		*Number*		*log*	
	(1) Any	(2) Growth	(3) Any	(4) Growth	(5) Any	(6) Growth
Finishing school$_{it}$	-0.043	-0.017	1.805	0.939	0.034	0.133
	(0.034)	(0.066)	(1.236)	(0.644)	(0.108)	(0.111)
City covariates $\times$ period FE	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates $\times$ period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates $\times$ period FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Main results using a fixed-effects estimation in a window of four centuries before and two centuries after the establishment of a finishing school (N= 1,421). All regressions include a full set of city and period fixed effects. We include full religious and educational covariates as defined in Table 3.2. As outcomes we consider all construction activity ("Any") in odd columns as well as growth-related construction activity ("Growth") in even columns, which excludes religious, military and palace buildings. In addition, we consider three transformations of these outcomes, namely indicators for building construction (columns 1 and 2), the raw number of buildings constructed (columns 3 and 4) and the log number of buildings constructed (columns 5 and 6). Standard errors clustered by city reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

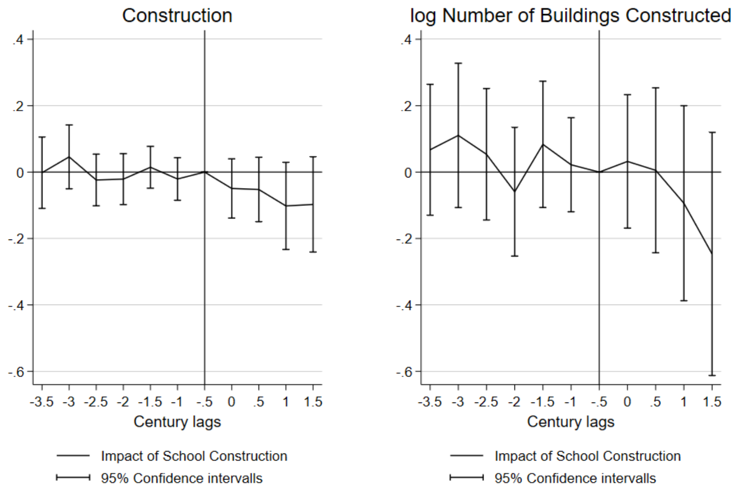


Figure 3.4: Impact of finishing schools on economic growth

Notes: The correlation between finishing schools and building construction. The outcome in the left panel is an indicator variable capturing construction activity in a given city and period, while the outcome in the right panel is the log number of buildings constructed plus one. Zero is the normalized opening period of the first finishing school in a city. The vertical line marks the reference period, which we choose to be 50 years prior to establishment of the first finishing school. Full set of controls included in both figures. 95%-confidence intervals reported.

3.6.3 Cultural change

In the last set of placebo exercises, we provide evidence against the premise that finishing schools are a reflection of broader cultural changes in society. To assess this alternative hypothesis, we exploit city and gender-specific changes in culture: the end of witch trials; the opening of female monasteries; the consecration of churches to a female saint; and the Protestant Reformation. Using event-study designs analogous to our analysis of finishing schools, we find no significant impacts on the prevalence of notable women from any of these cultural changes (Table 3.5 and Figure 3.5).

In panel A of Table 3.5, we use data on the end of witch trials in Germany from Leeson and Russ (2017). Witch trials disproportionately targeted widows living a more independent life as well as midwives and female folk healers (Ehrenreich and English 1973; Oster 2004).[32] We thus argue that the 'end of witch trials' in a city is informative of a change in local culture away from one of the most violent forms of discrimination against women. The threat of the stake forced midwives and folk healers to practice in secrecy. Then, the end of witch trials might have increased their likelihood of entering our sample. However, we see no impact of the end of witch trials on women becoming

[32]Leeson and Russ (2017) collect data on 3,080 witch trials in 121 German cities, with the first and last trials recorded in 1300 and 1792, respectively. Our inclusion is motivated by the fact that 76% of witch trials were conducted before 1648 and 23.5% of women were trialed between 1627–1633; a period in which finishing schools for girls first sprung up.

recognized for their achievements.

In panel B of Table 3.5, we exploit the opening of female monasteries taken from Cantoni, Dittmar, and Yuchtman (2018) as proxies for gender-specific cultural change. Female monasteries presented women with one of the few alternatives to "traditionally advocated marriage" (Frigo and Fernandez 2019) and household roles. The establishment of such monasteries could thus be considered reflective of local culture becoming more accepting towards women choosing a comparatively independent lifestyle.[33] However, we do not find significant impacts of the establishment of female monasteries on the number of notable women once we add economic, religious, and educational controls.

Table 3.5: Placebo estimates on the importance of finishing schools: Changing culture

	Non-Noble Secular		*Unmarried women*		*Teachers & Writers*		*Royals*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: End of witch trials								
End of Witch Trial$_{it}$	0.002	0.052	0.059*	0.062	0.014	0.005	0.031	-0.016
	(0.028)	(0.040)	(0.031)	(0.044)	(0.020)	(0.025)	(0.025)	(0.030)
Religious covariates × period FE		Yes		Yes		Yes		Yes
Panel B: Creation of a female monastery								
Female monastery opens$_{it}$	0.020**	0.012	0.027**	0.018	0.000	-0.004	-0.001	-0.006
	(0.009)	(0.008)	(0.012)	(0.012)	(0.003)	(0.004)	(0.007)	(0.010)
Religious covariates × period FE		Yes		Yes		Yes		Yes
Panel C: Church consecration to a female Saint								
Consecration to a female saint$_{it}$	0.047	0.031	0.019	-0.005	0.040*	0.041	-0.007	0.006
	(0.031)	(0.036)	(0.039)	(0.043)	(0.021)	(0.026)	(0.033)	(0.034)
Religious covariates × period FE		Yes		Yes		Yes		Yes
Panel D: Reformation happening in city								
Reformation in City$_{it}$	0.017	-0.028	0.069**	0.020	0.015	-0.009	0.030	0.036
	(0.025)	(0.019)	(0.027)	(0.033)	(0.015)	(0.016)	(0.033)	(0.041)
Religious covariates × period FE								
City covariates × period FE		Yes		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes		Yes

Notes: Main results using a fixed-effects estimation in a window of four centuries before and two centuries after the end of witch trials (panel A), the creation of a female monastery (panel B), a church consecration to a female Saint after 1650 (panel C), and the arrival of the Protestant reformation in a city (panel D). All outcomes are indicators equal to one if a notable woman from the respective group was born in a given city and period. All regressions include a full set of city and period fixed effects. Cities that ever had witch trials: 112; cities with a female monastery: 221; cities with a female church consecration: 152; cities that turned Protestant: 146. We include covariates as defined in Table 3.2 where indicated. We omit religious covariates in panel D, as our ruler fixed effects, the Catholicism (as of 1618) indicator, and the distance to Germany's denominational divide predict whether a city becomes Protestant. Difference-in-differences estimates confirm this picture and are presented in Table 3.F.3. Standard errors clustered by city reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Next, we turn to the consecration of churches to female saints in panel C of Table 3.5. We utilize data by Cantoni, Dittmar, and Yuchtman (2018) on 12,334 church construction events in Germany, and identify 1,610 events in which a church was conse-

[33] Cantoni, Dittmar, and Yuchtman (2018) have 414 female monasteries in Germany with the average year of foundation being 1275.

crated to honor a female saint.[34] We argue that since churches could be consecrated to any saint, using a female saint might indicate a cultural shift towards the inclusion of women and thus could be correlated with a higher status of women in society. Yet, we identify a precisely estimated null effect throughout all specifications.

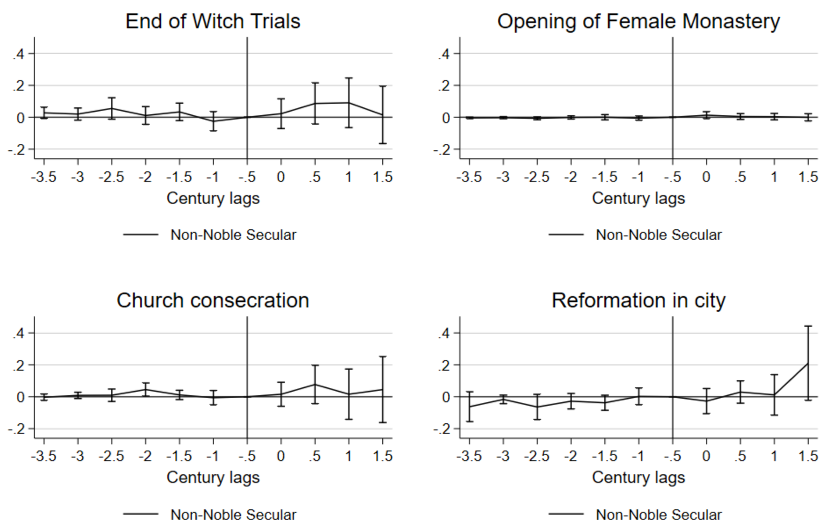


Figure 3.5: Impact of cultural change on notable women

Notes: The correlation between notable women and cultural change. The outcome in all panels is an indicator equal to one if a non-noble secular woman was born in a given city and period. The vertical line marks the reference period, which we choose to be 50 years prior to the respective event. Economic and educational controls included in all figures. Religious controls are omitted when identifying the impact of the Protestant Reformation. 95%-confidence intervals reported.

In panel D of Table 3.5, we use the timing of the Protestant Reformation in each city as an indicator of a potential shift in the status of women. We follow S. O. Becker and Woessmann (2008, 2009) who argue that, since Martin Luther suggested that women needed to be able to read, Protestantism had a positive impact on female education.[35] We utilize data by Cantoni (2015) on the timing of the Reformation in cities, to proxy for a cultural shift towards the inclusion and primary education of women following Luther's teachings. Our findings suggest that Protestantism, and the associated potential shift in gender roles, cannot explain the increase in notable women, teachers, or any other subcategory.[36]

[34]The average year of consecration in the data of Cantoni, Dittmar, and Yuchtman (2018) is 1452 in 260 cities.

[35]Note that this requirement to read was interpreted as providing basic primary schooling. Finishing schools provided secondary education that included French, arithmetic, and literature classes.

[36]We have 146 cities turning to Protestantism, 129 of which switched by the end of the sixteenth century. We substantiate our finding in Table 3.F.3 in which we use those cities in a standard difference-in-differences setup, and find weak results on non-noble secular women, but no results on teachers, activists, or nobility. We use the log distance to Wittenberg as an instrument (S. O. Becker and Woessmann 2009) and report insignificant reduced form impacts on notable women. The OLS estimates however, suffer from a pre-trend in which cities with more notable women are more likely to become Protestant.

The results presented in the event-study graph in Figure 3.5 support the findings from Table 3.5: it is unlikely that gender-specific cultural change contributed to the establishment of finishing schools and the following increase in notable women. We conclude that unobserved economic or cultural change is unlikely to bias our estimates on finishing schools. Instead, it is more likely that finishing schools were established by religious orders in response to religious competition or idiosyncratic shocks. Thus, finishing schools, conditional on fixed effects, can be interpreted as an exogenous shift in the supply of secondary education for women.

3.7 Mechanism

Based on the historical literature on finishing schools (Albisetti 1988) and the women's rights movement (Schraut 2019), we derive two complementary mechanisms that link the establishment of finishing schools to an emerging nucleus of the women's rights movement: access to critical ideas about women's role in society, and reduced costs of forming and accessing networks of like-minded peers. We interpret our results thus far as critical ideas about women's role in society taking hold in cities with finishing schools, as more unmarried women entered the human capital elite as teachers, writers and women's rights activists. In this section, we shed light on the second mechanism: finishing schools reducing the cost of forming and accessing networks of like-minded women. We document that the establishment of finishing schools positively impacted the emergence and size of networks between notable women and increased the immigration of notable women, further contributing to network formation.

3.7.1 Networks between notable women

We construct our measure of networks between women by analyzing their biographies in the *Neue Deutsche Biographie*. Here, we define a connection between two women if one is mentioned in the biographical text of the other, and the younger was at least 16 years old when the older woman died. A network thus exists in a city if at least one local woman is connected to another notable woman.[37] The size of a city's network in period t is then defined as the sum of notable women being mentioned in the biographies of all other women born in that city in period t.

In Table 3.6, we analyze the impact of finishing schools on networks between notable women. We find that finishing schools increase the likelihood of observing a network

[37] An example is Gertrud Bäumer: she attended the finishing school in Halle and became a teacher in Magdeburg. She was introduced to Helene Lange by an older colleague and joined the *Allgemeiner Deutschen Lehrerinnenverein* in Berlin 1898. Throughout their career, Bäumer and Lange closely collaborated on promoting women's rights, in particular women's access to education.

and its size four-fold (panel A). The estimated effect, however, predictably varies by the type of network constructed: in stark contrast to networks between non-noble secular (panel B) or politically active women (panel C), connections between religious or noble networks are unaffected by establishing a finishing school (panel D). The results on networks between notable women echo our main results: finishing schools increase networks only for politically active women, but not for the placebo group of the nobility.

Table 3.6: Fixed-effects results on the importance of finishing schools: Network formation within cities

	$\mathbb{I}[Connections > 0]$		log Connections	
	(1)	(2)	(3)	(4)
Panel A: Any network in city				
Finishing school$_{it}$	0.060***	0.043***	0.069***	0.052**
	(0.016)	(0.016)	(0.021)	(0.021)
Mean, untreated	0.015	0.015	0.020	0.020
Panel B: Network between non-noble secular women				
Finishing school$_{it}$	0.060***	0.043***	0.067***	0.052***
	(0.016)	(0.016)	(0.021)	(0.020)
Mean, untreated	0.012	0.012	0.016	0.016
Panel C: Network between politically active women				
Finishing school$_{it}$	0.016**	0.012	0.018**	0.015*
	(0.007)	(0.009)	(0.008)	(0.009)
Mean, untreated	0.003	0.003	0.003	0.003
Panel D: Network between religious women				
Finishing school$_{it}$	0.006	0.005	0.005	0.004
	(0.005)	(0.007)	(0.004)	(0.005)
Mean, untreated	0.004	0.004	0.004	0.004
Unit trend	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes
Religious covariates × period FE		Yes		Yes
Educational covariates × period FE		Yes		Yes
Observations	9,312	9,240	9,312	9,240

Notes: Main results using a fixed-effects estimation and all cities in all periods. All regressions include a full set of city and period fixed effects. We consider two types of dependent variables to capture the extensive and intensive margin of connections among notable women. $\mathbb{I}[Connections > 0]$ is an indicator equal to one if a city had at least one connected women born in this period, while 'log Connections' constitutes the natural logarithm of the number of women with connections plus one. We regress the number of connections between any women, non-noble secular women, politically active women, and religious women, as defined in the top row of each Panel, on our finishing school variable. We include covariates as defined in Table 3.2 where indicated. Standard errors clustered by city shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.7.2 Immigration of notable women

We provide further evidence on the formation of networks using the immigration of notable women. If finishing schools facilitated women to form and access networks of like-minded women, presumably they also increased the likelihood that women migrated to the city, acting as a pull factor. We document migration patterns using the difference between women's places of birth and death as recorded in the *Neue Deutsche Biographie*. A total of 507 women in our data have migrated at least 10 km between birth and death. We repeat our event-study for these immigrated non-noble secular women in Figure 3.6. Again, we observe no pre-trends and a distinct increase in the likelihood of immigration after the opening of the first finishing school (left panel); a finding robust to including control variables (right panel).

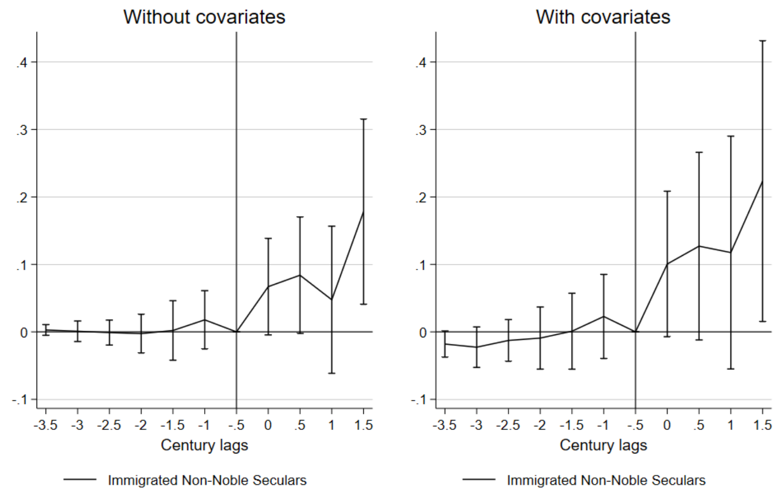


Figure 3.6: Impact of finishing school establishment on migrated women

Notes: Main results for women who migrated during their lifetimes, focusing on cities that ever established a finishing school. Zero is the normalized opening period of the first finishing school in a city. The vertical line marks the reference period, which we choose to be 50 years prior to establishment of the first finishing school. Full economic, religious, and educational controls added in the right panel. Point estimates reported in Table 3.G.1. 95%-confidence intervals reported.

To identify whether finishing schools attracted notable women, or the immigration of notable women instead facilitated the foundation of finishing schools (reverse causality), we provide two pieces of evidence: first, if immigration of notable women increased the likelihood of finishing school opening, Figure 3.6 would show differential pre-trends. The absence of such pre-trends suggests that finishing schools had a similar effect on immigrated women as on native women, and that finishing schools are likely not a result of immigration.

Second, we build on this result and provide further support for the idea of increased networking activity using the timing of immigration, or birth, of the first notable women as our source of variation. If finishing schools increased women's representation among the human capital elite, which in turn attracted notable women from other cities, we would observe that the first native notable woman increases immigration. If, however, immigration led to the opening of finishing schools, and therewith to the formation of a female human capital elite, the first immigration event would increase the number of notable women born in a city.

We explore these alternative hypotheses in Figure 3.7, using either the first woman who migrated to a city (left panel) or the first notable woman born in a city (right panel) as a shifter in the likelihood of observing future notable women being born. Using the first migration event as the 'treatment period' in the left panel, we report no impact on future non-noble secular women being born. In contrast, the right hand side of Figure 3.7 reveals that the first native-born notable woman induces a strong increase in immigration of other notable women from elsewhere.

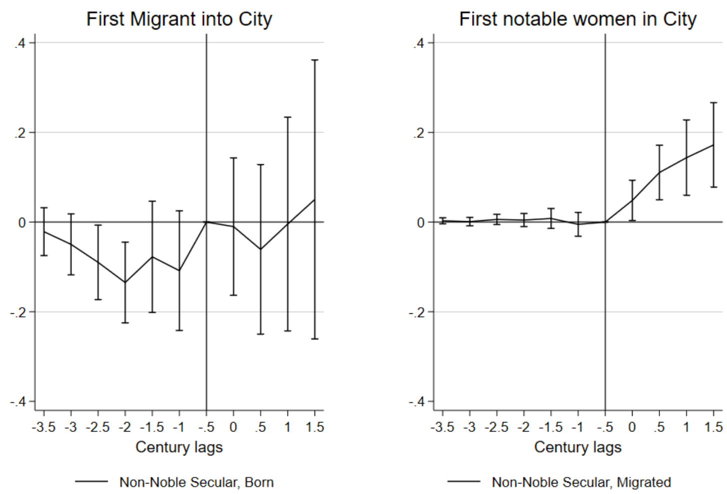


Figure 3.7: Impact of native and migrated women on subsequent notable women

Notes: The impact of the first notable female migrant on the birth of "native" notable women in a city is shown in the left panel. Conversely, the right panel shows the impact of the first "native" notable woman born in a city on the migration of notable women into the city. Zero is the normalized period of either the first migrated notable woman (left) or the first notable woman born in a city (right). Correspondingly, the outcome in the left panel is an indicator equal to one if a notable woman was born in a given city and period, while the outcome in the right panel is an indicator equal to one if at least one notable woman migrated to a city in a given period. The vertical line marks the reference period, which we choose to be 50 years prior to the respective event. Full controls included in both figures. 95%-confidence intervals reported.

Our results thus indicate that finishing schools increased women's representation among the human capital elite: women became teachers, writers and early activists, indicating that critical ideas about women's role in society took hold in cities with finishing schools. These women would eventually form networks with other women from the human capital elite and attracted other like-minded women from other cities. These early networks laid the foundation for the further dissemination of critical ideas and the institutionalization of the women's rights movement.

3.8 Finishing schools and the women's rights movement

When Dr. Martin Luther King Jr. and Susan B. Anthony spread their ideas and institutionalized their movement, they provided the social acceptance required for the civil rights and suffrage movements to succeed. German activists from the early phase of the women's rights movement pursued similar strategies to gain broader public appeal and turn their movement into a societal force (Berndt 2019; Nagelschmidt and Ludwig 1996; Schraut 2019). We measure the dissemination of critical ideas by digitizing all letters to the editor of the feminist newspaper *Frauen-Zeitung*, in which women's role in society was critically discussed. To capture the increasing institutionalization, we use establishment and membership data of local chapters of the women's rights movement in 1909. Lastly, we provide evidence that finishing schools, via accumulating human capital, disseminating critical ideas, and institutionalizing the movement, increased female representation in parliaments once suffrage was achieved.

3.8.1 Empirical approach

We document the link between finishing schools and the success of the women's rights movement in a cross-sectional setting. Specifically, we show that cities c with finishing schools in 1850 send more letters to the *Frauen-Zeitung* and have more local chapters of the women's rights movement in 1909. In doing so, we estimate cross-sectional regressions using specifications of the following type:

$$Y_c = \alpha + \beta \text{ finishing schools}_c + \boldsymbol{X_c}\, \gamma' + \varepsilon_c \tag{3.1}$$

In this cross-sectional setting, unobservable factors, previously captured by city fixed effects and linear time trends, potentially impact our interpretation. Even controlling for economic, religious and educational covariates ($\boldsymbol{X_c}$), unobservable factors could be correlated with the establishment of finishing schools and the women's rights movement. When schools were built in areas with greater appreciation of women's role in society or women's education, our point estimate would overstate the impact of finishing schools. We assess the magnitude of this potential bias using three complementary strategies: first, we report the bias-adjusted point estimate from a bounding exercise in the spirit of Oster (2019), comparing coefficients from a regression without any controls and restrictions to a regression with a full set of controls in areas of religious competition. Second, in Appendix 3.H we corroborate these findings and report point estimates from an instrumental variables strategy using monasteries in 1300 and religious competition as a shifter in the likelihood of establishing finishing schools. Third, we compare the effect of finishing schools using propensity score matching on all covariates in Appendix 3.H.1. All strategies reveal, if anything, a downward bias of our point estimates.

The historical literature on finishing schools suggests that religious competition was one determinant of the location of early finishing schools (Lewejohann 2014). Yet, religious competition may exhibit a direct effect on our measures, even when controlling for the distance to the denominational boundary. Thus, we limit our sample to cities within 10 km of the borders marking the religious divide in 1618, i.e. to regions where religious competition was particularly pronounced in the early phases of finishing school openings. Limiting our sample to cities within 10 km of the religious divide also enhances the comparability of cities. For instance, rather than comparing Berlin to Munich (600km due south), our strategy compares the neighboring cities of Hanover and Hildesheim.

We present our results linking finishing schools with the emergence of the women's rights movement in the late nineteenth century and with political representation of women throughout the twentieth century in Table 3.7. We start by examining the link between historical finishing schools established by 1850 and the dissemination of critical ideas of women's role in society to the general public (panel A), and the institutionalization of the women's rights movement by founding local chapters and recruiting female members (panels B and C). We then turn to an important outcome of the women's rights movement, female representation in parliaments after women achieved the right to both vote and stand for parliament in 1919 (panels D and E).

3.8.2 Dissemination of ideas

To measure dissemination of critical ideas, we digitize all letters to the editor of one of the first feminist newspapers in Germany, *Frauen-Zeitung* (1849-52), in panel A. We use the place of residence of all letter writers and link this to the presence of finishing schools in the nearest city. In Table 3.7 column (1), we estimate a bivariate regression without controls and restrictions, documenting an increase in the likelihood of sending a letter of 0.100 (s.e. 0.017), a 150% increase over the mean. Only 6.2% of cities without finishing schools by 1850 sent letters to *Frauen-Zeitung*, compared to 16.2% of cities with finishing schools. We interpret this increase as evidence that critical ideas are more common in cities with finishing schools.

To assess the potential severity of selection on unobservables, we report the bias-adjusted point estimate from a restricted estimation in column (2). Here, we include all previously defined controls and limit the sample to areas that, 200 years prior to the foundation of *Frauen-Zeitung*, had been religiously competitive. We estimate a similar point estimate of 0.122 (s.e. 0.037), a four-fold increase over the likelihood of

Table 3.7: Long-term impact of finishing schools on the women's rights movement and political representation

	$\mathbb{I}[> 0]$		log Number	
	(1)	(2)	(3)	(4)
Panel A: Leserbriefe, Frauenzeitung, 1849–1852				
Finishing schools	0.100***	0.122***	0.192***	0.241**
	(0.017)	(0.037)	(0.051)	(0.097)
R-squared	0.121	0.370	0.151	0.353
Mean, untreated	0.062	0.038	0.104	0.061
Bias-Adjusted β		0.132		0.266
Panel B: All women's rights organizations				
Finishing schools	0.150***	0.137***	1.419***	1.157***
	(0.027)	(0.050)	(0.179)	(0.306)
R-squared	0.101	0.362	0.211	0.483
Mean, untreated	0.367	0.275	444.355	155.802
Bias-Adjusted β		0.132		1.021
Panel C: Women's rights organizations to promote equal access to education				
Finishing schools	0.128***	0.074**	0.779***	0.496**
	(0.017)	(0.036)	(0.112)	(0.217)
R-squared	0.165	0.399	0.198	0.426
Mean, untreated	0.046	0.038	12.973	13.023
Bias-Adjusted β		0.046		0.337
Panel D: Member Parliament, 1919–1933				
Finishing schools	0.103***	0.101***	0.133***	0.105***
	(0.017)	(0.034)	(0.027)	(0.035)
R-squared	0.107	0.418	0.195	0.472
Mean, untreated	0.066	0.038	0.073	0.053
Bias-Adjusted β		0.100		0.091
Panel E: Member Parliament, 1949–2019				
Finishing schools	0.099***	0.091*	0.312***	0.268***
	(0.020)	(0.047)	(0.036)	(0.071)
R-squared	0.048	0.282	0.203	0.402
Mean, untreated	0.556	0.527	1.170	1.031
Bias-Adjusted β		0.088		0.241
City Covariates		Yes		Yes
Religious covariates		Yes		Yes
Educational covariates		Yes		Yes
Observations	388	183	388	183
Bandwidth		10		10

Notes: Cross-sectional results using all observations in odd columns and a sample limited to cities within 10 km of the denominational boundary as of 1618 in even columns. We use two transformations of our outcomes – an indicator version (columns 1 and 2) as well as a logarithmized version (columns 3 and 4) – and regress them on the number of finishing schools in a city. We include covariates as defined in Table 3.2 where indicated. "Bias-Adjusted β" is calculated following the procedure laid out in Oster (2019), assuming $R^{max} = 1.3\tilde{R}$ and $\delta = 1$. Standard errors clustered by city shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

sending a letter in cities without a finishing schools (0.038 in this sample). The bias-adjusted point estimate is of a similar magnitude to the baseline (0.132), indicating a slight downward bias stemming from selection on unobservable factors. In columns

(3) and (4) of Table 3.7, we repeat this exercise with the number of letters sent. Again, the bias-adjusted point estimate confirms the OLS point estimate and suggests a 24% increase in the number of letters sent to *Frauen-Zeitung*.[38]

3.8.3 Institutionalization of the movement

Next, we turn to studying the institutionalization of the German women's rights movement. To measure the institutionalization of networks in the second half of the nineteenth and the early twentieth century, we digitize novel data on local chapters of women's rights associations from the Imperial Statistical Office (Kaiserliches Statistisches Amt 1909). This source provides detailed establishment and membership data on more than 1,200 local chapters in 1909. The average local chapter in our dataset was established in 1898 and counted approximately 1,600 members. This source also allows us to differentiate between different types of associations – for example, female suffrage associations and associations dedicated to improving women's educational opportunities.

We exploit this unique micro data in panels B and C of Table 3.7. Controlling for covariates in column (2), we find that an additional finishing school by 1850 increases the likelihood that a city has any local women's rights association by 14 percentage points (panel B), equivalent to a 50% increase over the mean in cities without finishing schools. In particular, associations dedicated to promoting equal access to education for women exhibited stronger public support: if cities had established finishing schools by 1850, the number of members in these organizations exceeded that in cities without schools by 50% (panel C, column 4).[39]

3.8.4 Female representation in parliament

Our results suggest that critical ideas took hold in cities with finishing schools, leading to more members in women's rights organizations than in cities without finishing schools. First, the increasing representation of women among the human capital elite (Table 3.2) contributed to the creation of networks between cities that attracted other notable women (Figure 3.7). Second, these women were up to three times more likely to disseminate their critical ideas using the first female-led newspaper, the *Frauen-Zeitung*, as an outlet (Table 3.7, panel A). Finally, they organized in women's rights

[38]We use the transformation $\log(y+1)$ in columns (3) and (4). Due to the sparsity of our outcome data, we refer to columns (1) and (2) for inference. We only record 242 letters from 40 cities, with five cities sending over half of the letters.

[39]In Appendix 3.I, we directly correlate the number of non-noble secular women in 1850 with political activity at the turn of the century: a 10% increase in the number of notable women increases political activity by 15%.

groups (Table 3.7, panel B) and jointly lobbied for the core demands of the women's rights movement: equal access to education and female suffrage.

Thus, by educating young women and teachers, finishing schools contributed to the formation of a human capital elite that ultimately succeeded in achieving suffrage in 1919. Once suffrage was gained, this larger representation of women among the human capital elite should have translated into greater female political representation in parliaments.

We explore this hypothesis in panels D and E of Table 3.7. To measure political representation, we collect the place of birth of all female members of parliament in the Weimar Republic (1919–1933, panel D) and the Federal Republic of Germany (1949–2019, panel E).[40] We report positive and significant coefficients when regressing an indicator for and the number of female politicians in all parliamentary elections since 1919 on the number of finishing schools in 1850.[41]

While during the Weimar Republic, only 4% of cities without finishing schools sent women to parliament, this figure rose to 53% in the Federal Republic of Germany (panel D, column 2). In contrast, cities with historical finishing schools were 10 percentage points more likely to have sent women to parliament, equivalent to a 250% increase during the Weimar Republic and a 25% increase during the Federal Republic. panel D and E thus highlight cities' historical advantage as early movers towards a more gender-equal society, gained by the establishment of finishing schools more than 300 years earlier.[42]

3.9 Conclusion

We set out to determine conditions for the emergence and success of social movements using the example of the women's rights movement in Germany. Following the literature on social movements (Markoff 2015; Tilly, Castañeda, and Wood 2020) and the history of successful movements (Dr. Martin Luther King Jr. for the civil rights movement or Susan B. Anthony for the suffrage movement) we identify three key milestones. First, future leaders are educated and develop critical ideas. Second, these leaders disseminate their ideas using available mass media. Third, leaders institution-

[40]Germany uses a list-based electoral system in which voters vote for a party list. Thus, female representation on this list is relatively more likely driven by a woman's preference to be nominated, than by her electorate's preference, compared to a system where voters directly choose their representative.

[41]The findings are robust to estimating the impact in every period separately or jointly. The findings are not driven by large cities as the top 5 cities with the most finishing schools are Munich, Berlin, Obertaunuskreis, Landshut, and Dresden. Estimates increase without the largest 10 percent of the sample in 1600.

[42]We explore this early movers hypothesis in more detail in Appendix Table 3.H.2. Here, a city with 50 more years of exposure to finishing schools would imply 14% more letters to *Frauen-Zeitung*, twice the number of women's rights organizations in 1909 and 23% more women in parliament today.

alize their movement as their ideas take root in society.

We study the importance of one form of educational institution at these three milestones, using the example of the arrival of finishing schools and the women's rights movement in Germany. In this setting, newly collected panel and cross-sectional data allow us to draw out the effect of education on the success of social movements at every step of their development. First, after cities established finishing schools, women started to represent a larger share of the political, intellectual, and economic elite ("human capital elite"), forming an activist nucleus of the women's rights movement. Second, women born in such cities also sent a disproportionate share of editorial letters to female-led newspapers, important platforms for early women's rights activism. Third, cities with historical finishing schools hosted more and larger women's rights organizations, key forces in the advancement of women's empowerment.

Using a wide range of empirical specifications our paper highlights the role of education in contributing to the emergence and success of the German women's rights movement. Further, our empirical results suggest that a world without educational institutions but significant economic and cultural changes would not see the level or pace of social change we observe throughout history.

Taken together, our findings indicate that educational institutions, which foster the exchange of critical ideas and provide the space to form networks, can function as important catalysts for the formation of a human capital elite critical of their status quo. Yet, education does not only benefit those receiving it; to the contrary, society as a whole can benefit when committed activists fight for and bring about social change.

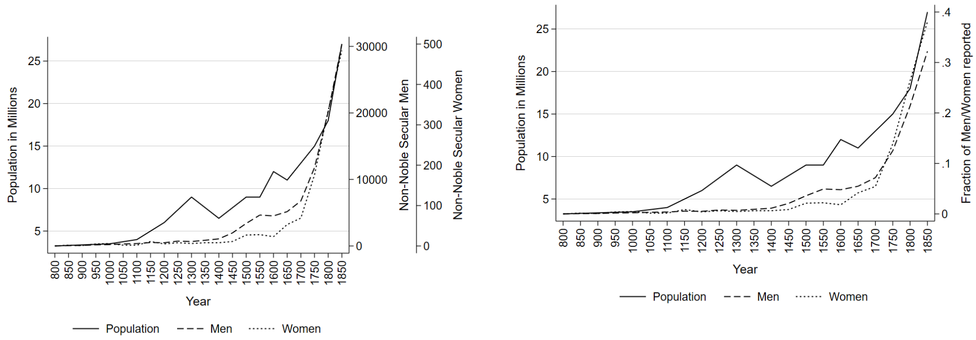
APPENDIX 3

3.A Record keeping in the *Neue Deutsche Biographie (NDB)*

Our main results show an increase in the representation of women among the human capital elite – as measured by notable women recorded in the NDB – following the establishment of finishing schools. In this Appendix we explore whether this increased representation of women is driven by changes in reporting. If women's inclusion in the NDB increased disproportionately over time, estimates of the impact of finishing schools might be confounded by a general time effect. In Figure 3.A.1, we provide direct evidence against this concern: the recording of notable women and men in the NDB followed the same trend, which is, moreover, in line with general population growth. This motivates our use of the *share* of notable women among all notable individuals as dependent variable and our interpretation of the data in the main text.

In Figure 3.A.1, we compare the trends of total population in Germany based on McEvedy and Jones (1978) to the trends in the number of men and women recorded in the NDB. While the levels are different, all time series follow the same trend over time suggesting no change in reporting that could affect our data. The right panel in Figure in 3.A.1 shows that also the fraction of non-noble secular women among all women in our data increased similar to the corresponding increase among notable men: women's non-noble secular shares went up from 10% to 80% with the men's increase being 35% to 90%. Again, the pattern closely follows population, so that calculating the share of women born in each city and period, relative to all notable women and men in that city and period, provides a good measure of women's representation among the human capital elite as it explicitly controls for trends.

A related concern is differential reporting between cities with and cities without finishing schools in the NDB. Specifically, finishing schools may have improved record keeping on notable women rather than increased women's share among the human capital elite. We offer two arguments against this interpretation: first, as shown in Figure 3.2 in the main text, we find no impact of finishing schools on notable women from the nobility; if finishing schools merely improved record keeping on notable women, one might reasonably expect this to manifest also in an increased representation of women from the nobility. Second, if finishing schools merely improved record keeping in the NDB, this ought to show up in differential pre-trends, as a purported record-keeping effect would presumably also extend to the women who contributed to the founding of finishing schools. However, as shown in Figure 3.2 and as emphasized in Appendix 3.E we find strong evidence against differential pre-trends.



(a) Women and men in our data.

(b) Share of non-noble secular women and men.

Figure 3.A.1: Number of women and men in the NDB relative to total population

Notes: The left panel depicts the population of Germany in its modern boundaries (solid line), the number of notable men (right axis, dashed line), and the number of notable women born in each period (right axis, dotted line). All lines follow the same trend, suggesting that our estimated impacts are not driven by a change in reporting. The right panel again depicts the population of Germany in its modern boundaries as well as the share of all *non-noble secular* women (men) among all notable women (men) born in each period. This indicates that also in the subcategory of *non-noble secular* individuals the NDB exhibits no differential time trends in reporting between women and men.

3.B Alternative empirical specifications and economic growth

We continue by documenting the robustness of our results presented in Table 3.2 in the main text. To this end, we start by the most basic two-way fixed effect design, only including period and city fixed effects in column (1) of Table 3.B.1. In the four subsequent columns we individually add and remove a city-specific trend as well as city, educational, and religious covariates. As expected, the largest drop originates from city covariates, and specifically, controlling for population. These covariates are responsible for almost the entire difference between the baseline and full specifications. This effect is largely an extensive margin effect, as when we drop all cities without population figures in 1600, we do not observe a change in the point estimates. The city-specific trend, while changing the point estimate significantly between columns (1) and (2), does not affect the point estimates when already controlling for covariates (columns (6) vs (7)). We thus conclude that our estimates do not rely on the inclusion of city-specific trends or a specific specification.

Table 3.B.1: Fixed-effects results on the importance of finishing schools - Sensitivity to covariates

	Baseline	with trends	with covariates				Full
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Non-Noble Seculars, $\mathbb{I}[Women > 0]$							
Finishing school$_{it}$	0.300***	0.230***	0.177***	0.298***	0.274***	0.181***	0.164***
	(0.029)	(0.029)	(0.030)	(0.030)	(0.031)	(0.032)	(0.033)
Panel B: Non-Noble Seculars, log Women							
Finishing school$_{it}$	0.464***	0.355***	0.235***	0.460***	0.423***	0.246***	0.204***
	(0.063)	(0.053)	(0.046)	(0.063)	(0.063)	(0.048)	(0.045)
Panel C: Non-Noble Seculars, Share Women							
Finishing school$_{it}$	0.022***	0.019***	0.019***	0.022***	0.023***	0.021***	0.021***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Panel D: Unmarried women, $\mathbb{I}[Women > 0]$							
Finishing school$_{it}$	0.276***	0.194***	0.167***	0.272***	0.260***	0.173***	0.147***
	(0.029)	(0.030)	(0.031)	(0.029)	(0.031)	(0.032)	(0.034)
Panel E: Unmarried women, log Women							
Finishing school$_{it}$	0.422***	0.302***	0.215***	0.415***	0.388***	0.226***	0.173***
	(0.060)	(0.049)	(0.045)	(0.061)	(0.060)	(0.047)	(0.043)
Panel F: Unmarried women, Share Women							
Finishing school$_{it}$	0.015***	0.011**	0.015***	0.015***	0.017***	0.016***	0.014**
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.006)
Panel G: Teachers & Writers, $\mathbb{I}[Women > 0]$							
Finishing school$_{it}$	0.196***	0.151***	0.111***	0.196***	0.176***	0.117***	0.104***
	(0.026)	(0.027)	(0.024)	(0.026)	(0.026)	(0.024)	(0.026)
Panel H: Teachers & Writers, log Women							
Finishing school$_{it}$	0.220***	0.174***	0.116***	0.220***	0.194***	0.120***	0.103***
	(0.037)	(0.034)	(0.027)	(0.037)	(0.035)	(0.028)	(0.029)
Panel I: Teachers & Writers, Share Women							
Finishing school$_{it}$	0.024***	0.019***	0.018***	0.024***	0.023***	0.019***	0.017***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Panel J: Activists, $\mathbb{I}[Women > 0]$							
Finishing school$_{it}$	0.077***	0.076***	0.044***	0.077***	0.078***	0.051***	0.053***
	(0.017)	(0.018)	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)
Panel K: Activists, log Women							
Finishing school$_{it}$	0.066***	0.064***	0.038***	0.067***	0.066***	0.043***	0.043***
	(0.017)	(0.017)	(0.014)	(0.018)	(0.017)	(0.015)	(0.015)
Panel L: Activists, Share Women							
Finishing school$_{it}$	0.011***	0.013***	0.008**	0.011***	0.012***	0.009**	0.011**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Unit trend		Yes					Yes
City covariates × period FE			Yes			Yes	Yes
Educational covariates × period FE				Yes		Yes	Yes
Religious covariates × period FE					Yes	Yes	Yes
Observations	9,312	9,312	9,312	9,288	9,264	9,240	9,240

Notes: Results from fixed-effects estimations using all cities and periods reported. We consider three types of dependent variables for several categories of notable women: (i) $\mathbb{I}[Women > 0]$ is an indicator taking value 1 if a city had at least one notable woman born in this period; (ii) "log Women" constitutes the natural logarithm of the number of women born plus one; (iii) "Share Women" divides the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians instead. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Column (1) denotes the absolute baseline, only including period and city fixed effects. Column (2) adds linear time trends to ascertain their impact on the point estimate. In columns (3)-(6), we add our full set of controls (as defined in Table 3.2) interacted with period fixed effects, first individually then jointly, without the linear time trends. In column (7), we then add linear time trends to show that linear time-trends do not impact the precision of our estimates. Standard errors clustered by city reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In a final step, we try to identify pairs of cities that only differ in the presence of finishing schools. Instead of classical matching procedures, which are usually done in cross-sectional settings, we employ increasingly parsimonious fixed effects to create smaller and smaller "cells" for cities in Table 3.B.2. We start with the full-specification including city-specific trends and all covariates interacted with period fixed effects. In column (2), we include fixed effects grouping cities into 3,244 cells according to their similarity regarding population, membership in the Hanseatic League, occurrence of anti-Jewish pogroms and religious battles within a given period. In columns (3) and (4) we slowly add similar cells for religious and educational covariates, before exactly matching on educational and economic covariates resulting in 6,580 different cells for cities to fall into. The results remain robust throughout the entire set of specifications.

Table 3.B.2: Fixed-effects results on the importance of finishing schools - Exactly matching on covariates in 1600

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $\mathbb{I}[Women > 0]$						
Finishing school$_{it}$	0.164***	0.203***	0.214***	0.159***	0.164***	0.171***
	(0.033)	(0.038)	(0.040)	(0.045)	(0.050)	(0.047)
Panel B: log Women						
Finishing school$_{it}$	0.204***	0.224***	0.238***	0.163***	0.167***	0.175***
	(0.045)	(0.047)	(0.050)	(0.055)	(0.059)	(0.058)
Panel C: Share Women						
Finishing school$_{it}$	0.021***	0.021***	0.021***	0.014**	0.015*	0.015**
	(0.005)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Exact match on economic covariates		Yes	Yes	Yes	Yes	Yes
Exact match on religious covariates			Yes	Yes	Yes	Yes
Exact match on educational covariates				Yes	Yes	Yes
Exact match on educational and economic covariates					Yes	Yes
Exact match on educational and religious covariates						Yes
Observations	9,312	9,312	9,312	9,312	9,312	9,312
Number of Fixed Effects	1,300	3,244	3,484	5,284	6,580	5,956

Notes: Results from fixed-effects regressions reported. We consider three types of dependent variables: (i) $\mathbb{I}[Women > 0]$ is an indicator taking value 1 if a city had at least one notable woman born in this period; (ii) "log Women" constitutes the natural logarithm of the number of women born plus 1; (iii) "Share Women" divides the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians instead. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. We include controls as defined in Table 3.2. In column (2) we interact all economic covariates with each other to compare cities within one population/Hanseatic League/Jewish presence/anti-Jewish pogrom/confessional battle/period cell. In column (3) we additionally interact all religious covariates with each other. In column (4) we additionally interact all educational covariates with each other. In column (5) we additionally control for the interaction of (2) and (4). In column (6) we additionally control for the interaction of (2) and (3). All covariates are interacted with period fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.B.1 Sensitivity to dropping observations

In a recent paper, Broderick, Giordano, and Meager (2020) stressed the importance of assessing the validity of results by analyzing their robustness to outliers. We implement this robustness test as follows: we drop entire sets of cities belonging to one ruling house rather than dropping individual cities (1 out of 388). With this procedure, we drop on average 18 cities, with the two largest sets of cities being ruled by the Catholic clergy (114) and the House of Hohenzollern (52). Since these two sets of cities also capture the distinction between Catholic and Protestant cities almost perfectly, the results of this analysis also document that our findings are not driven by cities from either denomination alone.

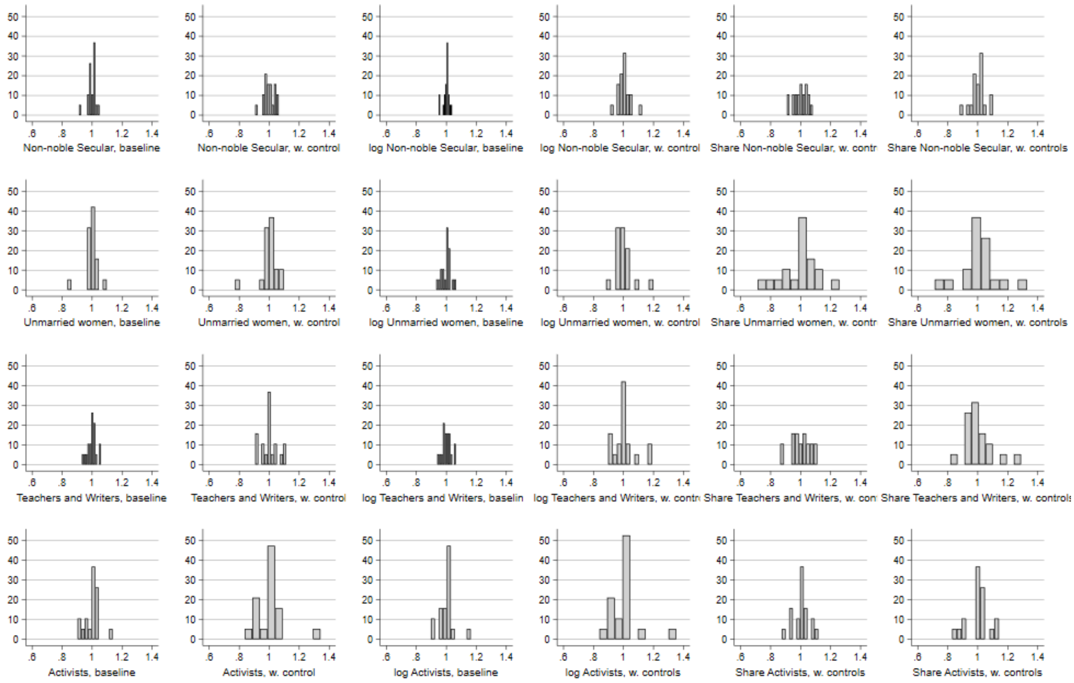


Figure 3.B.1: Senstivity to dropping sets of cities: Panel outcomes

Notes: The x-axis measures the ratio between the restricted point estimate when dropping one of 22 sets of cities and the corresponding original estimate in Table 3.2. This ratio is one if the restricted estimate is unchanged, 1.5 if the restricted estimate is 50% larger than the original, and 0.5 if the restricted estimate is 50% smaller than the original. We present all outcomes (rows) in all specifications (columns) corresponding to Table 3.2. The sum of all bars is 100%.

In Figure 3.B.1 and 3.B.2, we present all outcomes (rows) in all specifications (columns) corresponding to Tables 3.2 and 3.7. The x-axis measures the ratio between a restricted estimate when a set of cities is dropped and the original estimate from the corresponding table. If the restricted estimate remains unchanged, this ratio is one. It is 1.5 if the restricted estimate is 50% larger than the original, and 0.5 if the restricted estimate is 50% smaller than the original. We do this for 22 sets of cities belonging to different rulers and find a minimum of 0.7 (for the share of unmarried women) and a maximum of 1.3 (for the log number of activists) in the panel setting. These figures suggest that our panel estimates are highly robust to potential outliers as they only vary within 30% of the original effect size. The corresponding numbers for the cross-sectional regressions are 0.7 (for the log number of educational women's rights associations, with controls) and 1.6 (for the members of parliament 1949-2017, with controls).

Overall, the density plots reveal a stable pattern around the estimated mean, suggesting that our results are not driven by individual cities or sample selection.

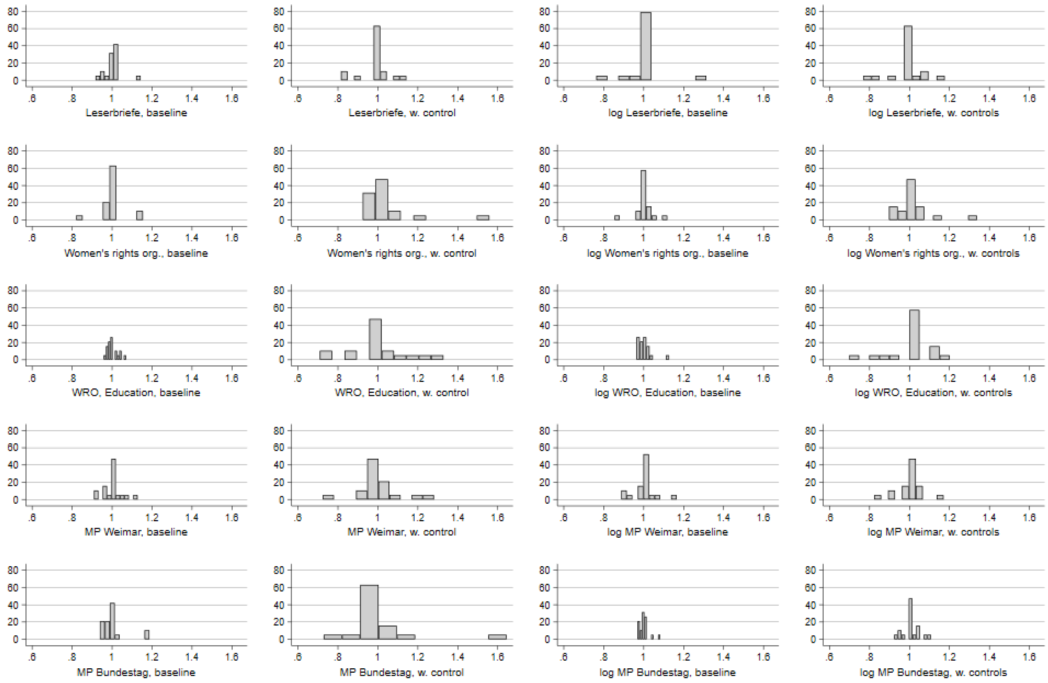


Figure 3.B.2: Senstivity to dropping sets of cities: Long-run outcomes

Notes: The x-axis measures the ratio between the restricted point estimate when dropping one of 22 sets of cities and the corresponding original estimate in Table 3.7. This ratio is one if the restricted estimate is unchanged, 1.5 if the restricted estimate is 50% larger than the original, and 0.5 if the restricted estimate is 50% smaller than the original. We present all outcomes (rows) in all specifications (columns) corresponding to Table 3.7. The sum of all bars is 100%. "WRO" in the third row stands for women's rights organization. "MP" stands for Member of Parliament.

3.B.2 The role of economic growth: flexibly controlling for construction

Finally, we address the possibility that our city covariates do not adequately capture economic growth by including construction data from Cantoni, Dittmar, and Yuchtman (2018). Neither using construction activity in 1650 (prior to the establishment of the first finishing school), nor the potentially endogeneous time-varying construction activity data change the point estimates significantly, as shown in Table 3.B.3. We thus conclude our identification is robust to including or excluding different sets of cities, city-specific trends, or economic activity.

Table 3.B.3: Fixed-effects results on the importance of finishing schools - Controlling for construction

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.161***	0.169***	0.208***	0.214***	0.020***	0.021***
	(0.034)	(0.034)	(0.045)	(0.045)	(0.006)	(0.005)
Mean, untreated	0.148	0.147	0.138	0.137	0.018	0.018
Panel B: Unmarried women						
Finishing school$_{it}$	0.150***	0.149***	0.182***	0.183***	0.014**	0.014**
	(0.035)	(0.035)	(0.042)	(0.042)	(0.006)	(0.006)
Mean, untreated	0.152	0.152	0.142	0.141	0.022	0.022
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.104***	0.109***	0.106***	0.112***	0.018***	0.019***
	(0.027)	(0.027)	(0.027)	(0.028)	(0.006)	(0.006)
Mean, untreated	0.075	0.075	0.059	0.059	0.019	0.019
Panel D: Activists						
Finishing school$_{it}$	0.065***	0.053***	0.049***	0.044***	0.015***	0.012**
	(0.018)	(0.018)	(0.015)	(0.016)	(0.005)	(0.005)
Mean, untreated	0.016	0.016	0.012	0.012	0.005	0.005
Panel E: Nobility						
Finishing school$_{it}$	-0.017	-0.015	-0.007	-0.006	-0.002	-0.002
	(0.017)	(0.016)	(0.017)	(0.015)	(0.009)	(0.008)
Mean, untreated	0.038	0.037	0.030	0.030	0.017	0.017
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Construction in 1650 × period FE	Yes		Yes		Yes	
Construction in every period × period FE		Yes		Yes		Yes
Observations	9,096	9,144	9,096	9,144	9,096	9,144

Notes: Results from fixed-effects regressions reported. We consider three types of dependent variables: (i) $\mathbb{I}[Women > 0]$ is an indicator taking value 1 if a city observed the birth of at least one notable woman in a given period; (ii) "log Women" constitutes the natural logarithm of the number of women born plus 1; (iii) "Share Women" divides the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians instead. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. All columns control for city and period fixed effects as well as city-specific linear trends in addition to interacting our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.C Dataset construction choices and timing of school establishment

In this Appendix, we discuss the construction of the Thiessen Polygons around each city that existed in 1300 CE, as taken from from Voigtländer and Voth (2012), and show that the results are robust to only using cities that already existed in 800 CE (Appendix 3.C.1). As the cities in Voigtländer and Voth (2012) might have oversampled cities with a Jewish presence, we instead use the territories and rulers in 1618 as our baseline and reproduce the main findings of the paper, concluding that neither dataset construction nor sample selection introduced a bias into our estimates (Appendix 3.C.2). We then highlight the impact of different school establishment periods (Appendix 3.C.3).

3.C.1 Structure of the data

We take the city-level data by Voigtländer and Voth (2012) as a starting point and construct Thiessen Polygons around the center of each city in their dataset. Thiessen Polygons are constructed such that every village or town inside the polygon around city i is closer to city i than to any other city $j \neq i$. Figure 3.C.1 shows the resulting polygons alongside the location of finishing schools and the number of notable women born within each area. By construction, the city lies in the center of its polygon.

We use this data structure and the set of cities used by Voigtländer and Voth (2012) to include their rich city-level covariates and to avoid relying on county boundaries. From the entire set of cities in Voigtländer and Voth (2012), we only select those cities that are mentioned before 1300 and are the oldest towns within a region. For example: Aachen has four recorded 'cities' in Voigtländer and Voth (2012): town_id 1, mentioned in 830, 13.45 km from Aachen; town_id 3, mentioned in 1118, 10.74 km from Aachen; town_id 4, mentioned in 870, 5.12 km from Aachen; and Aachen itself (town_id 5, mentioned in 400). Since these other cities are likely suburbs or dependent on Aachen's existence, we use the location of Aachen and merge all variables to Aachen. This has the advantage that our estimates are not biased by a potential rural-urban bias when including suburbs. We arrive at 388 cities by only using the oldest city within each region that lies in present-day Germany.

As the NDB starts recording notable individuals born from the year 800 onwards, using cities with recorded population levels by 800 is a natural alternative, which, however, reduces the sample of cities to 101. In Table 3.C.1 we document that results for both choices (1300 vs. 800) are similar across all specifications and outcomes.

The next choice concerns the length of periods. We choose to assign notable individuals to 50-year periods based on their year of birth. There are two reasons for our

50-year period choice: first, by choosing 50-years, we ensure that on average a woman that is born in this period either did or did not have access to a finishing school. Second, the scarce number of women recorded in the NDB prior to the 15th century implies a trade-off between statistical power and assignment accuracy. If we used every birth year separately, and thus matched schools most precisely, we would end up with no variation within most city by birth-year cells. Thus, to increase power, we rely on 50-year periods, and show robustness to using 25 year intervals in Table 3.C.2. Again, our point estimates remain unaffected.

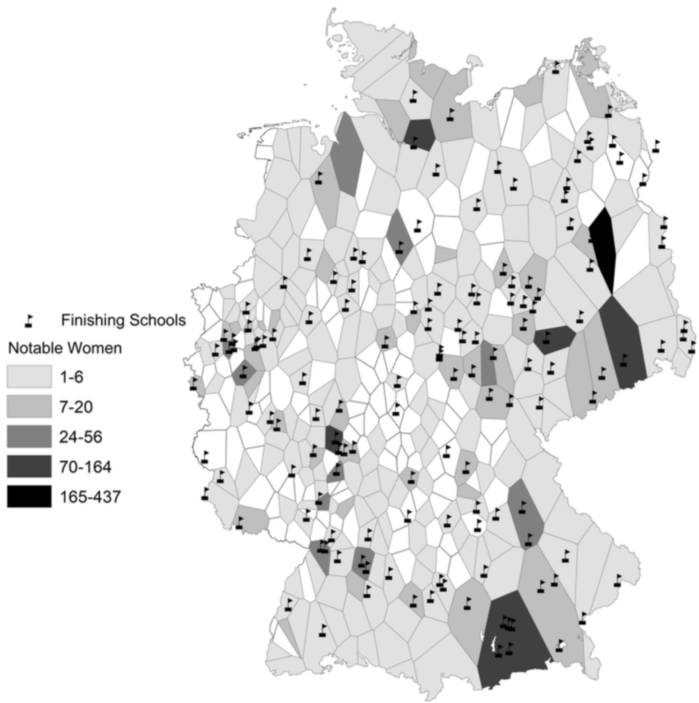


Figure 3.C.1: Thiessen Polygons, finishing schools and notable women

Notes: This figure shows our unit of observation, Thiessen polygons created around cities included in the data by Voigtländer and Voth (2012). By construction, the cities lie in the center of each Thiessen polygon. For simplicity we continue to refer to our unit of observation as "city". The figure also shows the location of finishing schools as well as the number of notable women born in each city.

The final choice concerns the classification of notable women into different (occupational) groups: *Non-Noble Seculars, Unmarried, Teachers & Writers, Activists*, and the *Nobility*. We grouped women together to ensure enough variation within every city-period-occupation-cell. In Table 3.C.3, we show the consistent impact across most occupational groups. In addition to our baseline results, we show that finishing schools increase the share of unmarried women (Panel A), artists (Panel D), writers (Panel E), politicians (Panel G), academics (H), but not the share of nuns (Panel J). This evidence, especially the impact on academics, artists, and writers, reinforces the notion that finishing schools increased the share of women among the human capital elite.

Table 3.C.1: Fixed-effects results on the importance of finishing schools - Changing the unit of observation to cities that existed in 800

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.251***	0.230***	0.465***	0.356***	0.016***	0.017*
	(0.049)	(0.064)	(0.098)	(0.100)	(0.005)	(0.009)
Mean, untreated	0.201	0.180	0.214	0.189	0.020	0.018
Panel B: Unmarried women						
Finishing school$_{it}$	0.134***	0.142**	0.356***	0.272***	0.007	0.010
	(0.048)	(0.067)	(0.086)	(0.097)	(0.007)	(0.009)
Mean, untreated	0.242	0.226	0.252	0.227	0.024	0.023
Panel B: Teachers & Writers						
Finishing school$_{it}$	0.183***	0.154**	0.257***	0.179**	0.019**	0.016
	(0.048)	(0.062)	(0.067)	(0.072)	(0.008)	(0.011)
Mean, untreated	0.103	0.090	0.091	0.076	0.019	0.016
Panel C: Activists						
Finishing school$_{it}$	0.104***	0.077*	0.100***	0.058	0.016**	0.016*
	(0.032)	(0.046)	(0.031)	(0.039)	(0.006)	(0.009)
Mean, untreated	0.029	0.026	0.023	0.020	0.005	0.005
Panel D: Nobility						
Finishing school$_{it}$	-0.018	-0.056	-0.001	-0.036	0.002	-0.033
	(0.037)	(0.044)	(0.039)	(0.043)	(0.019)	(0.023)
Mean, untreated	0.105	0.098	0.092	0.083	0.045	0.041
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	2,424	2,232	2,424	2,232	2,424	2,232

Notes: Results from fixed-effects regressions reported. Instead of building our dataset from cities that existed by 1300, we now consider all cities that exist in 800, resulting in a drop in the number of cities from 388 to 101. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.2: Fixed-effects results on the importance of finishing schools - Changing the Unit of observation to 25 year intervals

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.149***	0.096***	0.212***	0.110***	0.016***	0.014***
	(0.021)	(0.021)	(0.038)	(0.030)	(0.003)	(0.004)
Mean, untreated	0.094	0.093	0.142	0.142	0.015	0.015
Panel B: Unmarried women						
Finishing school$_{it}$	0.124***	0.088***	0.182***	0.097***	0.011***	0.012***
	(0.020)	(0.022)	(0.034)	(0.028)	(0.003)	(0.004)
Mean, untreated	0.097	0.097	0.143	0.142	0.017	0.017
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.088***	0.059***	0.096***	0.057***	0.015***	0.013***
	(0.017)	(0.016)	(0.022)	(0.019)	(0.004)	(0.004)
Mean, untreated	0.044	0.043	0.050	0.050	0.013	0.013
Panel D: Activists						
Finishing school$_{it}$	0.042***	0.030***	0.034***	0.023***	0.008***	0.007**
	(0.011)	(0.011)	(0.010)	(0.009)	(0.003)	(0.003)
Mean, untreated	0.009	0.009	0.009	0.009	0.003	0.003
Panel E: Royals						
Finishing school$_{it}$	-0.014*	-0.009	-0.007	-0.005	-0.003	-0.002
	(0.008)	(0.009)	(0.008)	(0.010)	(0.004)	(0.005)
Mean, untreated	0.021	0.021	0.025	0.025	0.010	0.010
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	18,624	18,480	18,624	18,480	18,624	18,480

Notes: Results from fixed-effects regressions reported. Instead of 50-year periods, we now employ 25-year periods. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.3: Fixed-effects results on the importance of finishing schools - All occupations

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Unmarried women						
Finishing school$_{it}$	0.194***	0.147***	0.302***	0.173***	0.011**	0.014**
	(0.030)	(0.034)	(0.049)	(0.043)	(0.005)	(0.006)
Mean, untreated	0.155	0.153	0.275	0.274	0.022	0.022
Panel B: Non-Royal women						
Finishing school$_{it}$	0.224***	0.164***	0.350***	0.201***	0.018***	0.018***
	(0.030)	(0.034)	(0.053)	(0.045)	(0.004)	(0.005)
Mean, untreated	0.156	0.154	0.285	0.284	0.018	0.018
Panel C: Occupation						
Finishing school$_{it}$	0.055***	0.025	0.058***	0.021	0.004	0.004
	(0.017)	(0.018)	(0.018)	(0.016)	(0.003)	(0.004)
Mean, untreated	0.019	0.019	0.020	0.020	0.004	0.004
Panel D: Artists						
Finishing school$_{it}$	0.137***	0.062**	0.187***	0.071**	0.027***	0.016**
	(0.027)	(0.028)	(0.043)	(0.033)	(0.007)	(0.007)
Mean, untreated	0.056	0.056	0.085	0.085	0.013	0.013
Panel E: Writers						
Finishing school$_{it}$	0.147***	0.099***	0.159***	0.096***	0.023***	0.020***
	(0.025)	(0.025)	(0.032)	(0.027)	(0.006)	(0.006)
Mean, untreated	0.067	0.067	0.084	0.083	0.020	0.020
Panel F: Doctors						
Finishing school$_{it}$	0.021*	-0.003	0.020**	-0.003	0.003	-0.000
	(0.011)	(0.011)	(0.009)	(0.009)	(0.003)	(0.003)
Mean, untreated	0.009	0.009	0.009	0.009	0.003	0.003
Panel G: Politicians						
Finishing school$_{it}$	0.058***	0.025	0.054***	0.018	0.011**	0.007
	(0.017)	(0.018)	(0.016)	(0.015)	(0.004)	(0.005)
Mean, untreated	0.018	0.018	0.020	0.020	0.005	0.005
Panel H: Academics						
Finishing school$_{it}$	0.080***	0.056***	0.069***	0.037***	0.009***	0.009**
	(0.015)	(0.016)	(0.014)	(0.014)	(0.003)	(0.004)
Mean, untreated	0.014	0.014	0.016	0.016	0.003	0.003
Panel I: Teachers						
Finishing school$_{it}$	0.041***	0.018	0.036***	0.014	0.006*	0.005
	(0.014)	(0.012)	(0.012)	(0.010)	(0.003)	(0.003)
Mean, untreated	0.011	0.011	0.012	0.012	0.003	0.003
Panel J: Nuns						
Finishing school$_{it}$	0.001	0.002	0.001	0.000	-0.002	-0.000
	(0.011)	(0.013)	(0.008)	(0.009)	(0.003)	(0.004)
Mean, untreated	0.012	0.012	0.012	0.013	0.004	0.004
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	9,312	9,240	9,312	9,240	9,312	9,240

Notes: Results from fixed-effects regressions using all cities and periods reported. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses.$*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

3.C.2 Sample selection: Using a different starting point for the analysis

In our baseline data, we created a balanced panel for each city in Voigtländer and Voth (2012) using Thiessen Polygons as a starting point (see 3.C.1 above). This procedure has the advantage that it does not rely on any administrative boundary, past or present, and any covariate from Voigtländer and Voth (2012) can easily be used. However, as the focus of their paper was on historical roots of antisemitism, the original data might have oversampled cities with occurrences of the Black Death and pogroms. We thus show the robustness of our results to using an alternative baseline source to create a balanced panel: the territories of Germany in 1619.

In Figure 3.C.2, we depict the territories of 21 different rulers, 91 ecclesiastical cities, 96 free cities and 57 imperial cities in Germany on the eve of the Thirty Years' war. We then use these administrative boundaries to create a balanced panel from 800 until 1950. The implicit assumption here is that people migrate disporporitonately within a ruler's territory and only rarely migrate between competing territories. We avoid this assumption using the Voigtländer and Voth (2012) cities in combination with Thiessen polygons.

The event-study results in Figure 3.C.3 and the fixed effects results in Table 3.C.4, however, confirm our initial results. We conclude that choosing the cities from Voigtländer and Voth (2012) to create Thiessen polygons did not introduce a bias into our setting.

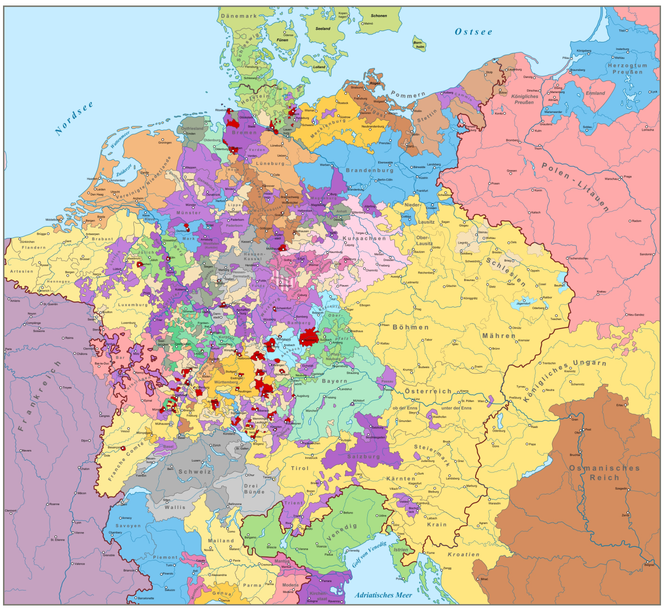
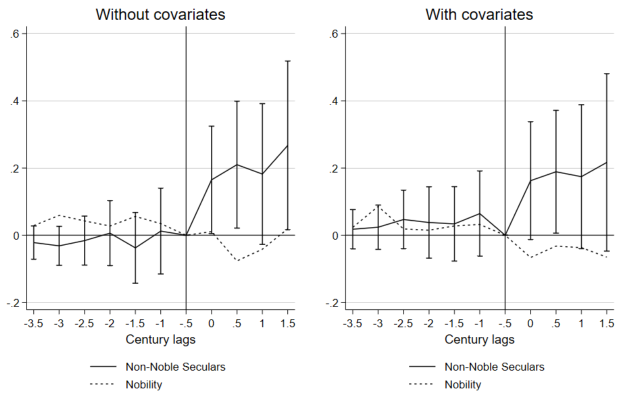
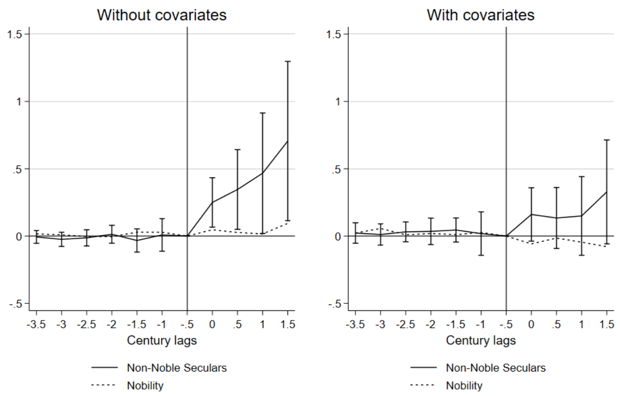


Figure 3.C.2: German territorial belongings and rulers in 1618

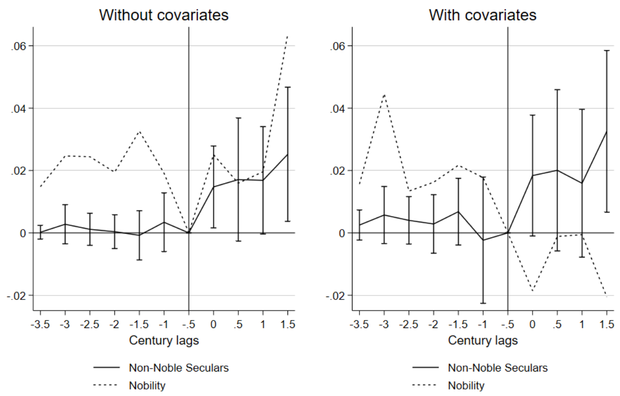
Notes: This figure shows the territories of rulers, ecclesiastical cities, free cities, and imperial cities in 1618, which we use as a baseline for the results in this section. License notice: Sir Iain. This W3C-unspecified vector image was created with Inkscape. (https://bit.ly/3LwzRos), https://bit.ly/3GQx06h.



(a) Indicator: Notable woman born in city



(b) Log. number of notable women born in city



(c) Female share of notable individuals born in city

Figure 3.C.3: Event-study: Impact of finishing school establishment on notable women using territories as of 1619 as the unit of observation

Notes: Event study results for *non-noble secular* women and women from the *nobility*. In Figure a, the outcome is an indicator equal to one if a notable woman from the respective group was born in a given city and period. Figure b uses the natural logarithm of number of women born plus one. Figure c denotes the number of notable women by the number of notable individuals of all genders. Zero is the normalized opening period of the first finishing school in a city. The vertical line marks the reference period, which we choose to be 50 years prior to establishment of the school. City and period fixed effects included in the left figure and full economic, religious, and educational controls added in the right. 95%-confidence intervals shown only for non-noble secular women.

Table 3.C.4: Fixed-effects results on the importance of finishing schools - Changing the unit of observation to territories in 1619

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.376***	0.188***	0.728***	0.270***	0.024***	0.017***
	(0.039)	(0.046)	(0.096)	(0.065)	(0.004)	(0.006)
Mean, untreated	0.081	0.078	0.141	0.128	0.011	0.011
Panel B: Unmarried women						
Finishing school$_{it}$	0.283***	0.154***	0.603***	0.244***	0.017***	0.018***
	(0.040)	(0.045)	(0.081)	(0.070)	(0.005)	(0.007)
Mean, untreated	0.086	0.082	0.157	0.134	0.013	0.013
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.283***	0.097**	0.381***	0.105***	0.033***	0.018**
	(0.039)	(0.038)	(0.065)	(0.036)	(0.007)	(0.008)
Mean, untreated	0.037	0.035	0.057	0.048	0.008	0.008
Panel D: Activists						
Finishing school$_{it}$	0.151***	0.060**	0.146***	0.057**	0.021***	0.014**
	(0.029)	(0.028)	(0.032)	(0.023)	(0.005)	(0.007)
Mean, untreated	0.007	0.006	0.008	0.006	0.002	0.002
Panel E: Nobility						
Finishing school$_{it}$	-0.070**	-0.052*	-0.035	-0.044*	-0.009	-0.018
	(0.029)	(0.028)	(0.025)	(0.024)	(0.012)	(0.012)
Mean, untreated	0.029	0.025	0.043	0.031	0.012	0.011
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	6,360	6,216	6,360	6,216	6,360	6,216

Notes: Results from fixed-effects regressions reported. Instead of using the cities in Voigtländer and Voth (2012), we use the territories (as of 1618) shown in Figure 3.C.2 as the unit of observation. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses.$*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

3.C.3 Using the exact opening time of finishing schools

In our baseline data, we created a balanced panel for each city to include never-treated cities and covariates. This decision is in line with the recent literature on event-study validity, as discussed in Appendix 3.E. In the resulting panel, we merged individuals to the closest of 50-year periods in cities. That is, if an individual is born in 1640, we merge her to the city's 1650 period, regardless of treatment status. In that setting, we have cities that switch into treatment, as well as pure-control cities in every period and can compare the three groups.

However, an event-study usually uses the exact timing to estimate the treatment effect. Ignoring never-treated cities, our data allows for such a fine-grained distinction. In this Appendix, we normalize the time period for every city to zero at the exact time the first school was opened. That is, if the first school opens in 1626 for the city of Aachen, we create city-specific period lags of arbitrary length. Yet, there are two disadvantages associated with this: first, we are unable to merge control cities to this framework, and thus the comparison is strictly among treated cities only. Second, the choice of omitted period is not innocuous: Women that are born 10 years prior to the opening of a Finishing schools still benefit from its construction, while not having had any say in its establishing. We thus need to normalize at an earlier period at which women could not have benefited from the future presence of finishing schools. While these considerations average out at 50-year intervals, they matter greatly at smaller intervals.

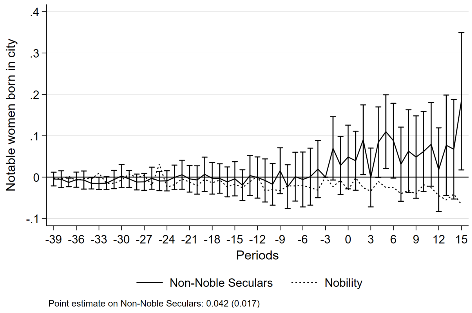
In Figure 3.C.4, we use the opening time of the first finishing school in our 129 cities with schools and create various lags around it. In all Panels, we aim to reference the estimates to a previous generation of women who could no longer benefit from education: parents. In Panel a, we create 10-year lag windows around each school and omit women born between 30 and 39 years prior to school opening. We omit women born between 20 and 39 years before in the 20-year Panel b, 25 and 50 years before in the 25-year Panel c and 50-100 years before in Panel d. We find no evidence for a pre-trend in any specification, a significant uptick after the opening, and point estimates that are not statistically different from our baseline.

Yet, as we discuss in Appendix 3.E, the inclusion of never-treated cities allows for a clean comparison between treatment and control, as well as a classical difference-in-differences setup (Appendix 3.F). These benefits, along with the possibility to merge covariates and the unchanged point estimates, motivate our choice to match women and schools to a balanced panel of cities, instead of using this exact-timing setup.

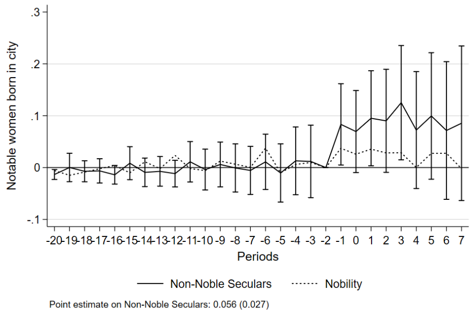
3.C.4 Timing of school construction

When taking historical accounts at face value, the establishment of *early* finishing schools by foreign Catholic women's orders constituted a shift in the supply of women's education as opposed to a local shift in the demand for education.

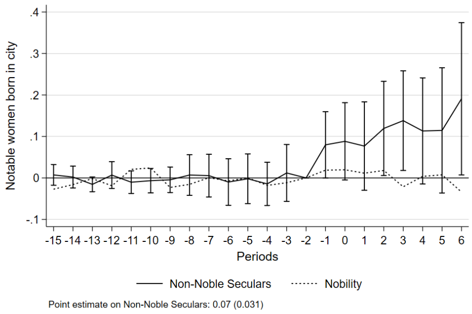
In this Appendix, we assess the severity of a potential bias in our estimates that would arise if the establishment of the *later* finishing schools in our data were largely driven by increasing demand for women's education. If the *later* schools (constructed between 1800 and 1850, i.e. after the fall of the Holy Roman Empire) accounted for



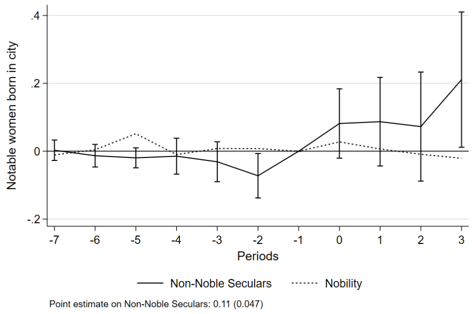
(a) 10 year periods



(b) 20 year periods



(c) 25 year periods



(d) 50 year periods

Figure 3.C.4: Event-study:
Impact of finishing school establishment on notable women

Notes: Event study graphs using the exact timing of the first finishing school in every city to create 10-year periods (3.C.4a), 20-year periods (3.C.4b), 25-year periods (3.C.4c) and 50-year periods (3.C.4d). We include fifty-year period fixed effects in all regressions for comparability across figures. Results are robust to using year fixed effects, amounting to $\geq$ 645 fixed effects. As a result of the exact matching on birthyears, we observe a significant increase in period -1 in Figures a-c): if a woman was 10 years old when the first finishing school opened, she attended this school and became notable for her achievements, we assign her to the -10 years bin. Thus, this “lead” does not reflect any anticipation effects, but is an artifact of data construction and fully intended.

all the impact on women’s representation among the human-capital elite, this would call into question our interpretation that the establishment of finishing schools constituted a supply-side shift. However, our results largely remain robust when only using schools constructed before 1800 in the odd columns of Table 3.C.5. In addition, the point estimates on *early* and *late* schools are not statistically different from each other in most specifications.

Moreover, in Table 3.C.6 we compare the impact of the first versus the second school constructed in a city and show that most of the impact indeed comes from the first established school. Combined with the impact of multiple schools shown in Figure 3.C.5, this suggests that indeed the first, arguably exogenous school opening, is responsible for the increase in the share of women among the human capital elite of German cities. This finding is confirmed in the difference-in-differences setting, where all periods produce similar estimates (Figure 3.F.2 and Table 3.F.2).

Table 3.C.5: Fixed-effects results on the importance of finishing schools - Early vs. late Schools

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1) Early	(2) Late	(3) Early	(4) Late	(5) Early	(6) Late
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.095**	0.185***	0.279***	0.246***	0.016**	0.023***
	(0.041)	(0.044)	(0.100)	(0.057)	(0.007)	(0.007)
Mean, untreated	0.147	0.148	0.137	0.138	0.019	0.018
Panel B: Unmarried women						
Finishing school$_{it}$	0.050	0.180***	0.231**	0.217***	0.004	0.018***
	(0.046)	(0.044)	(0.092)	(0.053)	(0.010)	(0.007)
Mean, untreated	0.148	0.152	0.137	0.141	0.022	0.022
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.095**	0.129***	0.166**	0.124***	0.011	0.022***
	(0.041)	(0.032)	(0.081)	(0.032)	(0.008)	(0.007)
Mean, untreated	0.074	0.074	0.058	0.059	0.019	0.019
Panel D: Activists						
Finishing school$_{it}$	0.053*	0.066***	0.070	0.051***	0.004	0.014**
	(0.029)	(0.022)	(0.043)	(0.018)	(0.004)	(0.006)
Mean, untreated	0.018	0.016	0.013	0.012	0.006	0.005
Panel E: Nobility						
Finishing school$_{it}$	-0.022	-0.014	0.004	-0.012	-0.002	-0.003
	(0.039)	(0.019)	(0.035)	(0.016)	(0.018)	(0.009)
Mean, untreated	0.031	0.037	0.024	0.030	0.015	0.017
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,984	8,400	6,984	8,400	6,984	8,400

Notes: Results from fixed-effects regressions following our main specification reported, comparing effect sizes between *early* (1650–1750) and *late* (1800–1850) finishing schools. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. In all columns we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 3.C.6: Fixed-effects results on the importance of finishing schools - Comparing the impact of the first to the second school

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
First finishing school$_{it}$	0.164***	0.156***	0.204***	0.147***	0.021***	0.020***
	(0.033)	(0.034)	(0.045)	(0.044)	(0.005)	(0.006)
Second finishing school$_{it}$		0.040		0.279**		0.003
		(0.058)		(0.109)		(0.008)
Mean, untreated	0.149	0.149	0.139	0.139	0.018	0.018
Panel B: Unmarried women						
First finishing school$_{it}$	0.147***	0.154***	0.173***	0.137***	0.014**	0.016**
	(0.034)	(0.035)	(0.043)	(0.043)	(0.006)	(0.006)
Second finishing school$_{it}$		-0.039		0.180*		-0.008
		(0.056)		(0.097)		(0.007)
Mean, untreated	0.153	0.153	0.143	0.143	0.022	0.022
Panel C: Teachers & Writers						
First finishing school$_{it}$	0.104***	0.082***	0.103***	0.065**	0.017***	0.015**
	(0.026)	(0.027)	(0.029)	(0.027)	(0.006)	(0.006)
Second finishing school$_{it}$		0.110**		0.191**		0.010
		(0.047)		(0.077)		(0.011)
Mean, untreated	0.075	0.075	0.059	0.059	0.019	0.019
Panel D: Activists						
First finishing school$_{it}$	0.053***	0.053***	0.043***	0.039**	0.011**	0.015***
	(0.018)	(0.020)	(0.015)	(0.016)	(0.005)	(0.005)
Second finishing school$_{it}$		-0.004		0.019		-0.017***
		(0.036)		(0.037)		(0.006)
Mean, untreated	0.016	0.016	0.012	0.012	0.005	0.005
Panel E: Nobility						
First finishing school$_{it}$	-0.013	-0.010	-0.007	-0.001	-0.002	-0.000
	(0.017)	(0.018)	(0.018)	(0.018)	(0.009)	(0.009)
Second finishing school$_{it}$		-0.015		-0.027		-0.008
		(0.035)		(0.026)		(0.014)
Mean, untreated	0.038	0.038	0.031	0.031	0.018	0.018
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,240	9,240	9,240	9,240	9,240	9,240

Notes: Results from fixed-effects regressions following our main specification reported, comparing effect sizes between the first and the second finishing school. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. In all columns we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses. $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

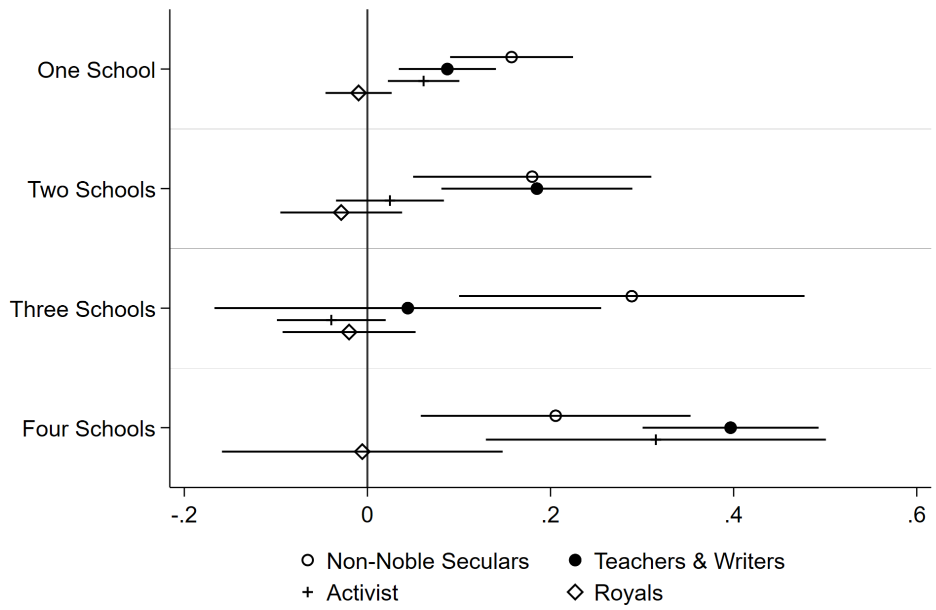


Figure 3.C.5: The impact of multiple schools

Notes: The cumulative impact of cities having one, two, three, or more school in the fixed effect estimation on the occurrence of notable women. The outcome is an indicator equal to one if a notable woman from the respective group was born in a given city and period. All covariates from Table 3.2 column (2) included.

3.D Spatial dependence and SUTVA

In this Appendix, we address the potential threat of spatial correlation, possible violations of the Stable Unit Treatment Value Assumption (SUTVA), and discuss spatial noise (Kelly 2020).

We show that standard errors accounting for spatial correlation are slightly smaller than cluster-robust standard errors at the city level (Table 3.D.1). To address potential violations of SUTVA, we exclude all cities that border a city with finishing schools in Table 3.D.2. If migration from cities without finishing schools to cities with such schools drove our findings, an increase in the 'cost of migration' by increasing control cities' distance to the next school city should result in significantly smaller estimates. As expected , we find no evidence that migration impacts our point estimates.

A recent literature has focused on how estimates indicating persistent effects of past events on more recent outcomes can be driven by spatial noise (Kelly 2020). To address the potential severity arising from this line of thought, we report a low Moran's I of 0.002 with a p-value of 0.156. In addition, we conduct an exercise where we randomly distribute schools across Germany in each period, holding the number of schools constant. The results in Figure 3.D.1 reveal that our results are clear outliers in this distribution, with the largest fraction of absolute values greater than our estimate at a mere 0.02 (for the results on Activists).

Taken together, the results presented in this Appendix suggest that our estimates are unlikely to be driven by spatial dependence and potential violations of SUTVA.

Table 3.D.1: Fixed-effects results on the importance of finishing schools - Standard errors corrected for spatial dependence

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.230***	0.164***	0.355***	0.204***	0.019***	0.021***
	(0.026)	(0.028)	(0.033)	(0.030)	(0.004)	(0.004)
Mean, untreated	0.150	0.149	0.140	0.139	0.018	0.018
Panel B: Unmarried women						
Finishing school$_{it}$	0.194***	0.147***	0.302***	0.173***	0.011**	0.014***
	(0.028)	(0.029)	(0.033)	(0.029)	(0.004)	(0.005)
Mean, untreated	0.155	0.153	0.144	0.143	0.022	0.022
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.151***	0.104***	0.174***	0.103***	0.019***	0.017***
	(0.019)	(0.020)	(0.018)	(0.020)	(0.005)	(0.005)
Mean, untreated	0.076	0.075	0.060	0.059	0.019	0.019
Panel D: Activists						
Finishing school$_{it}$	0.076***	0.053***	0.064***	0.043***	0.013***	0.011***
	(0.015)	(0.014)	(0.012)	(0.010)	(0.004)	(0.004)
Mean, untreated	0.016	0.016	0.012	0.012	0.005	0.005
Panel E: Nobility						
Finishing school$_{it}$	-0.018	-0.013	-0.009	-0.007	-0.002	-0.002
	(0.014)	(0.015)	(0.012)	(0.013)	(0.006)	(0.007)
Mean, untreated	0.039	0.038	0.031	0.031	0.018	0.018
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	9,312	9,240	9,312	9,240	9,312	9,240

Notes: Results from fixed-effects regressions following our main specification reported. Standard errors corrected for spatial dependence within 100 km as in Hsiang, Burke, and Miguel (2013) reported in parentheses. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses.$*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table 3.D.2: Fixed-effects results on the importance of finishing schools - Comparing towns with schools to non-neighboring towns without schools

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school$_{it}$	0.164***	0.167***	0.204***	0.156***	0.018**	0.021***
	(0.033)	(0.047)	(0.045)	(0.055)	(0.007)	(0.005)
Mean, untreated	0.149	0.164	0.139	0.158	0.017	0.016
Panel B: Unmarried women						
Finishing school$_{it}$	0.147***	0.140***	0.173***	0.122**	0.011	0.014**
	(0.034)	(0.049)	(0.043)	(0.056)	(0.008)	(0.006)
Mean, untreated	0.153	0.180	0.143	0.176	0.021	0.021
Panel C: Teachers & Writers						
Finishing school$_{it}$	0.104***	0.108**	0.103***	0.090**	0.023**	0.017***
	(0.026)	(0.044)	(0.029)	(0.043)	(0.009)	(0.006)
Mean, untreated	0.075	0.085	0.059	0.066	0.017	0.017
Panel D: Activists						
Finishing school$_{it}$	0.053***	0.039	0.043***	0.026	0.012*	0.011**
	(0.018)	(0.026)	(0.015)	(0.019)	(0.006)	(0.005)
Mean, untreated	0.016	0.011	0.012	0.008	0.003	0.003
Panel E: Nobility						
Finishing school$_{it}$	-0.013	-0.010	-0.007	0.007	0.001	-0.002
	(0.017)	(0.028)	(0.018)	(0.034)	(0.015)	(0.009)
Mean, untreated	0.038	0.068	0.031	0.056	0.029	0.029
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Non-Spillover sample		Yes		Yes		Yes
Observations	9,240	3,696	9,240	3,696	3,696	9,240

Notes: Results from fixed-effects regressions following our main specification reported, comparing effect sizes between the full sample and a sample where all neighboring cities without finishing schools are dropped. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. In all columns we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses. $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$.

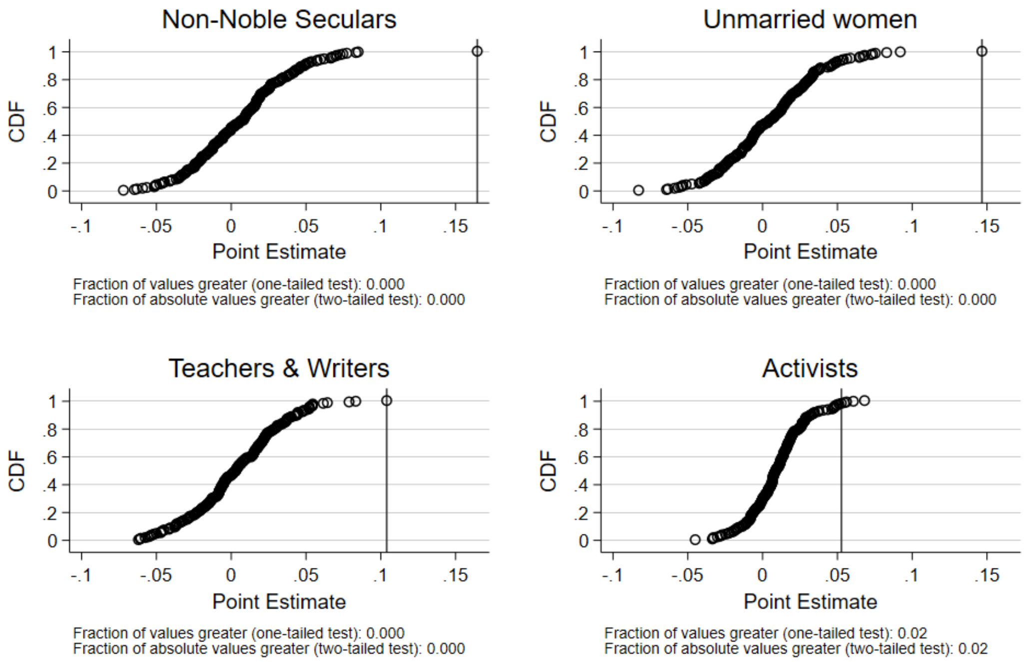
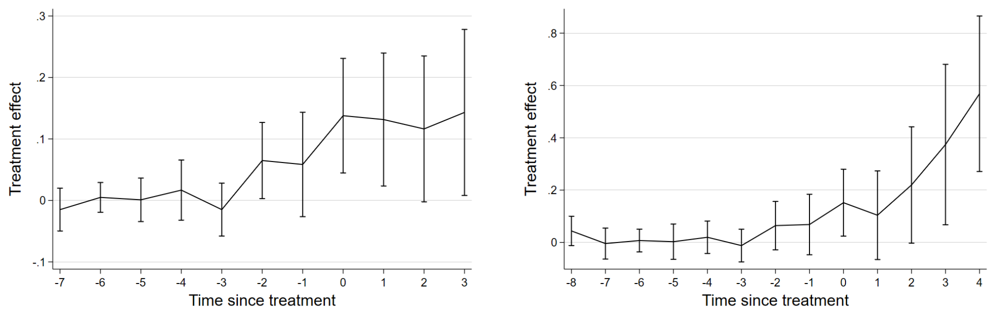


Figure 3.D.1: Placebo treatments:
Distributing schools across Germany and centuries

Notes: Each figure reports the point estimates from 200 randomization exercises that proceed as follows: We use the number of schools in every period and randomly distribute them across Germany. This is repeated for every period and used as a new explanatory variable in a regression with full controls. The outcome is an indicator equal to one if a city had at least one notable woman from the respective occupational group born in this period. The vertical line marks the baseline estimate in Table 3.2 column (2).

3.E Recent advances in event-study designs: DID with multiple time periods or heterogeneous treatment effects

There has been a rich recent debate in the literature on how to interpret the average treatment effect on the treated in event-study designs. Following these developments, Baker, Larcker, and Wang (2021) argue that "staggered treatment timing and treatment effect heterogeneity, either accross groups or over time, leads to biased Two-Way-Fixed-Effects DID [TWFE] estimates for the ATT", and propose three methods to assess the severity of this bias. First, show the event-study graph without controls (Figure 3.2) and by treatment group (Figure 3.F.2). Second, implement the method by Chaisemartin and D'Haultfœuille (2020) to assess whether heterogeneous treatment effects bias the estimate (Figure 3.E.1a). Third, implement the method by Callaway and Sant'Anna (forthcoming) to assess whether treatment effect heterogeneity by treatment period biases estimates (Figure 3.E.1b). Finally, show the implied weights following Goodman-Bacon (forthcoming), showing that the main effect is derived from the comparison of treatment versus control (Figure 3.E.2). All methods provide no evidence of differential pre-trends and provide similar point estimates, highlighting the validity of our empirical approach.



(a) Chaisemartin and D'Haultfœuille (2020)

(b) Callaway and Sant'Anna (forthcoming)

Figure 3.E.1: Alternative treatment effect aggregators

Notes: Implementing the approaches listed for the variable indicating whether a notable woman was born in city c. The average treatment effect on the treated (ATT) in Figure 3.E.1a (0.146, s.e. 0.052) is slightly smaller than the ATT in the right Figure 3.E.1b (0.284, s.e. 0.054). These point estimates are very similar to the baseline ATT in Figure 3.2 (0.146, s.e. 0.049).

Another way to assess the validity of our approach is by estimating the implied weight of each treatment period. In a classical event study design where one focuses on cities that are ever treated, late treatment cities are the implied control cities for early treatment cities (Goodman-Bacon forthcoming). Then, TWFE estimates are a weighted sum of individual treatment effects estimated for every city and period. Since these weights can be negative, inference can be affected. Using the approach suggested by Goodman-Bacon (forthcoming), we show in Table 3.E.1 that the weight of our effect comes from the comparison between treated and never-treated. This result is confirmed in Figure 3.E.2, where the DID estimate is almost exclusively derived from the differences between cities without and with finishing schools, thus validating our approach.

Table 3.E.1: Goodman-Bacon (forthcoming) decomposition of difference-in-differences estimation with variation in treatment timing

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	Weight	Av. DID Est.	Weight	Av. DID Est.	Weight	Av. DID Est.
Earlier Treatment vs. Later Control	0.071	0.160	0.071	0.227	0.071	0.015
Later Treatment vs. Earlier Control	0.013	0.028	0.013	-0.171	0.013	0.007
Treatment vs Never treated	0.915	0.315	0.915	0.492	0.915	0.023
Difference-in-differences estimate:		0.300		0.464		0.022

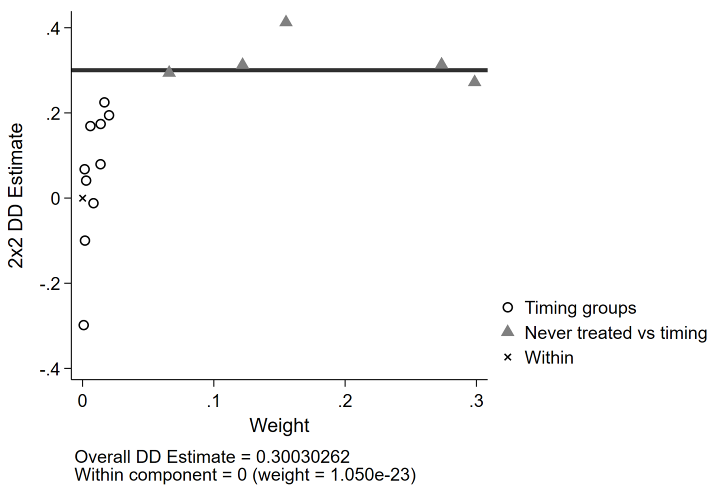
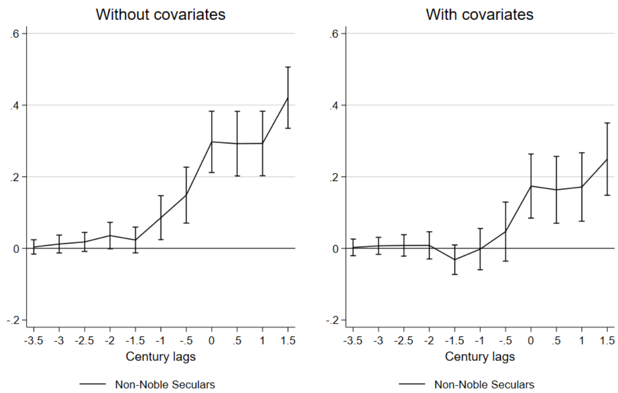


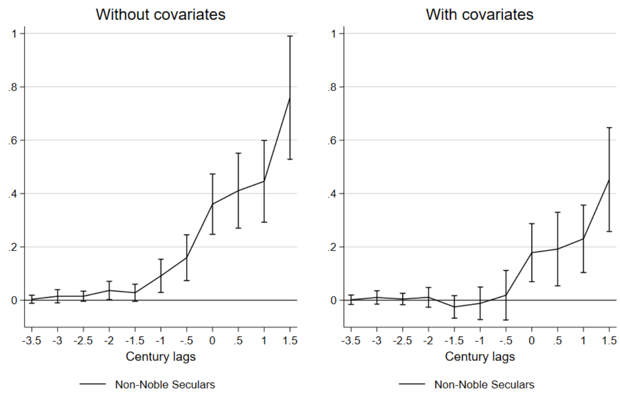
Figure 3.E.2: Goodman-Bacon (forthcoming) decomposition of difference-in-differences estimation with variation in treatment timing

Notes: Showing the implied weights against the treatment effect when using the indicator $\mathbb{I}[Women > 0]$. The treatment effect is amost entirely estimated from the comparison of treated cities to non-treated cities.

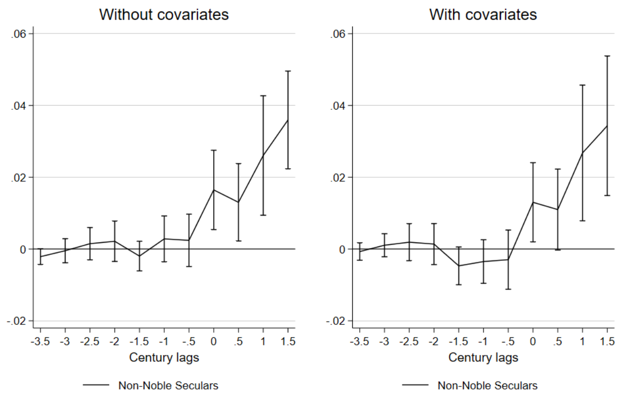
Figure 3.E.2 suggests that the point estimate in our TWFE estimation stems from the difference between never-treated cities and cities with finishing schools. We thus provide additional evidence for the parallel trends assumption including all cities. In our main Figure 3.2, we show parallel trends in the set of cities that ever established finishing schools. In Figure 3.E.3, we complement this evidence by including cities that never established a finishing school. The results speak in favor of the parallel trends assumption: When controlling extensively for economic, religious and educational covariates, the estimated leads are centered around zero and show no difference between cities with and without finishing schools.



(a) Indicator: Notable woman born in city



(b) Log. number of notable women born in city



(c) Female share of notable individuals born in city

Figure 3.E.3: Event-study using never-treated cities: Impact of finishing school establishment on notable women

Notes: Additional results for *non-noble secular* women, including all control cities. The outcome is an indicator equal to one if a notable woman from the respective group was born in a given city and period. In contrast to Figure 3.2 in the main text, here we also include cities which never established finishing schools to improve precision. Zero is the normalized time of establishment of finishing schools in the city; -4 is the omitted period and includes all never-treated cities. When extensively controlling for city characteristics in the right panel, all leads are insignificant. City and period fixed effects included in the left panel and full economic, religious and educational controls included in the right. 95%-confidence intervals reported.

3.F Standard difference-in-differences estimates and possible instruments in the panel setting

In this Appendix, we show results from a standard difference-in-differences estimator, comparing cities without finishing schools (control group) with cities that establish a finishing school by 1850 (treatment group) to complement our assessment of pre-trends in the event-study setting and assess whether specific periods impact the estimates disproportionately. We then continue and analyze whether the diffusion of Protestantism threatens the interpretation of our findings (S. O. Becker and Woessmann 2009). We conclude this Appendix with a complementary empirical strategy using monasteries established before 1300 as an instrument for finishing schools. We document local average treatment effects that are very similar to the main results presented in the paper.

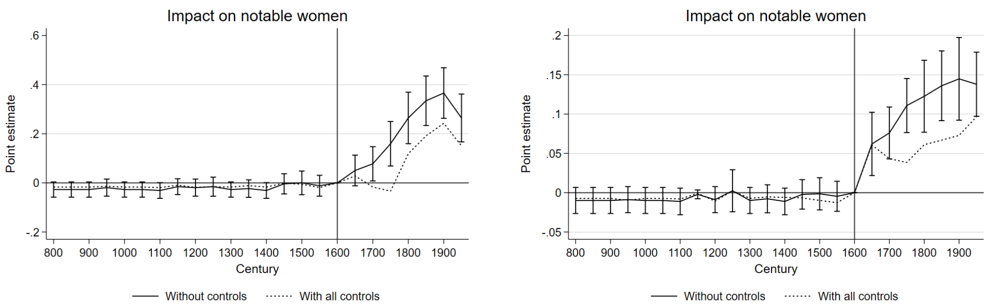
3.F.1 Standard difference-in-differences

We start by splitting our sample into cities that established finishing schools by 1850 and cities which did not and compare women's representation among the human capital elite in these two sets of cities before and after 1650, the period in which the first finishing school was founded. While this strategy allows for a more standard analysis of pre-trends than an event-study strategy, it also combines many treament periods into one, and thus likely underestimates the true impact. In Figure 3.F.1, we document the absence of significant pre-trends for both the extensive margin (establishing a school) and the intensive margin (number of schools). Yet, both panels reveal an increase in women's representation among the human capital elite in the periods after the first finishing school was established (1626). Point estimates are reported in Table 3.F.1 for both margins. First, the point estimates are very similar to the baseline results reported in Table 3.2 and are stable across specifications. Second, the point estimates on the intensive and extensive margin do not differ in most cases.

We continue and analyze the pre-trends for each treatment period separately in Figure 3.F.2. Again, we see no differential pre-trend in any pre-treatment period and significant impacts of schools only after the schools have been established. The results are somewhat stronger for the first and last schools, yet reveal no differential DID-estimate in Table 3.F.2. Here, we jointly estimate all treatment periods as compared to cities that never establish schools and find similar impacts across all types of schools. The only insignificant period is 1750, in which only three schools were established. Yet, even here the point estimate is statistically indistinguishable from the other periods.

We take this as evidence that our conclusion that finishing schools increase the share of women among the human capital elite is not driven by a specific functional form, identification strategy, or period. Also, while one could reasonably assume that the lack of variation in the outcome in the periods leading up to 1650 makes a pre-trend assessment problematic, the pre-trends are also insignificant in periods with more outcome variation such as the years 1600-1800 for the cities that establish finishing schools only in the 1850 period.

The effects in Figure 3.F.2 also indicate that the main effect in our baseline estimate is not driven by unobserved characteristics of the set of cities ever receiving finishing schools, which generally affect women's representation among the human capital elite in these cities after 1600. The temporal correspondence between the establishment of finishing schools and the timing of the effects (and the absence of pre-trends) certainly cannot alleviate all concerns about the potential endogeneity of the timing of school opening; however, it clearly points to an important nexus between the opening of finishing schools and the subsequent increase in women's representation among the human capital elite.



(a) Any finishing school **(b) Number of schools**

Figure 3.F.1: Difference-in-differences estimation: Comparing cities with and without finishing schools over time

Notes: These graphs split the data into cities that ever establish at least one finishing school and those without and compare those before and after 1650. The outcome is an indicator equal to one if a notable woman was born in a given city and period. The left Panel reports the point estimates from the interaction between period fixed effects and whether the city ever established a finishing school $\in \{0,1\}$. The right Panel reports the point estimates from the interaction between period fixed effects and the number of schools that have been established in this city by 1850 $\in \{0,1,2,3,4,5,8,10\}$. The omitted period is 1600, the period before the first schools were opened. Estimates without (solid line) and with (dashed line) all controls all indicate no pre-trends and an increase in the likelihood of women becoming notable only after the opening of the first school. While the left Panel can be interpreted as the extensive margin of finishing schools ("whether cities were different before"), the right Panel represents "how different these cities were before".

Table 3.F.1: Difference-in-differences estimation: Establishing finishing schools in cities

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school × Post 1650	0.182***	0.073***	0.264***	0.101***	0.010***	0.005*
	(0.024)	(0.022)	(0.041)	(0.030)	(0.003)	(0.003)
# Finishing schools × Post 1650	0.103***	0.063***	0.192***	0.148***	0.006***	0.004***
	(0.009)	(0.011)	(0.025)	(0.031)	(0.001)	(0.001)
Panel B: Unmarried women						
Finishing school × Post 1650	0.131***	0.044*	0.219***	0.079**	0.001	-0.001
	(0.024)	(0.025)	(0.038)	(0.032)	(0.004)	(0.004)
# Finishing schools × Post 1650	0.069***	0.033***	0.164***	0.126***	0.000	-0.001
	(0.013)	(0.012)	(0.023)	(0.028)	(0.001)	(0.002)
Panel C: Teachers & Writers						
Finishing school × Post 1650	0.113***	0.043***	0.122***	0.046***	0.014***	0.007*
	(0.018)	(0.016)	(0.022)	(0.016)	(0.003)	(0.004)
# Finishing schools × Post 1650	0.064***	0.039***	0.092***	0.073***	0.007***	0.004*
	(0.010)	(0.011)	(0.018)	(0.022)	(0.002)	(0.002)
Panel D: Activists						
Finishing school × Post 1650	0.036***	0.017*	0.032***	0.015*	0.004*	0.001
	(0.009)	(0.010)	(0.009)	(0.008)	(0.002)	(0.002)
# Finishing schools × Post 1650	0.027***	0.021***	0.028**	0.025**	0.002*	0.000
	(0.007)	(0.007)	(0.011)	(0.012)	(0.001)	(0.001)
Panel E: Nobility						
Finishing school × Post 1650	-0.012	-0.020	-0.000	-0.006	-0.002	-0.007
	(0.016)	(0.017)	(0.014)	(0.015)	(0.007)	(0.008)
# Finishing schools × Post 1650	0.004	0.003	0.015*	0.016*	0.006	0.005
	(0.008)	(0.009)	(0.009)	(0.009)	(0.004)	(0.005)
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Religious covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	9,312	9,240	9,312	9,240	9,312	9,240

Notes: Results from a "standard" difference-in-differences setup reported. We present extensive (at least one finishing school by 1850) and intensive (number of finishing schools by 1850) margin effects by interacting our finishing school variable with a post-1650 indicator (the first period with finishing schools). We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses.$*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

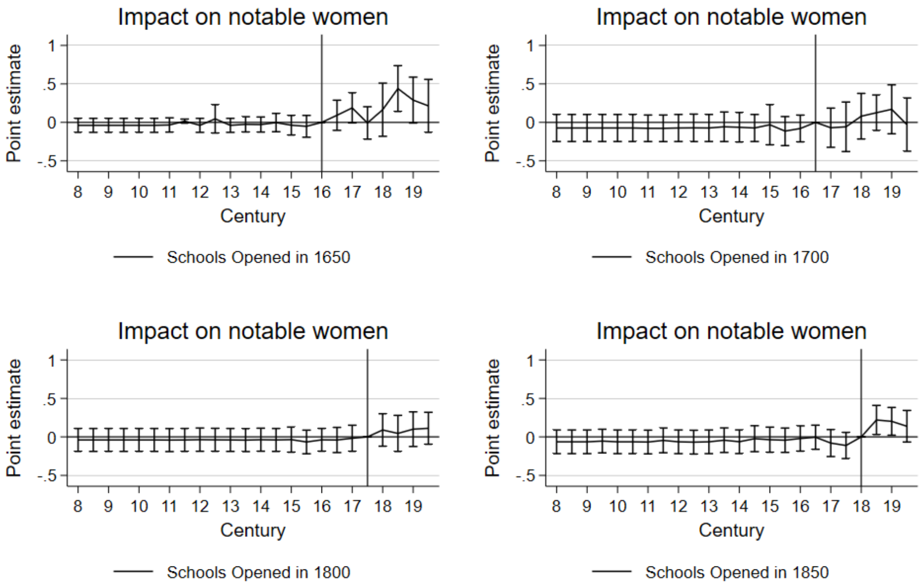


Figure 3.F.2: Parallel trends analysis: Lead-lag figure by treatment cohort

Notes: Each figure represents the lead-lag graph for the indicated treatment group relative to the never-treated control group. The outcome is an indicator equal to one if a notable woman was born in a given city and period. No controls included. 95% confidence intervals reported.

Table 3.F.2: Difference-in-differences estimation: Establishing finishing schools in different periods

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Finishing school by 1650 × post 1650	0.294***	0.190***	0.415***	0.228**	0.023***	0.021**
	(0.060)	(0.058)	(0.114)	(0.105)	(0.008)	(0.008)
Finishing school by 1700 × post 1700	0.248***	0.138**	0.350***	0.180**	0.014***	0.014**
	(0.076)	(0.061)	(0.105)	(0.084)	(0.005)	(0.006)
Finishing school by 1750 × post 1750	0.159*	0.069	0.855*	0.699*	0.025*	0.024
	(0.083)	(0.072)	(0.437)	(0.366)	(0.015)	(0.017)
Finishing school by 1800 × post 1800	0.195***	0.134**	0.347***	0.190**	0.015**	0.019**
	(0.047)	(0.052)	(0.099)	(0.087)	(0.007)	(0.008)
Finishing school by 1850 × post 1850	0.249***	0.203***	0.248***	0.137*	0.023***	0.023***
	(0.050)	(0.052)	(0.067)	(0.074)	(0.009)	(0.009)
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes		Yes
Educational covariates × period FE		Yes		Yes		Yes
Observations	9,312	9,240	9,312	9,240	9,312	9,240

Notes: Results from a "standard" difference-in-differences setup reported. We divide the sample with respect to whether a city had a finishing school in the indicated year and interact this variable with a post-year indicator to obtain the corresponding difference-in-differences estimate. All coefficients are jointly estimated. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses.$*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

3.F.2 Protestantism as a confounding factor

Next, we turn to the diffusion of Protestantism as a potential confounding factor. Martin Luther advocated the education of women to enable their independent study of the Bible (S. O. Becker and Woessmann 2009). It is important to note, however, that he only argued for primary education (particularly reading), and not the secondary education and teacher training provided by finishing schools. We thus do not expect a significant impact of the Protestant Reformation on women's representation among the human capital elite. In order to obtain a causal estimate that is not confounded by the potentially endogeneous decision to adopt Protestantism, we also provide estimates using an instrumental variables strategy based on a city's distance to Wittenberg, the Reformation's epicenter.

We assess the impact of the Protestant Reformation on women's representation among the human capital elite in Figure 3.F.3. In the right-hand Panel, we report estimates from an OLS regression of an indicator whether a notable woman was born in a given city and period on an indicator for whether a certain city adopted Protestantism by 1650. The lead-lag estimates suggest no consistently significant and positive effect of the Protestant Reformation on women's representation among the human capital elite until 1900. In the right hand Panel, we report estimates from a reduced form exercise where we replace the indicator for having adopted Protestantism by 1650 with the distance to Wittenberg, the city from which Protestantism spread across Germany. Again, we find no consistent positive effect on notable women. Taken together, Figure 3.F.3 suggests that our main results on the nexus between finishing schools and women's increasing representation among the human capital elite are unlikely to merely reflect the effects of the Protestant Reformation. The difference-in-differences estimates (odd columns) and reduced form estimates (even columns) in Table 3.F.3 confirm this pattern as they do not reveal a significant impact of the Reformation on women among the human capital elite.[43]

[43] We also find no evidence of a heterogeneous effect of the Reformation on the number of notable women.

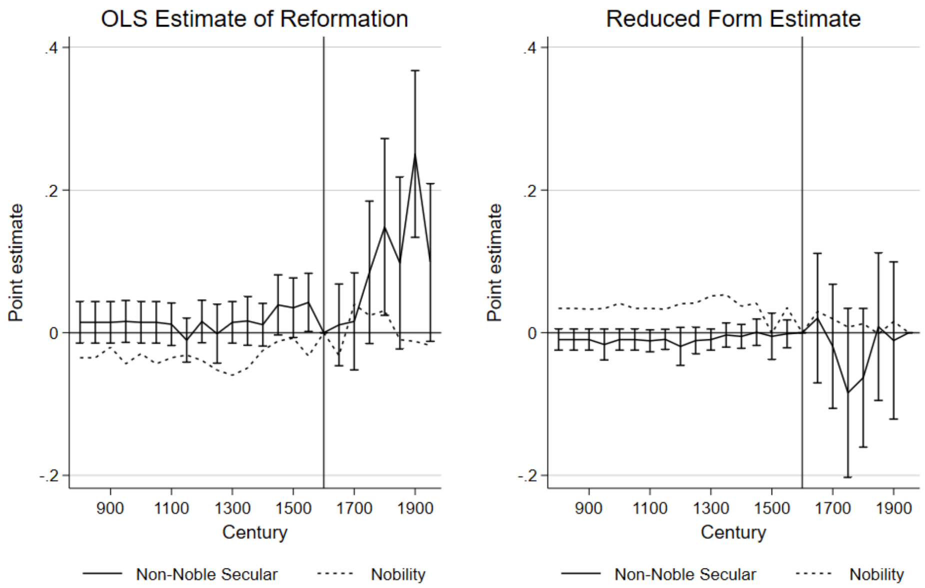


Figure 3.F.3: Using the Protestant Reformation as explanatory variation

Notes: Estimating the impact of switching to Protestantism and the reduced form impact of the log distance to Wittenberg across all time periods in our data for non-noble secular women and women from the nobility. The outcome is an indicator equal to one if a notable woman from the respective group was born in a given city and period. We exclude religious controls in all estimations. 95%-confidence intervals shown only for non-noble secular, the impact of nobility is indistinguishable from zero in all periods and specifications.

Table 3.F.3: Difference-in-differences estimation: Switch to Protestantism as a cultural shock to the role of women in society

	$\mathbb{I}[Women > 0]$		log Women		Share Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Noble Seculars						
Reformation in City × post 1600	0.056**		0.068*		0.003	
	(0.023)		(0.035)		(0.003)	
log Distance to Wittenberg × post 1600		-0.041*		-0.046		-0.003
		(0.022)		(0.039)		(0.003)
Panel B: Unmarried women						
Reformation in City × post 1600	0.083***		0.088**		0.003	
	(0.027)		(0.037)		(0.004)	
log Distance to Wittenberg × post 1600		-0.009		-0.023		-0.001
		(0.028)		(0.040)		(0.003)
Panel C: Teachers & Writers						
Reformation in City × post 1600	0.030*		0.028		0.004	
	(0.018)		(0.018)		(0.004)	
log Distance to Wittenberg × post 1600		-0.032		-0.031		-0.004
		(0.022)		(0.024)		(0.003)
Panel D: Activists						
Reformation in City × post 1600	0.014		0.010		0.001	
	(0.010)		(0.007)		(0.003)	
log Distance to Wittenberg × post 1600		-0.005		-0.005		-0.000
		(0.008)		(0.008)		(0.002)
Panel E: Nobility						
Reformation in City × post 1600	0.026		0.026		0.012	
	(0.022)		(0.019)		(0.010)	
log Distance to Wittenberg × post 1600		0.017		0.010		0.005
		(0.012)		(0.010)		(0.005)
Unit trend	Yes	Yes	Yes	Yes	Yes	Yes
City covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,288	9,288	9,288	9,288	9,288	9,288
F-Test		16.227		16.227		16.227

Notes: Results from a difference-in-differences estimation reported. In odd columns, we use an indicator variable whether a city has adopted Protestantism by 1650 to compute the difference-in-differences estimate. In even columns, we use the log distance to Wittenberg, the Reformation's epicenter, as a proxy for whether a city switched to Protestantism in a reduced form exercise. The corresponding first-stage effect exhibits an F-Stat of 16.23. We consider three types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator equal to one if a city observed the birth of at least one notable woman in a given period. "log Women" constitutes the natural logarithm of the number of women born plus one. "Share Women" denotes the number of women by the number of men and women in the same category, except for Activists, where we use the number of male politicians. Columns (1), (3), and (5) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2), (4), and (6) we interact city controls with period fixed effects to capture variation from economic and educational differences. We include the following controls measured in the 13th century: Whether the city was a Hanseatic League or bishopric city and whether it had a Jewish presence and a pogrom. Additionally, we include the following controls from 1600: confessional battle in the vicinity. In addition we control for the average temperature in 1650 to capture differential agricultural productivity, and hence income. City-level population in 1600 is included to capture different population effects and pre-existing male schools, universities in 1650 to capture differential educational preferences. All covariates are interacted with period fixed effects. Standard errors clustered by city shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.F.3 Monasteries as an instrument

Finally, we discuss a potential instrument for the establishment of finishing schools. From historical accounts we know that most of the early finishing schools were founded by Catholic nuns (Albisetti 1988). These nuns were often invited by rulers of German states and settled in available space in and around existing monasteries. We use monasteries that were established by 1300, more than 300 years prior to the opening of the first finishing school, as an instrument for finishing school establishment. With this instrument we exploit variation in the supply of buildings which could be converted to (or expanded to include) finishing schools at fairly low cost. By additionally limiting our analysis to cities in close vicinity to the inner-German denominational divide between Protestants and Catholics as of 1618, we hold religious competition constant. Thus, we estimate effects net of any direct impact of religious competition which the historical literature on finishing schools suggests as an important determinant of finishing school establishment (Lewejohann 2014). The key identification assumption is then that the number of monasteries established by 1300 in areas which were to become religiously competitive around the year 1600 only affects women's representation among the human capital elite via the construction of finishing schools. Figure 3.F.4 summarizes our findings. Using monasteries as an instrument provides reliable reduced form estimates that suggest a relevant instrument that is independent of the chosen bandwidth around the religious divide.

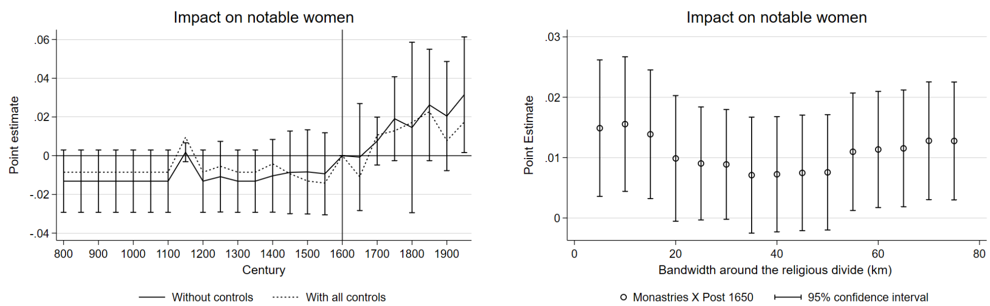


Figure 3.F.4: Reduced form estimates: Using monasteries in 1300 as an instrument

Notes: Estimating the reduced form impact of monasteries in 1300 on Non-Noble Secular women across all time periods in our data within 10km of the religious divide (left). The outcome is an indicator equal to one if a notable woman was born in a given city and period. Estimates with and without all controls all indicate no pre-trends and an increase in the likelihood of women becoming notable after the opening of the first school in 1626. Sensitivity of the point estimate comparing pre- and post-treatment periods to various bandwidths shown in the right figure. All controls included.

3.G Accumulation and role-model hypothesis

In this Appendix, we discuss whether finishing schools served as a pull factor motivating women from the human capital elite to migrate into a city. In contrast to the rest of the paper, where we link notable women to cities based on their place of *birth*, for this exercise, we leverage information on notable women's place of *death* to measure whether finishing schools attracted notable women from elsewhere. We thus investigate whether finishing schools contributed to a local accumulation of notable women, potentially via the mechanism that local notable women served as role-models in attracting others. In Table 3.G.1 we show that upon the establishment of the first finishing school in a city, more women from the human capital elite born in other cities moved to the city with the newly established finishing school. It is important to note that our rich data on notable women's places of birth and places of death allow us to distinguish the in-migration of notable women born elsewhere from spillover effects, which we discuss in Appendix 3.D. Our data also allow us to document that finishing schools attracted the in-migration of women from the human capital elite to these cities while ruling out that finishing schools were established in response to the in-migration of women from the human capital elite as evidenced by the clear absence of differential pre-trends in Figure 3.6 in the main text.

A further concern is that most of the positive effect of finishing schools on the in-migration of women from the human capital elite might be mechanical since finishing schools were primary employers for notable women. We test for this in the second Panel of Table 3.G.1: we find that once we add our control variables and thus adequately control for initial differences between cities, we see no significant effect of finishing schools on the number of notable teachers who migrated to a city with a finishing school. This suggests that a potential mechanical effect for teachers alone cannot account for the main effect shown in the first Panel of Table 3.G.1.

Taken together, the evidence presented in this Appendix suggests that finishing schools indeed served as a pull factor which attracted notable women born elsewhere.

Table 3.G.1: Testing role-model and accumulation hypotheses

	$\mathbb{I}[Women > 0]$		log Women	
	(1)	(2)	(3)	(4)
Panel A: Immigration of Non-Noble Seculars				
Finishing school$_{it}$	0.114***	0.059**	0.134***	0.049*
	(0.023)	(0.024)	(0.033)	(0.028)
Mean, untreated	0.042	0.042	0.034	0.034
Panel B: Immigration of Teachers & Writers				
Finishing school$_{it}$	0.049***	0.016	0.052**	0.015
	(0.018)	(0.019)	(0.022)	(0.019)
Mean, untreated	0.020	0.020	0.016	0.016
Unit trend	Yes	Yes	Yes	Yes
City covariates × period FE		Yes		Yes
Religious covariates × period FE		Yes		Yes
Educational covariates × period FE		Yes		Yes
Observations	9,312	9,240	9,312	9,240

Notes: Results from fixed-effects regressions using cities and periods reported. We consider two types of dependent variables: $\mathbb{I}[Women > 0]$ is an indicator taking value 1 if a city observed the immigration of at least one notable woman born elsewhere in a given period. "log Women" constitutes the natural logarithm of immigrated women plus 1. We regress our dependent variables on a dummy taking value 1 if a finishing school existed in a given city and period. Columns (1) and (3) constitute the baseline and include city and period fixed effects as well as city-specific linear trends. In columns (2) and (4) we interact our full set of control variables (as defined in Table 3.2) with period fixed effects to capture variation from economic, religious, and educational differences. Standard errors clustered at the city level reported in parentheses.$*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

3.H Specification and robustness in the cross-sectional setting

In this Appendix, we want to highlight that our cross-sectional setting is robust to using an instrumental variables estimation, to estimating effects of a city's length of exposure to finishing schools, and to matching on observables.

First, we discuss a potential instrument for the establishment of finishing schools. From historical accounts we know that most of the early finishing schools were founded by Catholic nuns (Albisetti 1988). These nuns were often invited by rulers of German states and settled in available space in and around existing monasteries. We use monasteries that were established by 1300, more than 300 years prior to the opening of the first finishing school, as an instrument for finishing school establishment. With this instrument we exploit variation in the supply of buildings which could be converted to (or expanded to include) finishing schools at fairly low cost. By additionally limiting our analysis to cities in close vicinity to the inner-German denominational divide between Protestants and Catholics as of 1618, we can hold religious competition constant and thus estimate effects net of any direct impact of religious competition which the historical literature on finishing schools suggests as an important determinant of finishing school establishment (Lewejohann 2014). The key identification assumption is then that the number of monasteries established by 1300 in areas which were to become religiously competitive around the year 1600 only affects women's representation among the human capital elite via the construction of finishing schools.

In Table 3.H.1, we show that indeed using the number of monasteries existing in 1300 as an instrument for the number of finishing schools in 1850 produces consistent estimates throughout all outcomes and main specifications (columns 1 and 4). Changing the cutoff year for pre-existing monasteries closer to 1648, the end of the Thirty Years' War, produces similarly sized estimates, yet smaller F-statistics (columns 2, 3, 5, and 6).

Finally, we estimate effects of a city's length of exposure to finishing schools (instead of the absolute number of finishing schools). In Table 3.H.2, we show that changing the independent variable to years since first opening produces very similar results in a wide range of specifications. Here, we define zero as having no school in 1850, and progressively move back in time to '224', indicating the school was build in 1626. In Table 3.H.2 we thus investigate whether more time elapsed since the establishment of the first finishing school in a city – and thereby a greater representation of women among the human capital elite – is associated with stronger support of the women's rights movement.

At a mean of 20 years of exposure to finishing schools, increasing the number of years by 10% (2 years), increases the number of letters to *Frauenzeitung* by 0.56%, the number of women's rights associations by 5% and the number of female members of parliament by 0.25% and 0.95% respectively. Or to put it differently, had a city opened a finishing school in 1800 (instead of never) and thus had 50 years more exposure to such a school, this would imply a 250% increase in exposure compared to the mean of 20 years. This city would have seen 14% more letters, twice the number of women's rights organizations, and 24% more women in postwar parliaments. These are sizeable effects, for a relatively small change in exposure.

Table 3.H.1: Long-term impact of finishing schools on political outcomes - IV estimates using different timings of the monastery instrument

	$\mathbb{I}[>0]$			log Number		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Leserbriefe, Frauenzeitung, 1849–1852						
Finishing schools	0.249**	0.274**	0.297**	0.412***	0.492***	0.444**
	(0.098)	(0.108)	(0.121)	(0.158)	(0.187)	(0.192)
Mean, untreated	0.038	0.038	0.038	0.061	0.061	0.061
Panel B: All women's rights organizations						
Finishing schools	0.378*	0.378	0.258	2.868*	2.844	2.308
	(0.223)	(0.241)	(0.219)	(1.680)	(1.835)	(1.678)
Mean, untreated	0.275	0.275	0.275	155.802	155.802	155.802
Panel C: Women's rights organizations to promote equal access to education						
Finishing schools	0.333**	0.340*	0.393**	2.099**	2.123**	2.504**
	(0.159)	(0.178)	(0.178)	(0.851)	(0.940)	(0.966)
Mean, untreated	0.038	0.038	0.038	13.023	13.023	13.023
Panel D: Member Parliament, 1919–1933						
Finishing schools	0.164*	0.122	0.137	0.227**	0.193**	0.226**
	(0.093)	(0.090)	(0.104)	(0.090)	(0.093)	(0.093)
Mean, untreated	0.038	0.038	0.038	0.053	0.053	0.053
Panel E: Member Parliament, 1949–2019						
Finishing schools	0.237	0.236	0.179	0.471**	0.480*	0.524*
	(0.174)	(0.189)	(0.192)	(0.223)	(0.247)	(0.269)
Mean, untreated	0.527	0.527	0.527	1.031	1.031	1.031
City Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates	Yes	Yes	Yes	Yes	Yes	Yes
Observations	183	183	183	183	183	183
Bandwidth	10	10	10	10	10	10
Monastery Year	1300	1500	1648	1300	1500	1648
F-Stat first stage	8.906	7.177	8.435	8.906	7.177	8.435

Notes: Results from two-stage least-squares (2SLS) regressions reported. We instrument the number of finishing schools in 1850 by the number of monasteries by 1300 (Columns 1 and 4), by 1500 (Columns 2 and 5), or by 1648 (Columns 3 and 6). We further limit the sample to cities within 10 km of the inner-German denominational divide to hold religious competition constant. In each panel, we employ two types of outcomes: (i) an indicator variable taking value 1 if at least one letter (Panel A), women's rights organization (Panels B and C), or female member of parliament (Panels D and E) was recorded for a given city; and (ii) the natural logarithm of these variables plus 1. We employ our full set of educational, economic, and religious control variables as defined in Table 3.2 in all columns. Standard errors clustered at the city level reported in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 3.H.2: Long-term impact of finishing schools on political outcomes: Years since opening of the first finishing school as explanatory variable

	$\mathbb{I}[>0]$				log Number			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Leserbriefe, Frauenzeitung, 1849–1852								
Years since first opening	0.001*	0.001			0.002**	0.001		
	(0.000)	(0.001)			(0.001)	(0.001)		
log Years since first opening			0.024**	0.025*			0.056***	0.061*
			(0.011)	(0.015)			(0.021)	(0.035)
Mean, untreated	0.062	0.038	0.062	0.038	0.105	0.061	0.105	0.061
Panel B: All women's rights organizations								
Years since first opening	0.002***	0.002***			0.015***	0.015***		
	(0.000)	(0.001)			(0.003)	(0.005)		
log Years since first opening			0.072***	0.096***			0.535***	0.634***
			(0.015)	(0.023)			(0.098)	(0.148)
Mean, untreated	0.366	0.275	0.366	0.275	447.696	155.802	447.696	155.802
Panel C: Women's rights organizations to promote equal access to education								
Years since first opening	0.001*	0.000			0.005*	0.003		
	(0.001)	(0.001)			(0.003)	(0.004)		
log Years since first opening			0.029**	0.017			0.169***	0.100
			(0.011)	(0.015)			(0.062)	(0.081)
Mean, untreated	0.047	0.038	0.047	0.038	13.074	13.023	13.074	13.023
Panel D: Member Parliament, 1919–1933								
Years since first opening	0.001**	0.001			0.001***	0.001*		
	(0.000)	(0.001)			(0.000)	(0.001)		
log Years since first opening			0.023**	0.022			0.025**	0.020*
			(0.011)	(0.014)			(0.010)	(0.011)
Mean, untreated	0.066	0.038	0.066	0.038	0.074	0.053	0.074	0.053
Panel E: Member Parliament, 1949–2019								
Years since first opening	0.001	0.001*			0.002*	0.004**		
	(0.001)	(0.001)			(0.001)	(0.002)		
log Years since first opening			0.044***	0.063***			0.095***	0.137***
			(0.015)	(0.022)			(0.025)	(0.036)
Mean, untreated	0.556	0.527	0.556	0.527	1.163	1.031	1.163	1.031
City Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	385	183	385	183	385	183	385	183
Bandwidth	400	10	400	10	400	10	400	10

Notes: Results from cross-sectional regressions reported. We employ two different explanatory variables: (i) time elapsed since the opening of the first finishing school in a city until 1850 (measured in years) and (ii) the natural logarithm plus 1 of time elapsed since the first finishing school. Hence, time elapsed since the first school opening is zero for cities which have never established a finishing school. In each panel, we employ two types of outcomes: (i) an indicator variable taking value 1 if at least one letter (Panel A), women's rights organization (Panels B and C), or female member of parliament (Panels D and E) was recorded for a given city; and (ii) the natural logarithm of these variables plus 1. We employ our full set of educational, economic, and religious control variables as defined in Table 3.2 in all columns. We report results using two different bandwidths around the inner-German denominational divide: 400km in odd columns and 10km in even columns. Standard errors clustered at the city level reported in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

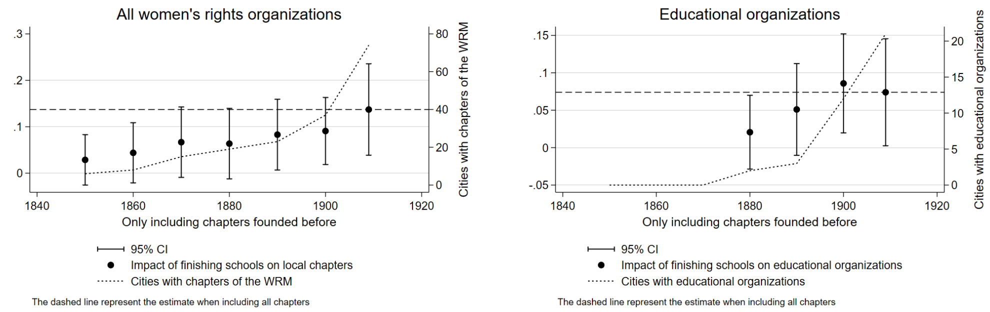


Figure 3.H.1: The impact of finishing schools on chapters of the women's rights movement - Time varying effects

Notes: Results from cross-sectional regressions reported. In both figures, the sample is limited to cities within 10km of the inner-German denominational divide. The left figure shows the impact of finishing schools on whether any local chapter of the women's rights movement was founded in a city by 1850, 1860, 1870, 1880, 1890, 1900, and 1909, respectively. The right figure shows the same impact on local chapters devoting their efforts to promoting equal access to education for women. All economic, educational, and religious controls included in both figures. 95-percent confidence intervals derived from standard errors clustered at the city level reported.

3.H.1 Comparison to propensity score matching

As a final step, we show robustness of our results to matching each city to its closest counterparts based on observable characteristics in Table 3.H.3. The point estimates in columns (3) and (6) are not statistically different from the OLS (columns 1 and 4) or the sample of cities that lie within 10 km of the religious divide (columns 2 and 5). In addition, the matched sample shows no signs of imbalances across all covariates (Table 3.H.4).

Table 3.H.3: Long-term impact of finishing schools on political outcomes: Comparison to matching estimators

	$\mathbb{I}[>0]$			log Number		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Leserbriefe, Frauenzeitung, 1849–1852						
Finishing schools	0.095***	0.122***	0.144***	0.183***	0.241**	0.209***
	(0.020)	(0.037)	(0.018)	(0.055)	(0.097)	(0.076)
Panel B: All women's rights organizations						
Finishing schools	0.064**	0.137***	-0.003	0.800***	1.157***	0.532***
	(0.027)	(0.050)	(0.023)	(0.160)	(0.306)	(0.194)
Panel C: Women's rights organizations to promote equal access to education						
Finishing schools	0.083***	0.074**	0.055*	0.549***	0.496**	0.510***
	(0.017)	(0.036)	(0.029)	(0.113)	(0.217)	(0.194)
Panel D: Member Parliament, 1919–1933						
Finishing schools	0.067***	0.101***	0.042	0.100***	0.105***	0.107**
	(0.018)	(0.034)	(0.027)	(0.029)	(0.035)	(0.049)
Panel E: Member Parliament, 1949–2019						
Finishing schools	0.060**	0.091*	0.012	0.246***	0.268***	0.280***
	(0.024)	(0.047)	(0.025)	(0.040)	(0.071)	(0.055)
City Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Religious covariates	Yes	Yes	Yes	Yes	Yes	Yes
Educational covariates	Yes	Yes	Yes	Yes	Yes	Yes
Propensity score matching			Yes			Yes
Observations	385	183	318	385	183	318
Bandwidth		10			10	

Notes: Results from cross-sectional regressions reported. We presents results from three sets of specifications: (i) using the full set of controls (columns 1 and 4); (ii) using the full set of controls when we limit the sample to cities within 10 km of the inner-German denominational divide (columns 2 and 5); and (iii) using propensity score matching (columns 3 and 6). In each panel, we employ two types of outcomes: (i) an indicator variable taking value 1 if at least one letter (Panel A), women's rights organization (Panels B and C), or female member of parliament (Panels D and E) was recorded for a given city; and (ii) the natural logarithm of these variables plus 1. We employ our full set of educational, economic, and religious control variables as defined in Table 3.2 in all columns. Standard errors clustered at the city level reported in parentheses. $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$.

Table 3.H.4: Balance in the matched sample

	Unmatched sample			Matched sample		
	β	s.e.	p-value	β	s.e.	p-value
log(Distance Wittenberg)	-0.080	0.040	0.046	0.084	0.079	0.291
log(Distance religious divide)	0.214	0.065	0.001	0.029	0.098	0.769
log(Population in 1650)	0.422	0.058	0.000	0.047	0.034	0.167
Temperature in Spring 1650	0.011	0.041	0.783	0.011	0.068	0.871
Temperature in Summer 1650	0.079	0.048	0.097	0.010	0.075	0.892
Temperature in Fall 1650	-0.002	0.036	0.947	-0.022	0.052	0.668
Temperature in Winter 1650	-0.119	0.048	0.014	-0.078	0.065	0.227
Hanse city	0.044	0.020	0.031	-0.016	0.039	0.689
Bishop seat	0.036	0.017	0.033	-0.030	0.022	0.184
Jewish settlement	0.081	0.025	0.001	0.021	0.039	0.598
Progrom	0.044	0.023	0.057	0.036	0.039	0.350
Battle during 30-years war	0.062	0.021	0.003	-0.049	0.049	0.314
Boy school in 1605	0.018	0.017	0.279	0.036	0.030	0.221
University in 1650	0.005	0.008	0.557	-0.004	0.011	0.701
Catholic region	0.012	0.023	0.597	0.014	0.042	0.746

Notes: Results from balance test of covariates in 1650 reported. Balancing is assessed using the regression $X_c = \alpha + \beta \cdot \text{Schools}_{c,1850} + \varepsilon_c$. The unmatched sample contains all cities in 1650, whereas the matched sample selects a nearest neighbor – that is a city that is comparable with respect to observables – for each city with at least one finishing school. While cities with finishing schools are closer to Wittenberg, further away from the inner-German denominational divide and have larger population in 1650, these differences disappear when matching cities to their nearest neighbor.

3.I Impact of notable women in 1850 on local political activity

In this Appendix, we directly ask what is the correlation between an additional non-noble secular women in 1850 and subsequent political activity in the next 100 years. To this end, we estimate the following equation in Table 3.I.1:

$$Y_c = \alpha + \beta \cdot \log(\text{Number Non-Noble Seculars+1})_{c,1850} + \boldsymbol{X}_c \, \gamma' + \varepsilon_c \tag{2}$$

Recognizing the endogeneity concerns associated with this equation, we nevertheless present estimates for their interpretability: a 10% increase in the number of notable women in a city is associated with a 2% increase in correspondence (Panel A), a 15% increase in women's rights associations (Panels B&C), and a 2% (4.6%) increase in the number of female members of parliament during the Weimar Republic (Federal Republic).

We conduct two exercises to judge the reliabilty of these correlations. First, we present point estimates with (odd columns) and without (even columns) controls, limited to 10 km of the religious boundary. The estimates remain stable throughout all specifications. Second, we instrument the number of notable women by the number of existing monastries in 1300 and provide the 2SLS coefficient, the p-value and F-statistic below the OLS estimates. However, as the exclusion restriction, monastries only affect political outcomes through their influence on finishing schools' impact on notable women, is likely to fail, we interpret these estimates with caution. All 2SLS estimates are significant and larger than the OLS estimates with a strong first stage of 14: a 10% increase in the number of notable women in each city is associated with an 8% increase in correspondence (Panel A), a 40% increase in women's rights associations for education (Panel C), and a 4% increase in the number of female members of parliament during the Weimar Republic (Panel D).[44]

[44] A similar exercise using finishing schools as an instrument can be conducted. It yields qualitatively similar results with a stronger first stage of 22.

Table 3.I.1: Impact of notable women in 1850 on political activity of the women's rights movement

	$\mathbb{I}[>0]$		log Number	
	(1)	(2)	(3)	(4)
Panel A: Leserbriefe, Frauenzeitung, 1849–1852				
log(Number Non-Noble Seculars)	0.221***	0.190***	0.381***	0.227***
	(0.028)	(0.058)	(0.078)	(0.076)
Implied 2SLS coefficient	0.246	0.483	0.285	0.800
P-value 2SLS coefficient	0.005	0.013	0.038	0.018
First stage F-statistic	29.640	14.916	29.640	14.916
Panel B: All women's rights organizations				
log(Number Non-Noble Seculars)	0.262***	0.133*	2.598***	1.511***
	(0.030)	(0.076)	(0.178)	(0.488)
Implied 2SLS coefficient	0.695	0.734	5.737	5.563
P-value 2SLS coefficient	0.000	0.074	0.000	0.052
First stage F-statistic	29.640	14.916	29.640	14.916
Panel C: Women's rights organizations to promote equal access to education				
log(Number Non-Noble Seculars)	0.300***	0.263***	1.813***	1.516***
	(0.025)	(0.055)	(0.144)	(0.303)
Implied 2SLS coefficient	0.531	0.646	2.840	4.073
P-value 2SLS coefficient	0.000	0.000	0.000	0.000
First stage F-statistic	29.640	14.916	29.640	14.916
Panel D: Member Parliament, 1919–1933				
log(Number Non-Noble Seculars)	0.206***	0.204***	0.257***	0.198***
	(0.030)	(0.050)	(0.044)	(0.049)
Implied 2SLS coefficient	0.422	0.319	0.461	0.440
P-value 2SLS coefficient	0.000	0.036	0.000	0.004
First stage F-statistic	29.640	14.916	29.640	14.916
Panel E: Member Parliament, 1949–2019				
log(Number Non-Noble Seculars)	0.197***	0.199***	0.610***	0.469***
	(0.027)	(0.070)	(0.051)	(0.109)
Implied 2SLS coefficient	0.418	0.460	0.935	0.914
P-value 2SLS coefficient	0.000	0.120	0.000	0.005
First stage F-statistic	29.640	14.916	29.640	14.916
City Covariates		Yes		Yes
Religious covariates		Yes		Yes
Educational covariates		Yes		Yes
Observations	388	183	388	183
Bandwidth		10		10

Notes: Results from cross-sectional regressions reported. Our explanatory variable in these regressions is the natural logarithm of the number of non-noble secular notable women in a given city by 1850. We also report the point estimate, p-value and F-statistic from an 2SLS regression below the OLS coefficient for convenience. To obtain the 2SLS estimate, we instrument the log number of notable women in city c with the number monasteries existing by 1300. We further enhance the comparability of cities, in particular with respect to historical levels of religious competition, by limiting our sample to cities within 10 km of the inner-German denominational divide in odd columns. In each panel, we employ two types of outcomes: (i) an indicator variable taking value 1 if at least one letter (Panel A), women's rights organization (Panels B and C), or female member of parliament (Panels D and E) was recorded for a given city; and (ii) the natural logarithm of these variables plus 1. We employ our full set of educational, economic, and religious control variables as defined in Table 3.2 in all columns. Standard errors clustered at the city level reported in parentheses. Significance levels: $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$.

3.J Additional history on finishing schools

In Figure 3.J.1, we depict the spatial distribution of finishing schools in Germany separately by denomination – that is, by indicating which school was Catholic and which was Protestant. As becomes apparent from the figure, the first schools were exclusively Catholic. In fact, the first Protestant school opened in 1698. The first school funded by city authorities opened in 1800. The observed acceleration in the roll-out of finishing schools in the period 1800-1850 is likely driven by the dissolution of the Holy Roman Empire (800-1806), freeing up resources from previous inner-German conflicts. More than 100 schools were built between 1825 and 1850 alone, most of them in Prussia. Interestingly, Prussia recruited many of its female teachers from Catholic Bavaria. Comparing treatment effects for early and late periods (Table 3.C.5) and treatment periods (Table 3.F.2) suggest no differential treatment effects with respect to time.

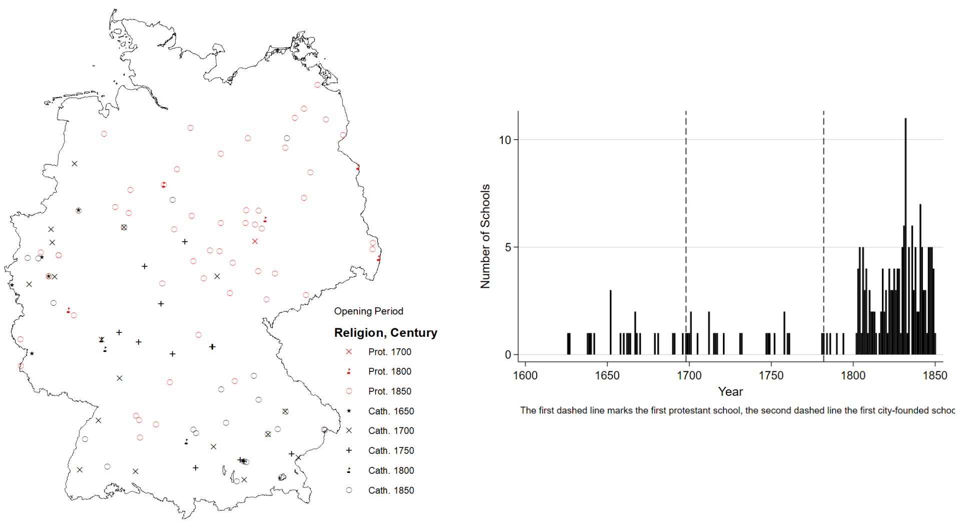


Figure 3.J.1: Opening years of finishing schools in Germany

Notes: The left map displays the spatial distribution Catholic and Protestant finishing schools in Germany. The right figure reports the opening year of each finishing school in our dataset.

BIBLIOGRAPHY

ABADIE, A., S. ATHEY, G. W. IMBENS, and J. WOOLDRIDGE (2017). *When Should You Adjust Standard Errors for Clustering?* Working Paper 24003. National Bureau of Economic Research.

ÅKERLUND, D., B. H. H. GOLSTEYN, H. GRÖNQVIST, and L. LINDAHL (2016). "Time Discounting and Criminal Behavior". *Proceedings of the National Academy of Sciences*, 113 (22), 6160–6165.

ALBISETTI, J. C. (1988). *Schooling German Girls and Women: Secondary and Higher Education in the Nineteenth Century*. Princeton: Princeton University Press.

ALTHOUSE, A. D. (2016). "Adjust for Multiple Comparisons? It's Not That Simple". *The Annals of Thoracic Surgery*, 101 (5), 1644–1645.

ANDERSON, L. R. and J. M. MELLOR (2008). "Predicting Health Behaviors with an Experimental Measure of Risk Preference". *Journal of Health Economics*, 27 (5), 1260–1274.

ANDREONI, J. (1989). "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence". *Journal of Political Economy*, 97 (6), 1447–1458.

ANGRIST, J. D. and K. LANG (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program". *American Economic Review*, 94 (5), 1613–1634.

ARBAK, E. and M.-C. VILLEVAL (2013). "Voluntary Leadership: Motivation and Influence". *Social Choice and Welfare*, 40 (3), 635–662.

ARON, A., E. N. ARON, and D. SMOLLAN (1992). "Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness". *Journal of Personality and Social Psychology*, 63 (4), 596–612.

BAIROCH, P., J. BATOU, and P. CHÈVRE (1988). *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats*. Genève: Librairie Droz.

BAKER, A., D. F. LARCKER, and C. C. Y. WANG (2021). "How Much Should We Trust Staggered Difference-in-Differences Estimates?" *Unpublished*.

BAYERISCHE AKADEMIE DER WISSENSCHAFTEN, HISTORISCHE KOMMISSION (1953). *Neue Deutsche Biographie*. Vol. 1. Berlin: Duncker & Humblot.

— (2019). *Deutsche Biographie*. https://www.deutsche-biographie.de/home. Last accessed: 2021-09-03.

BAYERISCHES LANDESAMT FÜR STATISTIK (2019). *Demographie-Spiegel für bayerische Gemeinden bis zum Jahr 2037 veröffentlicht*. https://www.e-statistik.eu/presse/mitteilungen/2019/pm181/index.html. Last accessed: 2021-07-15.

BEAMAN, L., R. CHATTOPADHYAY, E. DUFLO, R. PANDE, and P. TOPALOVA (2009). "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics*, 124 (4), 1497–1540.

BECKER, A., T. DECKERS, T. DOHMEN, A. FALK, and F. KOSSE (2012). "The Relationship Between Economic Preferences and Psychological Personality Measures". *Annual Review of Economics*, 4, 453–478.

BECKER, S. O. and L. WOESSMANN (2008). "Luther and the Girls: Religious Denomination and the Female Education Gap in 19th Century Prussia". *Scandinavian Journal of Economics*, 110 (12), 777–805.

— (2009). "Was Weber Wrong? A Human Capital Theory of Protestant Economic History". *The Quarterly Journal of Economics*, 124 (2), 531–596.

BÉNABOU, R. and J. TIROLE (2006). "Incentives and Prosocial Behavior". *American Economic Review*, 96 (5), 1652–1678.

BERNDT, S. (2019). "Louise Otto-Peters. Ein Kurzporträt". *Aus Politik und Zeitgeschichte*, 69 (8), 11–17.

BERNHEIM, B. D. (1994). "A Theory of Conformity". *Journal of political Economy*, 102 (5), 841–877.

BERTOCCHI, G. and M. BOZZANO (2016). "Women, Medieval Commerce, and the Education Gender Gap". *Journal of Comparative Economics*, 44 (3), 496–521.

BERTRAND, M. and A. MORSE (2016). "Trickle-Down Consumption". *The Review of Economics and Statistics*, 98 (5), 863–879.

BESLEY, T., J. G. MONTALVO, and M. REYNAL-QUEROL (2011). "Do Educated Leaders Matter?" *The Economic Journal*, 121 (554), F205–227.

BETSCH, C., L. H. WIELER, and K. HABERSAAT (2020). "Monitoring Behavioural Insights Related to COVID-19". *The Lancet*, 395 (10232), 1255–1256.

BOOIJ, A. S., E. LEUVEN, and H. OOSTERBEEK (2016). "Ability Peer Effects in University: Evidence from a Randomized Experiment". *The Review of Economic Studies*, 84 (2), 547–578.

BRODERICK, T., R. GIORDANO, and R. MEAGER (2020). "An Automatic Finite-Sample Robustness Metric: Can Dropping a Little Data Change Conclusions?" *Unpublished*.

BUNDESMINISTERIUM FÜR GESUNDHEIT (2021). *Impfdashboard*. https://impfdashboard.de/. Last accessed: 2021-07-15.

BURSZTYN, L., G. EGOROV, and R. JENSEN (2018). "Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education". *The Review of Economic Studies*, 86 (4), 1487–1526.

BURSZTYN, L., T. FUJIWARA, and A. PALLAIS (2017). "'Acting Wife': Marriage Market Incentives and Labor Market Investments". *American Economic Review*, 107 (11), 3288–3319.

BURSZTYN, L., A. L. GONZÁLEZ, and D. YANAGIZAWA-DROTT (2020). "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia". *American Economic Review*, 110 (10), 2997–3029.

BURSZTYN, L. and R. JENSEN (2015). "How Does Peer Pressure Affect Educational Investments?" *The Quarterly Journal of Economics*, 130 (3), 1329–1367.

— (2017). "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure". *Annual Review of Economics*, 9 (1), 131–153.

CALLAWAY, B. and P. H. C. SANT'ANNA (forthcoming). "Difference-in-Differences with Multiple Time Periods". *Journal of Econometrics*.

CANTONI, D. (2015). "The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands". *Journal of the European Economic Association*, 13 (4), 561–598.

CANTONI, D., J. DITTMAR, and N. YUCHTMAN (2018). "Religious Competition and Reallocation: The Political Economy of Secularization in the Protestant Reformation". *The Quarterly Journal of Economics*, 133 (4), 2037–2096.

CAPPELEN, A. W., B.-A. REME, E. Ø. SØRENSEN, and B. TUNGODDEN (2016). "Leadership and Incentives". *Management Science*, 62 (7), 1944–1953.

CARRELL, S. E., R. L. FULLERTON, and J. E. WEST (2009). "Does Your Cohort Matter? Measuring Peer Effects in College Achievement". *Journal of Labor Economics*, 27 (3), 439–464.

CARRELL, S. E., B. I. SACERDOTE, and J. E. WEST (2013). "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation". *Econometrica*, 81 (3), 855–882.

CHAISEMARTIN, C. de and X. D'HAULTFŒUILLE (2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". *American Economic Review*, 110 (9), 2964–96.

CHATTOPADHYAY, R. and E. DUFLO (2004). "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India". *Econometrica*, 72 (5), 1409–1443.

CHEN, D. L., M. SCHONGER, and C. WICKENS (2016). "oTree – An Open-Source Platform for Laboratory, Online, and Field Experiments". *Journal of Behavioral and Experimental Finance*, 9, 88–97.

CIALDINI, R. B., S. L. BROWN, B. P. LEWIS, C. LUCE, and S. L. NEUBERG (1997). "Reinterpreting the Empathy–Altruism Relationship: When One into One Equals Oneness". *Journal of Personality and Social Psychology*, 73 (3), 481–494.

COLEMAN, J. S. (1966). *Equality of Educational Opportunity*. Washington, DC: National Center for Educational Statistics.

CONRAD, A. (1996). "Weibliche Lehrorden und katholische höhere Mädchenschulen im 17. Jahrhundert". *Geschichte der Mädchen- und Frauenbildung*. Ed. by E. KLEINAU. Vol. 1: "Vom Mittelalter bis zur Aufklärung". Frankfurt, Main: Campus Verlag, 252–262.

COSMO – COVID-19 SNAPSHOT MONITORING (2021). *COSMO – Wave 41, April 2021*. https://projekte.uni-erfurt.de/cosmo2020/web/. Last accessed: 2021-07-15.

DANNENBERG, A. (2015). "Leading by Example versus Leading by Words in Voluntary Contribution Experiments". *Social Choice and Welfare*, 44 (1), 71–85.

DE GIORGI, G. and M. PELLIZZARI (2014). "Understanding Social Interactions: Evidence from the Classroom". *The Economic Journal*, 124 (579), 917–953.

DE QUIDT, J., J. HAUSHOFER, and C. ROTH (2018). "Measuring and Bounding Experimenter Demand". *American Economic Review*, 108 (11), 3266–3302.

DELLA PORTA, D. and A. MATTONI (2016). "Social Movements". *The International Encyclopedia of Political Communication*. Wiley Online Library.

DELLAVIGNA, S., J. A. LIST, and U. MALMENDIER (2012). "Testing for Altruism and Social Pressure in Charitable Giving". *The Quarterly Journal of Economics*, 127 (1), 1–56.

DELLAVIGNA, S., J. A. LIST, U. MALMENDIER, and G. RAO (2017). "Voting to Tell Others". *The Review of Economic Studies*, 84 (1), 143–181.

DEMING, D. J. (2017). "The Growing Importance of Social Skills in the Labor Market". *The Quarterly Journal of Economics*, 132 (4), 1593–1640.

DIEBOLT, C. and F. PERRIN (2013). "From Stagnation to Sustained Growth: The Role of Female Empowerment". *American Economic Review*, 103 (3), 545–49.

DIPPEL, C. and S. HEBLICH (2021). "Leadership in Social Movements: Evidence from the "Forty-Eighters" in the Civil War". *American Economic Review*, 111 (2), 472–505.

DITTMAR, J. and R. MEISENZAHL (2019). "Public Goods Institutions, Human Capital, and Growth: Evidence from German History". *The Review of Economic Studies*, 87 (2), 959–996.

DOHMEN, T., A. FALK, D. HUFFMAN, and U. SUNDE (2009). "Homo Reciprocans: Survey Evidence on Behavioural Outcomes". *The Economic Journal*, 119 (536), 592–612.

DOHMEN, T., A. FALK, D. HUFFMAN, U. SUNDE, J. SCHUPP, and G. G. WAGNER (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences". *Journal of the European Economic Association*, 9 (3), 522–550.

DROUVELIS, M. and D. NOSENZO (2013). "Group Identity and Leading-by-Example". *Journal of Economic Psychology*, 39, 414–425.

DUFLO, E. (2012). "Women Empowerment and Economic Development". *Journal of Economic Literature*, 50 (4), 1051–1079.

DUFLO, E., P. DUPAS, and M. KREMER (2011). "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya". *American Economic Review*, 101 (5), 1739–74.

EHRENREICH, B. and D. ENGLISH (1973). *Witches, Midwives, and Nurses: A History of Women Healers*. New York: Feminist Press.

ENGEL, J. and E. W. ZEEDEN, eds. (1995). *Großer historischer Weltatlas*. München: Bayerischer Schulbuch-Verlag.

ENIKOLOPOV, R., A. MAKARIN, and M. PETROVA (2020). "Social Media and Protest Participation: Evidence from Russia". *Econometrica*, 88 (4), 1479–1514.

EPPER, T., E. FEHR, H. FEHR-DUDA, C. T. KREINER, D. D. LASSEN, S. LETH-PETERSEN, and G. N. RASMUSSEN (2020). "Time Discounting and Wealth Inequality". *American Economic Review*, 110 (4), 1177–1205.

EVANS, R. J. (1980). "German Social Democracy and Women's Suffrage 1891-1918". *Journal of Contemporary History*, 3 (3), 533–557.

FALK, A., A. BECKER, T. DOHMEN, B. ENKE, D. HUFFMAN, and U. SUNDE (2018). "Global Evidence on Economic Preferences". *The Quarterly Journal of Economics*, 133 (4), 1645–1692.

FALK, A., A. BECKER, T. DOHMEN, D. HUFFMAN, and U. SUNDE (forthcoming). "The Preference Survey Module: A Validated Instrument for Measuring Time, Risk, and Social Preferences". *Management Science*.

FALK, A., F. KOSSE, and P. PINGER (2020). "Re-Revisiting the Marshmallow Test: A Direct Comparison of Studies by Shoda, Mischel, and Peake (1990) and Watts, Duncan, and Quan (2018)". *Psychological Science*, 31 (1), 100–104.

FALK, A., F. KOSSE, P. PINGER, H. SCHILDBERG-HÖRISCH, and T. DECKERS (2021). "Socioeconomic Status and Inequalities in Children's IQ and Economic Preferences". *Journal of Political Economy*, 129 (9), 2504–2545.

FELD, J. and U. ZÖLITZ (2017). "Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects". *Journal of Labor Economics*, 35 (2), 387–428.

FERNÁNDEZ, R. (2013). "Cultural Change as Learning: The Evolution of Female Labor Force Participation Over a Century". *American Economic Review*, 103 (1), 472–500.

FRIGO, A. and E. R. FERNANDEZ (2019). "Roots of Gender Equality: The Persistent Effect of Beguinages on Attitudes Toward Women". *Unpublished*.

GÄCHTER, S., D. NOSENZO, E. RENNER, and M. SEFTON (2012). "Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-by-Example". *Economic Inquiry*, 50 (4), 953–967.

GÄCHTER, S. and E. RENNER (2018). "Leaders as Role Models and Belief Managers in Social Dilemmas". *Journal of Economic Behavior & Organization*, 154, 321–334.

GÄCHTER, S., C. STARMER, and F. TUFANO (2015). "Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale". *PLoS One*, 10 (6), Article e0129478.

GALOR, O. and D. N. WEIL (1996). "The Gender Gap, Fertility, and Growth". *American Economic Review*, 86 (3), 374–387.

GARCIA-JIMENO, C., A. IGLESIAS, and P. YILDIRIM (2020). "Women, Rails, and Telegraphs: An Empirical Study of Information Diffusion". *Unpublished*.

GELLER, J., A. R. TOFTNESS, P. I. ARMSTRONG, S. K. CARPENTER, C. L. MANZ, C. R. COFFMAN, and M. H. LAMM (2018). "Study Strategies and Beliefs about Learning as a Function of Academic Achievement and Achievement Goals". *Memory*, 26 (5), 683–690.

GERHARD, U. (1990). *Unerhört: Die Geschichte der deutschen Frauenbewegung*. Hamburg: Rowohlt.

GESIS – LEIBNIZ-INSTITUT FÜR SOZIALWISSENSCHAFTEN (2019). *ALLBUS/GGSS 2018 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2018)*. GESIS Data Archive, Cologne. ZA5270 Data File. Version 2.0.0.

GLAESER, E. L., G. A. M. PONZETTO, and A. SHLEIFER (2007). "Why Does Democracy Need Education?" *Journal of Economic Growth*, 12, 77–99.

GOETTE, L. and E. TRIPODI (2020). "Social Influence in Prosocial Behavior: Evidence from a Large-Scale Experiment". *Journal of the European Economic Association*, 19 (4), 2373–2398.

GOLDIN, C. (1990). *Explaining the Gender Gap*. New York: Oxford University Press.

— (2006). "The Quiet Revolution that Transformed Women's Employment, Education, and Family". *American Economic Review*, 96 (2), 1–21.

GOLDIN, C. and L. F. KATZ (2003). *The "Virtues" of the Past: Education in the First Hundred Years of the New Republic*. Working Paper 9958. National Bureau of Economic Research.

GOLSTEYN, B. H. H., H. GRÖNQVIST, and L. LINDAHL (2014). "Adolescent Time Preferences Predict Lifetime Outcomes". *The Economic Journal*, 124 (580), F739–F761.

GOLSTEYN, B. H. H., A. NON, and U. ZÖLITZ (2021). "The Impact of Peer Personality on Academic Achievement". *Journal of Political Economy*, 129 (4), 1052–1099.

GOODMAN-BACON, A. (forthcoming). "Difference-in-Differences with Variation in Treatment Timing". *Journal of Econometrics*.

GURYAN, J., K. KROFT, and M. J. NOTOWIDIGDO (2009). "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments". *American Economic Journal: Applied Economics*, 1 (4), 34–68.

GÜTH, W., M. V. LEVATI, M. SUTTER, and E. VAN DER HEIJDEN (2007). "Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments". *Journal of Public Economics*, 91 (5), 1023–1042.

HAALAND, I., C. ROTH, and J. WOHLFART (forthcoming). "Designing Information Provision Experiments". *Journal of Economic Literature*.

HAIGNER, S. D. and F. WAKOLBINGER (2010). "To Lead or Not to Lead". *Economics Letters*, 108 (1), 93–95.

HANUSHEK, E. A., L. KINNE, P. LERGETPORER, and L. WOESSMANN (2021). "Patience, Risk-Taking, and Human Capital Investment across Countries". *Unpublished*.

HARRIS, J. (2009). *The Nurture Assumption: Why Children Turn Out the Way They Do*. 2nd ed. New York: Free Press.

HAUCH, G. (2019). "Zum Verhältnis von Revolution und Geschlecht im langen 19. Jahrhundert". *Aus Politik und Zeitgeschichte*, 69 (8), 32–38.

HERMALIN, B. E. (1998). "Toward an Economic Theory of Leadership: Leading by Example". *The American Economic Review*, 88 (5), 1188–1206.

HILL, H. C. (2017). "The Coleman Report, 50 Years On: What Do We Know about the Role of Schools in Academic Inequality?" *The ANNALS of the American Academy of Political and Social Science*, 674 (1), 9–26.

HOXBY, C. M. (2000). *Peer Effects in the Classroom: Learning from Gender and Race Variation*. Working Paper 7867. National Bureau of Economic Research.

HOXBY, C. M. and G. WEINGARTH (2005). "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects". *Unpublished*.

HSIANG, S. M., M. BURKE, and E. MIGUEL (2013). "Quantifying the Influence of Climate on Human Conflict". *Science*, 341 (6151).

KAISERLICHES STATISTISCHES AMT (1909). *Statistik der Frauenorganisationen im Deutschen Reiche*. Berlin: Heymanns Verlag.

KARING, A. (2021). "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone". *Unpublished*.

KARLAN, D. and M. A. MCCONNELL (2014). "Hey Look at Me: The Effect of Giving Circles on Giving". *Journal of Economic Behavior & Organization*, 106, 402–412.

KELLY, M. (2020). "Understanding Persistence". *Unpublished*.

KIMBALL, M. S., C. R. SAHM, and M. D. SHAPIRO (2008). "Imputing Risk Tolerance From Survey Responses". *Journal of the American Statistical Association*, 103 (483), 1028–1038.

KÖNIGLICHE ACADEMIE DER WISSENSCHAFTEN, HISTORISCHE COMMISSION (1875). *Allgemeine Deutsche Biographie*. Vol. 1. Leipzig: Duncker & Humblot.

KOSSE, F., T. DECKERS, P. PINGER, H. SCHILDBERG-HÖRISCH, and A. FALK (2020). "The Formation of Prosociality: Causal Evidence on the Role of Social Environment". *Journal of Political Economy*, 128 (2), 434–467.

LAVY, V. and A. SCHLOSSER (2011). "Mechanisms and Impacts of Gender Peer Effects at School". *American Economic Journal: Applied Economics*, 3 (2), 1–33.

LEESON, P. T. and J. W. RUSS (2017). "Witch Trials". *The Economic Journal*, 128 (613), 2066–2105.

LEWEJOHANN, S., ed. (2014). *Köln in unheiligen Zeiten. Die Stadt im Dreißigjährigen Krieg*. Köln: Böhlau Verlag.

MARKOFF, J. (2015). *Waves of Democracy: Social Movements and Political Change*. 2nd ed. Milton Park: Routledge.

MARTIN, A. J. (2009). "Motivation and Engagement Across the Academic Life Span: A Developmental Construct Validity Study of Elementary School, High School, and University/College Students". *Educational and Psychological Measurement*, 69 (5), 794–824.

MCEVEDY, C. and R. JONES (1978). *Atlas of World Population History*. London: A. Lane.

MCINTYRE, S. H. and J. M. MUNSON (2008). "Exploring Cramming: Student Behaviors, Beliefs, and Learning Retention in the Principles of Marketing Course". *Journal of Marketing Education*, 30 (3), 226–243.

MELANDER, E. (2020). "Transportation Technology, Individual Mobility and Social Mobilisation". *Unpublished*.

MISCHEL, W., Y. SHODA, and M. RODRIGUEZ (1989). "Delay of Gratification in Children". *Science*, 244 (4907), 933–938.

MOKYR, J., C. VICKERS, and N. L. ZIEBARTH (2015). "The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?" *Journal of Economic Perspectives*, 29 (3), 31–50.

MORRIS, A. D. and S. STAGGENBORG (2004). "Leadership in Social Movements". *The Blackwell Companion to Social Movements*. Ed. by D. A. SNOW, S. A. SOULE, and H. KRIESI. Boulder, Colorado: Blackwell Publishing, 171–197.

NAGELSCHMIDT, I. and J. LUDWIG (1996). *Louise Otto-Peters. Politische Denkerin und Wegbereiterin der deutschen Frauenbewegung*. Dresden: Sächsische Landeszentrale für politische Bildung.

NEKOEI, A. and F. SINN (2021). "Herstory: The Rise of Self-Made Women". *Unpublished*.

OOSTERBEEK, H. and R. VAN EWIJK (2014). "Gender Peer Effects in University: Evidence from a Randomized Experiment". *Economics of Education Review*, 38 (C), 51–63.

OSTER, E. (2004). "Witchcraft, Weather and Economic Growth in Renaissance Europe". *Journal of Economic Perspectives*, 18 (1), 215–228.

— (2019). "Unobservable Selection and Coefficient Stability: Theory and Evidence". *Journal of Business & Economic Statistics*, 37 (2), 187–204.

PEREZ-TRUGLIA, R. and G. CRUCES (2017). "Partisan Interactions: Evidence from a Field Experiment in the United States". *Journal of Political Economy*, 125 (4), 1208–1243.

POTTERS, J., M. SEFTON, and L. VESTERLUND (2007). "Leading-by-Example and Signaling in Voluntary Contribution Games: An Experimental Study". *Economic Theory*, 33 (1), 169–182.

RIEDL-VALDER, C. (2020). *Die Geschichte des Klosters St. Joseph*. https://www.hdbg.eu/kloster/index.php/detail/geschichte?id=KS0182. Last accessed: 2021-02-09.

RINGER, F. (1987). "On Segmentation in Modern European Educational Systems". *The Rise of the Modern Educational System: Structural Change and Social Reproduction, 1870-1920*. Ed. by D. K. MÜLLER, F. RINGER, and B. SIMON. Cambridge: Cambridge University Press.

SACERDOTE, B. I. (2011). "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" *Handbook of the Economics of Education*. Ed. by E. A. HANUSHEK, S. MACHIN, and L. WOESSMANN. Vol. 3. Amsterdam: Elsevier, 249–277.

— (2014). "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?" *Annual Review of Economics*, 6, 253–272.

SCHASER, A. (2000). *Helene Lange und Gertrud Bäumer: Eine politische Lebensgemeinschaft*. Köln: Böhlau.

SCHÖTZ, S. (2019). "Emanzipationsvorstellungen bei Louise Otto-Peters". *Aus Politik und Zeitgeschichte*, 69 (8), 4–10.

SCHRAUT, S. (2019). "Frauen und bürgerliche Frauenbewegung nach 1848". *Aus Politik und Zeitgeschichte*, 69 (8), 25–31.

SHAN, X. (2021). "Does Minority Status Drive Women Out of Male-Dominated Fields?" *Unpublished*.

SQUICCIARINI, M. P. (2020). "Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France". *American Economic Review*, 110 (11), 3454–91.

SQUICCIARINI, M. P. and N. VOIGTLÄNDER (2015). "Human Capital and Industrialization: Evidence from the Age of Enlightenment". *The Quarterly Journal of Economics*, 130 (4), 1825–1883.

STRACHEY, R. (1928). *The Cause: A Short History of the Women's Movement in Great Britain*. London: George Bell & Sons.

SUNDE, U., T. DOHMEN, B. ENKE, A. FALK, D. HUFFMAN, and G. MEYERHEIM (forthcoming). "Patience and Comparative Development". *The Review of Economic Studies*.

TABELLINI, G. and M. SERAFINELLI (2020). "Creativity over Time and Space". *Unpublished*.

TILLY, C., E. CASTAÑEDA, and L. WOOD (2020). *Social Movements 1768-2018*. 4th ed. Milton Park: Routledge.

VESTERLUND, L. (2003). "The Informational Value of Sequential Fundraising". *Journal of Public Economics*, 87 (3), 627–657.

VOIGTLÄNDER, N. and H.-J. VOTH (2012). "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany". *The Quarterly Journal of Economics*, 127 (3), 1339–1392.

WEIDMANN, B. and D. J. DEMING (forthcoming). “Team Players: How Social Skills Improve Team Performance”. *Econometrica*.

WHITMORE, D. (2005). “Resource and Peer Impacts on Girls’ Academic Achievement: Evidence from a Randomized Experiment”. *American Economic Review*, 95 (2), 199–203.

WINSTON, G. and D. ZIMMERMAN (2004). “Peer Effects in Higher Education”. *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*. Ed. by C. M. HOXBY. Chicago: University of Chicago Press, 395–424.

WOLFF, K. (2018). “Auch unsere Stimme Zählt! Der Kampf der Frauenbewegung um das Wahlrecht in Deutschland”. *Aus Politik und Zeitgeschichte*, 68 (42), 11–19.

ZHURAVSKAYA, E., M. PETROVA, and R. ENIKOLOPOV (2020). “Political Effects of the Internet and Social Media”. *Annual Review of Economics*, 12 (1), 415–438.

ZIMMERMAN, D. (2003). “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment”. *The Review of Economics and Statistics*, 85 (1), 9–23.

ZYMEK, B., G. NEGHABIAN, and L. ZIOB (2005). “Sozialgeschichte und Statistik des Mädchenschulwesens in den deutschen Staaten 1800-1945”. *Datenhandbuch zur deutschen Bildungsgeschichte*. Ed. by D. K. MÜLLER. Vol. 2, part 3: “Höhere und mittlere Schulen”. Göttingen: Vandenhoeck & Ruprecht.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

München, 14. September 2021

Leonhard Vollmer