UNIVERSITAT DE BARCELONA

Essays on Economics of Education

Candan Erdemli

PhD in Economics

Essays on Economics of Education

Candan Erdemli

UNIVERSITAT DE BARCELONA

2024

PhD in Economics | Candan Erdemli

PhD in Economics

Thesis title:

Essays on Economics of Education

PhD candidate:

Candan Erdemli

Advisors:

Judit Vall Castelló

Jenifer Ruiz-Valenzuela

Date:

January 2024

UNIVERSITAT DE BARCELONA

Acknowledgments

I believe that a strong support system and being surrounded by wonderful people are crucial for realizing and performing at one's best capacities. Fortunately, my PhD experience was enriched by such exceptional people, and I am grateful to each one who made these years a period of immense learning and joy.

First and foremost, I extend my deepest gratitude to my supervisors, Judit Vall Castelló and Jenifer Ruiz-Valenzuela, for their endless support, patience, and guidance. Looking back to the beginning of my PhD journey, I am amazed by how much I have learned from both of you. Thank you for always being there for me, even during your busiest times, showing me the positive side, teaching me how to find it, and helping me find solutions to the problems that initially seemed impossible to solve. Your continuous guidance, both in research and life, in exploring different perspectives has been invaluable. To both of you, my role models, I owe my growth as a researcher, learning that a great academic career can perfectly coexist with a fulfilling life.

My sincere thanks go to each member of the IEB for creating a true sense of a supportive academic community. Thank you all, not only for your valuable contributions to my research but also for making me feel at home throughout my PhD journey. Elisabet, thank you for looking after me at every step, even before the PhD, for always believing in me and making me believe in myself. Daniel, I deeply appreciate your continuous guidance and everything you taught me throughout our co-authorship.

I am extremely grateful to Lauri for his generous help and warm welcome during my research stay in Helsinki and for his enduring guidance and support.

Jordi Roca, thank you for making every administrative step as easy and smooth as possible. Susana and M. Àngels, your help in ensuring that everything was in order is greatly appreciated.

I extend my deep appreciation for the financial support provided throughout my PhD studies by FI-AGAUR (Agència de Gestió d'Ajuts Universitaris i de Recerca), IEB (Institut d'Economia de Barcelona), and Horizon Europe – European Research Executive Agency (REA) HORIZON-CL2-2022-DEMOCRACY-01; project number 101095106: "Education for Responsible Democratic Citizenship (DEMOCRAT)."

I greatly appreciate the support for my research by the INVEST Research Flagship Center, funded by the Academy of Finland Flagship Programme [grant number: 320162]. Also, I sincerely thank Smartick Company for generously providing me with one of the main datasets I used in my studies.

To my Barcelona family, the *fishies*, thank you, Clàudia, Ghizlen, Lili, Gerard, Marie, Abhinav, Fernanda, Adelia, Grace, Joël, Gökhan, Aly, Filippo, Mattia, Jeff, Dani, Marianna, Marta, Ale, Ceci, Darío, Mengwei, Julia, and Alessandro, for the lunches, Friday beers, laughs, and shared moments of joy and worry. You made these years shine for me, and it wouldn't have been the same without you.

Zeynep and Özlem, thank you for your valuable friendships over the years, the hours of phone calls, and the virtual coffees we shared. Once again, I've learned that distances do not matter.

To my growing family, thank you for believing in me under any circumstances. Mom and Dad, there is no way I can thank you enough for all you have done for me. Mom and Dad, the ones I met seven years ago, I appreciate your unwavering support throughout this process. Kuthan and Büşra, my siblings that I gained later in life, thank you for the joy you bring. My dear sister Canan, one of the greatest supports in my life, thank you for never letting me feel alone in this journey, even from another corner of the world, and for always lighting the way a few steps ahead. Thank you, Tolga, for always being there and generously sharing your experience. My little Derinkuş, my greatest teacher ever, I am grateful for everything you taught me in your short time and for all the inspiration you gave me.

Last but not least, I express my deepest gratitude to my life partner. Thank you, Kaan, for never ceasing to believe in me and always encouraging me to dream for the best even in the darkest moments. It has been a great pleasure to complete another step of our lives together.

Contents

1. Introduction

Education serves as a significant factor in shaping individuals' life trajectories. As widely acknowledged, it plays a critical role in determining individuals' future economic prospects. In Becker's words (1992), "The earnings of more educated people are almost always well above average." However, its impact extends far beyond income and employment. Education contributes to enhancing living standards (Wantchekon et al., 2015), influencing health and health-related behaviors (Conti et al., 2010; Johnson, 2010), and even bolstering cognitive abilities in later life (Banks and Mazzonna, 2012). Moving beyond individual outcomes, from a societal standpoint, education fosters increased political participation (Wantchekon et al., 2015) and yields intergenerational effects on fertility and infant health (Currie and Moretti, 2003).

Various factors can shape educational outcomes and human capital accumulation, and those have been extensively studied by economists. Examples include natural disasters, weather conditions, conflicts and wars, financial crises, and family health shocks (see for example Di Pietro (2018), Agamile and Lawson (2021), Weldeegzie (2017), Thomas et al. (2004), and Sun and Yao (2010)). As the world undergoes ongoing changes such as conflicts, epidemics, climate change, and increased technology adoption, it becomes crucial to keep investigating the potentially evolving determinants of educational outcomes and human capital accumulation.

This dissertation investigates how changes in society and the family shape educational outcomes. I analyze these changes from the perspective of gender inequalities and spillover effects of health shocks. In recent times, digitalization and the widespread adoption of online learning technologies have ushered in a transformative era in education. The second chapter of this dissertation delves into this digital revolution, examining gender differences in the online learning environment. By investigating how children perform and engage in online math learning, it aims to understand and address gender disparities that may emerge in this evolving educational landscape.

Beyond the digital realm, social movements have become significant actors for societal change, influencing individual perspectives and norms. The third chapter of this dissertation explores the aftermath of the Arab Spring, a monumental series of movements advocating for democratization and social justice. Focusing on the educational outcomes of second-generation immigrant women in Spain from the Middle East and North Africa (MENA) region, it aims to unravel the impact of these widespread movements on women's empowerment and educational achievements through shifts in beliefs and aspirations.

Transitioning from societal influences to the intimate sphere of the family, the last part of this dissertation delves into the profound effects of health shocks on educational outcomes. The fourth chapter explores the consequences of one of the most devastating experiences a family can endure — a child's death. By examining the impact on surviving siblings' educational trajectories, mental well-being, and parental outcomes, it seeks to shed light on the dynamics of familial adversity and its effects on education. Together, these chapters offer a comprehensive exploration of some factors shaping educational outcomes, contributing to a deeper understanding of the interplay between society, family, and education.

Having a detailed look at the empirical chapters of this dissertation, chapter 2, titled *"Gender Differences in Online Education"*, analyzes the gender gaps in children's online learning and how these gaps are correlated with the gender of the parent who mainly supervises the children.[1] In this chapter, we use data for Spain at the individual level from an online math learning platform which is used by children from over 100 countries, to document the gender differences in the context of online learning. We quantify the gender gaps in effort and relative performance outcomes and analyze whether the gaps differ by the gender of the parent who mainly supervises the children. Our main results point towards significant gender gaps in the relative performance outcomes in favor of boys, while the evidence for the effort gender gaps is only significant and economically meaningful when we compare the siblings of the opposite gender (controlling for parent fixed effects). The effort gaps are narrower or positive in favor of girls for children mainly supervised by their mothers. Further, we find that living in municipalities with more egalitarian gender norms is associated with narrower or positive gender gaps in effort outcomes while we do not find such differences in the relative performance outcomes.

This chapter makes several contributions to the literature on gender differences in education. First, it addresses the understudied gender gap in online learning outcomes, offering evidence from a contemporary online math learning platform. By focusing on the completion of daily exercises under minimal or no pressure, the study considers the potential differential impact of high-stakes testing conditions on results based on gender and provides insights into student effort and motivation. Second, examining families where online learning is primarily supervised by mothers versus fathers, we descriptively explore the differential consequences of such division of

[1]Paper coauthored with Judit Vall Castelló.

childcare responsibilities on the academic outcomes of boys and girls. Third, we contribute to the broader literature on gender norms and education. While previous research has associated smaller gender gaps with more equal gender social norms at the country level, we adapt this approach to the municipal level and examine whether the gender gap in online learning varies across municipalities with different gender norm profiles.

Chapter 3, titled *"Arab Spring and Women's Economic Empowerment"* explores the impact of social movements on women's economic empowerment through advanced educational outcomes resulting from potential shifts in beliefs and aspirations in the context of the Arab Spring.[2] In this chapter, we investigate the impact of the Arab Spring movements — a series of pro-democracy uprisings and protests in the Middle East and North Africa (MENA) — on the economic empowerment of young immigrant women in Spain with MENA-origin parents. Women in MENA countries gained significant visibility during the uprisings by actively participating in the protests and effectively using digital communication channels. First, we show that female MENA immigrants become more progressive in their beliefs and aspirations, compared to their non-MENA counterparts, following the Arab Spring. However, we do not find any effect on the beliefs and aspirations of male MENA immigrants. Next, focusing on second-generation immigrants who have been exposed to economic and political institutions in Spain throughout their lives (i.e. using the so-called epidemiological approach), we explore the impact of the Arab Spring movements on their education and labor market outcomes, isolated from the institutional changes, and driven by shifts in beliefs and aspirations. We find an increase in educational attainment and the probability of being in formal education for second-generation MENA females living in Spain after the Arab Spring, substantially closing the gaps between second-generation female immigrants from MENA and non-MENA countries. Also, we find a decrease in the probability of being NEET (not in education, employment, or training), and in the probability of being employed for MENA females, while we do not observe any significant change in the outcomes of MENA males.

This chapter makes contributions to three key areas in the economics literature. First, it extends the literature on the intersection of political and social protests with economic outcomes, by analyzing the cultural spillover effects of the Arab Spring on education and the labor market outcomes of individuals. Diverging from previous studies that predominantly focus on outcomes within a specific country where the protests occur, we explore the impact of a widespread movement on the children of MENA immigrants in Spain, providing evidence of the impacts of social movements transcending regional borders. Secondly, within the broader literature on the impact

[2] Paper coauthored with Daniel Montolio, Jenifer Ruiz-Valenzuela, and Judit Vall Castelló.

of the Arab Spring movements, our study stands out by employing an epidemiological approach, going beyond the national and regional borders. While previous research primarily concentrates on Egypt and utilizes regional variation in protest violence to identify impacts, we extend our analysis to second-generation immigrants in Spain from the MENA region which enables us to isolate our findings from changing economic and political institutions in the MENA region and instead focusing on shifts in beliefs, and aspirations. Lastly, we contribute to the literature on the impact of culture and norms on economic outcomes, building upon approaches used by Nollenberger et al. (2016) and Rodríguez-Planas and Nollenberger (2018). Our study investigates whether the Arab Spring induced a progressive shift in cultural values and aspirations among immigrants, influencing their education and labor market outcomes.

Chapter 4, titled *"Consequences of an Early Grave: Losing a Sibling During Childhood"*, analyzes the impact of a sibling loss on the educational outcomes of surviving children. The death of a child is devastating and life-altering for the entire family. Although a growing literature documents its negative impact on parental outcomes, very little is known about its consequences for the human capital accumulation of surviving siblings. This paper examines the impact of sibling loss during childhood on the surviving siblings' educational outcomes, using detailed register data from the entire population of Finland, spanning 24 birth cohorts. By focusing on unexpected child deaths caused by traffic accidents and exploiting the timing of sibling loss relative to the time of 9th-grade GPA measurement, I find that losing a sibling 2 years before the 9th grade has a negative impact of 19% of a standard deviation on the 9th-grade GPA. The effect is more pronounced and prevalent across different ages at the time of sibling loss for children with a lower socioeconomic background. Findings also suggest a 12-14 percentage points decrease in the probability of general track choice in the upper-secondary school following a sibling loss. Examining potential mechanisms, I find significant increases in the probability of antidepressant prescriptions for the surviving children and their parents. Moreover, a child loss increases the probability of taking sick leave and decreases the probability of employment for mothers, potentially suggesting a shift in the time allocation and parental time investment, though the quality of time remains unclear.

This chapter contributes to the extensive literature on the spillover effects of health shocks, focusing specifically on the impact of child death on the surviving siblings. While numerous studies have explored the negative effects of children's health shocks on parental outcomes, there is a gap in the literature in understanding how such health shocks affect other children in the same family. This chapter addresses this gap by providing evidence on the effects of losing a sibling during childhood on human capital development. Overcoming the identification challenge posed by the non-random distribution of sibling loss across the population, I analyze the impact of an unexpected loss, diverging from previous studies that do not distinguish the cause of death. In this chapter, I provide the first evidence on the consequences of sibling loss during childhood from the entire population of a country. By utilizing Finnish administrative records, this study not only enhances statistical power but also allows the identification of plausibly exogenous deaths with minimal anticipation effects, specifically those caused by traffic accidents. The unique advantage of linking educational records to medical outcomes and parental labor market outcomes enables a comprehensive investigation into previously unexplored mechanisms, namely the mental well-being of surviving siblings and parents, as well as the parental labor market outcomes. This exploration contributes valuable insights for policy considerations, taking into account the severe social and economic implications of such adverse life experiences.

Finally, chapter 5 concludes the dissertation, highlighting the main results, policy implications, and potential future research.

2. Gender Differences in Online Education[1]

2.1. Introduction

During the last few years, the use of online learning tools in education has been on the rise in most developed countries. This trend was further accelerated by the Covid-19 outbreak and the subsequent school closures in 2020. Given the increasing prevalence of these tools and their likely integration into the mainstream education system, it is essential to understand the role played by parents and the efforts put forth by children. On the other hand, previous economics literature has documented gender differences in parental investment, not only in developing countries but also in the context of developed countries. In the United States, boys receive more paternal time than girls (Lundberg et al., 2007; Price, 2008). In Canada, the United Kingdom, and the United States, parental time investments in teaching activities, such as reading, tend to favor girls, while fathers invest significantly more time in boys (Baker and Milligan, 2016). Differential parental investment for boys and girls might vary between mothers and fathers as well. Mammen (2011) finds that in the US, fathers allocate more time to their children if they have at least one boy, whereas mothers' total time investment is the same regardless of the gender composition of their children. Furthermore, not only the time invested may be different according to the gender match between parents and their children but also other elements affecting the educational outcomes may be different depending on this gender match. If similar parental gender bias and differences are present when using online learning tools, it may lead to a future gender gap in educational achievement and labor market outcomes.

This paper analyzes the gender gaps in online learning and how these gaps are correlated with the gender of the main supervisor. In particular, we quantify the effort and performance gaps between girls and boys when using an online learning platform in Spain and analyze whether the direction and the magnitude of these

[1]This paper has been resubmitted for revisions of the R&R process to the Journal of the Spanish Economic Association (SERIEs).

gaps differ depending on the gender of the parent who mainly supervises the child. Furthermore, we analyze whether the results are heterogeneous by the gender of the eldest sibling using the platform (to test for any role model effect), and the gender norms in the municipality of residence.

We find evidence of significant gender gaps in the relative performance outcomes in favor of boys while the evidence for the effort gender gaps is only significant and economically meaningful when we compare the siblings of the opposite gender. However, we find the effort gaps are narrower - even positive in favor of girls for persistency outcome which indicates whether at least one session is completed in each month - when the main supervisor is the mother. For performance outcomes, on the other hand, we do not observe differences in the gender gap depending on the gender of the main supervisor. We conduct the same analysis within a subsample of siblings of the opposite gender, including parent fixed effects, and find very similar results. We find heterogeneity by the gender of the eldest sibling in both fathers' and mothers' samples. Specifically, while the gender gap in the extra time devoted to solving problems and the number of sessions completed per month come from families where the eldest is a girl, the gender gap in the ratio of correctly solved problems is more pronounced in families where the eldest is a boy. Additionally, we find that living in a municipality with more egalitarian gender norms is associated with positive gender gaps in favor of girls in effort outcomes. However, we do not observe such a difference in gender gaps in relative performance outcomes. The gender gap in delayed completion which represents the extra time devoted to completing the sessions decreases with age, while the gender gap in other outcomes does not significantly change across age groups. Lastly, we do not find significant differences in the gender gaps in outcomes by income levels of the municipality of residence.

This study contributes to the literature on gender differences in education in several ways. First, while previous research has documented gender gaps in traditional face-to-face education, the gender gap in online learning outcomes has been understudied. Our paper contributes to the literature by providing evidence from a contemporary online learning tool.[2] This allows us to document gender differences in completing the daily exercises, rather than the test outcomes, which is important in several ways. Balart et al. (2018) show that performance in cognitive tests, such as PISA,

[2]Online learning has become the interest of many researchers since the start of the Covid-19 outbreak, subsequent school closures, and the rise of online education. Among others, Chetty et al. (2020) find that children experienced a reduction in learning on a math learning platform used in US schools and Ikeda and Yamaguchi (2021) find that school closures during Covid-19 increased students' study time using an online learning service in Japan. However, Chetty et al. (2020) use school-level data, and they do not examine results by gender. On the other hand, Ikeda and Yamaguchi (2021) do not find any heterogenous effect by gender.

is influenced by non-cognitive skills (also named as personality traits, soft skills, or character skills), by using a decomposition in PISA test scores. Also, Anaya et al. (2022) show that the difficulty level at the beginning of tests may influence the success of the later questions. There is also some evidence in the psychology literature suggesting that under low-stakes testing conditions, some individuals try harder than others (Duckworth et al., 2011). This issue becomes more important if the non-cognitive effects differ by gender. An example of this is the study by Montolio and Taberner (2021), where they find that male university students outperform their female counterparts under high pressure. Considering these, it is essential to focus on outcomes that accumulate within a month as a result of daily exercises that are completed under no or little pressure. In addition, using data from an online learning platform helps us document gender gaps in student effort and motivation. Attending the test session, time spent on each test item, and self-reported effort have been used as primary measures of motivation in the previous literature documenting gender gaps in the motivation of students, where females are generally found to exert more effort compared to males (DeMars et al., 2013).[3] We add to this literature by analyzing the gender gap in effort when using an online learning platform, where we measure effort by the indicators created based on the number of completed online sessions in a month.

Second, understanding the source of gender discrimination or bias within the family is crucial for designing relevant and effective policies. Several studies show that mothers devote more time to childcare responsibilities than fathers, and this gap increased after the Covid-19 pandemic (Golin, 2021; Andrew et al., 2020; Del Boca et al., 2020). However, the consequences of this division remain understudied. To the best of our knowledge, there is no evidence in the literature on how this division of childcare or educational support differently affects the effort and academic performance of boys and girls. To shed light on this part of the literature, we compare the gender gap in educational outcomes across two types of families: the ones where online learning is mainly supervised by the mother and those by the father. Furthermore, our study differs from the previous literature by identifying the gender gap among siblings, while most of the existing studies focus on the gap among peers. This enables us to estimate the gender differences in learning outcomes between a boy and a girl raised in the same family.

Third, we contribute to the extensive literature on the relationship between the gender gap in education and gender norms. Several studies find that countries with more equal gender social norms tend to have smaller gender gaps in achievement or educational preferences (e.g., González de San Román and De La Rica, 2012;

[3] See DeMars et al. (2013) for a comprehensive literature review.

Nollenberger et al., 2016; Rodriguez-Planas and Nollenberger, 2018; Gevrek et al., 2020). Although our data is specific to one country, we adapt this approach in our context to examine whether the gender gap in academic effort and performance is larger or smaller in municipalities with more egalitarian gender norms, which we proxy using the relative shares of females' and males' employment rates and contributions to household chores.

2.2. Data and Methodology

2.2.1. Smartick

We use individual-level anonymous data from Smartick, an online learning platform that is used in over 100 countries, including but not limited to Spain, the United Kingdom, the United States, Mexico, Colombia, Peru, Brazil, and South Africa. The content is offered in different dialects of Spanish, English, and Portuguese, according to the user's choice. In our study, we use information on members residing in Spain since they constitute the largest proportion of Smartick users.

Smartick is one of the most widely used online learning platforms in Spain. It is a math learning tool that offers a 15-minute online math session every day.[4] Students engage in 4 main areas: mental calculation, reasoning, logic, and programming. The platform is designed for children aged 4 to 14 and incorporates artificial intelligence to create personalized sessions tailored to the student's knowledge and abilities. These sessions are interactive and guided, with each exercise corrected immediately. If an answer is incorrect, it explains how it should have been done correctly. Therefore, parents' assistance is not required and Smartick encourages children to work on the sessions independently.

Since Smartick is mainly parent-based, either the father or the mother registers the child to the platform and he/she keeps track of the child's progress by checking the daily emails sent by Smartick regarding the child's attendance and performance in the sessions. In addition to reviewing the daily emails, parents can (and are encouraged by Smartick to) log in to the platform using their account to access a detailed performance analysis and the child's progress. This individual relationship established between one of the parents and Smartick automatically assigns responsibility for the child's online learning process to that parent.[5]

[4]Although Smartick also offers reading exercises, we focus solely on the outcomes of math exercises due to data restrictions.

[5]Even though one parent registers the child and receives the emails about the learning process, we only use this relationship as a proxy of "being the main supervisor" since we cannot observe whether the actual supervisor is the one who registers on the platform.

After a 7-day free trial period, parents are offered and choose from three types of paid contracts: one-month, three-month, and one-year contracts. We observe registration dates for free trial and paid contract, as well as monthly averages of the outcomes for each child starting from their contract date. We focus only on the duration of the first contract.[6] Our dataset includes members registered for a paid contract between January 2019 and July 2021. We restrict our sample to the members who registered for a free trial period starting from January 2019 to observe their first contract outcomes, as well as those who registered for a paid contract before 20 June 2021, to ensure that we observe the outcomes of at least one complete month.

Our dataset includes children's age, gender, and the presence of a health condition.[7] We also have information about the municipality of residence.[8] Additionally, we observe the anonymous parent ID, which allows us to identify the siblings in the sample, as well as the gender of the parent who registered the child on the platform.

Since Smartick is a paid platform, the characteristics of its users are likely to differ from those of the overall Spanish population, raising concerns about the external validity of our results. Therefore, following Chetty et al. (2020), we present the demographic characteristics of Smartick users in our sample in Table 2.1 to show the extent of the selection. Since we do not have information on the characteristics of parents, we proxy the demographic characteristics using the income levels of the municipalities where they reside. Specifically, we compare the income quartiles of the Smartick municipalities that we define as those where at least 5 (and 1) Smartick members live, to the income quartiles of all Spanish municipalities. As we expect, the results reveal a selected sample in terms of income distribution.

Moreover, the public school participation rate of the Smartick users[9] is 49.4%, while the public school participation rate in the Spanish population at the primary school level is 68%. Overall, Table 2.1 shows the degree of selection in terms of income level and public school participation in our sample compared to the Spanish population. In this selected sample, we anticipate less gender bias since these parents are likely to be more educated and more aware of gender equality concerns compared

[6]We do not include the outcomes of the months of later contracts because we cannot clearly identify whether a missing value corresponds to a month without a paid contract or a month with no completed sessions.

[7]The health conditions include high intellectual capacity, dyscalculia, dyslexia, intellectual disability, hearing disability, cerebral palsy, maturational delay, Down syndrome, attention deficit hyperactivity disorder (ADHD), autism, Asperger's syndrome, and pervasive developmental disorder.

[8]The municipality of residence is detected by Smartick as the location where they first register on the platform. In some cases, we observe different locations for children registered by the same parent, or the location information is missing for a child but identified for his/her sibling. In these cases, we assign all children for a given parent the first identified location.

[9]According to a survey conducted by Smartick in September 2021 on a representative sample of Smartick members, consisting of 2894 responders.

to the average Spanish population. Therefore, we believe that our results represent a lower bound of gender bias in this context.

The outcomes of interest in our study include the total number of sessions completed in a month, the average ratio of time spent to complete the exercises to the expected time, and the average ratio of correctly solved problems. Using the total number of sessions completed in each month, we also create two binary outcome variables representing whether the child completes at least one session and at least twenty sessions in each month, respectively. We categorize these outcome variables into two groups: children's effort and relative performance. The measures of effort include persistency, completion, and sessions, while the measures of relative performance include delayed completion and accuracy. *Persistency* and *completion* are binary variables that take the value 1 if at least one session and twenty sessions are completed in each month, respectively. *Sessions* represents the average number of sessions completed per month. *Delayed completion* is the average ratio of time spent on problems to the expected time. Lastly, *accuracy* is the average ratio of correctly solved problems.

It is important to note that we only observe the accuracy and the delayed completion variables for the months with at least one completed session. We construct these outcomes as averages over the fully observed months of the first contract.[10] For consistency, we exclude a monthly outcome if we do not observe the full month in our dataset. For example, if the first contract is for 3 months but our data only covers the period of the first month and a half, we only consider the outcomes of the first month.

The first panel in Table A.1 shows the descriptive statistics in the full sample, which includes 28,236 children residing in Spain. Of these, 52% of them are girls, creating a balanced sample in terms of child gender. However, there is an imbalance in parent gender, with 66% of the parents being mothers. This ratio aligns with the difference in average minutes spent per day with children in teaching-related activities by mothers and fathers. According to the nationally representative Spain Time Use Survey 2009-2010, mothers and fathers spend an average of 5.26 and 2.68 minutes per day, respectively, on teaching-related activities with their children, indicating that 66% of the total teaching time is contributed by mothers.[11]

The registered children in our dataset range in age from 4 to 16, with an average age of 8.54. Approximately 6% of the children in our sample have a health condition. Since we do not have information on household income levels, we use the average

[10] We conducted a separate analysis focusing on the outcomes of the first month, regardless of the contract type. The results are very similar to the baseline analysis and are available upon request.

[11] A mother (father) is defined as a woman (man) who has a daughter or a son living in the same household, but not a grandparent.

Table 2.1.: Demographic characteristics of Smartick users

	Smartick municipalities		Spanish population
	≥ 5 members	≥ 1 member	
average HH net income			
25th percentile	27,442	25,248	21,864
Median	31,142	28,862	25,535
75th percentile	34,978	33,519	30,411
public school participation			
	49.40%		68%
number of people			
number of Smartick members	26,648	28,236	-
2019 population	31,952,187	38,733,271	47,026,208
number of municipalities			
	484	1,302	8,151

Notes: This table shows the demographic characteristics of Smartick users, following Chetty et al. (2020). Smartick municipalities are defined as municipalities where at least 1 or 5 Smartick members reside. Income and population information is based on 2018 statistics from the Spanish Statistical Office (INE).

household net income in their municipality of residence as a proxy for income, with a mean value of 37,628 Euros. A total of 70% of our sample registered to the platform after the Covid-19 outbreak in Spain (after March 9, 2020), and 48% registered in the first 3 months of the outbreak when the schools were closed, and people had to stay at home. On average, each child has a total of 4.28 contracts, and 52% of the first contracts have a duration of 1 month. An average Smartick user completes 21.47 sessions in a month. Furthermore, 95% of the users complete at least one session, and 56% complete at least 20 sessions each month during their first contract period. Children, on average, spend 1.12 times the expected time to complete the sessions and answer 84% of the exercises correctly.

The second panel in Table A.1 presents descriptive statistics for the sibling subsample, which consists of 7,533 children who have at least one sibling of the opposite gender registered on the platform. We create this subsample to analyze gender differences among siblings, allowing us to control for family or household fixed effects while documenting the gender gaps. The distributions of variables in this sample closely resemble those in the full sample. In the siblings subsample, the average number of children with the same parent is inherently higher, and we observe that 50% of the observations in this subsample are from households where the eldest child is a boy.

Table A.2 shows the mean differences of the variables by child gender. In the full sample, boys are on average 0.23 years younger, 4% less likely to be registered by their mothers, and 3% more likely to have a health condition compared to girls. In the siblings subsample, similarly, boys are on average 0.25 years younger and 1% more likely to have a health condition than girls. Boys in the siblings subsample complete on average 0.36 more sessions than girls per month. In both samples boys and girls come from similar backgrounds in terms of income level and gender-age composition. For the relative performance variables, we observe statistically significant gender gaps in favor of boys, without controlling for any characteristics, however the magnitudes are small.

Table A.3 shows the mean differences by the gender of the parent. In line with the information in the previous table, fathers are 4% less likely to register girls on the platform compared to the mothers in the full sample. On average, children registered by their fathers are less likely to have a health condition and the number of children in their family is slightly lower compared to those registered by their mothers, in both the full sample and siblings subsample. The probability of the eldest child being a boy is 2 pp lower in the families where the father registers children in the siblings subsample. In the full sample, children registered by their fathers complete 0.23 fewer sessions per month and they have on average 2 pp lower completion rate than those registered by their mothers. However, we do not observe significant differences in the relative performance outcomes.

Figure A.1 presents the evolution of the total number of new registrations per month. As indicated by the red line, the number of registrations increased disproportionately in the month when the Covid-19 outbreak started. However, starting from June 2020, the numbers returned to previous levels. Figure A.2 and Figure A.3 show the total number of registrations per month, categorized by the gender of the child and by the gender of the parent, respectively. In terms of percentages, Figure A.4 does not show a clear pattern of gender pairs over the months. As indicated in Figure A.5, the majority of members are primary school students, representing a constant trend. However, there is an increase in registrations by pre-school children along with a decrease in registrations by secondary school students over time. Figure A.6 shows the percentage of new registrations by income categories, defined based on the quartiles of average income in the municipality of residence. The majority of members reside in high-income municipalities, and this trend remains stable over time.

These figures suggest that there is no clear pattern of change in the gender composition of children and parents, as well as in the age and income distribution over time, including the period after the Covid-19 outbreak. In other words, the observable characteristics of our sample do not exhibit significant changes over time, except for a slight shift in the age categories.[12]

2.2.2. Identifying the Gender Gaps

In order to identify the gender gap in effort and performance outcomes conditional on the main characteristics of children, we estimate the following linear regression:

$$Y_{ipm} = \beta_0 + \beta_1 G_{ipm} + \beta_2 X_{ipm} + \gamma_p + \theta_m + \varepsilon_{ipm} \tag{2.1}$$

where G_{ipm} is the dummy variable for the girl, X_{ipm} includes the child characteristics (age, presence of a health condition, total number of contracts, and type of the first contract), γ_p captures province fixed effects and θ_m captures contract year-month fixed effects. Standard errors are clustered at the province level. Our coefficient of interest, β_1, captures the gender gap in the related outcome. We estimate this equation in the mothers' and fathers' samples separately, as well as in the full sample by controlling for the gender of the parent. We normalize the continuous outcome variables (sessions, delayed completion, and accuracy) before the analysis so that we can compare the estimated coefficients properly.[13]

Next, in order to control for the household fixed effects and identify the gender gap among siblings, we estimate the same regression on the siblings subsample. As distinct from the previous regression, we include parent fixed effects instead of province fixed effects and we do not control for the number of total contracts, type of the first contract, and contract year-month since they are likely to be the same across children for a given parent. Formally, in Equation 2, Z_{ipm} includes the child characteristics (age and presence of a health condition), α_p captures the parent fixed effects, and the rest of the variables are the same as in Equation 1.

$$Y_{ipm} = \beta_0 + \beta_1 G_{ipm} + \beta_2 Z_{ipm} + \alpha_p + \varepsilon_{ipm} \tag{2.2}$$

For the first part of the heterogeneity analysis, we estimate the regression in which we interact the girl dummy (G_{ipm}) with the EldestBoy dummy. The EldestBoy dummy takes the value of 1 if the eldest sibling registered on the platform is a boy (or the siblings of the opposite gender are of the same age), and 0 otherwise. The caveat of this approach is that we can only observe the children who are registered on the platform. We cannot observe an elder brother or sister who is not registered

[12] We are aware that the characteristics of new registrations in June 2021 seem different than in previous months. One potential reason is that we only include new registrations until June 21 since we restrict our sample to the fully-observed months (see section 2.1 for a detailed explanation), and related to this, we only observe 12 new registrations in this month.

[13] In our analysis sample, we normalize the continuous outcome variables within the full sample and siblings sample separately before splitting them as mothers' and fathers' subsamples.

on the platform.

Then, we aim to understand whether the gender gap in educational outcomes is more pronounced in municipalities where there is less gender equality. Therefore, as the second part of our heterogeneity analysis, we create two variables to proxy for gender equality at the municipality level, using information from the 2011 Spanish Census microdata obtained from INE (Instituto Nacional de Estadistica).[14] First, we define the employment gender ratio as the ratio of the female employment rate to the male employment rate in each municipality.[15] Then, we create a dummy variable (E_m) which takes the value one if the employment gender ratio of a given municipality is equal to or above the median, and zero otherwise. Second, we define chores gender ratio as the ratio of the share of females reporting that they take care of most of the household chores to that of males in a given municipality. Then, we assign value one to variable C_m if the chores gender ratio in municipality m is below the median, and zero otherwise. We define both E_m and C_m variables in a way that value 1 reflects more egalitarian gender norms in municipality m. Then we estimate the previous equation by interacting the G_{ipm} variable with E_m, and C_m in separate regressions.[16]

2.3. Results

2.3.1. Main Results

Before presenting our results for the mothers' and fathers' samples separately, we document the gender gaps in the full sample, conditional on the gender of the supervising parent, other characteristics of the child, province, and contract year-month fixed effects. We also analyze the gender gaps in the siblings' subsample, conditional on the characteristics of the child and parent fixed effects.

As shown in the first panel of Table A.4, we do not observe gender differences

[14]The 2011 Spanish Census is the most recent census available for our analysis. In the microdata, municipalities with a population of less than 20,000 are not identified for confidentiality reasons. We match information from 394 municipalities identified in the census to 24,896 observations (88%) in our dataset out of 28,236 observations in total.

[15]The female (male) employment rate in a municipality is defined as the ratio of the number of employed females (males) to the number of all females (males) living in the municipality and aged from 18 to 65 years old.

[16]We also perform heterogeneity analysis where we use employment gender ratio, labor force participation gender ratio, and labor force participation ratio for the women with children under 16 years old at the province level instead of municipality level, by using data from 2019 employment statistics provided by INE. However, we do not find statistically significant differences between the provinces with more and less egalitarian gender norms. We believe that this is due to the lack of variation in the indicators of gender norms across provinces. Results are not shown in this paper but are available upon request.

in the effort outcomes in the full sample, but there are significant differences in relative performance outcomes including accuracy (8% of a standard deviation) and delayed completion (15% of a standard deviation). The first panel shows that children supervised by their mothers are less likely to be persistent, more likely to complete at least 20 sessions each month and complete more sessions per month on average, while they solve fewer problems correctly. In the second panel of Table A.4, where we focus on the siblings of opposite gender and control for parent fixed effects, we observe that girls complete fewer sessions (4% of a standard deviation), solve fewer exercises correctly (5.4% of a standard deviation) and spend proportionally more time than expected in completing sessions (18% of a standard deviation) compared to their brothers.

Figure 2.1 shows estimated gender gaps in the full sample for the children registered by their mothers and fathers. The complete set of coefficient estimates is presented in Table A.5 for both the mothers' and fathers' samples in two panels.

Regarding the effort variables (persistency, completion, and sessions), the gender gap is consistently positive when children are supervised by their mothers but negative when they are supervised by their fathers, although not all the coefficients are precisely estimated. Girls are 0.7% more likely to be persistent than boys when supervised by their mothers, and they complete fewer sessions (3.2% of a standard deviation) than boys when supervised by their fathers. Table A.6 presents the results of full sample regressions where we include the interaction term of girl and father dummy variables. The third row indicates that the gender gap for children registered by their fathers is significantly different for persistency and sessions outcomes. The gender gaps in the relative performance variables are negative in both the mothers' and fathers' samples. Specifically, girls solve fewer exercises correctly (8% and 9% of a standard deviation with mothers and fathers, respectively) and they spend proportionally more time than expected in completing sessions compared to boys (16% and 13% of a standard deviation with mothers and fathers, respectively).

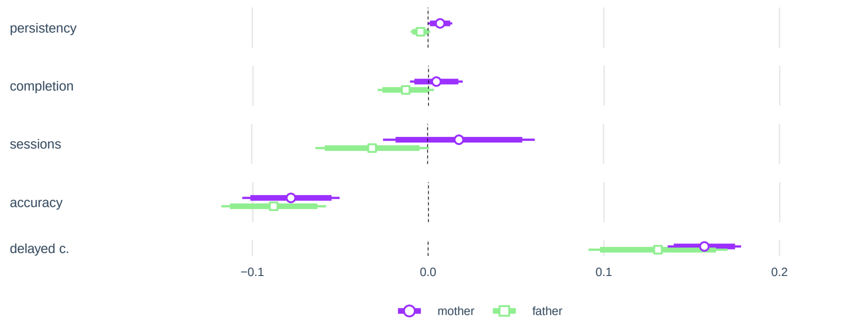
We repeat the same analysis in the siblings' subsample by controlling for parent fixed effects.[17] Figure 2.2 presents the estimated gender gaps among siblings, and Table A.7 shows the full regression results. Similar to the results in the full sample, we observe opposite directions of gender gaps in the persistency variable in the mothers' and fathers' samples, though neither of the estimates is significantly different from zero. While the gender gap in completion is very close to zero and imprecisely estimated in the mothers' sample, it is 3% of a standard deviation and significant at the 90% level in the fathers' sample. Girls complete 2.3% and 7.4% of a standard deviation fewer sessions than boys in the mothers' and fathers' samples,

[17]We exclude the number of contracts and type of the first contract variables when including the parent fixed effects, as they are the same for children with the same parent in most cases.

respectively.

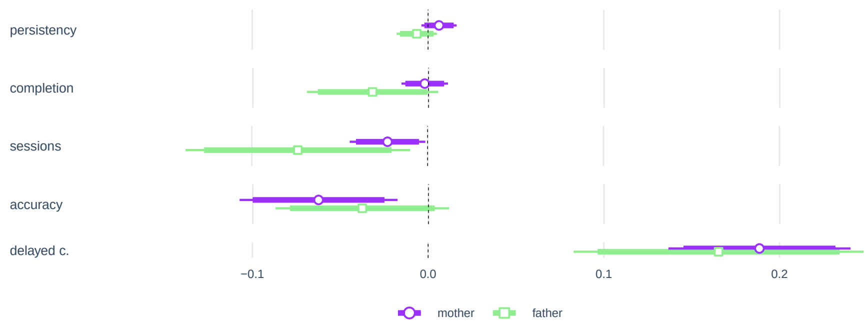
We observe significant gender gaps in the mothers' sample for both accuracy and delayed completion variables. However, in the fathers' sample, we only observe a significant gender gap in the delayed completion variable.

Figure 2.1.: Gender gap - Full sample



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers' and fathers' full samples separately, based on Equation 1. Control variables are the age of the child, presence of a health condition, total number of contracts, type of the first contract, province, and contract year-month fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

Figure 2.2.: Gender gap - Siblings' sample



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers' and fathers' siblings samples separately, based on Equation 2. Control variables are the age of the child, the presence of a health condition, and parent fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

2.3.2. Differences by Gender Composition of the Siblings and Gender Norms

First, we explore whether the gender gaps are more pronounced in families where the eldest sibling is a boy. Table 2.2 shows that in the families where the eldest sibling is a girl, girls complete fewer sessions (10% of a standard deviation), solve

Table 2.2.: Gender gap - Differences by gender of the eldest child

	persistency	completion	sessions	accuracy	delayed c.
Siblings Sample					
Girl	0.001	-0.021^{*}	-0.097^{***}	0.104^{***}	0.318^{***}
	(0.005)	(0.011)	(0.019)	(0.021)	(0.032)
	[0.86]	[0.05]	[0.00]	[0.00]	[0.00]
	[0.86]	*[0.07]*	*[0.00]*	*[0.00]*	*[0.00]*
EldestBoy	0.003	-0.013	-0.050	0.151^{***}	0.172^{***}
	(0.006)	(0.013)	(0.030)	(0.037)	(0.028)
	[0.62]	[0.32]	[0.10]	[0.00]	[0.00]
	[0.62]	*[0.40]*	*[0.17]*	*[0.00]*	*[0.00]*
GirlxEldestBoy	0.001	0.021	0.115^{***}	-0.275^{***}	-0.232^{***}
	(0.007)	(0.017)	(0.035)	(0.034)	(0.062)
	[0.86]	[0.22]	[0.00]	[0.00]	[0.00]
	[0.86]	*[0.27]*	*[0.00]*	*[0.00]*	*[0.00]*
R^2	0.087	0.180	0.120	0.102	0.108
Num. obs.	7533	7533	7533	7505	7502
N Clusters	50	50	50	50	50

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, health condition, number of contracts, type of the first contract.

more problems correctly (10% of a standard deviation), and are slower in completing sessions (32% of a standard deviation) than boys. In contrast, in families where the eldest sibling is a boy, girls complete more sessions (1.8% of a standard deviation), solve fewer problems correctly (17% of a standard deviation), and are slower in completing sessions (8.6% of a standard deviation) than boys.[18] Overall, these results suggest that the gender gap in accuracy is associated with the presence of an eldest brother using the platform, whereas the gap in sessions is associated with the presence of an elder sister. Although the delayed completion gap is consistently in favor of boys, we observe that it is narrower in the presence of an elder brother. This observation might be interpreted as girls taking more risks and being faster than their "role model" elder brothers.[19]

Next, we examine whether gender gaps in the outcomes are more pronounced in municipalities with lower gender equality. We interact the dummy variables constructed using employment gender ratio and chores gender ratio as proxies for

[18]Calculations by using the coefficient estimates from Table 2.2: 1.8% of a sd: -0.097 + 0.115 = 0.018; 17% of a sd: 0.104 - 0.275 = 0.170; 8.6% of a sd: 0.318 - 0.232 = 0.086.

[19]We acknowledge that, in principle, both elder brothers and younger brothers could potentially be seen as role models in our context. However, since we find narrower gender gaps in the presence of an elder brother, we interpret the results as a potential role model effect from the elder brothers.

gender equality in municipalities, with the girl dummy variable.[20] The first panel of Table 2.3 suggests that the gender gap in completion is not present in municipalities with more gender equality in employment. We observe similar patterns for the persistency and sessions outcomes, although the estimates are not precise. The second panel shows the differences in gender gaps by the "relative contribution to household chores" in the municipality of residence. Positive gender gaps in favor of girls in effort outcomes come from municipalities with higher gender equality in the relative contribution of females and males to the chores. However, we do not observe such differences in relative performance outcomes. Overall, these results suggest that municipalities with more egalitarian gender norms are associated with narrower or positive gender gaps in effort outcomes, while no such differences are found for relative performance outcomes.

2.3.3. Differences by Age and Income Levels

We then analyze whether gender gaps in effort and relative performance outcomes differ by the age of the children. To do so, we utilize both the continuous age variable and the categorical age variable that we define according to the school levels (pre-school (age 4-5), primary school (6-11), and secondary school (12-16)). We interact the age variable with the girl dummy variable in the full sample and estimate the regression conditional on the gender of the parent who supervises the child.

In Table A.8, Panel A shows that while the gender gap in the effort outcomes and accuracy variable does not change significantly by age, the gender gap in delayed completion decreases with age. In Table A.8, Panel B shows that the gender gap in completion outcomes primarily comes from pre-school kids and we observe a similar pattern in terms of the magnitudes for sessions outcome, although the estimations are not precise. While the accuracy gap comes from the primary school kids, the delayed completion gap is more pronounced for children in both pre-school and primary school. Overall, Table A.8 suggests that the gender gaps are stronger for younger girls, but most of them fade away with age.

Next, we create income categories within the Spanish population and among Smartick users based on the quartiles of average income in the municipality of residence. We then interact this categorical variable with the girl dummy variable to explore whether the sign and magnitude of the gender gaps differ depending on the income level. As shown in Table A.9 and Table A.10, we do not observe any significant differences in the gender gaps by income categories. We believe that the main reason for this is the high homogeneity of our sample in terms of income level, as the majority of the users come from high-income families.

[20] See Section 2 for definitions.

2.3.4. Robustness Checks

One concern in the context of a paid learning platform is the potential negative selection on ability in our sample. In other words, families might be more likely to register their children on this platform if their children are not performing well academically, and this selection might also vary by the gender of the child, which could potentially affect our results. To address this concern, we repeat the analysis in the "covid sample", which we define as the sample of children who registered on the platform between March 9 and May 31, 2020. This period corresponds to the first wave of the Covid-19 outbreak in Spain. During this period, parents were seeking educational tools to help their children stay on track because schools were closed unexpectedly, and distance education did not start immediately after the school closures. Therefore, we believe that potential selection on ability and gender in registrations plays a minimal role in this time period. This idea is supported by the increase in the total number of registrations and differences in the characteristics and outcome variables of the children registered during this period, as shown in Figure A.1 and Table A.11.

When we compare the characteristics and outcome variables of children in the covid sample to the rest of the sample (as shown in Table A.11), we find some notable differences. In this sample, the share of girls registered to the platform is 3% lower, while the share of mothers is 4% higher. The children in the covid sample are, on average, about half a year younger, less likely to have a health condition and live in municipalities with, on average, lower income levels. While in the full sample, the average number of children in the families is slightly higher, in the siblings sample it is slightly lower. In addition to the characteristics, we also observe differences in the distribution of outcome variables. Children in the covid sample complete, on average, 2.27 more sessions per month, are 1% more likely to be persistent and 12% more likely to complete at least 20 sessions per month compared to the others. They also spend less time on completing the sessions and solve slightly more exercises correctly. These differences suggest that if there is any selection on ability in our sample, this selection would be the smallest in this period. Furthermore, the income level of the municipalities where the children live is closer to the average Spanish population compared to the rest of the sample.

We estimate the gender gaps in the "covid sample" to examine whether our results are affected by a potential selection on ability and gender. As shown in Figure A.7 and Figure A.8, the magnitudes of estimated gender gaps in the covid sample are very similar to those in our baseline analysis. However, the estimates are not as precise as the baseline estimates in the siblings' subsample, which might be due to the lower number of observations. The magnitudes of the coefficient estimates are presented in Table A.12 and Table A.13.

Next, we divide the full sample into two groups: those who registered before and after the Covid-19 outbreak in Spain. In Table A.14, we present the regression results where we interact the girl dummy variable with the father dummy variable in these two subsamples.

We find that the gender gap in the persistency outcome and its negative association with the fathers' supervision come from the registrations made before the outbreak. Although the signs of the estimates are mostly in the same direction in the sample of registrations after the outbreak for the effort outcomes, they are not significantly different from zero. Regarding the relative performance outcomes, we observe that in the sample of registrations after the outbreak, the gender gap in the delayed completion outcome is slightly lower for the children with fathers.

We further examine whether the estimated gender gaps in the siblings' subsample are robust to including province fixed effects instead of parent fixed effects in the regressions, as shown in Table A.15. While the magnitudes of the estimates slightly change, they remain very similar to the baseline estimates and maintain the same statistical significant levels.

For the binary outcome variables (persistency and completion), we estimate logit regressions as an alternative to the linear probability model. Table A.16 presents the results for the full sample and siblings sample, respectively. The estimated gender gaps are consistent in direction with those estimated using linear probability models.

Next, we exclude children with a health condition from the sample to assess whether our results are affected by the presence of a health condition. Table A.17 shows the estimated gender gaps in the full sample of children without a health condition. The results closely resemble the baseline results reported in Table A.6.

We construct persistency and completion variables based on the number of sessions completed per month. We find that girls are 0.7% more likely to complete at least one session per month (persistency) than boys when the main supervisor is the mother. However, we do not observe gender gaps in the completion variable (see Table A.5).

To assess the stability of our results to different cut-offs of the number of sessions completed per month, we construct additional outcome variables with various cut-offs (1, 6, 11, ..., 26), where the outcome of the first cut-off is the same as the persistency variable. The results are presented in Figure A.9, where we estimate Equation 1 for different cut-off outcomes. In the mothers' sample, there are no significant gender gaps for different cut-offs, except for the cut-off of 1. However, in the fathers' sample, we observe that girls are 1.8% (2%) less likely than boys to complete at least 11 sessions (16 sessions) per month.

In our main specification, we use the continuous age variable to control for the differences in the gender gap across different ages. Additionally, in Figure A.10, we present our main results by including the age fixed effects instead of the continuous age variable. The significance levels and the magnitudes of the coefficient estimates in this model are very similar to those of our main specification.

To further control for differences by age, we repeat our analysis by normalizing all the outcome variables within each age. This is particularly important for the siblings' sample, where we compare siblings at different ages, and the difficulty level of exercises for different ages is likely to be different. The results presented in Figure A.11 and Figure A.12 are very similar to our baseline results. If anything, we observe larger magnitudes of coefficient estimates for the persistency variable, with the signs remaining the same as in the baseline.

2.3.5. Potential Mechanisms and Discussion

Our results suggest gender gaps in online math learning in favor of boys, especially in performance outcomes. However, we find no gender gap in favor of boys in effort outcomes when the main supervisor is the mother. While data limitations prevent us from analyzing the mechanisms behind these results, in this section we discuss the potential explanations for our findings.

In addition to the better math performance of boys, which is also evident in traditional education, another contributing factor may be that boys tend to be more interested in video games compared to girls. This interest could explain differences in effort outcomes, as boys might be more motivated to complete sessions they find enjoyable. However, it is important to note that we do not observe this pattern for children with their mothers as the main supervisors. This difference becomes more pronounced when examining the probability of completing at least 1, 11, and 16 sessions per month.

One possible explanation for these results is that mothers provide greater support to their daughters in engaging with math learning activities compared to fathers, which could lead to the disappearance of the difference in effort gaps. This may also explain why we do not observe similar differences in performance gaps.

Since Smartick is a platform where parental assistance is not required during the sessions, parents might mainly influence the child's decision to start the exercises. However, due to data limitations, we cannot directly observe the actual support provided by parents. Both parents may provide support even if only one has registered the child. Therefore, we cannot rule out the possibility that the differences in gender gaps depending on the gender of the supervisor might be driven by differences in family and household characteristics that are not observable to us.

One might expect that in more gender-egalitarian municipalities, fathers are more likely to supervise children. This expectation could create tension in our interpretation of the results since we find (i) larger gender gaps in effort when the father is the main supervisor, and (ii) living in a municipality with more egalitarian gender norms is associated with positive gender gaps in effort outcomes in favor of girls. To investigate this, we estimate a linear regression of the dummy variable, which takes the value of 1 if the mother supervises, on the gender norm measures we construct (as detailed in Section 3.2). We control for the gender of the child and include other control variables in our main specification (Equation 1). Table A.18 shows that there is no evidence to support a higher likelihood of fathers' supervision in more egalitarian municipalities.

Finally, it is important to emphasize that our results are specific to mathematics training. Given that performance and effort by gender have been observed to differ significantly across traditionally male or female subjects, the outcomes may vary when examining online learning in other subjects.

Table 2.3.: Gender gap - Differences by indicators of gender norms

	persistency	completion	sessions	accuracy	delayed c.
Employment Gender Ratio					
Girl	-0.010	-0.030^{**}	-0.067^{*}	-0.082^{**}	0.109^{**}
	(0.009)	(0.014)	(0.038)	(0.041)	(0.041)
	[0.27]	[0.04]	[0.08]	[0.05]	[0.01]
	[0.27]	*[0.08]*	*[0.10]*	*[0.08]*	*[0.05]*
E_m	-0.006	-0.041	-0.113^{**}	0.024	-0.054
	(0.005)	(0.030)	(0.050)	(0.026)	(0.033)
	[0.24]	[0.18]	[0.03]	[0.36]	[0.11]
	[0.30]	*[0.30]*	*[0.14]*	*[0.36]*	*[0.26]*
GirlxE_m	0.016	0.035^{*}	0.082	0.007	0.047
	(0.010)	(0.018)	(0.049)	(0.047)	(0.040)
	[0.12]	[0.05]	[0.10]	[0.88]	[0.25]
	[0.20]	*[0.20]*	*[0.20]*	*[0.88]*	*[0.31]*
R^2	0.071	0.147	0.110	0.119	0.092
Num. obs.	24896	24896	24896	24792	24783
N Clusters	50	50	50	50	50
Relative Contribution to Chores					
Girl	-0.005	-0.012	-0.036	-0.072^{**}	0.151^{***}
	(0.005)	(0.010)	(0.023)	(0.027)	(0.021)
	[0.34]	[0.21]	[0.13]	[0.01]	[0.00]
	[0.34]	*[0.26]*	*[0.22]*	*[0.03]*	*[0.00]*
C_m	-0.005	0.005	0.012	0.036^{*}	0.026
	(0.006)	(0.009)	(0.016)	(0.020)	(0.019)
	[0.34]	[0.55]	[0.48]	[0.07]	[0.17]
	[0.55]	*[0.55]*	*[0.55]*	*[0.36]*	*[0.43]*
GirlxC_m	0.012^{**}	0.018^{*}	0.056^{**}	-0.005	-0.000
	(0.006)	(0.011)	(0.021)	(0.026)	(0.019)
	[0.04]	[0.09]	[0.01]	[0.85]	[0.99]
	[0.09]	*[0.15]*	*[0.06]*	*[0.99]*	*[0.99]*
R^2	0.071	0.147	0.110	0.119	0.092
Num. obs.	24896	24896	24896	24792	24783
N Clusters	50	50	50	50	50

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Each regression includes a constant term, province, and contract year-month fixed effects, as well as the control variables: age, health condition, number of contracts, and type of the first contract.

2.4. Conclusion

This paper investigates gender gaps in effort and relative performance outcomes in the context of online education, focusing on children primarily supervised by their mothers and fathers. Utilizing data from a widely used online learning platform in Spain, we find that the gender gaps in the effort outcomes are more pronounced for the children who are supervised by their fathers. Our results hold when we compare siblings of opposite gender.

Since our study focuses on the context of online learning, our outcome variables may not directly correspond to those commonly used in previous literature, such as standardized test scores. Among our outcome variables, the most comparable to standardized test scores is accuracy, which represents the average ratio of correctly solved problems when using the online platform for each child. We observe an accuracy gender gap of 8% of a standard deviation[21] favoring boys. This estimated gap is of a similar magnitude to the math gender gap in PISA 2018, where boys outperformed girls with a difference of 6.8% of a standard deviation.[22]

We find that living in a municipality with more egalitarian gender norms is associated with positive gender gaps in effort outcomes in favor of girls, which is in line with the previous findings in the literature (e.g., by González de San Román and De La Rica, 2012; Nollenberger et al., 2016; Rodriguez-Planas and Nollenberger, 2018; Gevrek et al., 2020). However, the difference in effort outcomes does not translate into the gender gap in relative performance outcomes.

We contribute to the existing literature on gender gaps in learning outcomes by providing new evidence from an online learning context while most previous studies focus on traditional education settings.

The increasing adoption of online learning technologies by both parents and schools suggests that these tools will become an important part of regular education systems in the near future. In light of these developments, we believe that it is very important to document gender differences in the use of these tools, as these are likely to translate into differences in further education and labor market outcomes.

While our analysis focuses on Spain, it is important to note that the Smartick platform is utilized by children in many other countries, with the option for three different languages. Therefore, we believe that our study is relevant not only in the Spanish context but also in the context of other countries.

[21] In the siblings' subsample, where we compare opposite-gender siblings, the corresponding gender gap is 5.4% of a standard deviation.

[22] According to the authors' own calculations based on PISA 2018 descriptive statistics reported in OECD (2019).

A. Appendix: Additional Tables and Figures

Table A.1.: Descriptive statistics

1. Full Sample

Variable	N	Mean	Std. Dev.	Min	Max
female	28236	0.52	0.5	0	1
mother	28236	0.66	0.48	0	1
age	28236	8.54	2.33	4	16
health condition	28236	0.06	0.24	0	1
income	27985	37628	9587	16692	90902
registered after covid	28236	0.7	0.46	0	1
covid sample	28236	0.48	0.5	0	1
number of children	28236	1.72	0.88	1	9
total contracts	28236	4.28	4.16	1	31
first contract	28236				
... 1	14615	52%			
... 3	10806	38%			
... 12	2815	10%			
sessions	28236	21.47	7.91	0	32
persistency	28236	0.95	0.22	0	1
completion	28236	0.56	0.5	0	1
delayed completion	28105	1.12	0.28	0.3	3.26
accuracy	28116	0.84	0.06	0.22	1

2. Siblings Sample

Variable	N	Mean	Std. Dev.	Min	Max
female	7533	0.5	0.5	0	1
mother	7533	0.65	0.48	0	1
age	7533	8.56	2.37	4	16
health condition	7533	0.05	0.22	0	1
income	7459	38137	9987	17354	90902
registered after covid	7533	0.71	0.45	0	1
covid sample	7533	0.51	0.5	0	1
number of children	7533	2.45	0.83	2	9
eldest is a boy	7533	0.5	0.5	0	1
total contracts	7533	4.36	4.29	1	30
first contract	7533				
... 1	3926	52%			
... 3	2782	37%			
... 12	825	11%			
sessions	7533	22.14	7.76	0	32
persistency	7533	0.95	0.21	0	1
completion	7533	0.58	0.49	0	1
delayed completion	7502	1.09	0.26	0.3	2.84
accuracy	7505	0.84	0.06	0.42	1

Notes: The table shows the descriptive statistics for the full sample and siblings' sample. The income variable is measured as the average household net income of the municipality of residence, and the covid sample corresponds to the sample of users who registered between March 9, 2020, and May 30, 2020.

Table A.2.: Mean differences by child gender

	boys mean (sd)	girls mean (sd)	difference (t-value)
1. Full Sample			
mother	0.64 (0.48)	0.68 (0.47)	-0.04^{***} (-7.07)
age	8.42 (2.36)	8.65 (2.29)	-0.23^{***} (-8.38)
health condition	0.08 (0.27)	0.05 (0.22)	0.03^{***} (9.83)
income	37600 (9540)	37653 (9631)	-53.3 (-0.46)
number of children	1.72 (0.86)	1.72 (0.89)	0 (-0.25)
sessions	21.52 (7.93)	21.42 (7.89)	0.1 (1.09)
persistency	0.95 (0.22)	0.95 (0.21)	0 (-0.15)
completion	0.56 (0.5)	0.55 (0.5)	0.01 (1.59)
delayed completion	1.1 (0.28)	1.14 (0.28)	-0.04^{***} (-11.06)
accuracy	0.84 (0.06)	0.83 (0.06)	0.01^{***} (9.38)
N	13602	14634	-
2. Siblings Sample			
mother	0.65 (0.48)	0.65 (0.48)	0 (-0.06)
age	8.43 (2.38)	8.68 (2.35)	-0.25^{***} (-4.56)
health condition	0.06 (0.23)	0.04 (0.2)	0.01^{***} (2.72)
income	38137 (9999)	38136 (9975)	1.45 (0.01)
eldest boy	0.5 (0.5)	0.49 (0.5)	0.02 (1.42)
number of children	2.45 (0.83)	2.45 (0.84)	0 (0.11)
sessions	22.31 (7.75)	21.96 (7.76)	0.36^{**} (2.01)
persistency	0.95 (0.21)	0.95 (0.21)	0 (0.09)
completion	0.59 (0.49)	0.57 (0.49)	0.01 (1.23)
delayed completion	1.06 (0.26)	1.11 (0.26)	-0.05^{***} (-7.9)
accuracy	0.84 (0.06)	0.83 (0.06)	0.003^{**} (2.42)
N	3783	3750	-

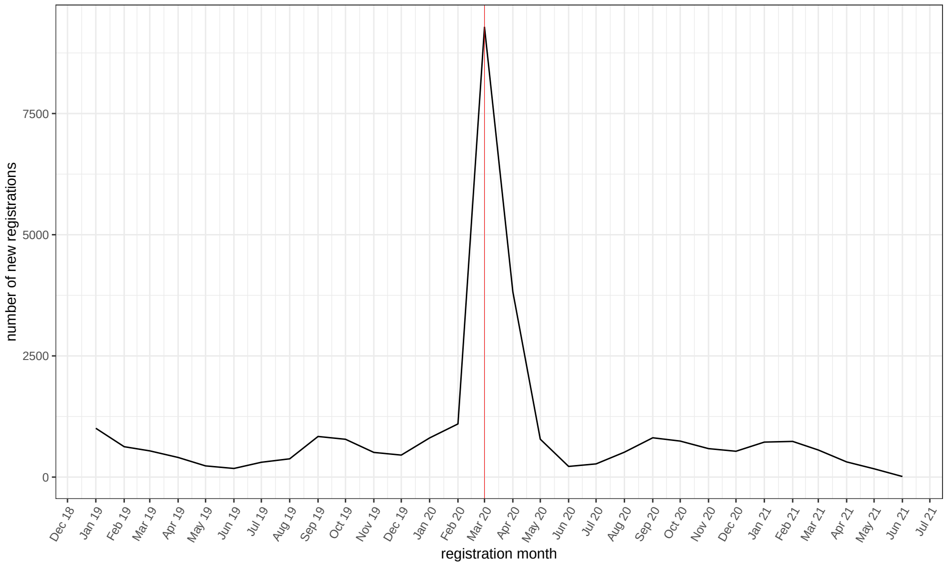
Notes: This table compares the characteristics and outcome variables of girls and boys. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table A.3.: Mean differences by gender of the parent

	fathers mean (sd)	mothers mean (sd)	difference (t-value)
1. Full Sample			
girl	0.49 (0.5)	0.53 (0.5)	-0.04^{***} (-7.07)
age	8.55 (2.35)	8.53 (2.32)	0.02 (0.59)
health condition	0.05 (0.22)	0.07 (0.25)	-0.01^{***} (-4.92)
income	37683 (9581)	37599 (9591)	84.12 (0.7)
number of children	1.69 (0.84)	1.73 (0.9)	-0.04^{***} (-3.24)
sessions	21.32 (7.94)	21.55 (7.89)	-0.23^{**} (-2.29)
persistency	0.95 (0.21)	0.95 (0.22)	0.00 (0.78)
completion	0.54 (0.5)	0.56 (0.5)	-0.02^{***} (-2.99)
delayed completion	1.12 (0.28)	1.12 (0.28)	-0.00 (-0.49)
accuracy	0.84 (0.06)	0.84 (0.06)	0.003*** (3.746)
N	9720	18516	-
2. Siblings Sample			
girl	0.5 (0.5)	0.5 (0.5)	0.00 (-0.06)
age	8.64 (2.42)	8.52 (2.34)	0.12^{**} (2.13)
health condition	0.04 (0.19)	0.06 (0.23)	-0.02^{***} (-3.51)
income	38193 (9778)	38106 (10098)	86.25 (0.36)
eldest boy	0.48 (0.5)	0.5 (0.5)	-0.02^{*} (-1.7)
number of children	2.42 (0.79)	2.47 (0.86)	-0.05^{**} (-2.24)
sessions	22.19 (7.7)	22.1 (7.79)	0.09 (0.49)
persistency	0.96 (0.2)	0.95 (0.22)	0.01 (1.59)
completion	0.57 (0.49)	0.58 (0.49)	-0.01 (-0.98)
delayed completion	1.09 (0.25)	1.08 (0.26)	0.00 (0.4)
accuracy	0.84 (0.06)	0.83 (0.06)	0.003^{*} (1.82)
N	2644	4889	-

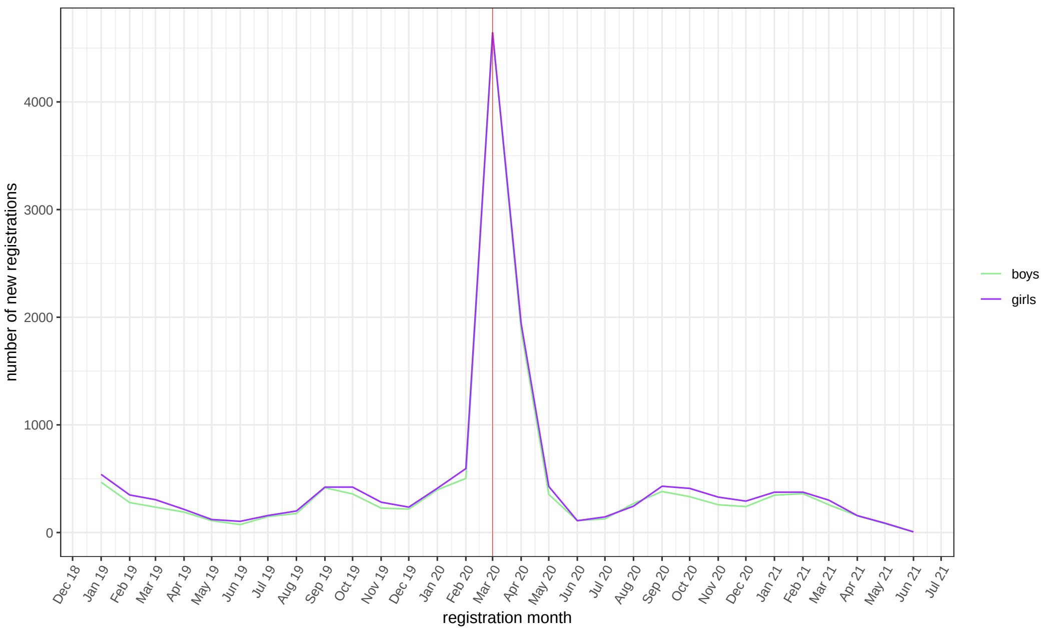
Notes: This table compares the characteristics and outcome variables of children in the fathers' and mothers' samples. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Figure A.1.: Total number of new registrations - monthly



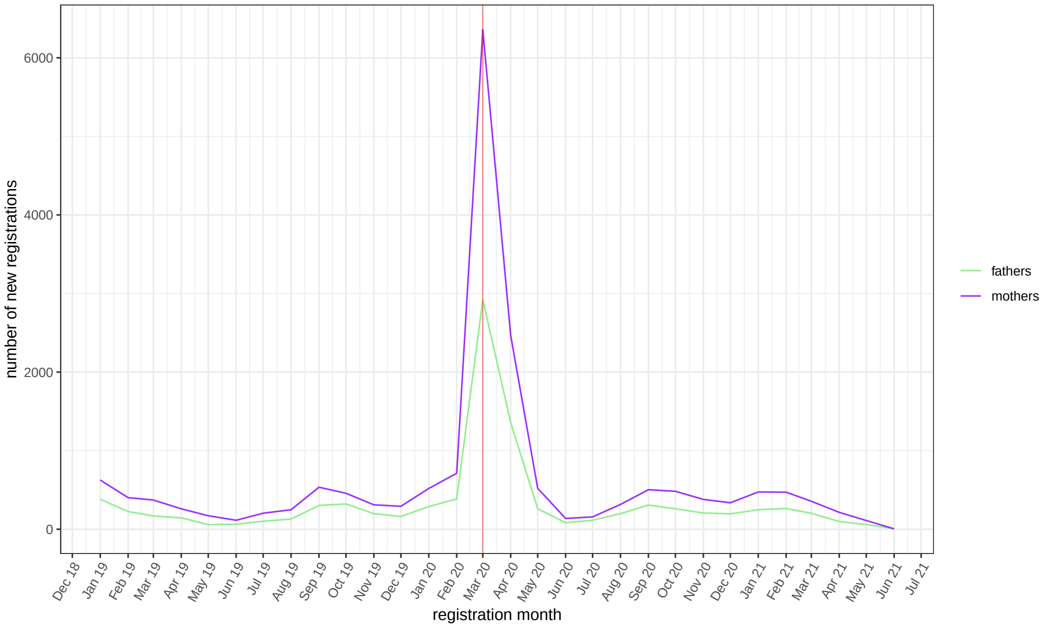
Notes: This figure shows the evolution of the total number of new registrations to the Smartick platform per month, by people residing in Spain.

Figure A.2.: Total number of new registrations by child gender - monthly



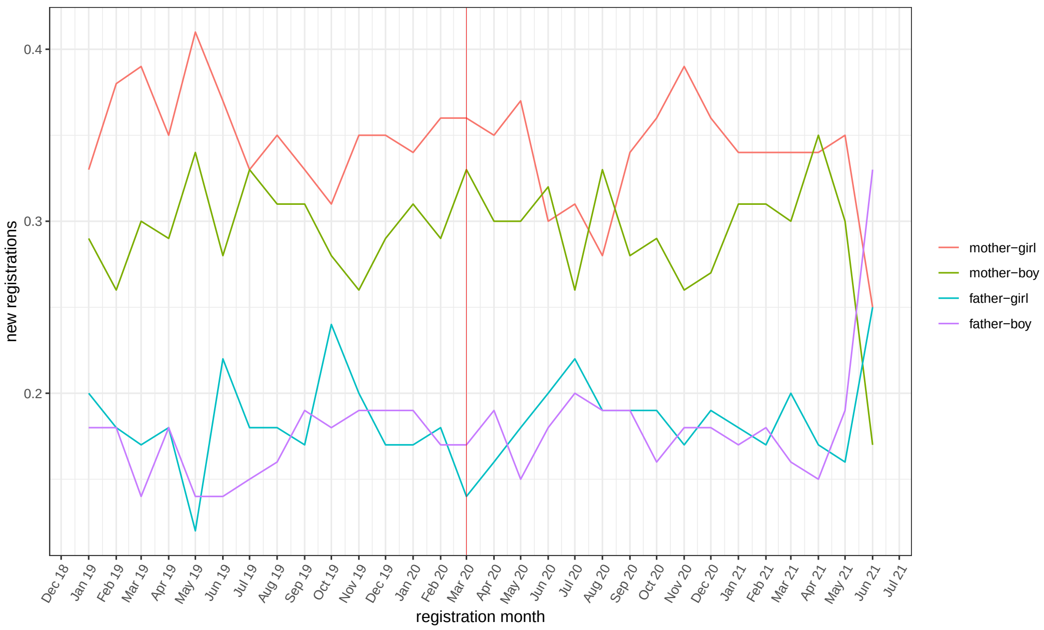
Notes: This figure shows, by child's gender, the evolution of the total number of new registrations to the Smartick platform per month, by people residing in Spain.

Figure A.3.: Total number of new registrations by parent gender - monthly



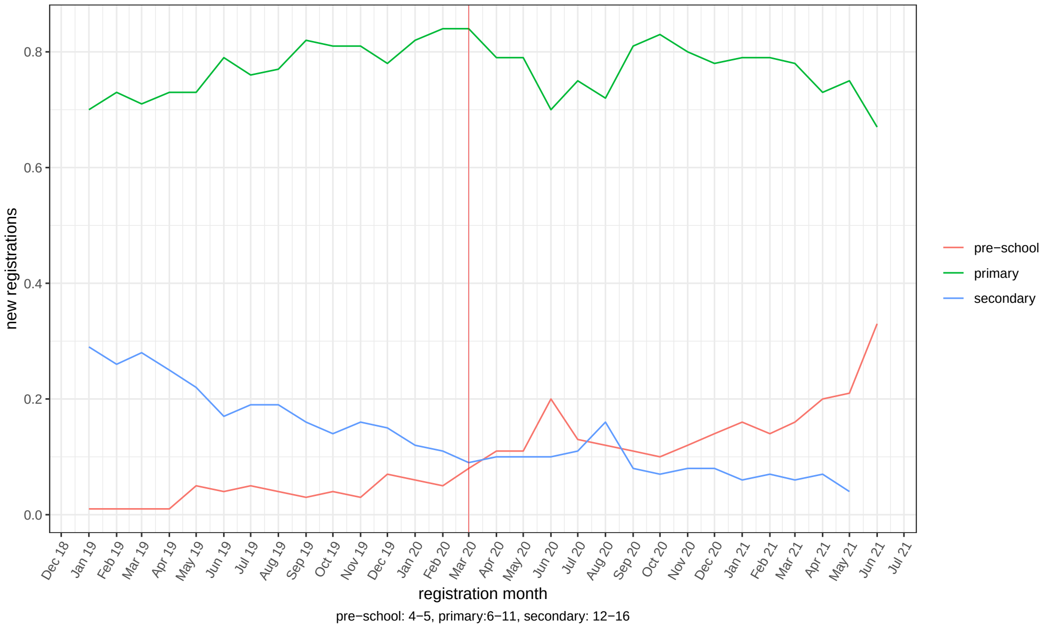
Notes: This figure shows, by parent's gender, the evolution of the total number of new registrations to the Smartick platform per month, by people residing in Spain.

Figure A.4.: Percentage of new registrations by gender pairs - monthly



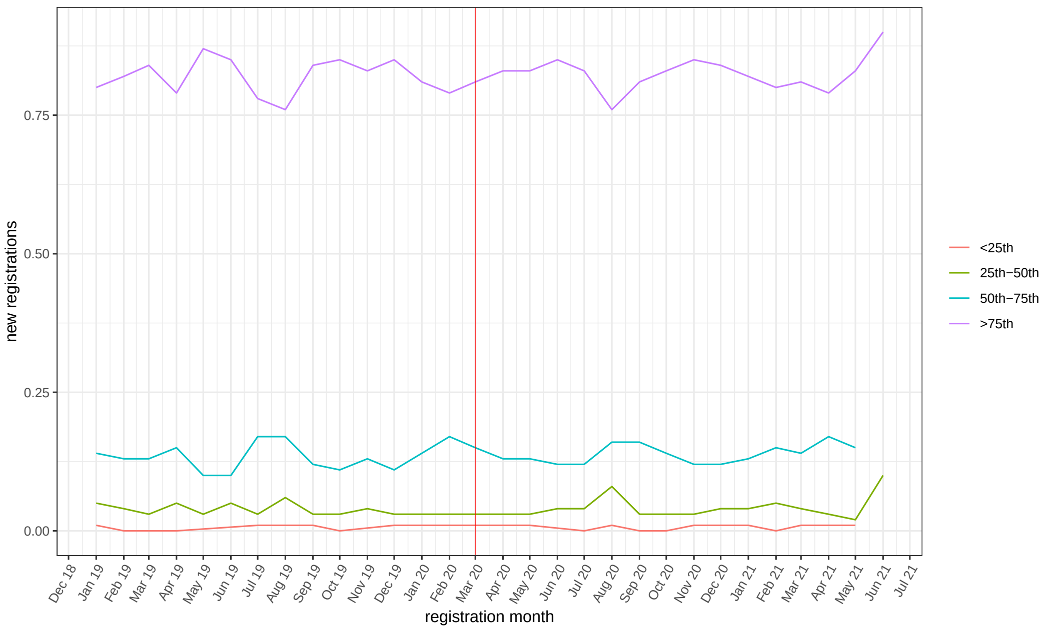
Notes: This figure shows the evolution of the percentages of new registrations by gender pairs of children and parents to the Smartick platform per month, by people residing in Spain.

Figure A.5.: Percentage of new registrations by age categories - monthly



Notes: This figure shows the evolution of the percentages of new registrations by age categories of children to the Smartick platform per month, by people residing in Spain.

Figure A.6.: Percentage of new registrations by income categories - monthly



Notes: This figure shows the evolution of the percentages of new registrations by income categories to the Smartick platform per month, by people residing in Spain.

Table A.4.: Gender gaps in the full sample and siblings' sample

	persistency	completion	sessions	accuracy	delayed c.
Full Sample					
Girl	0.003	-0.002	0.001	-0.080^{***}	0.148^{***}
	(0.003)	(0.005)	(0.014)	(0.009)	(0.011)
	[0.25]	[0.77]	[0.97]	[0.00]	[0.00]
	[0.41]	*[0.97]*	*[0.97]*	*[0.00]*	*[0.00]*
Mother	-0.005^{*}	0.011^{**}	0.024^{**}	-0.048^{***}	0.007
	(0.003)	(0.005)	(0.012)	(0.014)	(0.010)
Age	-0.004^{***}	0.002^{*}	0.004^{*}	-0.124^{***}	-0.068^{***}
	(0.000)	(0.001)	(0.003)	(0.004)	(0.002)
Health Cond.	-0.015^{*}	-0.017^{*}	-0.044^{**}	-0.174^{***}	0.346^{***}
	(0.007)	(0.008)	(0.021)	(0.033)	(0.026)
Num of contr	0.004^{***}	0.023^{***}	0.058^{***}	0.008^{***}	0.002
	(0.000)	(0.001)	(0.002)	(0.001)	(0.002)
1st C:12	-0.143^{***}	-0.241^{***}	0.092^{***}	-0.357^{***}	0.538^{***}
	(0.007)	(0.011)	(0.022)	(0.019)	(0.017)
1st C:3	-0.056^{***}	-0.205^{***}	-0.044^{**}	-0.118^{***}	0.112^{***}
	(0.004)	(0.013)	(0.020)	(0.012)	(0.013)
Constant	0.985^{***}	0.476^{***}	-0.418^{***}	1.472^{***}	0.631^{***}
	(0.013)	(0.039)	(0.081)	(0.103)	(0.051)
Province FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.068	0.144	0.107	0.121	0.092
Num. obs.	28236	28236	28236	28116	28105
N Clusters	50	50	50	50	50
Siblings Sample					
Girl	0.002	-0.013	-0.041^{***}	-0.054^{***}	0.181^{***}
	(0.003)	(0.008)	(0.011)	(0.016)	(0.020)
	[0.53]	[0.11]	[0.00]	[0.00]	[0.00]
	[0.53]	*[0.13]*	*[0.00]*	*[0.00]*	*[0.00]*
Age	-0.005^{***}	-0.007^{***}	-0.025^{***}	-0.037^{***}	-0.031^{***}
	(0.001)	(0.002)	(0.004)	(0.011)	(0.011)
Health Cond.	-0.016	-0.022	-0.028	-0.348^{***}	0.276^{**}
	(0.012)	(0.029)	(0.064)	(0.083)	(0.127)
Constant	1.069^{***}	1.096^{***}	0.975^{***}	0.162	-0.650^{***}
	(0.018)	(0.032)	(0.050)	(0.141)	(0.142)
Parent FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	No	No	No	No	No
R^2	0.718	0.814	0.869	0.685	0.635
Num. obs.	7533	7533	7533	7505	7502
N Clusters	50	50	50	50	50

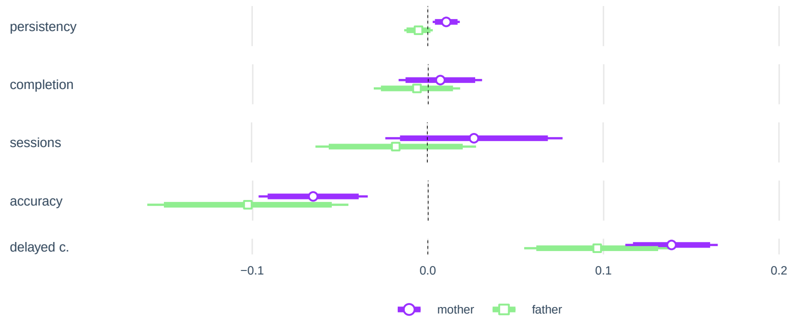
Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. The coefficient estimates are from the regressions of five different outcome variables given in different columns, on the indicator variables of the child's gender, parent's gender, and the control variables.

Table A.5.: Gender gaps in mothers' and fathers' (full) sample

	persistency	completion	sessions	accuracy	delayed c.
Mothers' Sample					
Girl	0.007^{*}	0.004	0.018	-0.078^{***}	0.157^{***}
	(0.003)	(0.007)	(0.021)	(0.014)	(0.010)
	[0.05]	[0.55]	[0.41]	[0.00]	[0.00]
	[0.09]	*[0.55]*	*[0.51]*	*[0.00]*	*[0.00]*
R^2	0.069	0.151	0.115	0.123	0.098
Num. obs.	18516	18516	18516	18435	18430
N Clusters	50	50	50	50	50
Fathers' Sample					
Girl	-0.004	-0.013	-0.032^{*}	-0.088^{***}	0.131^{***}
	(0.003)	(0.008)	(0.016)	(0.015)	(0.020)
	[0.12]	[0.11]	[0.06]	[0.00]	[0.00]
	[0.12]	*[0.12]*	*[0.09]*	*[0.00]*	*[0.00]*
R^2	0.074	0.141	0.103	0.126	0.088
Num. obs.	9720	9720	9720	9681	9675
N Clusters	50	50	50	50	50
Two sample t-test					
Difference	0.011	0.017	0.05	0.01	0.03
p-value	(0.01)	(0.11)	(0.07)	(0.63)	(0.24)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are (in parenthesis) clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, health condition, number of contracts, type of the first contract.

Figure A.7.: Gender gap - Covid sample



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers and fathers full covid samples separately, based on Equation 1. Covid sample includes the members registered between 9 March 2020 and 31 May 2020, which corresponds to the first wave of the Covid-19 outbreak in Spain. Control variables are age of the child, presence of a health condition, total number of contracts, type of the first contract, province and contract year-month fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

Table A.6.: Gender gaps in the full sample

	persistency	completion	sessions	accuracy	delayed c.
Girl	0.007^{*}	0.005	0.017	-0.078^{***}	0.157^{***}
	(0.003)	(0.007)	(0.021)	(0.014)	(0.011)
	[0.05]	[0.54]	[0.41]	[0.00]	[0.00]
	[0.09]	*[0.54]*	*[0.51]*	*[0.00]*	*[0.00]*
Father	0.011^{***}	-0.002	0.000	0.051^{**}	0.005
	(0.004)	(0.007)	(0.020)	(0.022)	(0.014)
	[0.01]	[0.71]	[1.00]	[0.03]	[0.71]
	[0.04]	*[0.89]*	*[1.00]*	*[0.06]*	*[0.89]*
GirlxFather	-0.011^{***}	-0.018	-0.049^{*}	-0.007	-0.025
	(0.004)	(0.011)	(0.027)	(0.023)	(0.017)
	[0.01]	[0.12]	[0.08]	[0.76]	[0.14]
	[0.04]	*[0.17]*	*[0.17]*	*[0.76]*	*[0.17]*
Age	-0.004^{***}	0.002^{*}	0.004^{*}	-0.124^{***}	-0.068^{***}
	(0.000)	(0.001)	(0.003)	(0.004)	(0.002)
Health Cond.	-0.014^{*}	-0.016^{*}	-0.044^{**}	-0.173^{***}	0.346^{***}
	(0.007)	(0.008)	(0.022)	(0.033)	(0.026)
Num of contr	0.004^{***}	0.023^{***}	0.058^{***}	0.008^{***}	0.002
	(0.000)	(0.001)	(0.002)	(0.001)	(0.002)
1st C:12	-0.143^{***}	-0.241^{***}	0.092^{***}	-0.357^{***}	0.538^{***}
	(0.007)	(0.012)	(0.022)	(0.019)	(0.017)
1st C:3	-0.056^{***}	-0.205^{***}	-0.044^{**}	-0.118^{***}	0.112^{***}
	(0.004)	(0.013)	(0.020)	(0.012)	(0.013)
Constant	0.978^{***}	0.484^{***}	-0.402^{***}	1.423^{***}	0.633^{***}
	(0.014)	(0.039)	(0.087)	(0.092)	(0.047)
Province FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.068	0.144	0.107	0.121	0.092
Num. obs.	28236	28236	28236	28116	28105
N Clusters	50	50	50	50	50

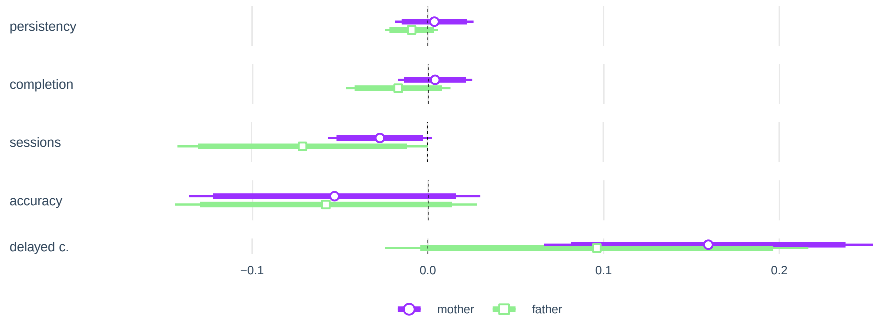
Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. The coefficient estimates are from the regressions of five different outcome variables given in different columns, on the indicator variables of the child's gender, the parent's gender, their interaction, and the control variables.

Table A.7.: Gender gaps in mothers' and fathers' (siblings) sample

	persistency	completion	sessions	accuracy	delayed c.
Mothers' Sample					
Girl	0.006	−0.002	−0.023**	−0.063***	0.189***
	(0.005)	(0.007)	(0.011)	(0.022)	(0.026)
	[0.22]	[0.75]	[0.04]	[0.01]	[0.00]
	[0.27]	*[0.75]*	*[0.06]*	*[0.02]*	*[0.00]*
R^2	0.703	0.824	0.870	0.687	0.634
Num. obs.	4889	4889	4889	4868	4867
N Clusters	49	49	49	49	49
Fathers' Sample					
Girl	−0.006	−0.032*	−0.074**	−0.038	0.165***
	(0.006)	(0.019)	(0.032)	(0.025)	(0.041)
	[0.27]	[0.09]	[0.02]	[0.13]	[0.00]
	[0.27]	*[0.16]*	*[0.06]*	*[0.17]*	*[0.00]*
R^2	0.753	0.796	0.868	0.684	0.636
Num. obs.	2644	2644	2644	2637	2635
N Clusters	50	50	50	50	50
Two sample t-test					
Difference	0.012	0.030	0.051	-0.025	0.023
p-value	(0.10)	(0.13)	(0.13)	(0.45)	(0.63)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are (in parenthesis) clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, health condition, number of contracts, type of the first contract.

Figure A.8.: Gender gap - Covid sample (siblings)



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers and fathers siblings covid samples separately, based on Equation 2. Covid sample includes the members registered between 9 March 2020 and 31 May 2020, which corresponds to the first wave of the Covid-19 outbreak in Spain. Control variables are age of the child, presence of a health condition and parent fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

Table A.8.: Gender gap in the full sample - Differences by age

	persistency	completion	sessions	accuracy	delayed c.
A. Continuous Age Variable					
Girl	0.012*	−0.004	−0.003	−0.091***	0.451***
	(0.006)	(0.016)	(0.038)	(0.031)	(0.050)
	[0.05]	[0.82]	[0.94]	[0.01]	[0.00]
	[0.09]	*[0.94]*	*[0.94]*	*[0.01]*	*[0.00]*
Age	−0.003***	0.002	0.004	−0.125***	−0.050***
	(0.001)	(0.001)	(0.003)	(0.004)	(0.003)
	[0.00]	[0.14]	[0.16]	[0.00]	[0.00]
	[0.00]	*[0.16]*	*[0.16]*	*[0.00]*	*[0.00]*
Girl*Age	−0.001	0.000	0.000	0.001	−0.035***
	(0.001)	(0.002)	(0.005)	(0.003)	(0.006)
	[0.16]	[0.91]	[0.94]	[0.69]	[0.00]
	[0.41]	*[0.94]*	*[0.94]*	*[0.94]*	*[0.00]*
R^2	0.068	0.144	0.107	0.121	0.094
Num. obs.	28236	28236	28236	28116	28105
N Clusters	50	50	50	50	50
B. Age Categories					
Girl	0.005	−0.041***	−0.051*	0.026	0.127**
	(0.006)	(0.014)	(0.028)	(0.040)	(0.058)
	[0.39]	[0.01]	[0.08]	[0.52]	[0.03]
	[0.48]	*[0.03]*	*[0.13]*	*[0.52]*	*[0.08]*
Primary	−0.007*	0.018	0.079***	−0.760***	0.215***
	(0.004)	(0.011)	(0.024)	(0.052)	(0.047)
	[0.10]	[0.11]	[0.00]	[0.00]	[0.00]
	[0.11]	*[0.11]*	*[0.00]*	*[0.00]*	*[0.00]*
Secondary	−0.025***	0.009	0.038	−1.101***	−0.160***
	(0.008)	(0.012)	(0.028)	(0.054)	(0.049)
	[0.00]	[0.46]	[0.19]	[0.00]	[0.00]
	[0.00]	*[0.46]*	*[0.23]*	*[0.00]*	*[0.00]*
Girl*Primary	−0.003	0.043**	0.054	−0.141***	0.030
	(0.008)	(0.016)	(0.034)	(0.047)	(0.066)
	[0.70]	[0.01]	[0.12]	[0.00]	[0.65]
	[0.70]	*[0.03]*	*[0.20]*	*[0.02]*	*[0.70]*
Girl*Secondary	−0.005	0.040*	0.065	−0.044	−0.176**
	(0.009)	(0.024)	(0.046)	(0.041)	(0.071)
	[0.60]	[0.10]	[0.16]	[0.28]	[0.02]
	[0.60]	*[0.24]*	*[0.27]*	*[0.35]*	*[0.08]*
R^2	0.068	0.145	0.108	0.109	0.096
Num. obs.	28236	28236	28236	28116	28105
N Clusters	50	50	50	50	50

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: mother dummy, health condition, number of contracts, type of the first contract. For the regressions in Panel B, the baseline category is preschool. (preschool: 4-5, primary:6-11, secondary: 12-16)

Table A.9.: Gender gap - Differences by income categories

	persistency	completion	sessions	accuracy	delayed c.
Girl	-0.024	-0.022	-0.047	-0.135	0.057
	(0.032)	(0.076)	(0.145)	(0.100)	(0.130)
	[0.45]	[0.77]	[0.75]	[0.18]	[0.66]
	[0.77]	*[0.77]*	*[0.77]*	*[0.77]*	*[0.77]*
25th-50th	-0.016	-0.063	-0.123	0.045	-0.093
	(0.018)	(0.042)	(0.092)	(0.110)	(0.124)
	[0.36]	[0.14]	[0.19]	[0.68]	[0.46]
	[0.57]	*[0.47]*	*[0.47]*	*[0.68]*	*[0.57]*
50th-75th	0.001	-0.058	-0.150^{**}	-0.062	-0.115
	(0.018)	(0.040)	(0.069)	(0.097)	(0.107)
	[0.96]	[0.15]	[0.04]	[0.53]	[0.29]
	[0.96]	*[0.39]*	*[0.18]*	*[0.66]*	*[0.48]*
>75th	-0.006	-0.057	-0.165^{**}	0.006	-0.155
	(0.017)	(0.043)	(0.080)	(0.094)	(0.112)
	[0.74]	[0.19]	[0.04]	[0.95]	[0.17]
	[0.92]	*[0.32]*	*[0.22]*	*[0.95]*	*[0.32]*
Girl*25th-50th	0.030	0.016	0.050	-0.064	0.127
	(0.036)	(0.073)	(0.147)	(0.121)	(0.171)
	[0.41]	[0.82]	[0.73]	[0.60]	[0.46]
	[0.82]	*[0.82]*	*[0.82]*	*[0.82]*	*[0.82]*
Girl*50th-75th	0.023	0.011	0.022	0.065	0.059
	(0.036)	(0.074)	(0.141)	(0.108)	(0.131)
	[0.53]	[0.89]	[0.88]	[0.55]	[0.65]
	[0.89]	*[0.89]*	*[0.89]*	*[0.89]*	*[0.89]*
Girl*>75th	0.027	0.023	0.054	0.020	0.076
	(0.032)	(0.076)	(0.145)	(0.101)	(0.131)
	[0.40]	[0.76]	[0.71]	[0.84]	[0.56]
	[0.84]	*[0.84]*	*[0.84]*	*[0.84]*	*[0.84]*
Province FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.067	0.144	0.107	0.047	0.070
Num. obs.	27985	27985	27985	27865	27854
N Clusters	49	49	49	49	49

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: mother dummy, health condition, number of contracts, type of the first contract. Baseline Category is "$<$ 25th percentile"

Table A.10.: Gender gap - Differences by income categories within Smartick

	persistency	completion	sessions	accuracy	delayed c.
Girl	-0.006	-0.005	-0.014	-0.195^{***}	0.143^{*}
	(0.012)	(0.030)	(0.060)	(0.062)	(0.082)
	[0.64]	[0.86]	[0.81]	[0.00]	[0.09]
	[0.86]	*[0.86]*	*[0.86]*	*[0.02]*	*[0.22]*
25th-50th	0.015	-0.023	-0.081^{*}	-0.128^{**}	-0.059
	(0.012)	(0.022)	(0.048)	(0.049)	(0.060)
	[0.24]	[0.29]	[0.10]	[0.01]	[0.34]
	[0.34]	*[0.34]*	*[0.24]*	*[0.06]*	*[0.34]*
50th-75th	0.016	-0.002	-0.037	-0.078^{**}	-0.101
	(0.010)	(0.026)	(0.049)	(0.034)	(0.070)
	[0.12]	[0.94]	[0.46]	[0.03]	[0.16]
	[0.26]	*[0.94]*	*[0.57]*	*[0.13]*	*[0.26]*
>75th	0.009	-0.050^{*}	-0.151^{***}	-0.067^{*}	-0.126
	(0.012)	(0.029)	(0.051)	(0.038)	(0.076)
	[0.47]	[0.09]	[0.00]	[0.09]	[0.11]
	[0.47]	*[0.13]*	*[0.02]*	*[0.13]*	*[0.13]*
Girl*25th-50th	0.012	-0.016	-0.004	0.133	-0.037
	(0.016)	(0.037)	(0.073)	(0.083)	(0.078)
	[0.48]	[0.67]	[0.96]	[0.11]	[0.63]
	[0.84]	*[0.84]*	*[0.96]*	*[0.57]*	*[0.84]*
Girl*50th-75th	-0.001	-0.005	-0.009	0.090	0.002
	(0.013)	(0.032)	(0.064)	(0.058)	(0.084)
	[0.91]	[0.89]	[0.88]	[0.13]	[0.99]
	[0.99]	*[0.99]*	*[0.99]*	*[0.63]*	*[0.99]*
Girl*>75th	0.012	0.012	0.032	0.077	-0.015
	(0.012)	(0.030)	(0.060)	(0.063)	(0.084)
	[0.33]	[0.70]	[0.59]	[0.23]	[0.86]
	[0.82]	*[0.86]*	*[0.86]*	*[0.82]*	*[0.86]*
Province FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.067	0.145	0.108	0.047	0.070
Num. obs.	27985	27985	27985	27865	27854
N Clusters	49	49	49	49	49

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: mother dummy, health condition, number of contracts, type of the first contract. Baseline Category is "< 25th percentile"

Table A.11.: Mean differences by covid sample

	non-covid mean (sd)	covid mean (sd)	difference (t-value)
1. Full Sample			
girl	0.53 (0.5)	0.5 (0.5)	0.03^{***} (4.92)
mother	0.64 (0.48)	0.67 (0.47)	-0.04^{***} (-6.39)
age	8.77 (2.41)	8.29 (2.2)	0.49^{***} (17.69)
health condition	0.07 (0.25)	0.05 (0.23)	0.02^{***} (5.39)
income	37978 (9930)	37246 (9186)	732.07^{***} (6.39)
number of children	1.69 (0.9)	1.75 (0.85)	-0.06^{***} (-6)
sessions	20.39 (8.09)	22.65 (7.53)	-2.27^{***} (-24.32)
persistency	0.94 (0.23)	0.96 (0.2)	-0.01^{***} (-5.37)
completion	0.5 (0.5)	0.62 (0.49)	-0.12^{***} (-19.89)
delayed comp	1.15 (0.29)	1.1 (0.27)	0.05^{***} (15.58)
accuracy	0.83 (0.06)	0.84 (0.06)	-0.01^{***} (-11.53)
N	14723	13513	-
2. Siblings Sample			
girl	0.5 (0.5)	0.5 (0.5)	0.00 (0.33)
mother	0.62 (0.49)	0.68 (0.47)	-0.06^{***} (-5.83)
age	8.76 (2.45)	8.36 (2.27)	0.4^{***} (7.36)
health condition	0.05 (0.23)	0.05 (0.21)	0.01 (1.63)
income	38703 (10616)	37582 (9297)	1121.79^{***} (4.86)
eldest boy	0.49 (0.5)	0.5 (0.5)	-0.02 (-1.47)
number of children	2.49 (0.86)	2.42 (0.81)	0.07^{***} (3.67)
sessions	21.03 (7.98)	23.22 (7.38)	-2.18^{***} (-12.35)
persistency	0.95 (0.23)	0.96 (0.2)	-0.01^{**} (-2.35)
completion	0.52 (0.5)	0.64 (0.48)	-0.11^{***} (-10.09)
delayed comp	1.11 (0.26)	1.06 (0.25)	0.04^{***} (7.46)
accuracy	0.83 (0.06)	0.84 (0.06)	0.005^{***} (-3.43)
N	3728	3805	-

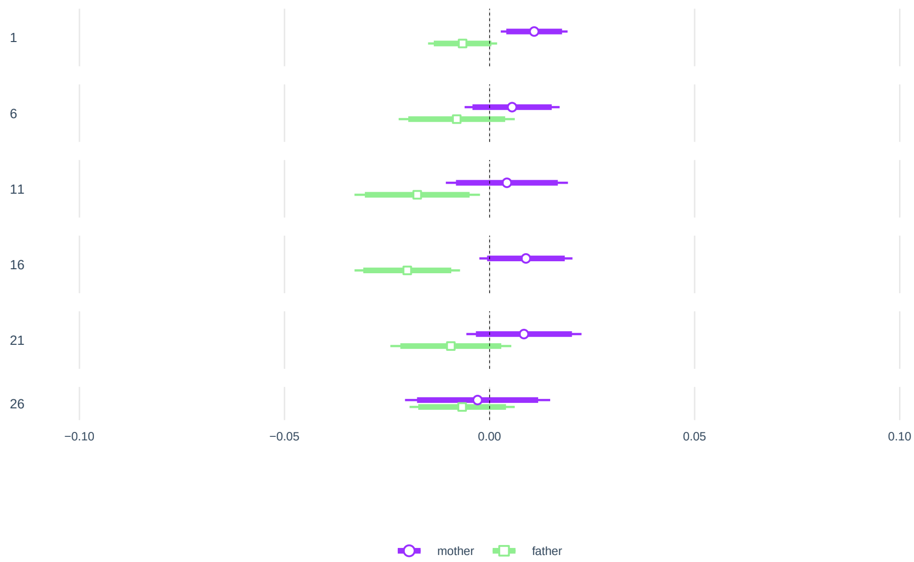
Notes: This table compares the characteristics and outcome variables of children in the covid sample to the rest of the sample. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table A.12.: Gender gaps in the mothers' and fathers' Covid sample

	persistency	completion	sessions	accuracy	delayed c.
Mothers' Sample					
Girl	0.010***	0.007	0.027	−0.066***	0.139***
	(0.004)	(0.012)	(0.025)	(0.015)	(0.013)
	[0.01]	[0.56]	[0.30]	[0.00]	[0.00]
	[0.01]	*[0.56]*	*[0.37]*	*[0.00]*	*[0.00]*
R^2	0.078	0.183	0.138	0.098	0.098
Num. obs.	9116	9116	9116	9088	9088
N Clusters	50	50	50	50	50
Fathers' Sample					
Girl	−0.005	−0.006	−0.018	−0.103***	0.096***
	(0.004)	(0.012)	(0.023)	(0.029)	(0.021)
	[0.20]	[0.60]	[0.43]	[0.00]	[0.00]
	[0.33]	*[0.60]*	*[0.54]*	*[0.00]*	*[0.00]*
R^2	0.096	0.197	0.136	0.099	0.100
Num. obs.	4397	4397	4397	4381	4381
N Clusters	50	50	50	50	50

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, health condition, number of contracts, type of the first contract.

Figure A.9.: Different number of sessions cut-offs for completion variable



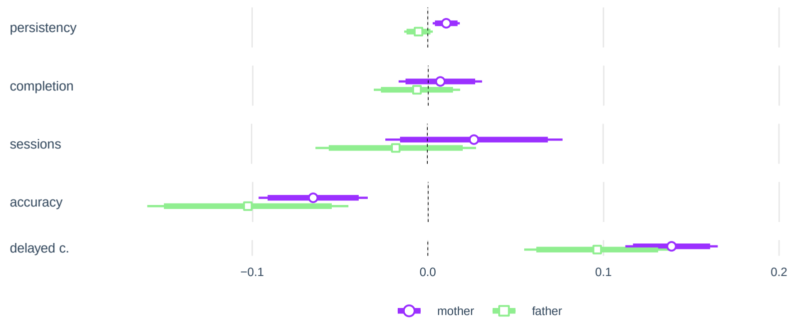
Notes: The Figure shows the estimated gender gaps in completion outcome for different numbers of session cut-offs in mothers' and fathers' full samples separately, based on Equation 1. Control variables are the age of the child, presence of a health condition, total number of contracts, type of the first contract, province, and contract year-month fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

Table A.13.: Gender gaps in the mothers' and fathers' Covid sample (siblings)

	persistency	completion	sessions	accuracy	delayed c.
Mothers' Sample					
Girl	0.004	0.004	−0.027*	−0.053	0.160***
	(0.011)	(0.011)	(0.015)	(0.041)	(0.047)
	[0.01]	[0.56]	[0.30]	[0.00]	[0.00]
	[0.01]	*[0.56]*	*[0.37]*	*[0.00]*	*[0.00]*
R^2	0.720	0.846	0.889	0.713	0.673
Num. obs.	2590	2590	2590	2581	2581
N Clusters	49	49	49	49	49
Fathers' Sample					
Girl	−0.009	−0.017	−0.071*	−0.058	0.096
	(0.008)	(0.015)	(0.035)	(0.043)	(0.060)
	[0.20]	[0.60]	[0.43]	[0.00]	[0.00]
	[0.33]	*[0.60]*	*[0.54]*	*[0.00]*	*[0.00]*
R^2	0.799	0.841	0.887	0.715	0.668
Num. obs.	1215	1215	1215	1213	1213
N Clusters	48	48	48	48	48

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term and parent fixed effects. Control variables: age, health condition.

Figure A.10.: Age fixed effects in the main specification - Full sample



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers' and fathers' full samples separately, based on Equation 1, where the regression includes age fixed effects instead of the continuous age variable. Control variables are the age fixed effects, presence of a health condition, total number of contracts, type of the first contract, province, and contract year-month fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

Table A.14.: Gender gaps in the full sample - Registrations before and after Covid-19 outbreak

	persistency	completion	sessions	accuracy	delayed c.
Registrations Before Covid-19 Outbreak					
Girl	0.011^{*}	0.003	0.026	-0.061^{**}	0.161^{***}
	(0.006)	(0.010)	(0.021)	(0.028)	(0.025)
	[0.07]	[0.77]	[0.21]	[0.04]	[0.00]
	[0.11]	*[0.77]*	*[0.27]*	*[0.09]*	*[0.00]*
Father	0.018^{***}	0.005	0.026	0.050^{**}	-0.035
	(0.005)	(0.018)	(0.030)	(0.024)	(0.031)
	[0.00]	[0.79]	[0.40]	[0.04]	[0.27]
	[0.00]	*[0.79]*	*[0.49]*	*[0.11]*	*[0.45]*
GirlxFather	-0.016^{**}	-0.002	-0.029	0.015	0.011
	(0.006)	(0.014)	(0.035)	(0.030)	(0.035)
	[0.01]	[0.90]	[0.42]	[0.62]	[0.75]
	[0.07]	*[0.90]*	*[0.90]*	*[0.90]*	*[0.90]*
Province FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.103	0.136	0.097	0.116	0.115
Num. obs.	8520	8520	8520	8477	8476
N Clusters	50	50	50	50	50
Registrations After Covid-19 Outbreak					
Girl	0.005	0.006	0.016	-0.082^{***}	0.153^{***}
	(0.004)	(0.010)	(0.027)	(0.013)	(0.011)
	[0.15]	[0.56]	[0.55]	[0.00]	[0.00]
	[0.25]	*[0.56]*	*[0.56]*	*[0.00]*	*[0.00]*
Father	0.008	-0.006	-0.012	0.055^{*}	0.020
	(0.005)	(0.011)	(0.029)	(0.028)	(0.017)
	[0.13]	[0.62]	[0.66]	[0.05]	[0.26]
	[0.32]	*[0.66]*	*[0.66]*	*[0.26]*	*[0.44]*
GirlxFather	-0.009	-0.024	-0.057	-0.022	-0.037^{*}
	(0.006)	(0.019)	(0.043)	(0.028)	(0.021)
	[0.14]	[0.19]	[0.19]	[0.42]	[0.09]
	[0.24]	*[0.24]*	*[0.24]*	*[0.42]*	*[0.24]*
Province FE	Yes	Yes	Yes	Yes	Yes
Cont. Month FE	Yes	Yes	Yes	Yes	Yes
R^2	0.057	0.152	0.118	0.130	0.090
Num. obs.	19716	19716	19716	19639	19629
N Clusters	50	50	50	50	50

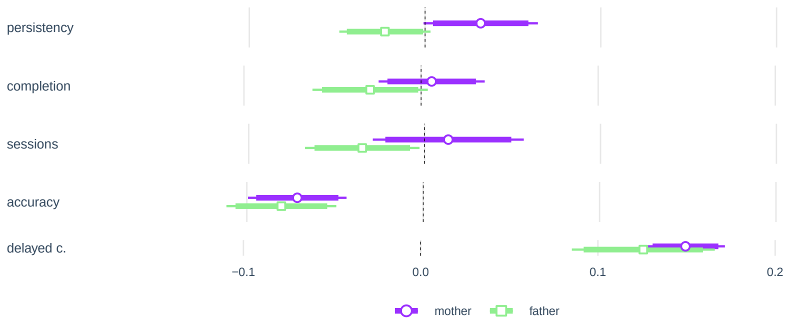
Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, health condition, number of contracts, type of the first contract.

Table A.15.: Gender gaps in the mothers' and fathers' siblings sample - Province FE

	persistency	completion	sessions	accuracy	delayed c.
Mothers' Sample					
Girl	0.005	−0.003	−0.028***	−0.047***	0.205***
	(0.004)	(0.005)	(0.009)	(0.016)	(0.020)
	[0.24]	[0.54]	[0.00]	[0.01]	[0.00]
	[0.31]	*[0.54]*	*[0.01]*	*[0.01]*	*[0.00]*
R^2	0.097	0.204	0.143	0.104	0.121
Num. obs.	4889	4889	4889	4868	4867
N Clusters	49	49	49	49	49
Fathers' Sample					
Girl	−0.005	−0.027*	−0.066**	−0.015	0.189***
	(0.004)	(0.015)	(0.027)	(0.019)	(0.029)
	[0.21]	[0.07]	[0.02]	[0.43]	[0.00]
	[0.26]	*[0.12]*	*[0.05]*	*[0.43]*	*[0.00]*
R^2	0.108	0.196	0.149	0.128	0.126
Num. obs.	2644	2644	2644	2637	2635
N Clusters	50	50	50	50	50

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, health condition, number of contracts, type of the first contract.

Figure A.11.: Normalization within age - Full sample



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers' and fathers' full samples separately, based on Equation 1, where the outcome variables are normalized within each age. Control variables are the age of the child, presence of a health condition, total number of contracts, type of the first contract, province, and contract year-month fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

Table A.16.: Logit Regressions

	Mothers		Fathers	
A. Full Sample	persistency	completion	persistency	completion
Girl	0.136	0.022	-0.117^{*}	-0.062^{*}
	(0.086)	(0.034)	(0.069)	(0.037)
AIC	5723.798	22488.430	2975.372	12078.474
BIC	6389.041	23153.673	3585.837	12688.939
Log Likelihood	−2776.899	−11159.215	−1402.686	−5954.237
Deviance	5553.798	22318.430	2805.372	11908.474
Num. obs.	18516	18516	9720	9720
B. Siblings Sample				
Girl	0.423	−0.079	−0.852	-0.681^{**}
	(0.277)	(0.124)	(0.574)	(0.320)
AIC	4649.769	5455.443	2476.539	3073.614
BIC	18542.024	19347.699	9362.076	9959.150
Log Likelihood	−185.884	−588.722	−67.270	−365.807
Deviance	371.769	1177.443	134.539	731.614
Num. obs.	4889	4889	2644	2644

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Control variables for the regressions in Panel A: age, health condition, number of contracts, type of the first contract, province, and contract year-month fixed effects. Control variables for the regressions in Panel B: age, health condition, and parent fixed effects.

Table A.17.: Gender gaps in the full sample of children without a health condition

	persistency	completion	sessions	accuracy	delayed c.
Girl	0.008^{*}	0.004	0.019	-0.071^{***}	0.159^{***}
	(0.004)	(0.008)	(0.025)	(0.015)	(0.012)
	[0.06]	[0.60]	[0.46]	[0.00]	[0.00]
	[0.10]	*[0.60]*	*[0.57]*	*[0.00]*	*[0.00]*
Father	0.011^{***}	−0.004	0.001	0.041^{*}	0.006
	(0.004)	(0.008)	(0.024)	(0.023)	(0.015)
	[0.01]	[0.64]	[0.98]	[0.08]	[0.70]
	[0.04]	*[0.88]*	*[0.98]*	*[0.19]*	*[0.88]*
GirlxFather	-0.011^{**}	−0.016	−0.046	0.009	−0.025
	(0.005)	(0.013)	(0.033)	(0.025)	(0.019)
	[0.04]	[0.20]	[0.17]	[0.73]	[0.19]
	[0.22]	*[0.26]*	*[0.26]*	*[0.73]*	*[0.26]*
R^2	0.067	0.144	0.107	0.117	0.091
Num. obs.	26476	26476	26476	26373	26363
N Clusters	50	50	50	50	50

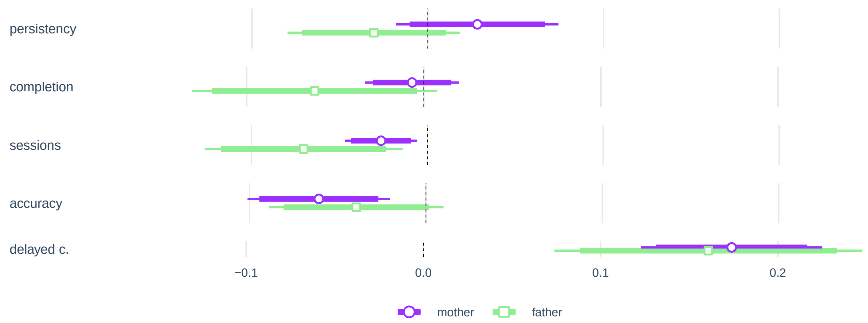
Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. Original p-values and Benjamini & Hochberg (1992) adjusted p-values *(italic)* are given in brackets. Each regression includes a constant term, province, and contract year-month fixed effects. Control variables: age, number of contracts, type of the first contract.

Table A.18.: Full sample - Registration probability by gender norms municipalities

	Mother	Mother
Chores	−0.005	
	(0.007)	
EmpGendRatio		0.005
		(0.013)
Girl	0.044***	0.044***
	(0.011)	(0.011)
Age	−0.002	−0.002
	(0.001)	(0.001)
Health Cond.	0.073***	0.074***
	(0.011)	(0.011)
Num of contr	−0.004***	−0.004***
	(0.001)	(0.001)
1st C:12	−0.086***	−0.086***
	(0.011)	(0.011)
1st C:3	−0.010	−0.010
	(0.009)	(0.009)
Constant	1.519***	1.514***
	(0.032)	(0.032)
Province FE	Yes	Yes
Cont. Month FE	Yes	Yes
R^2	0.013	0.013
Num. obs.	24896	24896
N Clusters	50	50

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are clustered at the province level. The coefficient estimates are from the regressions of the dummy variable which takes a value of 1 if the mother registers the child, and 0 otherwise on the gender norms measurement variable (chores and EmpGenRatio in two columns) and the control variables.

Figure A.12.: Normalization within age - Siblings sample



Notes: The Figure shows the estimated gender gaps for different outcomes in mothers' and fathers' siblings' samples separately, based on Equation 2, where the outcome variables are normalized within each age. Control variables are the age of the child, the presence of a health condition, and parent fixed effects. Standard errors are clustered at the province level. Thin and thick lines represent the 95% and 90% confidence intervals, respectively.

3. Arab Spring and Women's Economic Empowerment

3.1. Introduction

Democratization processes and social movements stand as powerful agents for change, driving societal transformation across various dimensions. They play a crucial role in advancing human rights, fostering political stability, promoting economic development, facilitating social inclusion, and advocating for equality and social justice. Additionally, the process of democratization and social movements often involve citizen activism, including women's participation in advocating for democratic reforms, the formation of women's rights movements, and the amplification of women's visibility and representation.

Arab Spring movements constituted a series of pro-democracy uprisings and protests across the Middle East and North Africa in late 2010, 2011, and beyond. Women actively participated in these movements, sometimes taking on leadership roles (Singerman, 2013). They were among the first to protest at iconic sites such as Avenue Habib Bourguiba in Tunisia, Tahrir Square in Egypt, and Pearl Roundabout in Bahrain (Shihada, 2011). The participating women represented a diverse spectrum, including individuals of various ages, religious affiliations, conservative and liberal beliefs, veiled and unveiled, and spanning different socioeconomic backgrounds (Khamis, 2011). Beyond physical protests, women's involvement in the Arab Spring extended into the digital realm. Female social media organizers and journalists played a crucial role in documenting and disseminating information about the revolutions. As noted by Shihada (2011), they utilized platforms like Twitter and other digital communication channels to tweet, film, and report on the unfolding events. This virtual activism provided a global audience with real-time insights and increased the visibility of the women during these uprisings. Another significant event during the protests was that Tawakkol Karman, a Yemeni journalist, became the first woman to win the Nobel Peace Prize from the Arab world 'for their non-violent struggle for the safety of women and for women's rights to full participation in peace-building work' (Nobel Prize Outreach AB, 2024).

In this paper, we explore the impact of the Arab Spring movements on women's economic empowerment through potential shifts in beliefs and aspirations. Although these movements might have potentially affected the beliefs and aspirations, as well as the economic outcomes of male immigrants as well, we expect to see a significant impact specifically for females since women gained much more visibility during the uprisings. We first analyze the effect of these movements on the aspirations of immigrants in Spain from the Middle East and North Africa (MENA) region. To do so, we construct an index of progressive values which proxy beliefs and aspirations, using information from the European Social Survey. Employing a difference-in-differences approach, we find an increase in this index of progressive values after the Arab Spring for female MENA immigrants residing in Spain, relative to their non-MENA counterparts. However, we do not observe any relative change in this index for male MENA immigrants after the Arab Spring.

Shifts in aspirations, triggered by the Arab Spring, could lead to changes in education and labor market outcomes. We follow the epidemiological approach which is used to identify the causal impacts of culture on economic outcomes by using the variation in cultural values across different immigrant groups residing in the same country and sharing a common institutional and economic environment (Fernandez, 2011). We go one step further and explore whether the shift in aspirations caused by the Arab Spring could result in changes in the education and labor market outcomes of MENA immigrants. Following the epidemiological approach literature, we restrict our attention to second-generation immigrants to minimize the role of institutions in the country of origin. Moreover, this restriction is important in our study since it eliminates the compositional change in the sample of immigrants that we focus on that could potentially be caused by an immigrant inflow after the Arab Spring movements.

To examine the impact on education and labor market outcomes of second-generation MENA immigrants, we use data from the Spanish Labor Force Survey spanning 15 years and employ a dynamic difference-in-differences approach. Comparing the evolution of outcomes of the second-generation immigrant women from MENA countries with other second-generation immigrant women, we find an increase in educational attainment and the probability of being in formal education, substantially closing the gap between second-generation female immigrants from MENA and non-MENA countries. Also, we find a decrease in women's probability of being NEET (not in education, employment, or training), and their employment probability. However, we do not find any significant changes in the outcomes of males following the Arab Spring.

This paper mainly contributes to three strands of the economics literature. First, we add to the literature on political and social protests and economic outcomes. A recent example of this literature from the developed country contexts is the one on the #MeToo movement (see, for example, Cheng and Hsiaw, 2022; Batut et al., 2021; Levy and Mattsson, 2023). This literature examines the impact of the political and social movements in a given country on the outcomes of individuals living there. Differently, we are analyzing the cultural spillover effects of one movement that spread across a wide region. Moreover, unlike the previous studies, we provide evidence on the impacts of social movements exceeding the regions they take place, by focusing on the children of the MENA immigrants residing in Spain. Second, we contribute to a specific part of this literature, on the impacts of the Arab Spring movements. Previous work examines its effect on women's say regarding decisions on household expenditures, tolerance towards domestic violence, and the intention to circumcise daughters (Bargain et al., 2019), labor market outcomes (El-Mallakh et al., 2018; Ghazalian, 2022), marriage and fertility (Ferhat et al., 2022), social trust, uncertainty, education, and health (Liu et al., 2019). Except for Ghazalian (2022), the previous work focuses on Egypt and exploits the regional variation in the intensity of violence during the protests to identify the impact. Differently, transcending the country borders and the borders of the MENA region, we use the epidemiological approach and focus on the second-generation immigrants in Spain from the MENA region. By using this strategy, we compare the outcomes of second-generation immigrants from MENA countries with the outcomes of second-generation immigrants from non-MENA countries. This allows us to isolate our results from the effect of changing economic and political institutions as well as the uncertainty caused by the uprisings within the region where the protests happen. In this way, we are able to identify the effect resulting from the potential shifts in the norms, beliefs, culture, and aspirations following the Arab Spring. Lastly, we contribute to the economics literature on the role of culture and norms. By using a similar approach to ours, Nollenberger et al. (2016) and Rodríguez-Planas and Nollenberger (2018) study the effect of culture and norms of the country of origin on the gender gap in the test performance of second-generation immigrant children in Europe. Similarly, Neuman (2018) examines the labor force participation of immigrant women depending on their country of birth. We move one step further on this approach and examine whether a potential change in the norms of the country of origin caused by the Arab Spring shifts the immigrants' cultural values and aspirations resulting in a progressive shift in their education and labor market outcomes.

3.2. Data

First, we investigate the potential shift in the norms and aspirations among MENA immigrants residing in Spain, utilizing data from the European Social Survey (ESS) spanning waves 1-9, corresponding to the years 2003 to 2019 on a biannual basis.[1] To do so, we construct an index of progressive values through Principal Component Analysis (PCA).

Table 3.1.: Descriptive statistics (ESS-Spain)

	All immigrants	First-generation	Second-generation
Number of observations			
MENA	244	206 (84%)	38 (16%)
non-MENA	1568	1281 (82%)	287 (18%)
Mean age (SD)			
MENA	37.28 (14.04)	38.08 (13.87)	32.95 (14.37)
non-MENA	40.02 (15.28)	39.97 (13.79)	40.23 (20.65)
Female (%)			
MENA	39	35	60
non-MENA	54	53	58
MENA immigrants' country of birth (%)			
Morocco	93.20	78.69	0

Notes: This table shows the number of observations, age and gender distributions, and country of birth share of immigrants in the ESS sample residing in Spain, observed in waves 1-9, constituting our analysis sample. The table includes the observations for which the index of progressive values is defined, i.e., the individuals whose answers for the questions used to construct the index are not missing (82% of all immigrants appearing in the sample).

In Table 3.1, we present the number of observations, age and gender distributions, and country of birth details for immigrants residing in Spain as observed in repeated cross-sections in the ESS sample during waves 1-9, which constitutes our main sample.[2] The table shows that 244 (13%) of these immigrants originate from the MENA region, with 82-84% being first-generation immigrants. Given the limited presence of 38 second-generation immigrants throughout waves 1 to 9, our analysis includes all immigrants, with separate results reported for first and second-generation immigrants as well. An additional rationale for not exclusively focusing on second-generation immigrants is their higher mean age compared to the primary age group that we focus on for the next part of our analysis, by using data from the Spanish

[1]We do not include wave 10 (2022) to avoid the potential effects of the COVID-19 pandemic.

[2]The table includes the observations for which the index of progressive values is defined, i.e., the individuals whose answers for the questions are used to construct the index are not missing (82% of all immigrants appearing in the sample).

Labor Force Survey. While our focus in that is the 20-24 age range, the mean ages for second-generation MENA and non-MENA immigrants are 33 and 40, respectively.

Table 3.1 further illustrates that 39% of all MENA immigrants in our sample are females, resulting in a slightly unbalanced gender distribution. This imbalance arises from males being more prevalent in migration from the MENA region, with female first-generation immigrants constituting 35%. Within our sample, 93% of all MENA immigrants were born in Morocco, comprising 79% of first-generation MENA immigrants. This percentage is naturally zero for second-generation immigrants as they were all born in Spain to first-generation immigrant parents.

We construct our main outcome variable, the index of the progressive values, as a measure of the norms and aspirations. We perform a principal component analysis with the six variables that we standardize, (detailed in Table A.1) and use the resulting first component as an index of progressive values.[3] Considering the discrete nature of these variables, we also construct an alternative index following the "polychoric PCA" method (Kolenikov and Angeles, 2009). Our analysis yields consistent findings using both indices.

For the next part of our analysis focusing on the implications of the changes in aspirations for education and labor market outcomes, we use data from the Spanish Labor Force Survey provided by the Spanish Statistical Office (INE). We pool the publicly available micro-datasets from the 2nd quarter of each year between 2005 and 2019.[4] The resulting dataset for our analysis is a pool of repeated cross-sections for 15 years.

This comprehensive dataset allows us to identify the second-generation immigrants who live in the same household as their parents since all household members are asked about their country of birth. To minimize a potential consequent selection bias, we focus on the age group "20-24", rather than older ages.[5] Table 3.2 shows the country of birth group distribution of mothers and fathers of the second-generation immigrants in our analysis sample. We define the MENA region as the combination of "Africa" and "Western Asia (Middle East)", and the Latin America region as the combination of "Central America and Caribbean" and "South America".

A "second-generation immigrant" throughout this paper is defined as an individual

[3] We select the six relevant variables in our context (outlined in Table A.1) from the "human values" section of the ESS which is consistently included in each wave. Although in principle we could consider other variables available in at least one wave before and after the Arab Spring to be able to examine changes, our preference is to incorporate information from as many waves as possible to comprehensively understand potential shifts in the norms and values, especially considering the low number of observations of MENA immigrants per each wave.

[4] We choose the 2nd quarter of each year to minimize labor market seasonality effects and do not include data from 2020 onwards to avoid the COVID-19 period.

[5] Publicly available datasets include the age groups and country of birth groups rather than each age and each country of birth.

Table 3.2.: Country of origin groups of the parents

	Mothers		Fathers	
	N	%	N	%
Africa and Western Asia (Middle East)	872	14.73	850	16.98
Central America and Caribbean	311	5.25	178	3.56
South America	1,531	25.86	1,070	21.37
EU-15	1,194	20.17	632	12.62
EU-28 (not EU-15)	34	0.57	20	0.40
Rest of Europe	117	1.98	42	0.84
North America	44	0.74	18	0.36
East Asia (Far East)	27	0.46	23	0.46
South and Southwest Asia	59	1	65	1.3
Oceania	6	0.1	4	0.08
Spain	1,726	29.15	2,104	42.03
Total	5,921	100	5,006	100

Notes: This table presents the country of origin group distribution of mothers and fathers of the second-generation immigrants in our analysis sample derived from the Spanish Labor Force Survey. We define the MENA region as the combination of "Africa" and "Western Asia (Middle East)", and the Latin America region as the combination of "Central America and Caribbean" and "South America".

who was born in Spain, with at least one parent born abroad. Therefore, as shown in Table 3.2, our sample includes individuals with one parent born in Spain in the case where the other parent was born abroad.

Our treatment group consists of second-generation immigrants with at least one parent born in the MENA region and the control group consists of second-generation immigrants with no parents from the MENA region. To address concerns about the comparability of these two groups, we repeat our analysis by restricting the control group to second-generation immigrants with at least one parent from the Latin American region. Nevertheless, the absence of pre-trends in our baseline results as well as the results with Latin America as the control group alleviates these concerns.

To measure the economic empowerment of young second generation MENA women, we focus on education and labor market outcomes. Specifically, we examine the rates of upper-secondary (*bachillerato* in Spain) or higher educational attainment, the current enrollment status in formal education, the prevalence of the "Not in Education, Employment, or Training" (NEET) category, and the probability of employment.

Table 3.3 provides key demographic and educational characteristics for the treatment group and both control groups, where Latin America is a subset of the main control group non-MENA. As shown in the table, 45-47% of the second-generation immigrants are females for each group, and only 32% of second-generation immigrants are born to both parents from the MENA region. In our further analysis, we restrict the treatment group to those individuals with both MENA-origin parents, as

Table 3.3.: Descriptive statistics (Spanish Labor Force Survey)

	MENA	non-MENA	Latin America
Second-generation immigrant (%):			
Female	47	46	45
Spanish citizen	94	96	96
Lower-secondary educational attainment	92	95	95
Parents' country of origin (%):			
Mother from the country group	65	-	68
Father from the country group	70	-	54
Both parents from the country group	32	-	16
Mother's educational attainment (%):			
Primary or lower	41	13	12
Lower-secondary or upper-secondary	39	54	50
Higher education	20	33	38
Father's educational attainment (%):			
Primary or lower	38	14	11
Lower-secondary or upper-secondary	38	52	47
Higher education	24	33	42
Number of observations	1058	5024	1301

Notes: This table presents the key demographic and educational characteristics for the treatment group and both control groups, where Latin America is a subset of the main control group non-MENA. The information is derived from the Spanish Labor Force Survey. All numbers represent percentages except for the number of observations.

well as those with a MENA-origin mother only, and those with a MENA-origin father only. It is noticeable from the table that the parents from MENA countries have lower educational attainment compared to the rest. Also, there is some variation at the individual level regarding holding Spanish citizenship which might be correlated to education and labor market outcomes. Therefore, we control for the educational attainment of both parents separately as well as the Spanish nationality in our main specification.

3.3. Empirical Approach

To investigate the impact of the Arab Spring on beliefs and aspirations, we estimate the following linear regression:

$$Y_{it} = \alpha + \beta MENA_i + \theta After_t + \gamma MENA_i \times After_t + \varepsilon_{it} \tag{3.1}$$

where Y_{it} is the index of the progressive values (calculated by PCA or polychoric PCA) of individual i, measured in year t. $MENA_i$ is the indicator variable that takes a value of 1 if individual i is a MENA immigrant, and 0 if the individual is a non-MENA immigrant. $After_t$ is an indicator that takes a value of 1 if $t \geq 2011$, and 0 otherwise. Our coefficient of interest is γ which is the differential change in the index of the progressive values for MENA immigrants compared to non-MENA immigrants after the Arab Spring.

Next, to analyze the impact of the Arab Spring on second-generation immigrants' education and labor market outcomes, we estimate the following linear regression for females and males, separately, using linear probability models:

$$Y_{ijt} = \alpha + \beta M_{ijt} + \theta_t + \sum_{t=2005-06, t\neq 2011-12}^{t=2017-2018} \gamma_t M_{ijt} \mathrm{I}_{\theta_t} + \delta X_{ijt} + \eta_j + \varepsilon_{ijt} \qquad (3.2)$$

where Y_{ijt} is the binary outcome variable (upper-secondary educational attainment, in formal education, NEET or employment) of individual i, residing in the region (autonomous community) j, and observed at the year pair t. M_{ijt} is the indicator of whether individual i is a second-generation immigrant from MENA (0 if the individual is a second-generation immigrant from another country), θ_t is the year pair fixed-effects[6], I_{θ_t} is the binary indicator for year pair θ_t, X_{ijt} are the individual characteristics (parental educational attainment[7] - for the mother and the father, separately - and Spanish citizenship), and η_j is the region (autonomous community) fixed-effects. Since our dataset does not have a panel structure, we are comparing different cohorts of second-generation immigrants in our analysis. The coefficient of interest is γ_t which is the differential evolution of the outcomes of the second-generation individuals with parents from MENA, compared to those second-generation individuals with parents from other countries.

The underlying assumption for a causal interpretation of our coefficients of interest ($\gamma_{2013-14}$, $\gamma_{2015-16}$, and $\gamma_{2017-18}$) is that in the absence of the Arab Spring, the evolution of the outcomes of MENA and non-MENA second-generation immigrants would not have been significantly different from each other. One supportive evidence for this assumption is the absence of pre-trends which would be implied by coefficient estimates of $\gamma_{2005-06}$, $\gamma_{2007-08}$, and $\gamma_{2009-10}$ that are insignificant, which we observe in our findings.

Determining the base year relative to which we estimate the impact of the Arab

[6]Since per each year we observe a very small sample of second-generation MENA immigrants for males and females, separately, we group two consecutive years over the period that we focus on to increase the precision of our estimates. For this reason, we drop the last year 2019 from the estimation sample. We also present our results for the baseline from yearly estimations where 2019 is included as well, in subsection 3.4.2.

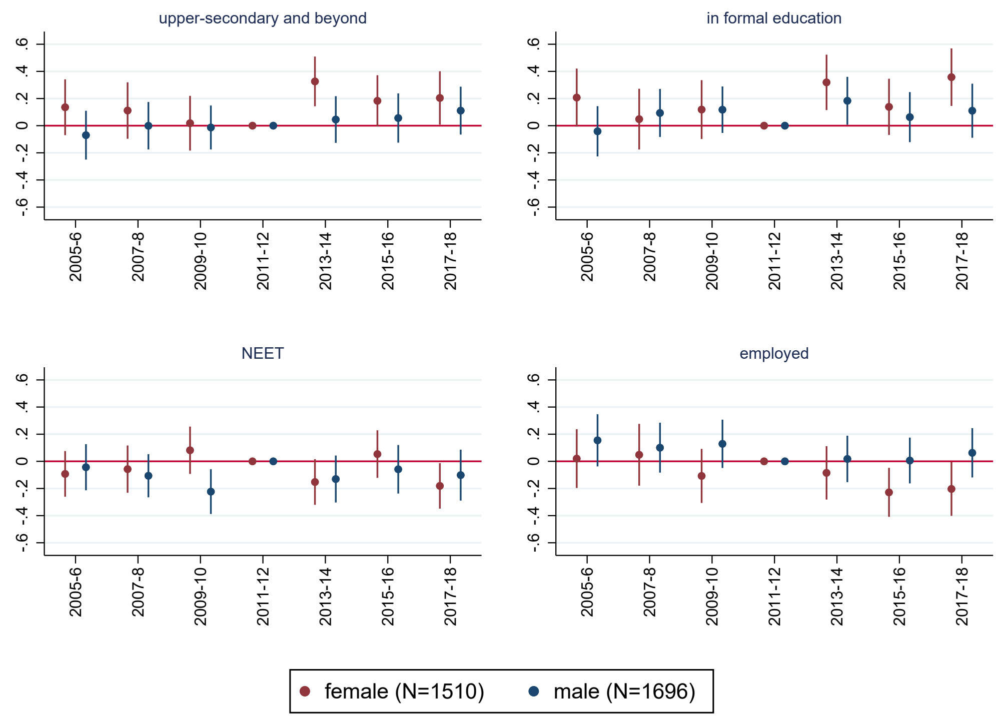
[7]Three categories of the educational attainment variable: (i) primary school or lower, (ii) lower-secondary or upper-secondary (including both vocational and general tracks), and (iii) higher education

Spring is not straightforward, given that it is a series of movements that spread across a wide region and lasted for years. The uprisings commenced on December 17, 2010, in Tunisia, and continued predominantly in 2011 and 2012 as they spread across the region. As depicted in Figure A.1, Google search trends for the Arab Spring topic globally show two peaks corresponding to October 2011 and September 2012. Similarly, in Spain, the term was most popular in 2011 and 2012. The figure suggests that the events garnered attention in these two years, both globally and in Spain. Therefore, for our empirical analysis, we consider the base year pair as 2011-12.[8]

Although the Arab Spring started at the end of 2010, and continued in 2011, 2012, and beyond, we would like to address the potential effect of the Spanish financial crisis that started in 2008 that our estimations might be partly capturing. For our coefficient estimates to capture the effect of this crisis, second-generation MENA immigrants or their parents should have been differently affected by the financial crisis. To examine whether this is the case, we conduct a robustness check in section 3.5 by comparing the changes in employment probabilities and occupation distributions of MENA and non-MENA first-generation immigrants. Moreover, we show in subsection 3.4.2 that the employment probability and labor force participation of the parents of second-generation MENA immigrants relative to that of other second-generation immigrants in our main analysis sample do not change significantly over time.

[8] In the yearly analysis that we present in subsection 3.4.2, the omitted year is 2011. Supporting our choice of the base year, and the choice of the base year pair in the main analysis, we have statistically insignificant coefficient estimates for years 2010 and 2012.

Figure 3.1.: Arab Spring and second-generation immigrants



Notes: This figure shows coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

3.4. Results

3.4.1. Beliefs and Aspirations

Given the increased visibility of women during the Arab Spring uprisings and their important role in the protests, in this section, we investigate potential shifts in beliefs, norms, and aspirations of the immigrants from the MENA region, proxied by the index of progressive values. Table 3.4 presents results from estimating Equation 3.1 separately for females and males. The outcome variables in different columns are the indices of progressive values, constructed using both PCA and polychoric PCA methods. In Panel A, the first column indicates a significant increase in the index of progressive values for female MENA immigrants compared to their non-MENA counterparts after the Arab Spring. These results remain robust when using the alternative polychoric PCA method, as shown in column 2. Conversely, for male

MENA immigrants, there is no significant change in the index of progressive values compared to male non-MENA immigrants after the Arab Spring.

Shifting our focus to Panel B which presents the coefficient estimates for the first-generation immigrant subsample, we observe results that are similar to those for the entire immigrant sample, holding for both females and males. However, in Panel C, examining the second-generation immigrant subsample, we find no statistically significant changes in the relative values of the index of progressive values after the Arab Spring for either gender. Despite the lack of statistical significance, possibly due to a limited number of observations, we note positive coefficients for both female and male second-generation immigrants concerning the interaction term. The estimates for female second-generation immigrants are in the same direction as those of female first-generation immigrants, however, it diverges for second-generation males compared to their first-generation counterparts. Yet, we do not interpret the results in Panel C as conclusive evidence, given the large standard errors and lack of statistical significance.

3.4.2. Education and Labor Market Outcomes

In this section, we take one step further from the impact of the Arab Spring movements on beliefs and aspirations and explore its effects on the education and labor market outcomes of second-generation immigrants residing in Spain.

Figure 3.1 presents the baseline results where we compare second-generation immigrants aged 20 to 24 with at least one parent from MENA to those second-generation immigrants in the same age group, with neither parent from MENA.[9] We estimate Equation 3.2 for females and males separately, and for both genders, we do not find significant differences between MENA and non-MENA second-generation immigrants in the years preceding the Arab Spring uprisings.[10]

We observe an increase in the probability of completing upper-secondary education or higher education for females, while there is no significant change for males. The impact is more profound for the years 2013-2014 where we also find an increase in the probability of being in formal education as well as a decrease in NEET (although it is not very precisely estimated.) For the next two pairs of years, our results suggest an increase in the probability of completing upper-secondary education or beyond where the magnitude of the estimate is smaller compared to the first period after the Arab Spring. Moreover, we find a decrease in the employment probability for these two periods, along with an increase in the probability of being in formal education

[9] We present the magnitudes of the coefficients in Table A.2.

[10] One exception is the NEET outcome where we observe a negative and significant coefficient estimate for males in the year pair 2009-10.

and a decrease in NEET for the last period. While we observe these changes for females, we do not observe any significant change in the outcomes of males except for the increase in the probability of being enrolled in formal education for the years 2013-2014 which is significant at the 10% level.[11]

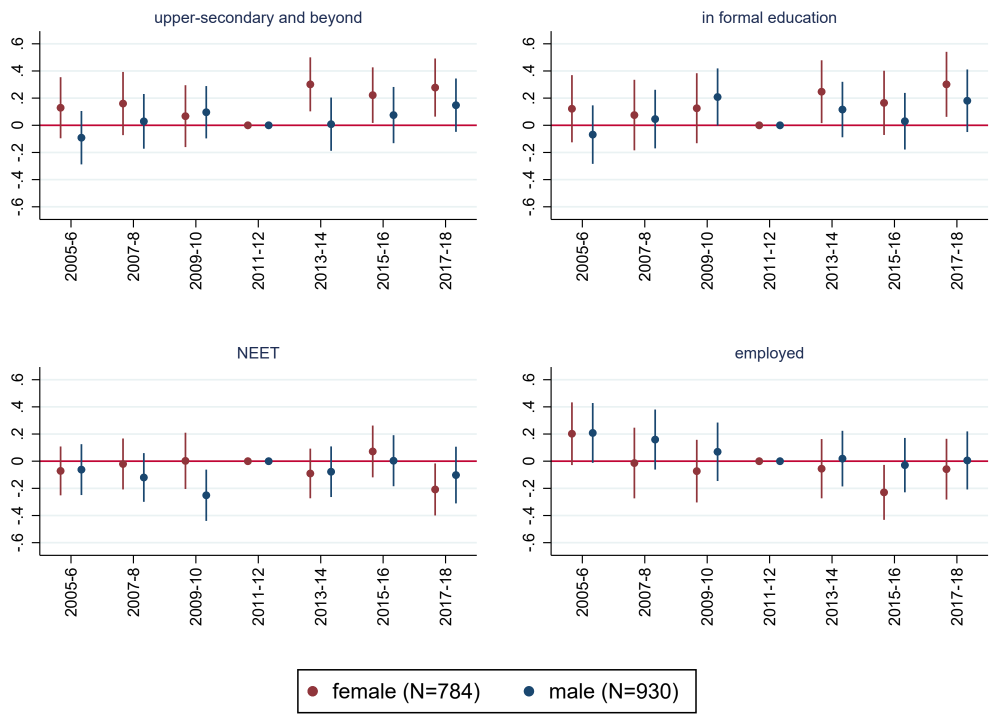
To interpret the magnitudes of the effects, we also estimate Equation 3.2 by using a logit model. Table A.3 shows that the signs of the coefficients are the same as the ones estimated by using a linear probability model shown in Table A.2, and they suggest very similar conclusions in terms of the significance levels. We also present the marginal effects estimated following the logit regressions for MENA and non-MENA groups, relative to the year pair 2011-12. As presented in Table A.4, the probabilities of completing upper-secondary education and beyond, and being in formal education increase by 21 and 23 percentage points, respectively, for second-generation female MENA immigrants in 2013-14. The effect on upper-secondary educational attainment decreases to 14-15 percentage points in the following periods, while the magnitude of the estimate remains similar in the third period for the probability of being in formal education. For these educational outcomes, the gaps between the MENA and non-MENA female immigrants substantially closes after the Arab Spring. We find 11 and 19 percentage points decreases for NEET in the third period and for employment probability in the second period, which could be considered as results of increased participation in educational activities. For males, as reported in Table A.5, there is only an increase of 21 percentage points in the probability of being in formal education in the year pair 2012-13 and a decrease in NEET as its consequence. However, we do not observe any significant change in other outcomes and other year pairs.

As seen in Table 3.2, most of the first-generation immigrants are from EU-15, Latin America, or MENA region, and our control group mostly includes second-generation immigrants with parents from EU-15 and Latin America. To alleviate concerns regarding the EU-15 immigrants being potentially far away from MENA immigrants in terms of cultural heritage and socio-economic background, we repeat our analysis with the control group restricted only to immigrants from Latin America. Since our focus is the second-generation immigrants who were raised in Spain, we do not have the concern that the Latin American second-generation immigrants would have a language advantage over the MENA second-generation immigrants. Figure 3.2 presents the estimates where the control group is the second-generation immigrants with at least one parent from Latin America.[12] As the graphs show,

[11]Figure A.2 presents the results from the yearly estimation, where we do not group two consecutive years. Results are very similar to the baseline, however very noisy because of the limited number of observations of second-generation MENA immigrants per year.

[12]We assign 4 individuals (0.2% of the whole sample) whose one parent is from MENA and the

Figure 3.2.: Arab Spring and second-generation immigrants - Control group: Latin America



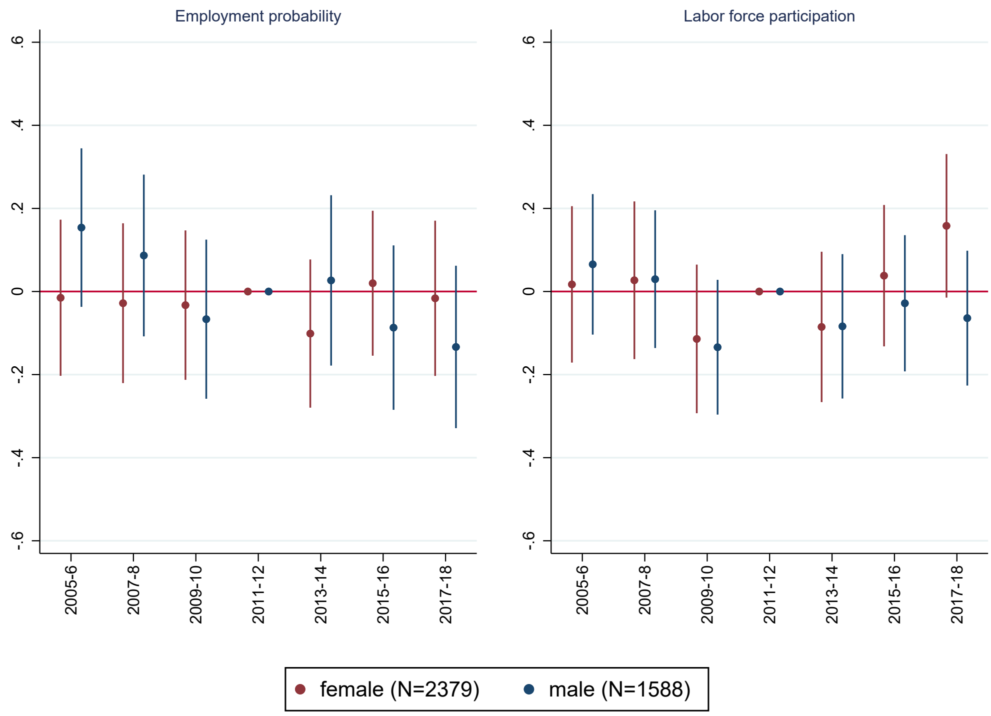
Notes: This figure shows coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation immigrant from the Latin America, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

our findings are very similar to the baseline results. As a further robustness check, we take the native Spanish individuals of the same age range as the control group. Figure A.3 in the Appendix shows that results are very close to the baseline results.

Our findings so far suggest an increase in upper-secondary educational attainment, an elevated likelihood of being in formal education, a decline in NEET status, and reduced employment probability for female second-generation MENA immigrants in comparison to their non-MENA counterparts, following the Arab Spring. The reduced employment probability could potentially caused by the increased participation in education within the same age group. However, it is also plausible to consider that the Arab Spring may have influenced labor market dynamics within households with one parent originating from the MENA region. This could be explained by

other parent is from Latin America to the treatment (MENA) group, since our treatment definition is having at least one parent from MENA.

Figure 3.3.: Arab Spring and first-generation immigrant parents



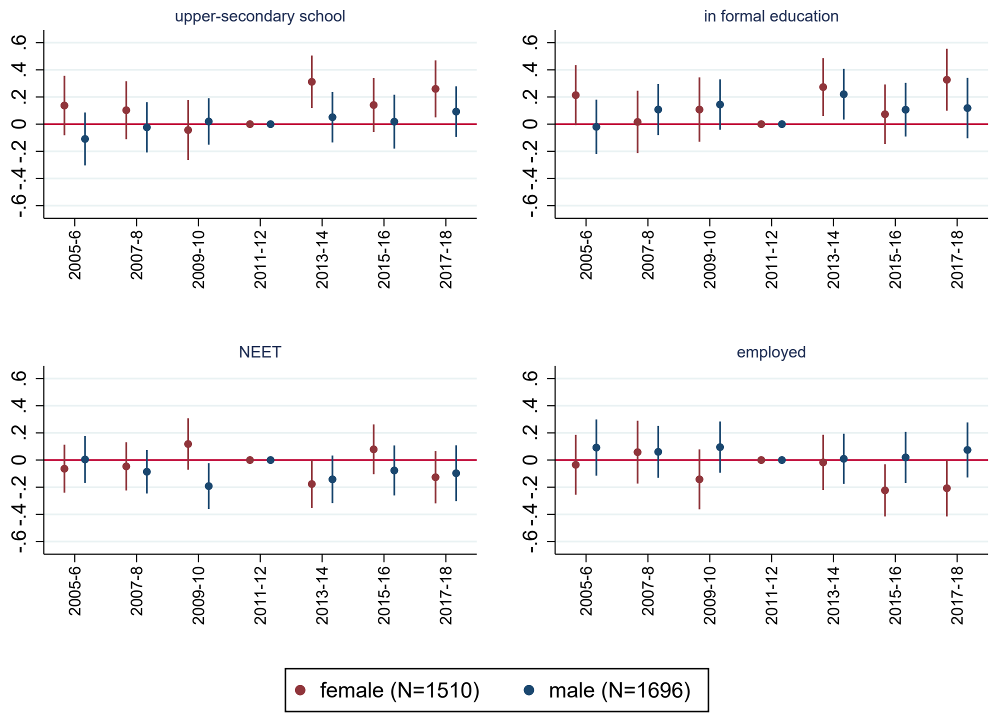
Notes: This figure shows coefficient estimates and the 95% confidence intervals from estimating regressions of employment probability or labor force participation on an indicator that takes a value of 1 if the immigrant is a parent of a second-generation MENA immigrant and 0 if they are a parent of a second-generation non-MENA immigrant, year pair fixed effects, and their interactions. The coefficient estimates presented in the graphs are the ones of the interaction term of the parent of a second-generation MENA immigrant indicator and the year pair, where the year pair 2011-12 is omitted.

motives such as an emerging need to remit funds or host family members leaving the region after the uprisings. While such a change might potentially impact the medium and long-term objectives of second-generation MENA immigrants, it seems less likely that such effects would be gender-specific.

To explore the possibility that our findings are responses to economic needs or household employment dynamics stemming from the Arab Spring, we analyze the evolution of outcomes for immigrant parents within our main analysis sample. Specifically, we compare the outcomes of parents who migrated from the MENA region with those of parents who migrated from other countries. It is crucial to note that these are not all immigrant parents in the survey sample but rather the parents of individuals in our main analysis sample — those with at least one child aged 20-24 during the relevant survey period. Figure 3.3 displays the results separately for males (fathers) and females (mothers) for the employment and labor force participation outcomes. As seen in the graphs, we find no significant effect of the Arab Spring on first-generation immigrant parents' employment probability and labor force participation.

Next, to analyze potential heterogeneities in our findings depending on which parent is a first-generation immigrant from the MENA region, we estimate Equation 3.2 for three different subsamples of second-generation immigrants. Figure A.4, Figure A.5, and Figure A.6 present results for second-generation immigrants with both parents from MENA, with the mother from MENA, and with the father from MENA, respectively. Figure A.4 shows that the results are not statistically significant for the subsample of the second-generation immigrants with both parents from the MENA region, and the coefficient estimates are smaller except for the outcome of the probability of being in formal education. Figure A.5 and Figure A.6 present results similar to the baseline for the subsamples of second-generation immigrants with mothers and fathers, respectively, from the MENA region. Although some coefficient estimates are statistically significant in the fathers' sample, but not in the mothers' sample, their confidence intervals mostly overlap, suggesting no significant differences between the two subsamples.

Figure 3.4.: Arab Spring and second-generation immigrants: compositional change



Notes: This figure shows coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, and region of residence (autonomous community) fixed effects. Additionally, these regressions include the interaction terms of parental educational attainment and year-pair dummies to account for a potential compositional change of immigrants over the years. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

Table 3.4.: Change in the index of the progressive values

	Females		Males	
	PCA	polychoric PCA	PCA	polychoric PCA
A. All immigrants				
MENA	-0.59^{***}	-0.52^{***}	-0.10	-0.04
	(0.19)	(0.17)	(0.17)	(0.13)
after	-0.16	-0.20^{**}	0.05	0.03
	(0.09)	(0.07)	(0.10)	(0.08)
MENA $\times$ after	0.61^{**}	0.57^{**}	-0.20	-0.17
	(0.24)	(0.22)	(0.22)	(0.18)
R^2	0.01	0.01	0.00	0.00
N	937	937	875	875
B. First generation immigrants				
MENA	-0.65^{***}	-0.56^{***}	-0.01	0.02
	(0.20)	(0.19)	(0.17)	(0.14)
after	-0.16	-0.17^{**}	0.11	0.08
	(0.10)	(0.08)	(0.11)	(0.09)
MENA $\times$ after	0.73^{***}	0.65^{***}	-0.32	-0.27
	(0.27)	(0.25)	(0.24)	(0.20)
R^2	0.01	0.01	0.01	0.00
N	747	747	740	740
C. Second generation immigrants				
MENA	-0.35	-0.32	-0.57	-0.42
	(0.50)	(0.42)	(0.83)	(0.63)
after	-0.13	-0.25	-0.20	-0.21
	(0.19)	(0.16)	(0.23)	(0.19)
MENA $\times$ after	0.11	0.24	0.47	0.39
	(0.58)	(0.50)	(0.90)	(0.69)
R^2	0.01	0.02	0.01	0.01
N	190	190	135	135

Notes: This table presents the coefficient estimates from estimating Equation 3.1, which regresses the index of progressive values constructed by PCA or the polychoric PCA method on the indicator of being a MENA immigrant, being observed after the Arab Spring, and their interaction for females and males, separately. Panel A presents the results for all immigrants while Panels B and C show results for the subsamples of the first and second-generation immigrants, respectively. P-values: < 0.01 ***, < 0.05 **, < 0.10 *.

3.5. Robustness checks

In this section, we perform several robustness checks for our results on education and labor market outcomes. First, to examine whether our results capture a part of the effect of the Spanish financial crisis that started in 2008, we present the evolution of employment probabilities and occupational distributions of MENA first-generation immigrants relative to non-MENA first-generation immigrants residing in Spain. Figure A.7 shows that after the 2008 financial crisis, the relative employment probability of male MENA immigrants decreased by 8-10 percentage points compared to their non-MENA counterparts. However, for females, we do not observe any effect on the employment probability of MENA immigrants relative to non-MENA immigrants. Figure A.8 shows the relative changes in occupation categories of the first-generation MENA immigrants relative to their non-MENA counterparts. Overall, there is not a significant trend of change, except for the shift from the service sector to unskilled roles for females, starting in 2014. Since this trend starts six years after the start of the financial crisis, we do not consider this as a concern for our identification.

Our main findings on education and labor market outcomes as well as the beliefs and aspirations are driven by females, and we do not observe changes in most of the outcomes of males. Therefore, the results presented in Figure A.7 and Figure A.8, partly mitigate concerns about the thread of the 2008 financial crisis to our identification. Moreover, as we show in Figure 3.3 and discuss in subsection 3.4.2, employment and labor force participation of the parents of individuals in our analysis sample do not differentially evolve within our period of interest, further supporting our causal interpretation. However, we cannot eliminate the possibility that the differential impacts of the financial crisis on the employment probability of male MENA and non-MENA first-generation immigrants who are not the parents of the second-generation immigrants in our sample have some spillover effects that drive part of our findings on the education and labor market outcomes.

Since our data does not have a panel structure and we focus on a specific age group (20-24) for each period, we would like to address the potential change in the composition of the individuals in different cohorts. For this reason, we control for the interaction of the parent's educational attainment and the year pair. Figure 3.4 shows the results which are very similar to the baseline for the two variables measuring educational outcomes. [13]

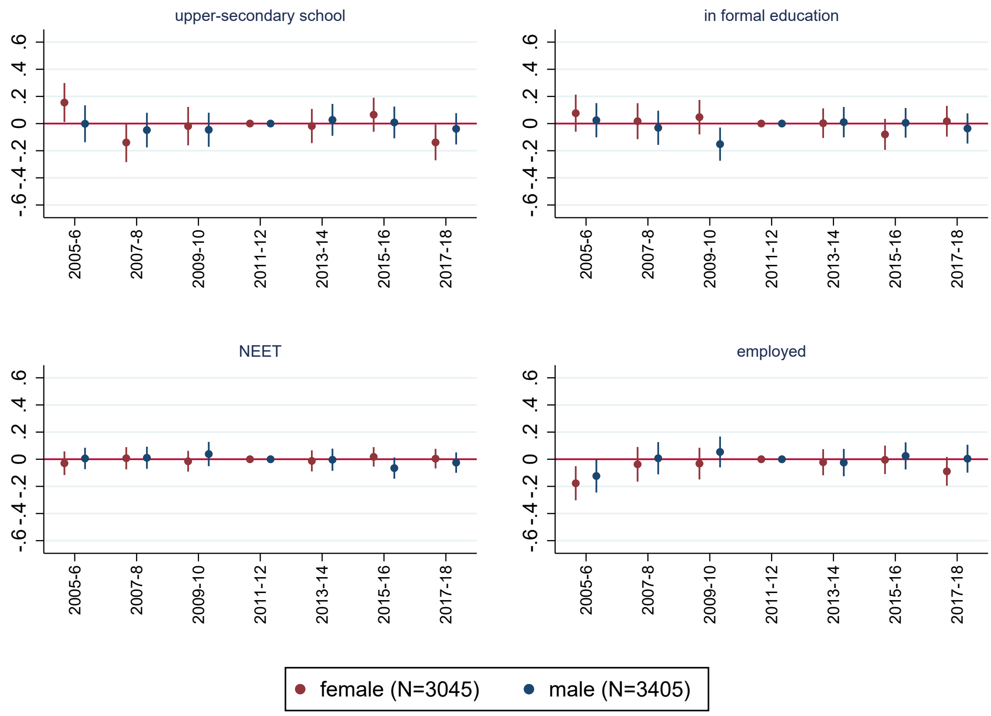
Second, we cluster the standard errors at the regional (autonomous community) level by using wild bootstrap clustering to address the potential shocks that might have occurred at the regional level, since in Spain, Autonomous Communities

[13]There are some changes in the level of statistical significance for the NEET variable where in the baseline the coefficient estimate for the last cohort (2017-18) was borderline significant.

are responsible for some educational policies. We report the resulting p-values in Table A.6 in brackets, along with the original p-values from our baseline specification given in the parentheses. Most of the coefficient estimates support very similar conclusions to the baseline results, while the precision is higher or lower for some cases that are within mostly a 10% significance level using both ways of standard error calculations. Only for the NEET outcome, the coefficients for females and males become statistically insignificant when using wild bootstrap clusters.

Lastly, we conduct a placebo analysis where the placebo treatment group is the second-generation immigrants from Latin America and the control group is the group with neither of the parents from Latin America (also excluding MENA.) Figure 3.5 shows that almost all coefficient estimates are around zero and they are not statistically significant.

Figure 3.5.: Arab Spring and second-generation immigrants: placebo test



Notes: This figure shows coefficient estimates and the 95% confidence intervals from a placebo test regressing 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation Latin American immigrant and 0 if they are a second-generation non-Latin American immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

3.6. Discussion and Conclusion

This study investigates whether the Arab Spring had an impact on the educational attainment and labor market outcomes of second-generation immigrant women in Spain with MENA-origin parents, through shifts in beliefs and aspirations. During these movements, awareness about women's empowerment was raised, influencing the educational and career aspirations and beliefs of young women.

We first show that there is an increase in the index of progressive values - proxying beliefs, and aspirations - of MENA female immigrants compared to their non-MENA counterparts after the Arab Spring, while there is no significant change in the index for males.

Next, we find an increase in educational attainment and the probability of being in formal education of female second-generation MENA immigrants, closing the gap between the MENA and non-MENA second-generation female immigrants to a large extent. The positive impact on educational outcomes is accompanied by a decrease in the probability of these women being NEET and in their employment probability. However, we do not find any significant change in the outcomes of male second-generation MENA immigrants. Our results are robust to using alternative control groups and different specifications.

Our findings shed light on the impact of a major social and political event on the aspirations and economic decisions of immigrant populations, highlighting its impact on subsequent generations. In line with our expectations based on the increased visibility and significant participation of women during the Arab Spring protests, our results suggest changes in women's beliefs and aspirations towards a progressive direction and economic empowerment. Further research in these lines can contribute to a deeper understanding of the long-term consequences of such transformative events on individuals and societies, in terms of both societal changes and economic outcomes such as educational choices and labor market trajectories from a perspective of gender equality.

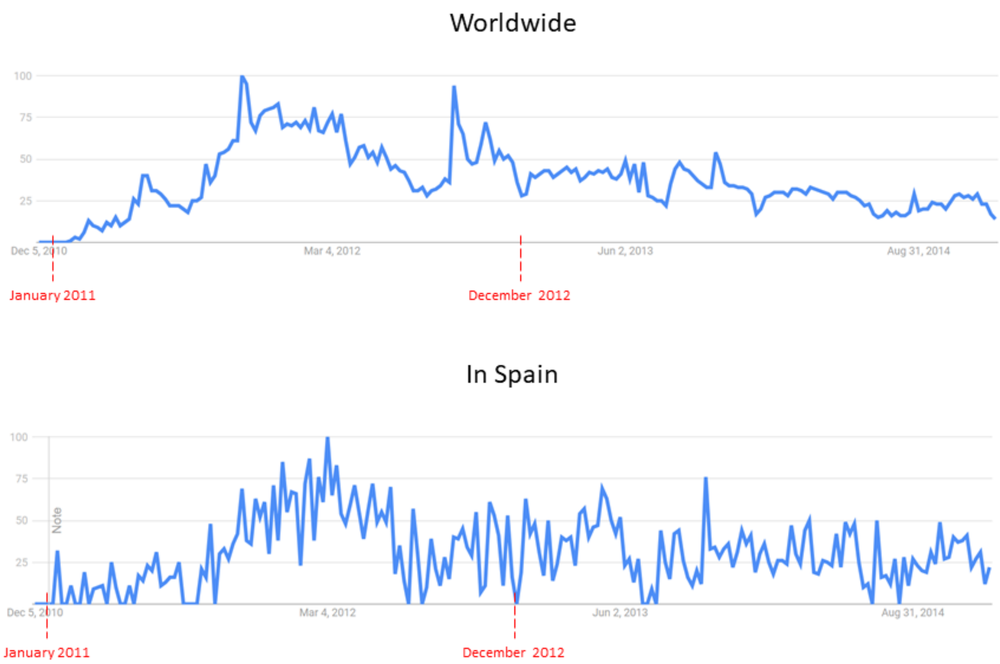
A. Appendix: Additional Tables and Figures

Table A.1.: Variables for the index of the progressive values

1. Variable descriptions
Important that people are treated equally and have equal opportunities *(reversed)*
Important to make own decisions and be free *(reversed)*
Important to be successful and that people recognize achievements *(reversed)*
Important to do what is told and follow rules
Important to be humble and modest, not draw attention
Important to follow traditions and customs
2. Value options
1: Very much like me
2: Like me
3: Somewhat like me
4: A little like me
5: Not like me
6: Not like me at all

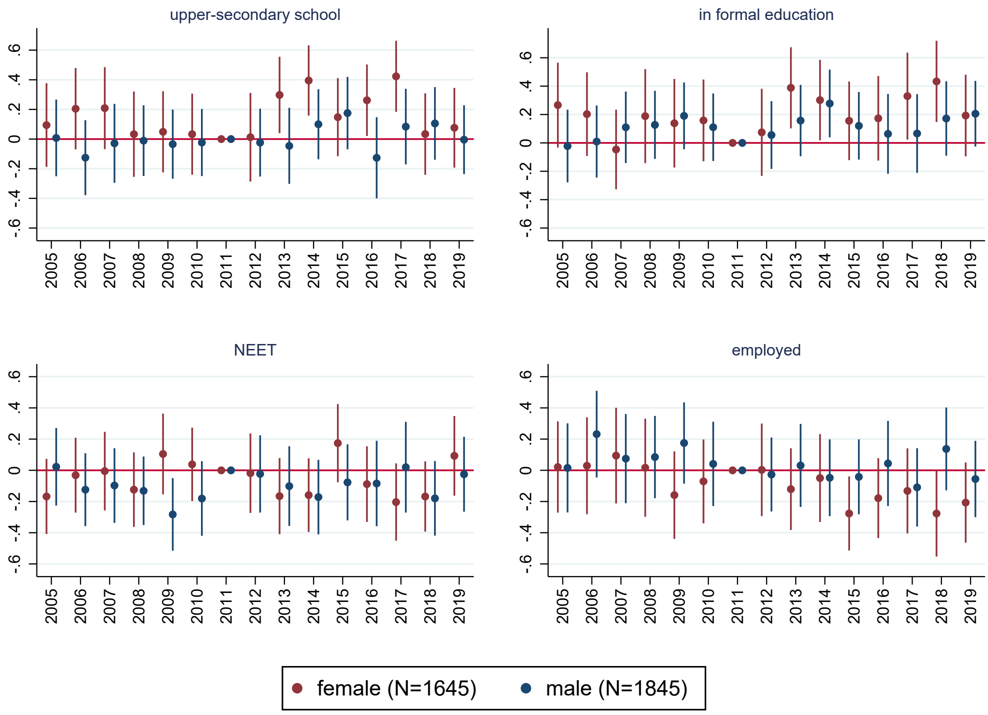
Notes: This table shows the variables that are used to construct the index of the progressive values. These variables are in the human values section of the ESS, and the values are reversed for the first three variables to keep the same progressive direction on the index, i.e., the most progressive for the highest value 6: "not like me at all". The higher values of the index, resulting from the principal component analysis, represent more progressive values according to these variables.

Figure A.1.: Timeline of Google Search Trends



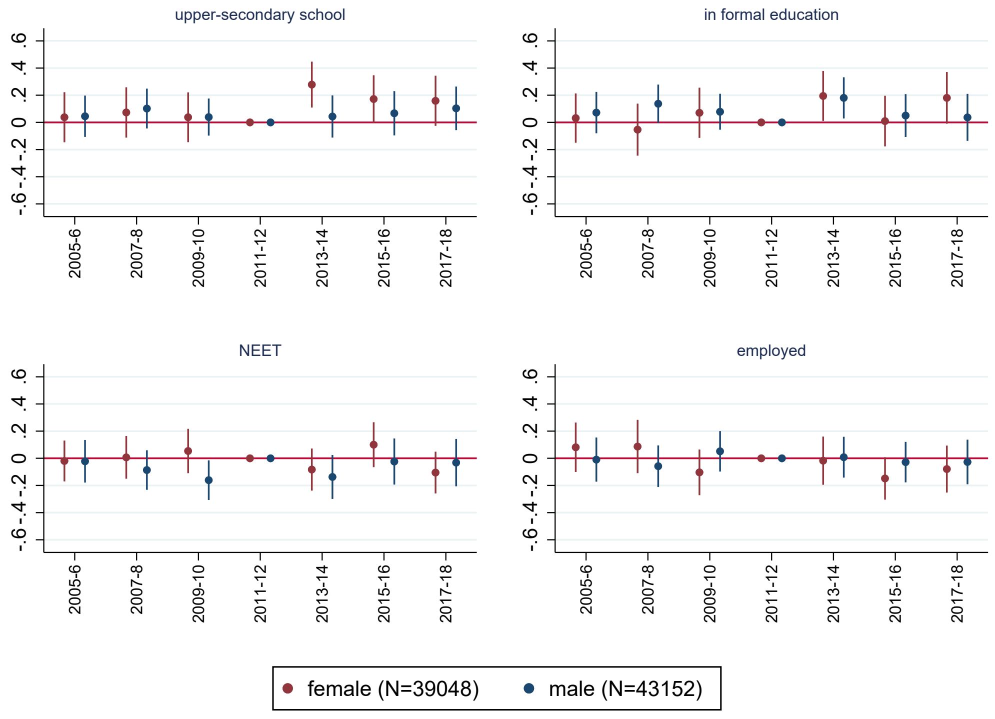
Notes: This figure shows the evolution of Google Search trends for the concept "Arab Spring". Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term. Source: Google Trends.

Figure A.2.: Arab Spring and second-generation immigrants



Notes: This figure shows coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year 2011 is omitted.

Figure A.3.: Arab Spring and second-generation immigrants - Control group: natives



Notes: This figure shows coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are native Spanish, year pair fixed effects, their interactions, mother's and father's educational attainments, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

Table A.2.: Arab Spring and the second-generation immigrants: LPM

	(1) upper-secondary and beyond	(2) in formal education	(3) NEET	(4) employed
Females				
MENA	-0.16**	-0.20**	0.06	0.10
	(0.08)	(0.08)	(0.06)	(0.08)
2013-14	-0.08**	-0.08*	0.08**	0.02
	(0.04)	(0.05)	(0.03)	(0.04)
2015-16	-0.03	-0.07	0.03	0.06
	(0.04)	(0.05)	(0.03)	(0.04)
2017-18	-0.03	-0.07	0.02	0.13***
	(0.04)	(0.05)	(0.03)	(0.05)
MENA × 2013-14	0.33***	0.32***	-0.15*	-0.08
	(0.09)	(0.10)	(0.09)	(0.10)
MENA × 2015-16	0.18*	0.14	0.05	-0.23**
	(0.10)	(0.11)	(0.09)	(0.09)
MENA × 2017-18	0.20**	0.36***	-0.18**	-0.20**
	(0.10)	(0.11)	(0.09)	(0.10)
Mean of the outcome in 2011-12	0.77	0.60	0.17	0.23
R^2	0.14	0.14	0.10	0.08
N	1510	1510	1510	1510
Males				
MENA	0.01	-0.10	0.14**	-0.09
	(0.06)	(0.06)	(0.07)	(0.06)
2013-14	0.02	0.02	-0.01	-0.03
	(0.04)	(0.05)	(0.03)	(0.04)
2015-16	0.04	0.01	0.01	-0.02
	(0.04)	(0.04)	(0.03)	(0.04)
2017-18	0.04	-0.07	0.02	-0.01
	(0.04)	(0.04)	(0.03)	(0.04)
MENA × 2013-14	0.05	0.18**	-0.13	0.02
	(0.09)	(0.09)	(0.09)	(0.09)
MENA × 2015-16	0.06	0.06	-0.06	0.01
	(0.09)	(0.09)	(0.09)	(0.09)
MENA × 2017-18	0.11	0.11	-0.10	0.06
	(0.09)	(0.10)	(0.10)	(0.09)
Mean of the outcome in 2011-12	0.58	0.54	0.21	0.29
R^2	0.21	0.19	0.14	0.07
N	1696	1696	1696	1696

Notes: This table shows coefficient estimates and standard errors in parentheses from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the table are the ones of the interaction term of the second-generation MENA indicator and the year pair for the period after the Arab Spring, where the year pair 2011-12 is omitted. P-values: < 0.01 ***, < 0.05 **, < 0.10 *.

Table A.3.: Arab Spring and the second-generation immigrants: Logit

	(1) upper-secondary and beyond	(2) in formal education	(3) NEET	(4) employed
Females				
MENA	-0.97**	-0.95**	0.47	0.56
	(0.44)	(0.38)	(0.47)	(0.42)
2013-14	-0.57**	-0.44*	0.72**	0.12
	(0.28)	(0.23)	(0.31)	(0.25)
2015-16	-0.22	-0.37	0.23	0.36
	(0.29)	(0.23)	(0.33)	(0.25)
2017-18	-0.20	-0.35	0.20	0.70***
	(0.30)	(0.24)	(0.34)	(0.25)
MENA × 2013-14	2.06***	1.57***	-1.18*	-0.54
	(0.62)	(0.51)	(0.63)	(0.58)
MENA × 2015-16	1.11*	0.69	0.19	-1.60***
	(0.58)	(0.50)	(0.59)	(0.62)
MENA × 2017-18	1.19*	1.73***	-1.42*	-1.14**
	(0.61)	(0.53)	(0.73)	(0.55)
N	1481	1510	1510	1510
Males				
MENA	0.11	-0.52	0.73**	-0.55
	(0.33)	(0.34)	(0.37)	(0.37)
2013-14	0.15	0.12	-0.12	-0.18
	(0.23)	(0.22)	(0.29)	(0.22)
2015-16	0.24	0.06	0.07	-0.10
	(0.23)	(0.22)	(0.27)	(0.22)
2017-18	0.23	-0.32	0.14	-0.03
	(0.23)	(0.22)	(0.28)	(0.22)
MENA × 2013-14	0.17	0.93**	-0.67	0.13
	(0.46)	(0.46)	(0.53)	(0.52)
MENA × 2015-16	0.29	0.36	-0.36	-0.01
	(0.51)	(0.51)	(0.52)	(0.55)
MENA × 2017-18	0.50	0.60	-0.60	0.40
	(0.48)	(0.52)	(0.54)	(0.51)
N	1696	1696	1696	1696

Notes: This table shows coefficient estimates and standard errors in parentheses from estimating Equation 3.2 by using a Logit model which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the table are the ones of the interaction term of the second-generation MENA indicator and the year pair for the period after the Arab Spring, where the year pair 2011-12 is omitted. P-values: < 0.01 ***, < 0.05 **, < 0.10 *.

Table A.4.: Logit marginal effects - Females

		(1) upper-secondary and beyond	(2) in formal education	(3) NEET	(4) employed
2005-06	non-MENA	-0.11**	-0.21***	0.02	0.18***
		(0.05)	(0.05)	(0.04)	(0.05)
	MENA	0.02	0.01	-0.06	0.21**
		(0.09)	(0.09)	(0.06)	(0.10)
2007-08	non-MENA	-0.06	-0.12**	-0.01	0.16***
		(0.05)	(0.06)	(0.04)	(0.05)
	MENA	0.05	-0.07	-0.05	0.21**
		(0.09)	(0.10)	(0.06)	(0.11)
2009-10	non-MENA	-0.02	-0.08	-0.04	0.07
		(0.04)	(0.05)	(0.04)	(0.05)
	MENA	-0.01	0.04	0.04	-0.04
		(0.09)	(0.10)	(0.07)	(0.09)
2013-14	non-MENA	-0.08**	-0.09*	0.09**	0.02
		(0.04)	(0.04)	(0.04)	(0.04)
	MENA	0.21***	0.23**	-0.05	-0.08
		(0.08)	(0.09)	(0.06)	(0.10)
2015-16	non-MENA	-0.03	-0.07	0.02	0.06
		(0.04)	(0.05)	(0.03)	(0.04)
	MENA	0.14*	0.07	0.06	-0.19**
		(0.08)	(0.10)	(0.07)	(0.09)
2017-18	non-MENA	-0.03	-0.07	0.02	0.13***
		(0.04)	(0.05)	(0.04)	(0.04)
	MENA	0.15*	0.27***	-0.11*	-0.08
		(0.08)	(0.09)	(0.06)	(0.09)
Mean of the outcome in 2011-12	non-MENA	0.83	0.65	0.13	0.22
	MENA	0.53	0.40	0.32	0.26
N		1481	1510	1510	1510

Notes: This table shows the estimated marginal changes relative to the baseline year pair 2011-12 for non-MENA and MENA female second-generation immigrants, following the estimation of Equation 3.2 by using a Logit model, and corresponding standard errors in parentheses which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. P-values: < 0.01 ***, < 0.05 **, < 0.10 *.

Table A.5.: Logit marginal effects - Males

		(1) upper-secondary and beyond	(2) in formal education	(3) NEET	(4) employed
2005-06	non-MENA	0.11**	0.04	-0.12***	0.08
		(0.05)	(0.05)	(0.04)	(0.05)
	MENA	0.02	0.00	-0.13*	0.25***
		(0.07)	(0.09)	(0.07)	(0.08)
2007-08	non-MENA	0.10**	-0.04	-0.10***	0.08
		(0.05)	(0.05)	(0.04)	(0.05)
	MENA	0.08	0.06	-0.17***	0.19**
		(0.07)	(0.08)	(0.06)	(0.08)
2009-10	non-MENA	0.04	-0.09*	0.01	0.04
		(0.05)	(0.05)	(0.04)	(0.05)
	MENA	0.02	0.03	-0.17***	0.17**
		(0.06)	(0.08)	(0.06)	(0.07)
2013-14	non-MENA	0.03	0.02	-0.02	-0.04
		(0.04)	(0.04)	(0.04)	(0.04)
	MENA	0.06	0.21***	-0.12*	-0.01
		(0.07)	(0.08)	(0.07)	(0.07)
2015-16	non-MENA	0.04	0.01	0.01	-0.02
		(0.04)	(0.04)	(0.04)	(0.04)
	MENA	0.09	0.09	-0.05	-0.02
		(0.08)	(0.09)	(0.07)	(0.08)
2017-18	non-MENA	0.04	-0.06	0.02	-0.01
		(0.04)	(0.04)	(0.04)	(0.04)
	MENA	0.12*	0.06	-0.08	0.06
		(0.07)	(0.09)	(0.07)	(0.08)
Mean of the outcome in 2011-12	non-MENA	0.61	0.57	0.16	0.32
	MENA	0.49	0.44	0.36	0.18
N		1696	1696	1696	1696

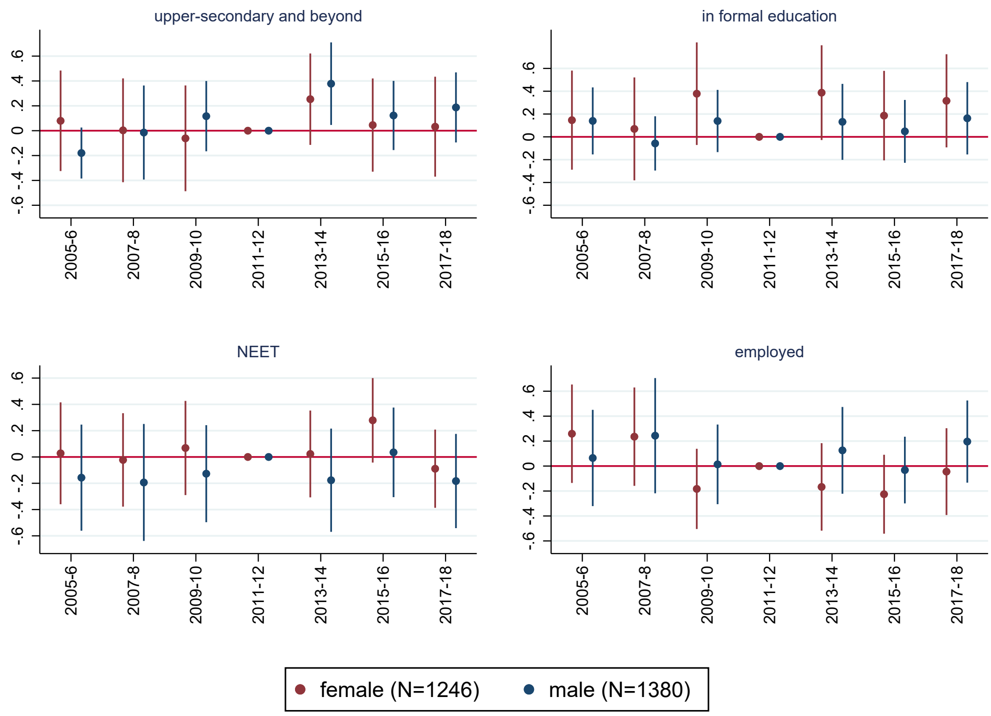
Notes: This table shows the estimated marginal changes relative to the baseline year pair 2011-12 for non-MENA and MENA male second-generation immigrants, following the estimation of Equation 3.2 by using a Logit model, and corresponding standard errors in parentheses which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. P-values: < 0.01 ***, < 0.05 **, < 0.10 *.

Table A.6.: Arab Spring and second-generation immigrants - Wild bootstrap clusters

	upper-secondary and beyond	in formal education	NEET	employed
Females				
MENA × 2005-06	0.14	0.21^{*}	-0.09	0.02
	(0.14)	(0.06)	(0.25)	(0.85)
	[0.14]	[0.01]	[0.22]	[0.86]
MENA × 2007-08	0.11	0.05	-0.06	0.05
	(0.25)	(0.67)	(0.50)	(0.66)
	[0.24]	[0.46]	[0.49]	[0.63]
MENA × 2009-10	0.02	0.12	0.08	-0.11
	(0.84)	(0.28)	(0.31)	(0.30)
	[0.89]	[0.32]	[0.45]	[0.11]
MENA × 2013-14	0.33^{***}	0.32^{***}	-0.15^{*}	-0.08
	(0.00)	(0.00)	(0.06)	(0.42)
	[0.05]	[0.00]	[0.14]	[0.41]
MENA × 2015-16	0.18^{**}	0.14	0.05	-0.23^{**}
	(0.04)	(0.19)	(0.49)	(0.02)
	[0.04]	[0.28]	[0.7]	[0.03]
MENA × 2017-18	0.20^{**}	0.36^{***}	-0.18^{**}	-0.20^{***}
	(0.03)	(0.00)	(0.03)	(0.06)
	[0.02]	[0.00]	[0.12]	[0.03]
N	1510	1510	1510	1510
Males				
MENA × 2005-06	-0.07	-0.04	-0.04	0.15
	(0.45)	(0.63)	(0.57)	(0.11)
	[0.45]	[0.62]	[0.56]	[0.19]
MENA × 2007-08	-0.00	0.09	-0.11	0.10
	(1.00)	(0.28)	(0.11)	(0.22)
	[0.99]	[0.28]	[0.15]	[0.24]
MENA × 2009-10	-0.01	0.12	-0.22^{**}	0.13
	(0.86)	(0.17)	(0.04)	(0.17)
	[0.86]	[0.23]	[0.1]	[0.21]
MENA × 2013-14	0.05	0.18^{**}	-0.13^{*}	0.02
	(0.54)	(0.03)	(0.08)	(0.75)
	[0.52]	[0.08]	[0.16]	[0.74]
MENA × 2015-16	0.06	0.06	-0.06	0.01
	(0.66)	(0.58)	(0.63)	(0.92)
	[0.62]	[0.58]	[0.65]	[0.91]
MENA × 2017-18	0.11	0.11	-0.10	0.06
	(0.32)	(0.43)	(0.39)	(0.46)
	[0.37]	[0.42]	[0.41]	[0.46]
N	1696	1696	1696	1696

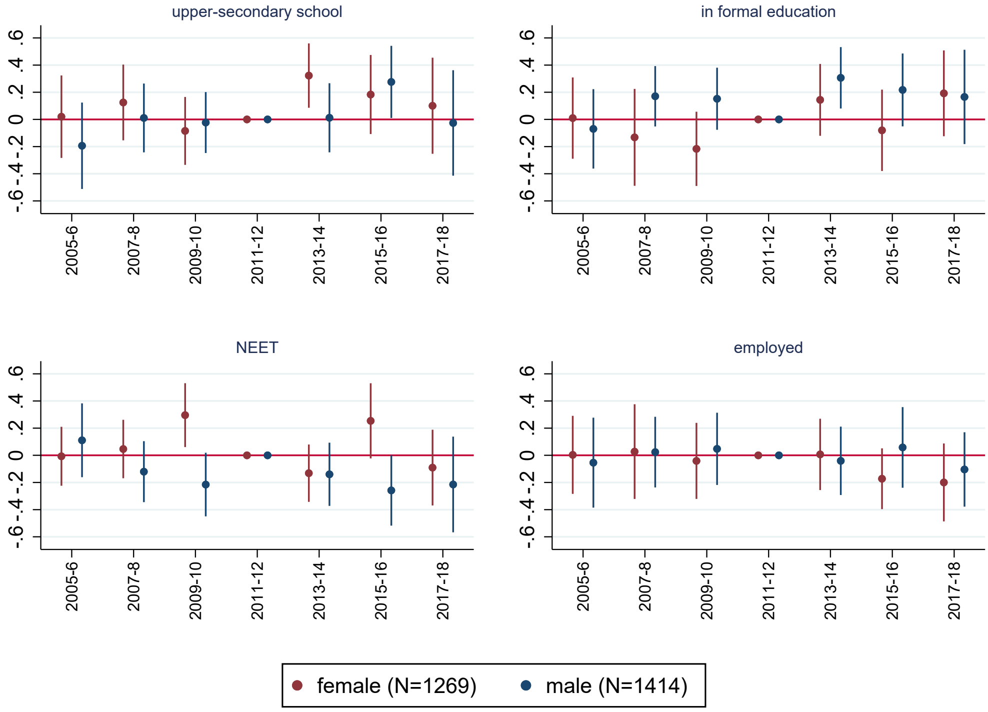
Notes: This table shows coefficient estimates, corresponding p-values in parentheses, and the p-values obtained from Wild bootstrap clusters at the regional level in brackets from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the table are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted. P-values: < 0.01 ***, < 0.05 **, < 0.10 *.

Figure A.4.: Arab Spring and second-generation immigrants with both parents from MENA



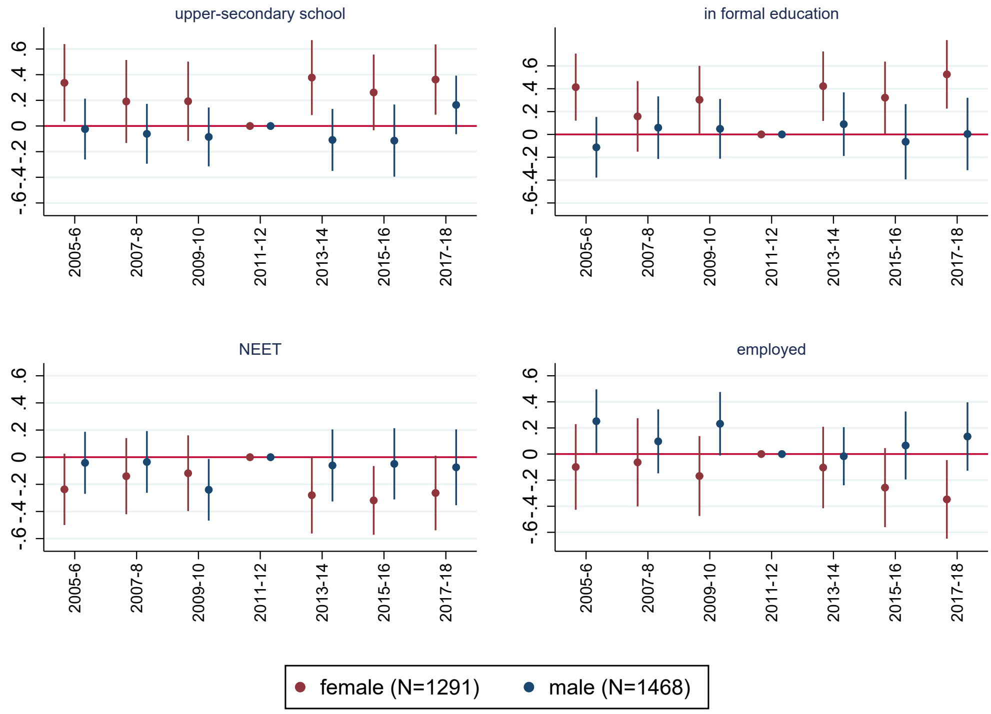
Notes: This figure shows, for the subsample of second-generation immigrants with both parents from MENA, coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

Figure A.5.: Arab Spring and second-generation immigrants with the mother from MENA



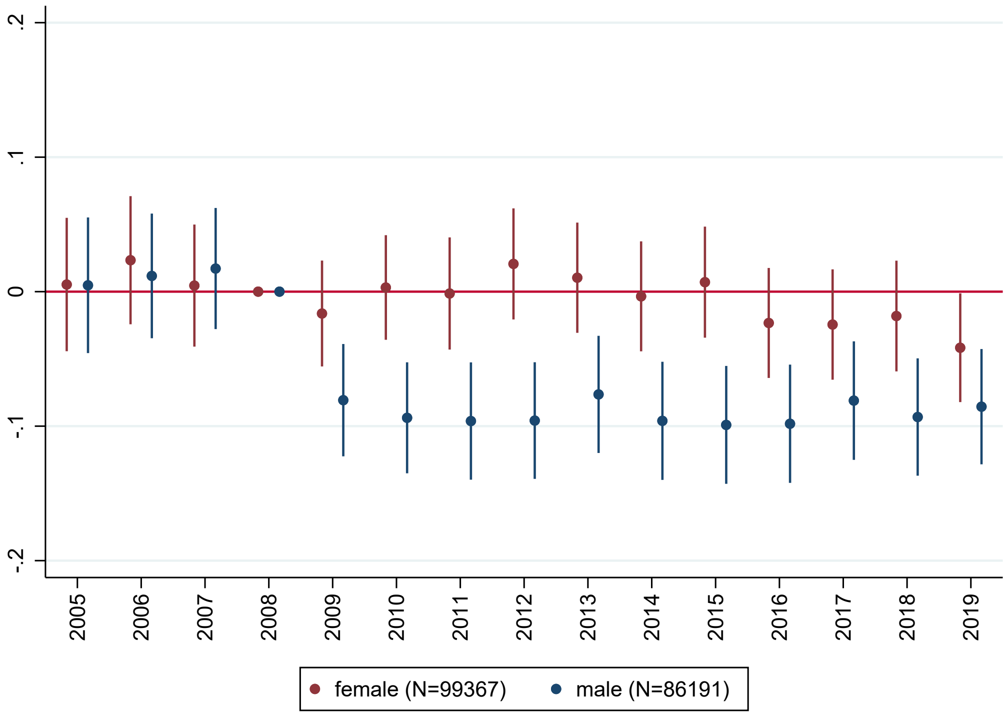
Notes: This figure shows, for the subsample of second-generation immigrants with the mother from MENA, coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

Figure A.6.: Arab Spring and second-generation immigrants with the father from MENA



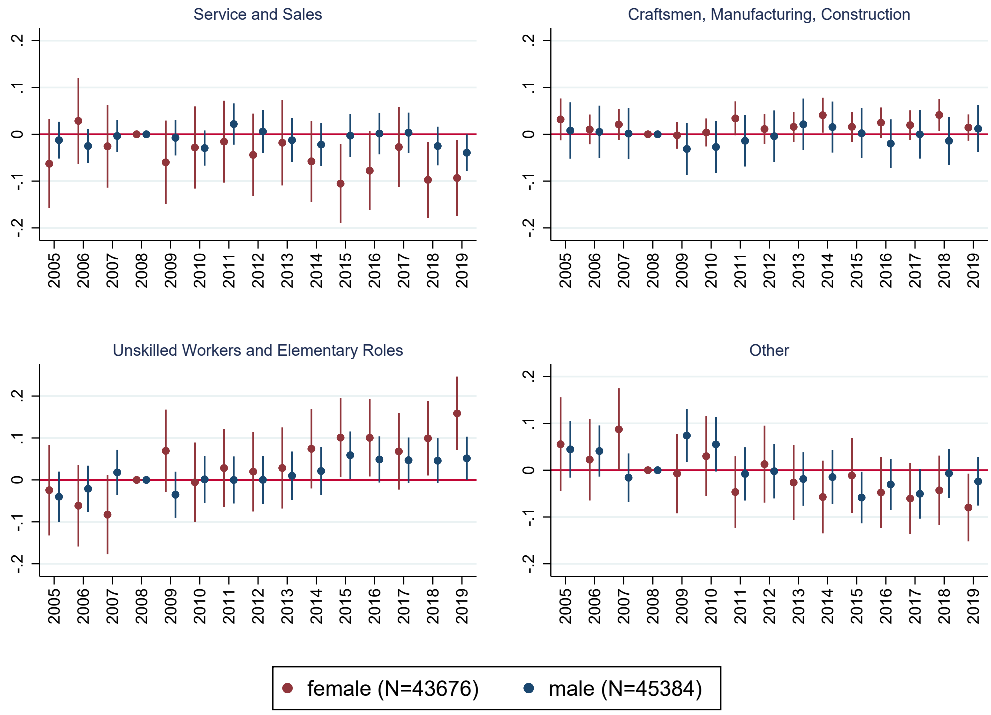
Notes: This figure shows, for the subsample of second-generation immigrants with the father from MENA, coefficient estimates and the 95% confidence intervals from estimating Equation 3.2 which regresses 4 different outcome variables (upper-secondary education and beyond, being in formal education, being in NEET, and probability of employment) on an indicator that takes a value of 1 if the individual is a second-generation MENA immigrant and 0 if they are a second-generation non-MENA immigrant, year pair fixed effects, their interactions, mother's and father's educational attainments, an indicator of holding Spanish citizenship, and region of residence (autonomous community) fixed effects. The coefficient estimates presented in the graphs are the ones of the interaction term of the second-generation MENA indicator and the year pair, where the year pair 2011-12 is omitted.

Figure A.7.: Employment probability of the first generation immigrants



Notes: This figure shows, for the first-generation male and female immigrants, coefficient estimates and the 95% confidence intervals from a linear regression of the binary indicator of employment on an indicator that takes a value of 1 if the individual is a first-generation MENA immigrant and 0 if they are a first-generation non-MENA immigrant, year pair fixed effects, and their interactions. The coefficient estimates presented in the graphs are the ones of the interaction term of the first-generation MENA indicator and the year, where the year 2008 is omitted.

Figure A.8.: Occupation categories of the first generation immigrants



Notes: This figure shows, for the first-generation male and female immigrants, coefficient estimates and the 95% confidence intervals from four separate linear regressions of the binary indicators of the occupation categories (service and sales; craftsmen, manufacturing, construction; unskilled workers and elementary roles; and other) on an indicator that takes a value of 1 if the individual is a first-generation MENA immigrant and 0 if they are a first-generation non-MENA immigrant, year pair fixed effects, and their interactions. The coefficient estimates presented in the graphs are the ones of the interaction term of the first-generation MENA indicator and the year, where the year 2008 is omitted.

4. Consequences of an Early Grave: Losing a Sibling During Childhood

4.1. Introduction

The death of a child is a very large stressor for both parents and surviving siblings. While children experience their own grieving process after such a loss, they may also receive less attention from their mourning parents, who are likely to face mental health problems, marital dissolution, and worsening labor market outcomes (Adhvaryu et al., 2022; Breivik and Costa Ramon, 2021; Vaalavuo et al., 2023; Van den Berg et al., 2017). Surviving children may encounter mental disorders, attempt suicide, require hospitalization more frequently, and face an increased mortality risk after losing a sibling (Bolton et al., 2016; Gerhardt et al., 2012; Yu et al., 2017). Additionally, the loss of a child might lead to changes in how parents allocate their time and financial resources among the surviving children. The aforementioned problems that parents may face after the loss can affect the quality of time they spend with surviving children due to changes in their mental well-being, marital dissolution, or the availability of financial resources resulting from worsening labor market outcomes.

Despite these significant effects, surprisingly, little is known about the consequences in terms of educational outcomes and human capital accumulation of surviving children. Contributing to filling this gap in the economics literature, this paper presents the first evidence from the entire population of a country, as well as in a European context, on the impact of losing a sibling during childhood on educational achievement. To study this question, I use individual-level data from Finnish administrative records spanning 24 birth cohorts. It is unlikely that experiencing a sibling loss during childhood is randomly distributed across the population. For instance, variables measuring concepts such as the socio-economic background of a child could potentially predict both educational achievement and the likelihood of experiencing the loss of a sibling. To address this challenge, I employ an identification approach that exploits the variation in timing of an unexpected sibling loss relative to the 9th grade. Specifically, children who suffer a sibling loss before 9th grade form the treatment group, while those suffering sibling loss after 9th grade act as the control group. The underlying assumption for causal interpretation is that the relative timing of the loss is as good as random across affected families, for which I provide supporting evidence.

I find a negative impact of 19% of a standard deviation on the 9th-grade GPA of surviving children after experiencing a sibling loss at age 14, two years before the 9th grade. However, there is no significant effect for those who experienced sibling loss at other ages, hiding some heterogeneity by socioeconomic status and gender. That is, the impact is stronger for children whose mothers have lower educational attainment and for girls. Yet, the gender difference is not prominent and not statistically significant for every age group. Further, I find a significant decrease of 12-14 percentage points (25-33%) in the probability of choosing a general upper-secondary school track following a sibling loss.

The impact on the 9th-grade GPA is not substantial, and not homogenous across different socio-economic groups or ages at the time of sibling loss, suggesting potential compensatory mechanisms. On the other hand, a 12-14 percentage points decrease in the probability of choosing a general upper-secondary school track is sizable. Delving into the potential mechanisms, I examine the impact of sibling loss on several other outcomes. First, considering the Finnish welfare state, surviving children and their parents are likely to receive mental health and grief support, facilitating a smoother bereavement process. To examine this potential mechanism, I employ an event study framework and estimate the effect of the loss on antidepressant prescriptions for surviving children and their parents. I find a substantial increase in the probability of antidepressant prescriptions for surviving children as well as their parents, suggesting a help-seeking behavior of the affected families and a corresponding response from healthcare professionals to mitigate the mental health challenges posed by this traumatic event.

Another potential channel could involve a reallocation of parental time and financial resources toward surviving siblings following the loss. Although specific information on parents' actual time allocation is unavailable, I use sick leave and unemployment as proxies for time use. Following the event study framework by Kleven et al. (2019), I find an increase in the probability of mothers taking sick leaves in the year of the loss and the subsequent year, accompanied by a persistent decline in their employment probability starting from the year after the loss. On the other hand, there is no evidence of a shift in fathers' probability of taking sick leave after the loss, nor in their employment probability. These findings suggest an increase in the time free from work for mothers after loss that could be potentially invested in surviving children. However, given the poor mental well-being indicated by the increase in antidepressant intake, the quality of the potentially increased parental time investment is debatable.

Lastly, another mechanism could be the teacher's compensation through increased attention or different grading behavior towards bereaving students. The 9th-grade GPA is an average of the teacher-assessed subject grades where teachers are expected to follow certain guidelines, but it is not a standardized test result. Therefore, this channel for the 9th-grade GPA outcome remains a possibility with no chance of being examined by using the data in this study. However, track choice outcome is much less likely to be affected by teachers' behavior, for which I find substantial negative effects.

There is a broad literature on the spillover effects of health shocks providing evidence on the effects of child death in the family on surviving children. A recently growing number of studies find negative impacts of children's mild and severe health shocks such as ADHD (Kvist et al., 2013), type 1 diabetes (Eriksen et al., 2021), hospitalizations (Breivik and Costa-Ramon, 2021), disabilities (Burton et al., 2017; Gunnsteinsson and Steingrimsdottir, 2019), cancer diagnosis (Adhvaryu et al., 2022; Vaalavuo et al., 2023), and death (Van Den Berg et al., 2017) on parents' labor market outcomes and mental well-being. Similarly, children's behavioral and educational outcomes are found to be negatively affected by health shocks within the family (Kristiansen, 2021; Le and Nguyen, 2017; Alam, 2015; Bratti and Mendola, 2014; Dhanaraj 2016; Luca and Bloon, 2018; Mendolia et al, 2019; Sun and Yao, 2010; Johnson and Reynolds, 2013; Aaskoven et al., 2022; Morefield, 2010; Stans, 2020). Our knowledge, however, is more limited about how children's health shocks affect other children in the family. While having a sibling with ADHD or a disability has been found to have a negative impact on academic achievement and behavioral outcomes (Breining, 2014; Black et al., 2021; Fletcher et al., 2012), in-utero and early health shocks lead to increased parental investment to older and healthy siblings (Parman, 2013; Yi et al, 2015).

To the best of my knowledge, four studies examine how losing a sibling during childhood affects educational outcomes. While Thamarapani et al. (2020) and Gautier (2021) study this question in developing country contexts, Fletcher et al. (2013) and Fletcher et al. (2018) focus on the context of the US. Thamarapani et al. (2020) use the Indonesian Family Life Survey, which is representative of 83% of the population, to examine the impact of losing a sibling across ages at the time of death on years of schooling, secondary school enrollment, and fertility. They find a negative impact on surviving brothers' years of schooling, compared to those who were born after death. On the contrary, in the context of a conflict, Gautier (2021) finds positive effects of losing a sibling during the 1994 genocide in Rwanda on the education and later life outcomes of surviving women. The positive impact on the years of schooling is explained by the increased parental investment and relief programs for genocide survivors. The two studies using USA data from surveys find negative effects.[1] While Fletcher et al. (2013) compare the educational attainment, academic performance, and later life outcomes of children with and without a sibling loss, Fletcher et al. (2018) compare the cognitive and socio-emotional outcomes of children who lost a sibling at different ages, before and after death for 121 deceased children in total.

The contribution of this paper is to the broad literature on the spillovers of health shocks within the family. Specifically, I contribute to the limited literature exploring the impact of losing a sibling during childhood on human capital development. This topic presents a significant identification challenge due to the non-random distribution of sibling loss across the population. Fletcher et al. (2018) have made strides in identifying causal effects in a peaceful setting,[2] by using survey data from the US and exploiting the timing of death. However, the sample Fletcher et al (2018) use is very small, and additionally, not all deaths can be considered to be exogenous or unanticipated, such as those following a prolonged illness. To overcome this challenge, I analyze the impact of an unexpected loss, a departure from previous studies that do not differentiate the cause of death.[3]

Using Finnish administrative records, I provide the first evidence from the whole population of a country, which not only enhances the statistical power to detect significant differences but also enables the identification of plausibly exogenous deaths with minimal anticipation effects —specifically, those caused by traffic accidents.

Utilizing administrative records from the entire population offers a unique advantage by allowing the linkage of educational records to medical outcomes for both surviving siblings and parents, as well as parental labor market outcomes. Through these links, I investigate previously unexplored mechanisms, such as the mental well-being of surviving siblings and their parents, as well as parental time use measured by sick leave spells and unemployment. This comprehensive approach provides valuable insights for policy considerations.

It is crucial to comprehensively analyze the consequences of sibling loss during childhood. The impact of such an experience extends beyond the immediate grief, influencing both short and long-term life outcomes for the surviving children and other family members. As shown in this study, the educational achievement and choices of the surviving siblings, the labor market outcomes of parents - especially

[1] Fletcher et al (2013) use data from the National Longitudinal Study of Adolescent Health (Add Health) and the Wisconsin Longitudinal Study (WLS); and Fletcher et al. (2018) use data from the National Longitudinal Survey of Youth.

[2] Gautier (2021) provides evidence from the 1994 Rwanda genocide by using an IV approach, during a civil conflict.

[3] Fletcher et al. (2013) conduct a heterogeneity analysis by cause of death in four broad categories: infant death, accident or suicide, sudden illness, and long-term illness.

mothers - and the mental health of all family members are affected negatively. These effects can carry severe social and economic implications, highlighting the need for dedicated policy attention. Implementing appropriate policies can provide bereaved family members with the necessary support to navigate these challenging times, minimizing potential lasting damage and societal costs on a broader level.

4.2. Data

4.2.1. Datasets and Sample Selection

To examine how losing a sibling affects children's educational achievement, I use register data on the entire population of Finland provided by Statistics Finland.[4] Information from several separate datasets is matched by the unique personal identification number. Appendix B summarizes each dataset used in this study, as well as the outcome variables and how they have been constructed.

I focus on the birth cohorts who are supposed to apply for upper-secondary schools in years available in application registers (1989, 1991-2013[5]) Since children apply for upper-secondary education at age 16 (end of the 9th grade), the sample is restricted to 1973, 1975-1997 birth cohorts, with known parents, at least one sibling and exactly one sibling loss before the deceased was younger than 25 years old. I exclude those who lost more than one sibling to be able to assign a certain year of shock to each child. I focus on surviving siblings of children with an unexpected death to eliminate any anticipation effect as well as the deaths more related to the health behavior or lifestyle of the family. Therefore, I restrict my attention to deaths caused by traffic accidents. These are potentially correlated with the lifestyle of the family or the risk-taking behavior. However, as long as this behavior or lifestyle does not differ depending on the surviving child's age at the time of sibling loss, it does not pose a threat to my identification strategy. I identify the deaths caused by traffic accidents by using the statistical cause of death documented in ICD-9 and ICD-10 classifications for years between 1988-1995 and 1996-2016, respectively.[6]

Siblings born after the death are excluded from the analysis since they were not exposed to the loss, and the fertility decision after losing a child tends to systematically differ across families with different characteristics. To observe children's and their parents' background characteristics before the loss and in the 9th grade, I construct a sample of children who themselves as well as their parents were alive and

[4] These datasets are not publicly available. Details of data access conditions can be found on Statistics Finland's website. (Statistics Finland, n.d.)

[5] I exclude 1990 from the analysis since it does not include information for all applicants in 1990.

[6] V01-89 for ICD-10, and E800-804 for ICD-9.

present in Finland one year before the death of the sibling and when the surviving children were 16 years old.

In the whole period of available years with the cause of death information (1988-2016), I observe 16530 deaths of individuals younger than 25 years old, 3260 (20%) of which died because of traffic/land transport accidents. After applying the sample restrictions mentioned above, my main analysis sample of surviving siblings consists of 1530 children, with 1105 deceased siblings in total. Descriptive statistics for this sample as well as the deceased siblings and parents are shown in Table 4.1. All time-variant variables are measured one year before death. 72% of the deceased children and 47% of their surviving siblings are males. Notably, surviving children live on average in relatively crowded houses with a 5.3 household size. Outcome variables except for the 9th-grade GPA (antidepressant prescription, employment, and sick leave) are measured for the subsample of individuals included in the event study analysis (those who are observed each year in the corresponding period for the relevant event study design.)

Figure 4.1 shows the age-at-death distribution of surviving children who lost a sibling between ages 10 and 20 and their deceased siblings. The first graph shows the surviving children's age-at-death distribution for our main analysis sample, where those who experienced the loss at or before age 16, time of measurement of the 9th-grade GPA, are considered as treated.[7] In the second graph, where the age-at-death distribution of the deceased siblings of children in our sample is depicted, the minimum age at death is one since children born and died within the same year are not included in cause of death registers.[8] Notably, there is an increase in the incidents of death starting from age 15 which is the minimum legal age for driving a moped in Finland.

4.2.2. Education in Finland and the 9th grade GPA

Compulsory education in Finland starts at age 7 and lasts for 9 years.[9] The first 6 years comprise the primary education stage, whereas the last 3 years comprise the lower-secondary education stage. In the 9th grade, in February-March, children apply for the upper-secondary schools through the central Joint Application System. They have a right to apply up to 5 academic and/or vocational high schools in their

[7] Number of observations of surviving siblings in this graph is 1560, while number of observations in our main analysis is 1530. This difference is caused by the missing values of the 9th-grade GPA for some observations.

[8] Age values less than or equal to 7 are grouped in one age group category, in line with the data confidentiality rules of Statistics Finland.

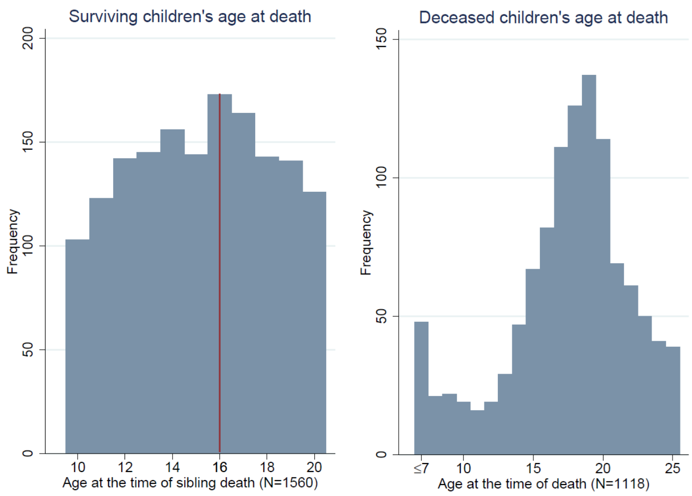
[9] Compulsory education has been extended to upper-secondary school (age 18) in 2021, not including the period of interest in this study.

Table 4.1.: Descriptive statistics

	Mean	SD	N
Deceased Siblings			
Male	0.72	0.40	1530
First-born	0.47	0.50	1530
Surviving Children			
Male	0.47	0.50	1530
Native Finnish speaker	0.96	0.19	1530
Firstborn	0.18	0.38	1530
Household size	5.29	2.67	1521
Urban	0.67	0.47	1530
9th-grade GPA	7.40	1.14	1530
9th-grade GPA (age 17)	7.51	1.17	163
Antidepressant prescription	0.01	0.09	1084
Mothers			
Age	42.42	4.98	1530
Primary / lower-sec	0.26	0.44	1530
Upper-secondary non-tertiary	0.48	0.50	1530
Mother tertiary or higher	0.26	0.44	1530
Antidepressant prescription	0.09	0.28	1003
Employed	0.79	0.41	939
Sick-leave	0.08	0.27	530
Fathers			
Age	44.82	5.66	1530
Primary / lower-sec	0.34	0.47	1530
Upper-secondary non-tertiary	0.46	0.50	1530
Tertiary or higher	0.21	0.40	1530
Antidepressant prescription	0.05	0.22	1062
Employed	0.81	0.39	939
Sick-leave	0.32	0.47	530

Notes: This table shows background characteristics and some outcome variables of deceased children, surviving siblings, and their parents, for our main analysis sample. All time-variant variables are measured one year before the child's death.

Figure 4.1.: Age distributions of the surviving children and the deceased siblings



Notes: The figures show the number of surviving children in my analysis sample and their deceased siblings who died in a traffic accident by age at death. The sample of surviving children is restricted to those who were 10-20 years old at the time of sibling loss, with known parents, exactly one sibling loss, observed at age 16 and one year before the loss. Their deceased siblings are individuals who died in a traffic accident between the ages of 1-25. In the second figure, ages at death less than or equal to 7 are censored at 7 in line with data confidentiality rules of Statistics Finland.

preferred order. They are admitted by schools based mainly on their 9th grade GPA which is the average of their grades from individual school subjects completed in the 9th grade.[10]

The main outcome variable in this paper is the 9th-grade GPA, which is obtained from Joint Application Registers for the years 1989, and 1991-2013. Since this outcome relies on teacher evaluations, it is important to note that they may be subjective. The 9th grade GPA ranges from 4 (failing) to 10 (the best grade). As shown in Table 4.1, the mean value of the 9th-grade GPA in my analysis sample is 7.40, with a standard deviation of 1.14. For a reasonable comparison over the years, I standardize this variable within each graduation year. To eliminate the effects that might be caused by a delay in graduation, I only include the children who graduated from the 9th grade on time (by age 16) in the analysis.

A potential concern about this restriction would be that the least affected children would select themselves into graduating on time, which might cause a downward bias in the coefficient estimates. To address this concern, I present evidence that losing a sibling does not affect the probability of graduation on time in Section 4.1.

[10]The average of grades in mathematics, physics, chemistry, biology, native language and literature, other domestic language, foreign languages, geography, social studies, history, religion, and health information.

4.3. Empirical Approach

I estimate the impact of losing a sibling during childhood on educational outcomes, by exploiting the variation in the timing of sibling loss relative to one year after the 9th-grade GPA measurement time, age 17. Formally, I estimate the following equation:

$$Y_i = \alpha + \sum_{t=10, t\neq 17}^{t=20} \gamma_t \mathrm{I}_{it} + \theta_y + \delta X_i + \varepsilon_i \tag{4.1}$$

where Y_i is the outcome variable ((1) 9th grade GPA - standardized within each year, and (2) general upper-secondary school track) of child i; I_{it} is an indicator of whether the age of the surviving sibling at the time of death is t for individual i; θ_y captures the graduation year (calendar year of the 9th grade) fixed effects; and X_i is a vector of child, household and parental characteristics, namely child's gender, being native Finnish speaker, birth order, living in an urban area, household size and highest educational attainment of both parents, as well as the deceased child's age at death. All time-variant characteristics are measured one year before death.

The coefficients of interest are γ_{10}-γ_{16} which captures the differences in the outcome of interest of children with ages of sibling loss 10-16, till six years before or at the 9th grade, in comparison to those losing their sibling when the surviving children are 17 years old, one year after the 9th grade. Coefficient estimates of γ_{18}, γ_{19} and γ_{20} could be considered as a placebo test since the performance of children who lost a sibling 2-4 years after obtaining the 9th grade GPA are expected not to be different than the performance of those who lost a sibling 1 year after obtaining the 9th grade. This hypothesis is confirmed if the point estimates for γ_{18}, γ_{19} and γ_{20} are not significantly different from zero.

The underlying assumption for this identification approach to be able to identify causal effects is that the timing of sibling loss relative to age 17 (of surviving child) is as good as random. I test an implication of this assumption that is the background characteristics of children before the loss do not significantly differ depending on the timing of the loss. Specifically, I test whether the characteristics of children (measured one year before death) who lost a sibling before age 17 are different from the characteristics of children (measured at the same age as the former group) who lost a sibling at age 17.

Consider two groups of children who lost their siblings at age 16 and 17. If the timing of the loss is as good as random, it implies that the background characteristics of these two groups of children were not different from each other when they were 15 years old, the common closest age before the loss for both groups. To test this implication, I run separate regressions of several variables capturing demographic and socioeconomic characteristics of the surviving children, on an indicator that takes value 1 if the child experienced the loss at age 16, and value 0 if the loss happened when the surviving child was at age 17. Importantly, the time-variant dependent variables are measured at age 15 for both groups. Then, I do the same comparison for another two groups of children with a loss at age 15 (14, 13, 12, 11, 10) and age 17, where the time-variant variables are measured at age 14 (13, 12, 11, 10, 9) for both groups, always keeping the control group as the ones with a sibling loss at age 17, which is the control group in my main analysis.

Formally, I estimate separate regressions of these background variables for each $t \in \{10, 11, .., 15\}$, on the indicator that takes a value of 1 if the age of surviving child at the time of death is t, and 0 if the age is 17; as follows:

$$Y_i^{t-1} = \gamma \mathrm{I}_i^t + \varepsilon_i^t \tag{4.2}$$

where

$$\mathrm{I}_i^t = \begin{cases} 1, & \text{if age of } i \text{ at death}=t \\ 0, & \text{if age of } i \text{ at death}=17 \end{cases}$$

and Y_i^{t-1} is the child's or parent's background characteristics measured at age $t-1$.

Figure A.1 and Figure A.2 present the results of this test for the outcomes of parents and children, respectively. Both figures suggest that the children with sibling loss at different ages are not different from each other, especially in terms of socioeconomic background. Specifically, differences in parental educational attainment, employment and earnings (adjusted for 2019 prices) fluctuate around zero and most of them are not statistically significant.[11] It is important to note that in Figure A.2 one significant difference across ages at the sibling loss is the age difference between the surviving child and the deceased sibling, for ages at sibling loss 10-14. This difference is reasonable given the higher probability of death after age 15 of the deceased child, as depicted in Figure 4.1. Nonetheless, by including age at death fixed effects in Equation 4.1, I control for these age differences.

As a second control for the identification assumption, Figure A.3 shows the employment and earnings outcomes of parents of children who lost their siblings at different ages, measured one year before death and relative to the corresponding outcome of the parents of children who lost their siblings at age 17. As depicted in the graphs, there is no significant difference in these time-variant characteristics measured one year before death, either.

[11] Since several characteristics for several years are measured for this exercise, it is plausible to expect a few coefficients to be statistically significant.

Event Study Design for Mental Health Outcomes: To estimate the impact of child death on mental health outcomes, proxied by the probability of being prescribed antidepressants, I exploit the variation in the timing of child death within an event study framework. Specifically, I construct a balanced panel of children, along with their mothers and fathers, who experienced a loss between the ages of 10 and 20, corresponding to 71% of my main analysis sample.[12] The observations cover 3 years before and 7 years after the loss. Then, I estimate the coefficients of indicator variables for years relative to the death year ("event time") using the following equation separately for siblings, mothers, and fathers:

$$Y_{it}^{g} = \alpha_{i}^{g} + \sum_{t=-3, t\neq -1}^{t=7} \gamma_{t}^{g} \mathrm{I}_{t} + \varepsilon_{it}^{g} \tag{4.3}$$

where Y_{it}^{g} is the outcome of interest for individual i of group g (sibling, mother, father) at event time t, α_{i}^{g} represents individual fixed effects, and I_t denotes event time fixed effects. Omitting the event time dummy at $t = -1$, the coefficients of interest are γ_{t}^{g} which measure the impact of the loss relative to one year before the loss.

Event Study Design for Parents' Labor Market Outcomes: To estimate the impact of child death on parents' labor market outcomes, I employ an event study framework following Kleven et al. (2019). The specification I use is analogous to theirs, with the event being child death instead of childbirth. I construct balanced panels of mothers and fathers of the surviving children in my main analysis sample, observed each year between 4 years before and 6 years after the loss.[13] Subsequently, I estimate the following equation separately for mothers and fathers:

$$Y_{iyt}^{g} = \sum_{t=-4, t\neq -1}^{t=6} \gamma_{t}^{g} \mathrm{I}_{it} + \omega_{iy}^{g} + \delta_{y}^{g} + \varepsilon_{iyt}^{g} \tag{4.4}$$

Here, Y_{iyt}^{g} represents the outcome of interest (sick leave, employment, or earnings) for individual i of group g (mother or father) in year y and at event time t, and I_t are event time dummies. Following Kleven et al. (2019), I control for underlying life-cycle trends by including age at year y fixed effects denoted by ω_{iy}^{g} and for time trends of macroeconomic conditions by including year fixed effects denoted by δ_{y}^{g}.

[12] Information on antidepressant prescriptions is only available starting from 1993. This constraint, together with the balanced sample requirement between 3 years before and 7 years after the loss, causes the sample used for the event study design to be smaller than the main analysis sample.

[13] Since the balanced sample requires each individual to be observed each of these 11 years, the sample for this event study exercise is 61% of the main analysis sample for employment and income outcomes. For the sick leave outcome, it is 35% of the main analysis sample since this information is available starting from 1995.

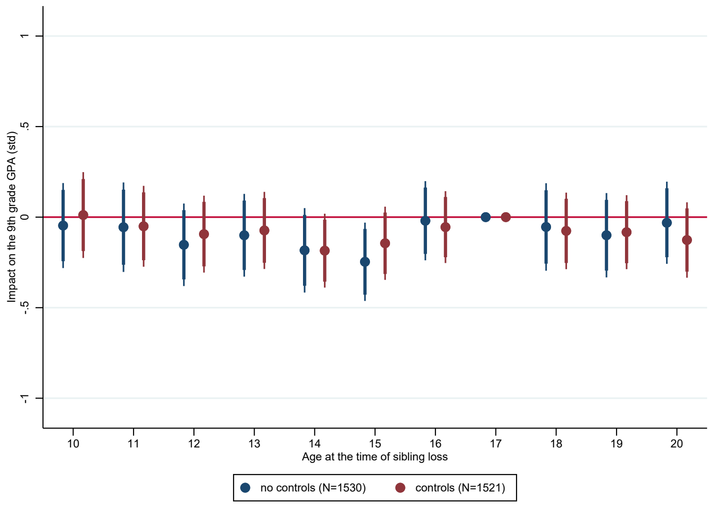
The coefficients of interest are γ_t^g, which measure the impact of the loss on parents' labor market outcomes relative to one year before the loss.

4.4. Results

4.4.1. Educational Outcomes

This section investigates the impact of sibling loss during childhood on educational outcomes. Before delving into the main findings concerning 9th-grade GPA, I analyze the impact on the probability of timely graduation (by age 16). As depicted in Figure A.4, this probability is not affected significantly by the loss of a sibling.

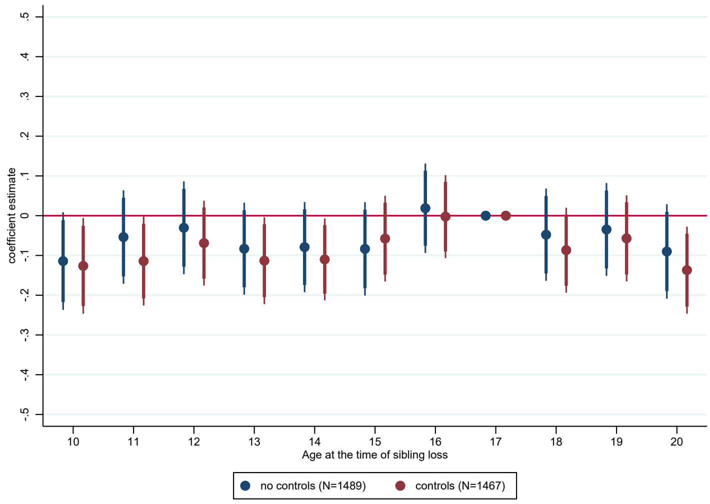
Figure 4.2.: Impact of losing a sibling on the 9th grade GPA



Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure 4.2 presents the impact of losing a sibling on the standardized 9th grade GPA for each age at sibling loss, by estimating Equation 4.1 with and without control variables. Results suggest no significant effect for children losing their siblings at the ages between 10 and 13. However, the 9th grade GPA of those who lost their sibling at age 14 is 19% of a standard deviation (0.22 points, 2.9%) below compared to those who lost their sibling at age 17, just after the 9th grade. As for the age 15, I find an effect of a similar magnitude (14% of a standard deviation), however, the coefficient is not precisely estimated. The 9th-grade GPA of those losing their siblings at age 16 is not affected by the loss. For ages between 18 and 20, coefficient estimates

Figure 4.3.: Impact of losing a sibling on general track choice



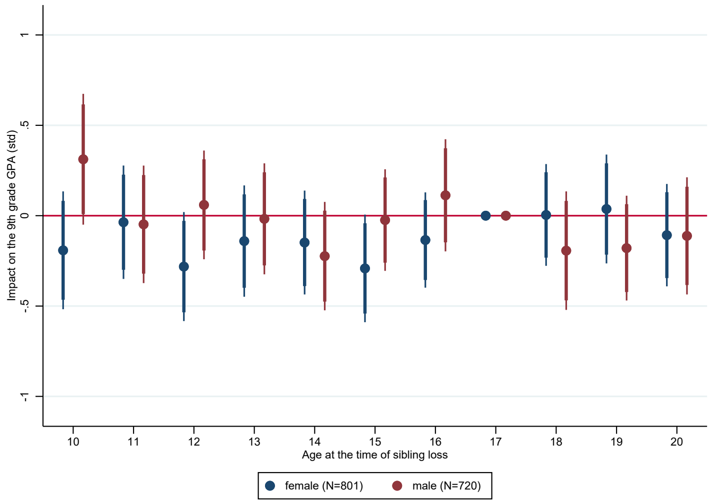
Notes: This figure shows coefficient estimates from the regression of the indicator that takes a value of 1 if the child is enrolled in a general track in upper-secondary school, and 0 if enrolled in a vocational track at age 16 on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, age at death, and graduation year fixed effects.

are small in magnitude and not statistically significant. This observation works as a placebo check, confirming the expectation that the 9th-grade GPA of those who lost their siblings after the 9th grade at different ages are not systematically different from each other, providing support for my identification approach.

Next, I analyze the impact on the upper-secondary school enrollment on time (by age 16) and the track choice of the surviving siblings. First, on the extensive margin, Figure A.5 shows that there is no impact on enrollment in upper-secondary school education. Then, Figure 4.3 presents results from estimating Equation 4.1 with and without control variables, where the outcome of interest is the indicator of whether the child is enrolled in a general or vocational track upper-secondary education. Results indicate a significant decrease of 12-14 percentage points (28-33%) in the probability of choosing a general track over a vocational track, for ages of sibling loss 10, 11, 13, and 14.

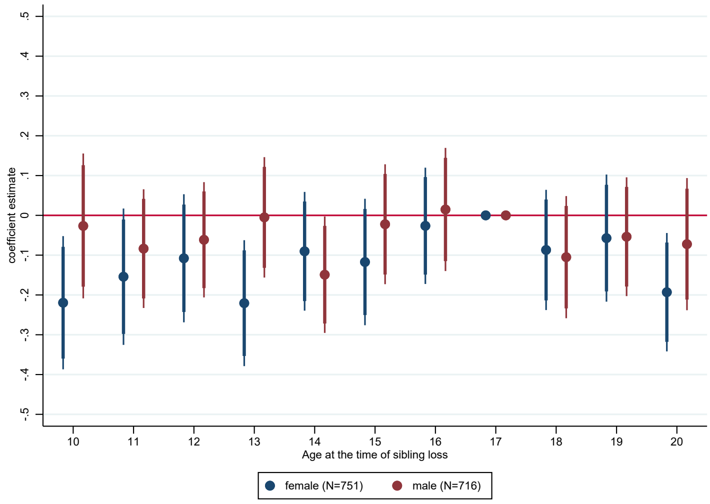
Figure 4.4 presents the results for the 9th grade GPA estimated for the subsample of females and males, separately. For females who lost their siblings, the coefficient estimates consistently exhibit a negative trend, proving statistically significant at a 90% confidence level for ages 12 and 15. However, no analogous trend is observed for males. The coefficient estimates for males fluctuate around zero, lacking statistical significance. These findings suggest that the 9th-grade GPA of females tends to be more adversely affected compared to males, except for ages of sibling loss 11, 13,

Figure 4.4.: Heterogeneity by gender (GPA)



Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for females and males. Control variables include being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, age at death, and graduation year fixed effects.

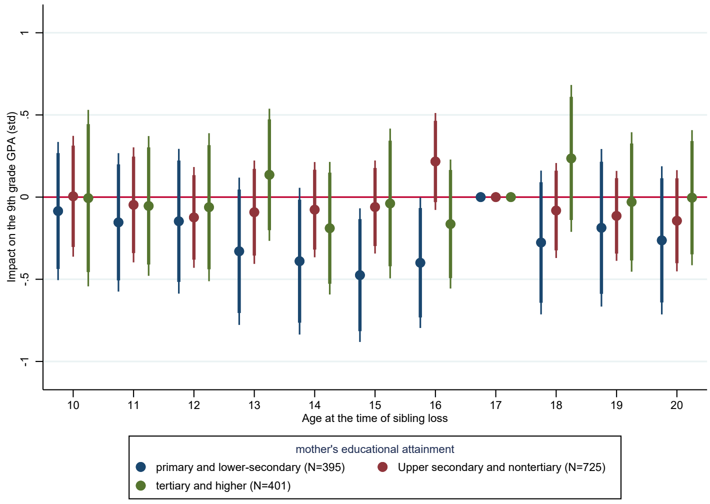
Figure 4.5.: Heterogeneity by gender (track choice)



Notes: This figure shows coefficient estimates from the regression of the indicator that takes a value of 1 if the child is enrolled in a general track in upper-secondary school, and 0 if enrolled in a vocational track at age 16 on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for females and males. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, age at death, and graduation year fixed effects.

and 14 where the confidence intervals overlap considerably. A similar pattern is observed for the probability of choosing the general track for the upper-secondary school. As depicted in Figure 4.5, for ages 10 and 13, females' general track choice probability decreases by 24 and 22 percentage points, respectively, while the effect is almost zero and not statistically significant for males. However, there is not a clear pattern for this difference, nor any evidence of such differential effect for those who lost their siblings at other ages.

Figure 4.6.: Heterogeneity by mother's educational attainment (GPA)



Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for surviving siblings with mothers having (i) primary and lower-secondary, (ii) upper-secondary and non-tertiary, and (iii) tertiary and higher educational attainment. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, father's highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Moving forward, I investigate the heterogeneity in results based on the socioeconomic status of children, proxied by parents' educational attainment. In Figure 4.6, I present results from separate regressions for subsamples of children with mothers having different educational attainment levels. The findings suggest that the negative effect is more pronounced for surviving children with lower-educated mothers. Specifically, for ages of sibling loss 13 to 16, the effect ranges between 33-47% of a standard deviation, proving significant for ages of sibling loss 14 to 16. Results based on the educational attainment of the father, as presented in Figure A.6, mirror those of the mother.

As depicted in Figure A.7, Figure A.8, and Figure A.9, no discernible heterogeneity is observed based on the presence of other siblings, the age difference between the surviving and deceased children, and birth order.[14]

[14]Figure A.9 suggests a significant negative impact for ages 12, 13, and 15 for the subsample of children where the deceased sibling is not the first-born. However, it is important to note that the observed effect is not solely driven by this particular subsample, as its confidence intervals overlap

4.4.2. Mechanisms

The findings so far indicate a negative impact for ages of sibling loss 12 to 15 within certain demographic groups. However, this impact is neither substantial nor uniform across all ages and socio-economic groups. On the other hand, I find a decrease of 12-14 percentage points in the probability of choosing a general upper-secondary school track, which is notable in size. These findings suggest the presence of potential compensatory mechanisms for the 9th-grade GPA, but not for the general track choice.

While there could be many changes following a child's death in the family which could potentially explain the effects on the educational outcomes, I focus on two main channels that can be proxied by using the administrative records. Specifically, in the subsequent sections delving into potential mechanisms, I examine the effect of child loss on the mental health of both surviving siblings and their parents, as well as its impact on parental labor market outcomes.

Mental Health

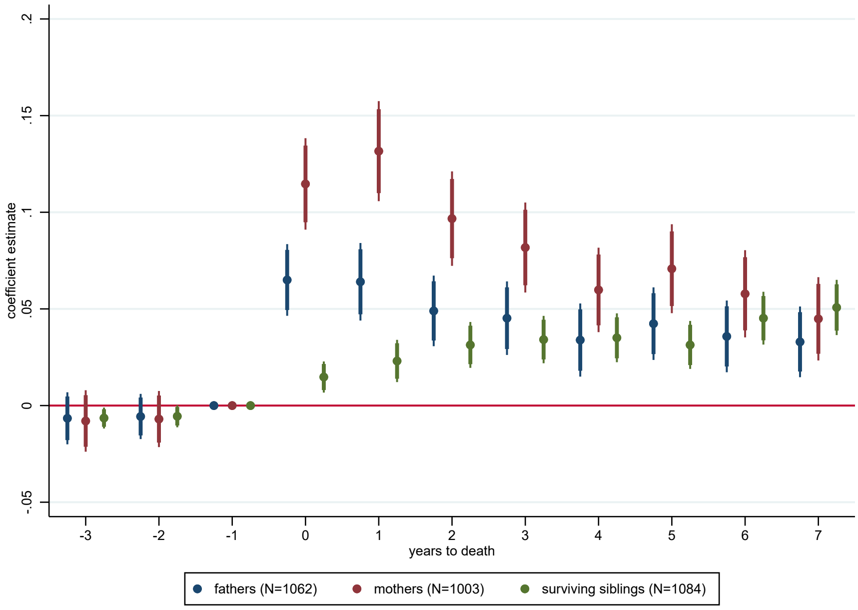
In this section, I present the estimated impact of sibling loss on mental health, using antidepressant prescriptions as a proxy. It is important to note that being prescribed antidepressants indicates an individual's initiative to seek mental health support. However, not all individuals seeking mental health support are necessarily prescribed antidepressants. Therefore, this variable serves as a conservative proxy, representing a lower bound of help-seeking behavior for mental well-being. Additionally, those prescribed antidepressants may represent cases with more severe consequences of the loss or individuals experiencing a relatively smoother bereavement period due to the mental health support they receive. It's essential to consider these nuances when interpreting the results.

As detailed in Appendix B, prescription data is available only starting from 1993, and to construct a balanced sample from 3 years before to 7 years after, I restrict my attention to losses between 1996 and 2013. Figure 4.7 displays results from estimating Equation 4.3 for surviving children, their mothers, and fathers separately. Across all groups, there is no evidence of a pre-trend.[15] For surviving children, there is an increase of 1 percentage point in the probability of being prescribed antidepressants in the same year as the sibling loss. Considering the age range of the surviving siblings at the time of the event (10-20), this effect is notable. Furthermore,

with those of other subsamples. (Please note that a subsample in this graph is not necessarily exclusive of others, as surviving children might have additional siblings besides the deceased one.)

[15] While the coefficient estimate for event time -3 is statistically significant for surviving siblings, its magnitude is very small.

Figure 4.7.: Antidepressant prescription probability



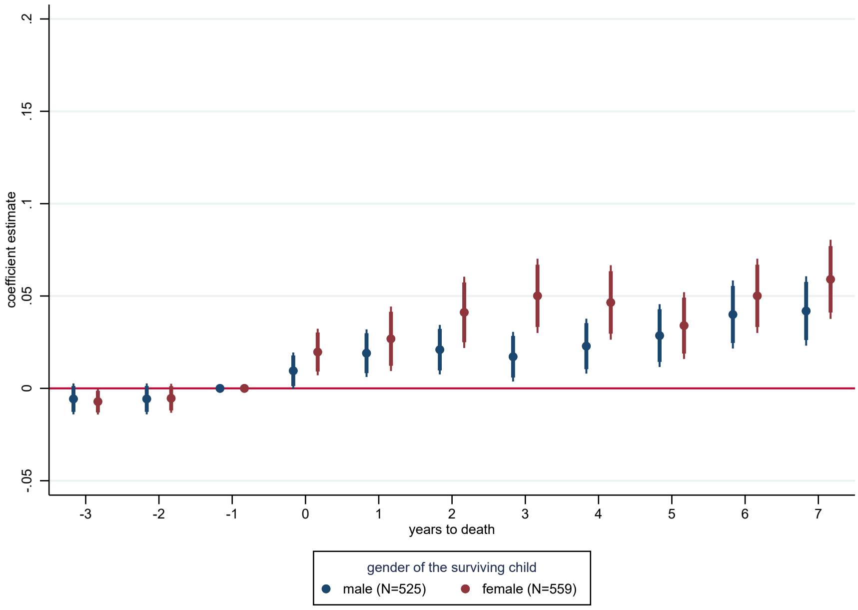
Notes: This figure shows coefficient estimates from the regression of the probability of being prescribed antidepressants on the "event time", and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include individual fixed effects.

there is an upward trend until 7 years after the loss, where the magnitude of the impact reaches 5 percentage points at the end of the observable period.

For parents, there is a notable increase in the probability of being prescribed antidepressants in the year of the child's death. Mothers experience a substantial increase of 11 percentage points, followed by 13 percentage points in the subsequent year. Then, there is a decreasing trend, resulting in a 4 percentage point impact at the end of the observable period. Fathers experience a smaller effect at the year of loss compared to mothers, with a 6 percentage point increase at event times 0 and 1. The impact shows a slight downward trend, converging to 3 percentage points at the end of the observable period. Considering the mean of the antidepressant prescriptions of mothers and fathers one year before death (9% and 5%, respectively, as shown in Table 4.1), the percentage increase of mothers is slightly higher than that of fathers. All in all, for all family members, there is a persistent impact that does not disappear even 7 years after loss.

Figure 4.8 shows the change in the antidepressant prescriptions for the surviving siblings, estimated by females and males, separately. For girls, the magnitude of the coefficient estimate is larger than that of boys, with significant differences in the 3rd and the 4th years following the sibling loss.

Figure 4.8.: Antidepressant prescription probability by gender of the surviving child



Notes: This figure shows coefficient estimates from the regression - separately estimated for females and males - of the probability of being prescribed antidepressants on the "event time", and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include individual fixed effects.

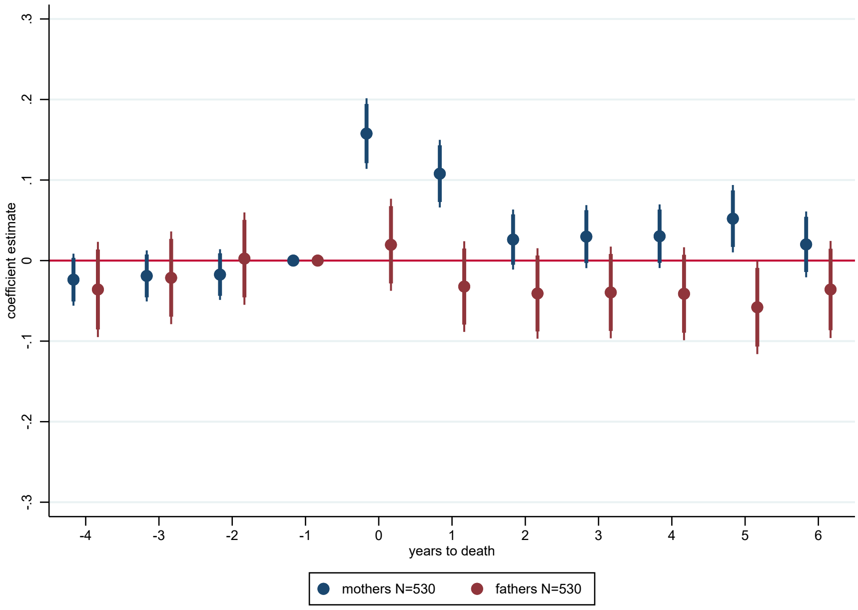
Labor Market Outcomes of Parents

This section examines the impact of child loss on the labor market outcomes of parents with surviving siblings in the main analysis sample. Figure 4.9 presents coefficient estimates from estimating Equation 4.4, separately for mothers and fathers, where the outcome of interest is the probability of receiving sickness allowances in a given year. For fathers, there is no discernible effect of child loss on the probability of taking sick leave. However, for mothers, there is a significant increase of 16 percentage points in the probability of taking sick leave in the year of loss, followed by an 11 percentage point increase in the subsequent year. Starting from the second year after the loss, the coefficient estimate becomes smaller and statistically insignificant, except for year 5.

Next, Figure 4.10 illustrates coefficient estimates for the impact on the probability of employment. Similar to the sick leave outcome, there is no significant effect for fathers until the last observable year (6 years after death), though a downward trend emerges from year 4. In contrast, mothers experience a persistent decrease of 4-5 percentage points in the probability of employment from the year following the loss.

Finally, the impact of the loss on the earned income of parents is presented in Figure A.10. For mothers, there is a persistent decline in income by approximately 2000

Figure 4.9.: Parental sick leave

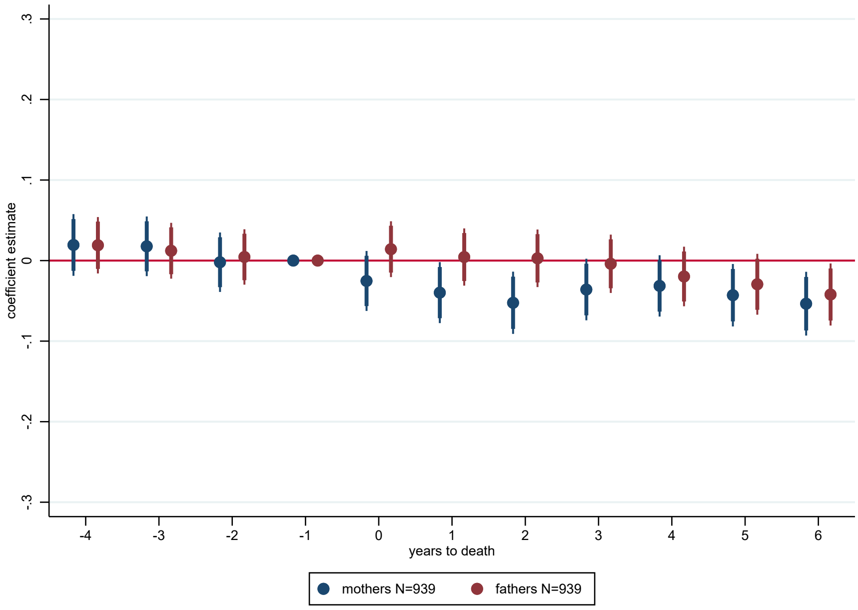


Notes: This figure shows coefficient estimates from the regression - separately estimated for mothers and fathers of the surviving children - of the probability of receiving sickness allowance on the "event time", and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include age and year fixed effects.

Euros per year, starting from one year after the loss. Fathers also exhibit a decreasing trend, but the estimates are less precise compared to those of mothers. Starting from year 3, coefficient estimates are statistically significant, with magnitudes ranging from around 2000 to 3500 Euros.

These findings suggest worsening labor market outcomes for mothers following the death of a child, while fathers' labor market outcomes are not affected significantly, except for the wage drop that emerges starting 3 years after the child's death. The increase in the probability of mothers' sick leave-taking and the decrease in their employment could suggest increased parental time investment for the surviving children. However, given the increased antidepressant intake shown in Section 4.2.1, the quality of this time investment is not clear.

Figure 4.10.: Parental employment



Notes: This figure shows coefficient estimates from the regression - separately estimated for mothers and fathers of the surviving children - of the probability of employment on the "event time", and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include age and year fixed effects.

4.5. Robustness Checks

To analyze the robustness of the findings, I perform several checks in this section.[16] First, I standardize the 9th-grade GPA outcome within each school-year pair - while I standardize it within each year for the main specification - to account for potential variations between schools, such as differences in grading behavior or the geographical setting. This involves standardization based on the anonymous lower-secondary school ID for each child. As shown in Figure A.11, the results closely align with the baseline findings showed in Figure 4.2.

The main specification (Equation 4.1) incorporates controls for the educational attainment of both parents, capturing the socio-economic background of the children. As an alternative measure for socio-economic status, Figure A.12 includes controls for the income of both parents measured one year before the child's death. The findings are very close to the baseline.

To investigate potential distinctions in the impact for full siblings who share the same biological father in addition to the same biological mother as the deceased child, I estimate Equation 4.1 for this specific subsample. As demonstrated in Figure A.13,

[16] I present the results for the 9th-grade GPA outcome in this section. However, results for the track choice outcome are also robust to all of these changes and are available upon request.

the results remain unchanged for this group of siblings.

Given that the cause of death in this study is traffic accidents, there is a possibility that surviving children might have been involved in the same accident that resulted in the loss of their sibling. To address this concern, I repeat the analysis for a subsample of surviving siblings who were not hospitalized within the same month as the loss or the following month. The aim of including the following month is to take into account the potential physical effects of the accident that might be realized later. The causes of hospitalization include various reasons, including psychiatric causes or injuries. Results, as presented in Figure A.14, closely align with the baseline, suggesting that the effect on 9th-grade GPA is not driven by the worsening health status of the surviving child resulting from the accident.[17]

4.6. Discussion and Conclusion

This paper analyzes the impact of losing a sibling during childhood on educational outcomes in Finland, exploiting the time variation of sibling death relative to the 9th grade. Results suggest that losing a sibling has a negative impact of 19% of a standard deviation (0.22 points) on the 9th-grade GPA of surviving children with a loss at age 14, 2 years before the 9th grade. Considering that a sibling loss is a major adverse event, the effect is rather small. Examining potential compensation mechanisms, I find substantial increases in the probability of antidepressant prescriptions. One interpretation of these results could be that mental health support to children and their parents helps with their grieving process, resulting in small effects on grades. On the other hand, the negative effect on the general track choice (12-14pp, 28-33%) is not small, suggesting that not all educational outcomes remain unaffected by the loss.

Another compensation mechanism might be an increased time investment from the mothers, given that I find an increase in their probability of receiving sick leave and a decrease in their employment probability. Yet, this finding should be interpreted cautiously since it might instead suggest that mothers have more difficulty coping with the loss compared to fathers and they are more likely to take sick leaves and get unemployed. Although these questions might be out of the scope of this paper, they should be carefully examined in future research for potential policy implications.

Another channel that I cannot explore with the available data is potential compen-

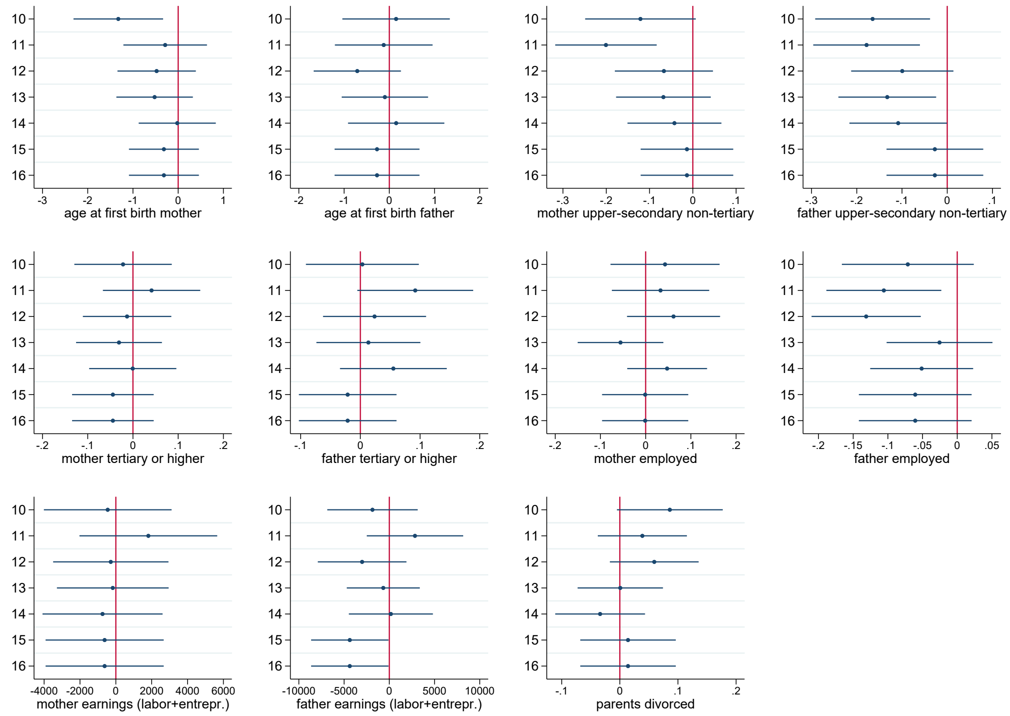
[17]Given the availability of hospital discharge registers from 1994 onwards, this analysis is restricted to sibling deaths occurring between 1994 and 2016, excluding the period spanning 1986 to 1993. Consequently, there is a substantial reduction in the number of observations. For comparability, Figure A.15 presents the main results derived from this restricted sample, focusing exclusively on losses between 1994 and 2016.

sation from the teachers. Although in Finland teachers are very thoroughly selected and educated and are expected to follow certain rules in grading, this remains a possibility given that the 9th-grade GPA is not a standardized test result. However, it is important to note that the track choice is expected to be less affected by potential compensation from teachers.

The negative impact on the general track choice suggests a medium-run effect of the loss on the human capital accumulation of surviving siblings. Concerning this, such a devastating life experience during childhood might potentially have impacts on other outcomes, including long-term mental health, marriage and fertility decisions, and labor market trajectories that could affect the next generations. It is crucial to causally analyze the long-term effects of sibling loss during childhood to derive accurate policy implications. Therefore, the long-term impacts of sibling loss will be the subject of future research.

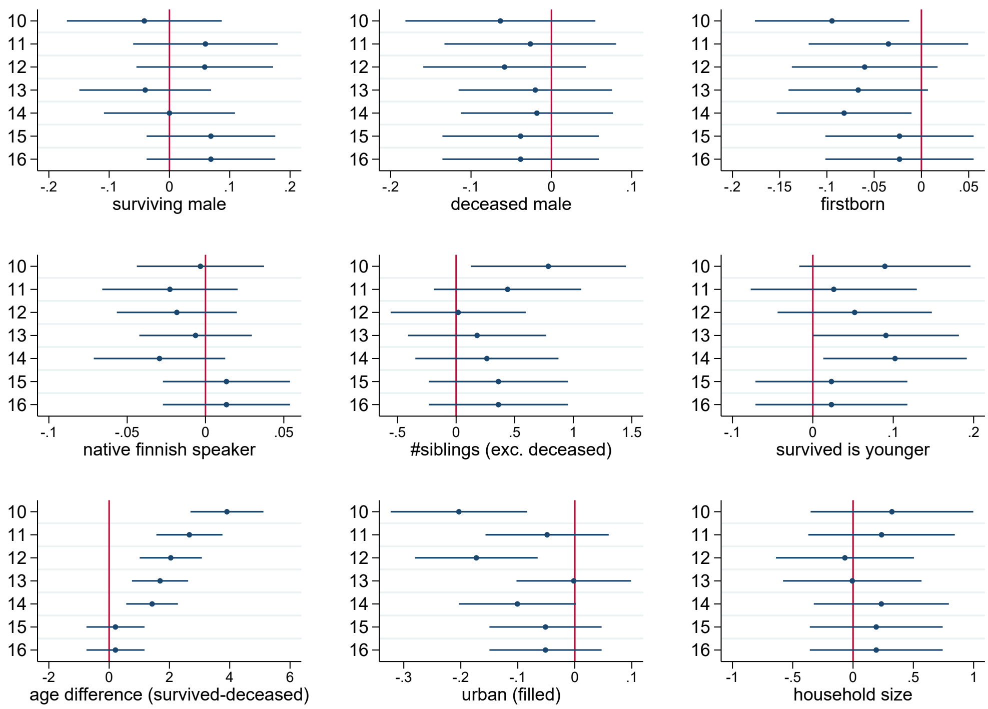
A. Appendix A: Additional Figures

Figure A.1.: Parents' background characteristics



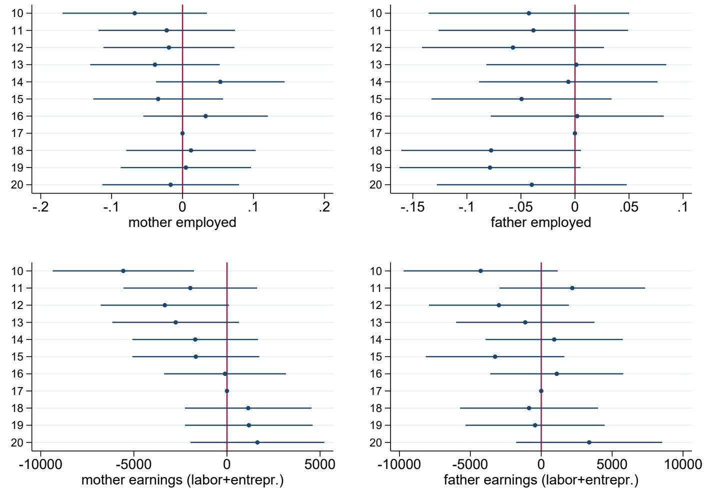
Notes: Figures show the coefficient estimates and 95% confidence intervals from the separate regressions of the several variables for each $t \in \{10, 11, .., 16\}$, on the indicator that takes a value of 1 if the age of surviving child at the time of death is t, and 0 if the age is 17. Earnings are adjusted for 2019 prices, and all time-variant variables are measured in the closest common age of the surviving children before the sibling loss.

Figure A.2.: Surviving children's background characteristics



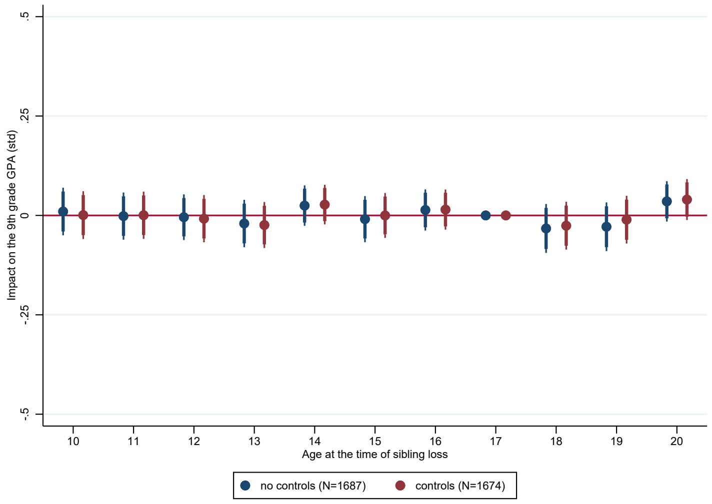
Notes: Figures show the coefficient estimates and 95% confidence intervals from the separate regressions of the several variables for each $t \in \{10, 11, .., 16\}$, on the indicator that takes a value of 1 if the age of surviving child at the time of death is t, and 0 if the age is 17. Earnings are adjusted for 2019 prices, and all time-variant variables are measured in the closest common age of the surviving children before the sibling loss.

Figure A.3.: Parents' background characteristics - one year before child's death



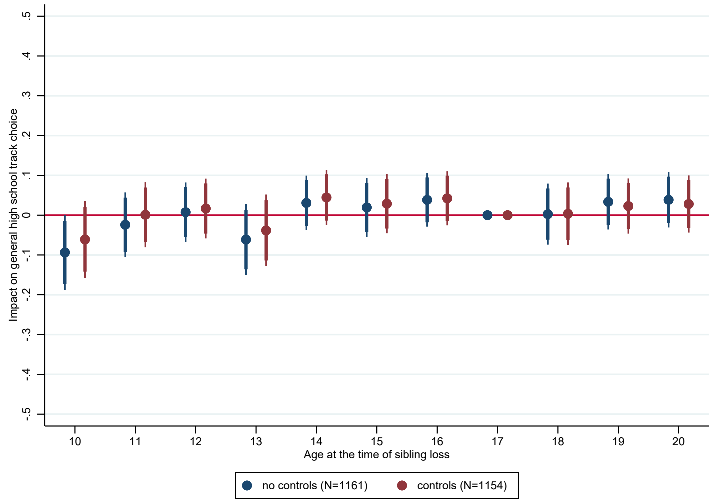
Notes: This figure shows the employment and earnings outcomes of parents of children who lost their siblings at different ages measured one year before the child's death and relative to the corresponding outcome of the parents of children who lost their siblings at age 17.

Figure A.4.: Probability of graduating on time



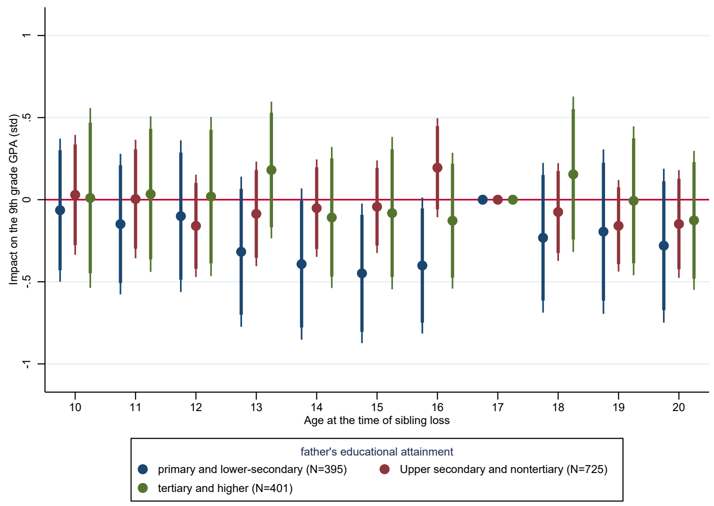
Notes: This figure shows coefficient estimates from the regression of the probability of graduating on time on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, and age at death.

Figure A.5.: Probability of upper-secondary school enrollment on time



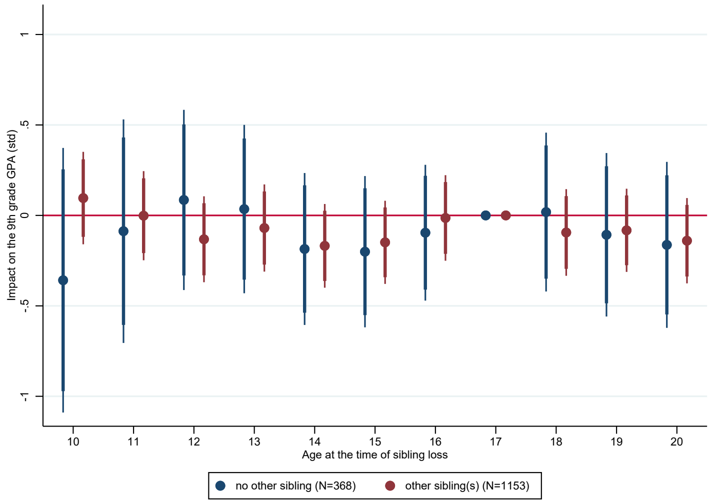
Notes: This figure shows coefficient estimates from the regression of the probability of upper-secondary school enrollment on time on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, age at death, and graduation year fixed effects.

Figure A.6.: Heterogeneity by father's educational attainment (GPA)



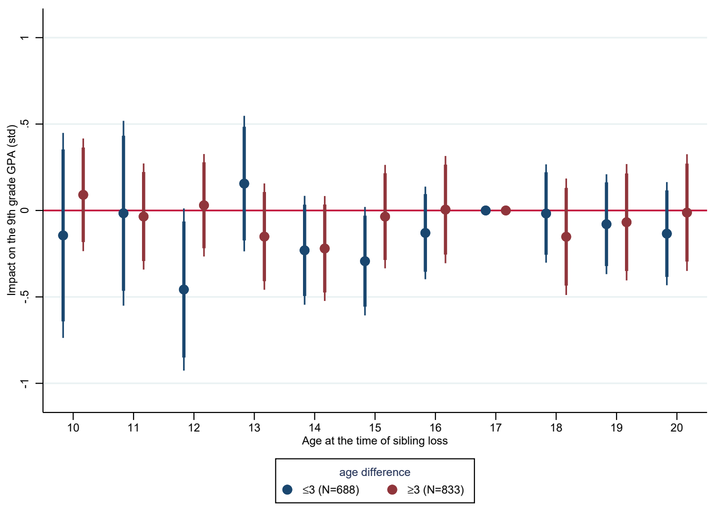
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for surviving siblings with fathers having (i) primary and lower-secondary, (ii) upper-secondary and non-tertiary, and (iii) tertiary and higher educational attainment. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, mother's highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.7.: Heterogeneity by presence of other siblings (GPA)



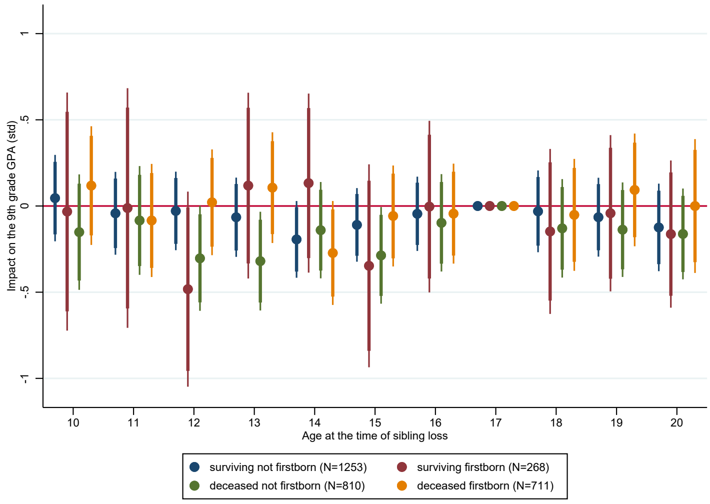
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for those with no other siblings and those with other siblings. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.8.: Heterogeneity by age difference (GPA)



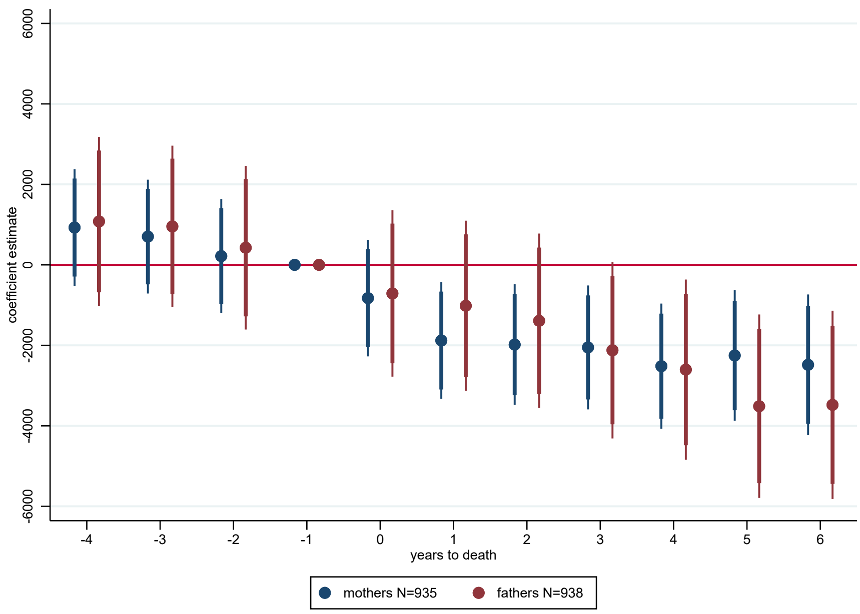
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for those with an absolute age difference ≤ 3, and the rest. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.9.: Heterogeneity by birth order (GPA)



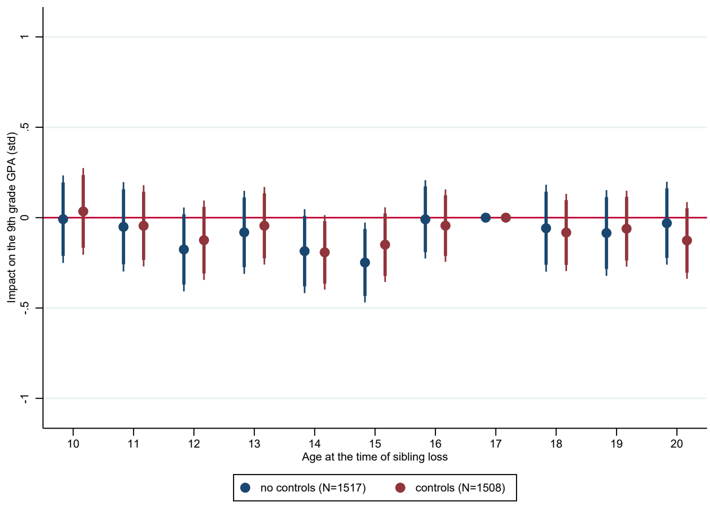
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, separately for (i) later-born surviving siblings, (ii) firstborn surviving siblings, (iii) surviving children losing a later-born deceased sibling, and (iv) surviving children losing a firstborn deceased sibling. Control variables include the child's gender, being a native Finnish speaker, being firstborn (only for (iii) and (iv)), living in an urban area, HH size, both parents' highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.10.: Parental income



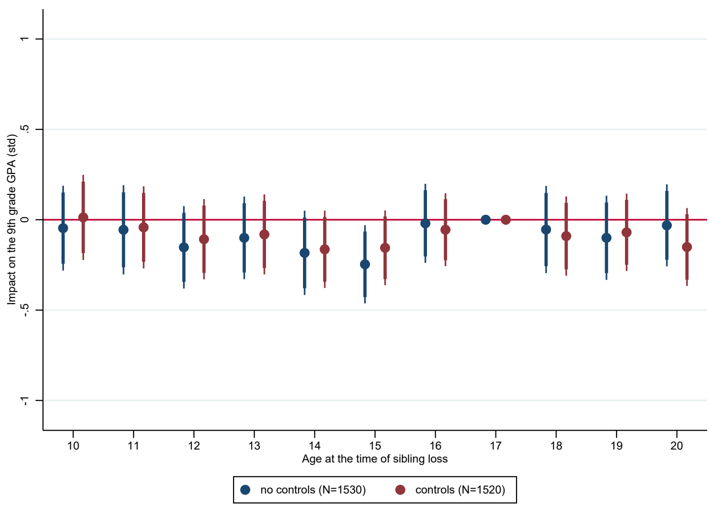
Notes: This figure shows coefficient estimates from the regression - separately estimated for mothers and fathers of the surviving children - of income on the "event time", and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include age and year fixed effects.

Figure A.11.: GPA standardized within schools



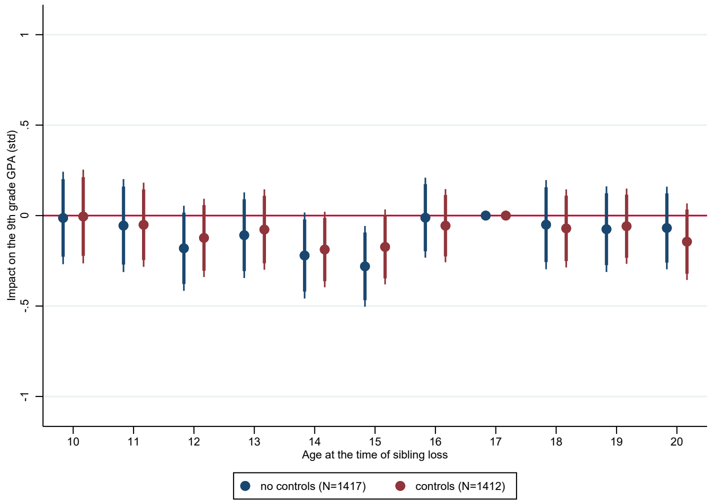
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA (within year-school) on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' highest educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.12.: Impact of losing a sibling on the 9th grade GPA - Parents' income



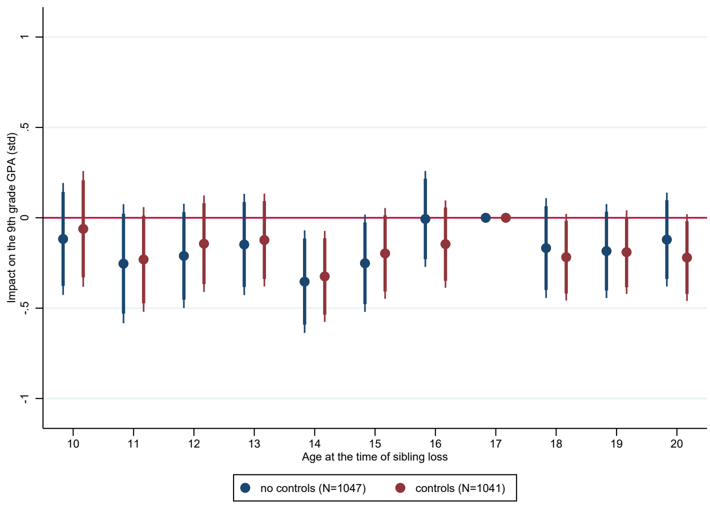
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' income measured one year before the child's death, deceased child's age at death, and graduation year fixed effects.

Figure A.13.: Impact of losing a sibling on the 9th grade GPA - Full siblings



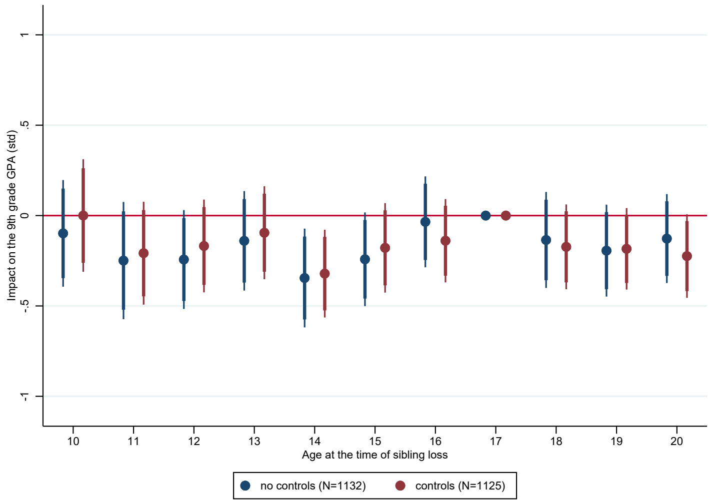
Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, for the subsample of the full siblings who has the same biological father in addition to the same biological mother as the deceased child. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.14.: Surviving children not hospitalized (GPA)



Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, for the subsample of children who were not hospitalized within the same month as the loss or the following month. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' educational attainment, deceased child's age at death, and graduation year fixed effects.

Figure A.15.: Impact of losing a sibling on the 9th grade GPA - 1994 onwards



Notes: This figure shows coefficient estimates from the regression of standardized 9th grade GPA on the categorical variable of the age at sibling loss and corresponding confidence intervals of 95% and 90% in thin and thick lines, respectively, for the subsample of children who lost a sibling between 1994 and 2016. Control variables include the child's gender, being a native Finnish speaker, being firstborn, living in an urban area, HH size, both parents' educational attainment, deceased child's age at death, and graduation year fixed effects.

B. Appendix B: Details of the Data

B.1. Main Datasets

This section summarizes the datasets used in this study and the outcome variables.

Background characteristics of the deceased and surviving children as well as their parents and relevant information on the household level are obtained from the FOLK basic, FOLK household-dwelling, and FOLK income modules. To identify siblings, I use the biological mother links from the FOLK child-parents dataset.

Year and cause of death information comes from Cause of Death Registers which cover the deaths of all individuals (excluding those born and died within the same year) with a Finnish personal identification number and with a known cause of death, for years between 1988 and 2016.

Joint application registers provide information on the 9th grade GPA as well as the application and admission to upper secondary schools of all students who apply for an upper-secondary educational institution. Students Data Module includes individual-level data on students who have been enrolled in upper secondary school and higher levels of education, starting from 1995.

I use hospitalization information for the robustness checks from the Register of Healthcare collected by the Finnish Institute for Health and Welfare (THL) which has been available starting from 1994. Antidepressant Prescriptions are reported in Prescriptions Registers.

B.2. Outcome Variable Descriptions

The **9th-grade GPA** is reported in the Joint Application Registers for Upper-secondary Schools. For the analysis, I standardize it within each graduation year. The indicator for **graduating on time** is a dummy variable that I construct using the year of graduation certificate reported in the same dataset.

Enrollment in upper-secondary school on time is constructed from the Students Data Module, by using the age of the individual and the year. Similarly, the indicator for being enrolled in a **general track** in the upper-secondary school is reported in this dataset. Both variables are only available from 1995.

Antidepressant prescriptions of surviving siblings and their parents are obtained from Prescription Records which are collected by Kela, the Social Insurance Institution of Finland. The dummy variable takes the value of 1 if the person is prescribed antidepressants at least once during the year, and 0 otherwise. This variable is only available from 1993.

Earned income for parents is the sum of labor and entrepreneurial income reported in the FOLK Income Module.

Sick Leave for parents is a dummy variable that I construct by using the sickness allowance reported in the FOLK Income Module. It takes the value of 1 if the person received a positive sickness allowance during the year, and 0 otherwise. This variable is only available from 1995 onwards.

The **employment** variable for parents considers entrepreneurs or wage-earners during the last week of the year, and it is reported in the FOLK Basic Module.

5. Concluding Remarks

This dissertation explores the ways in which some relevant societal and familial changes influence educational outcomes and individual human capital accumulation, with a specific focus on gender inequality and the spillover effects of health shocks. Considering recent global improvements such as the wide adoption of technology use, prevalent challenges like conflicts and epidemics, and ever-evolving family dynamics, I examine relevant contemporary factors shaping educational decisions. The three empirical chapters not only offer important policy implications but also contribute valuable insights for future research.

Chapter 2 addresses one of the most significant global transformations of recent years — the widespread adoption of online learning technologies. In this chapter, we document the gender gaps in effort and performance outcomes when using an online learning platform, and how these gaps differ by the gender of the parent mainly supervising the children. Using data at the individual level from a widely used online learning platform in Spain, we find significant gender gaps in the relative performance outcomes in favor of boys. On the other hand, the effort gender gaps are only significant and economically meaningful when comparing the siblings of the opposite gender. The effort gaps are narrower or positive in favor of girls for children mainly supervised by their mothers. Further, we observe narrower or positive gender gaps in effort outcomes for children living in municipalities with more egalitarian gender norms, while for the relative performance outcomes, there are no such differences.

This chapter contributes to the economics literature on gender gaps in education mainly by addressing the understudied gender gap in online learning outcomes, gender gaps in motivation and effort, and the association between gender social norms and gender gaps in education. Taking into account the increase in the use of online learning tools and their progressive integration into the regular educational system, these results provide important information to minimize gender biases in these new settings. Moreover, the findings on narrower or positive gender gaps in favor of girls in effort outcomes for children supervised by their mothers suggest that the differential parental investments based on the gender of the children might not be the same for mothers and fathers. Even though these results do not imply causal effects, and there could be potentially confounding factors affecting both the gender gap among siblings and the gender of the parent supervising them, they highlight the need for future research on potential differential parental investments based on the gender of the child and the investing parent.

Delving into the backbone of societal changes, Chapter 3 explores the social protests and democratization movements. In this chapter, we analyze how such movements affect the economic empowerment of women through enhanced educational achievements driven by shifts in beliefs and aspirations, in the context of the Arab Spring. We show that female MENA immigrants, after the Arab Spring, exhibit a more progressive stance in their beliefs and aspirations compared to their non-MENA counterparts. This effect, however, is not observed among male MENA immigrants. Shifting our focus to second-generation immigrants and economic outcomes, we find an increase in educational attainment and participation in formal education for second-generation MENA females in Spain after the Arab Spring. Additionally, we observe a decrease in the probability of being NEET (not in education, employment, or training) and employed for second-generation MENA females, while no significant changes are observed for the outcomes of second-generation MENA males.

The main contribution of this chapter to the economics literature is to provide new evidence on the economic effects of social movements and protests by analyzing the impacts beyond the region where the social movement is rooted. We reveal the effect of a significant social and political event on the aspirations and economic choices of immigrants. Aligned with our expectations, increased visibility, and active involvement of women in the Arab Spring protests, we observe a shift in women's beliefs and aspirations towards progressivism and economic empowerment. These results indicate how strong the effects of the social protests and democratization movements could be in shifting beliefs and aspirations, and consequently affecting economic outcomes, even getting beyond the borders of countries and regions. They also provide insights for policy implications, highlighting that policies aiming to progressively improve the aspirations of women could potentially lead to increased education and economic empowerment, creating more equal societies. In this direction, future research can contribute to a deeper understanding of the long-term implications of such transformative events and shifts in beliefs and aspirations on individuals and societies, especially from a gender equality perspective. Moreover, not only educational decisions but also other relevant life outcomes such as marriage and fertility decisions are potentially to be affected by such events and to be examined in future research.

Finally, Chapter 4 shifts the focus from the changes in society to those in the family. In this chapter, I explore the impact of sibling loss during childhood on the surviving siblings' educational outcomes, using detailed administrative data from the entire population of Finland and focusing on 24 birth cohorts. It exploits the timing of an unexpected sibling loss caused by a traffic accident relative to the time of 9th-grade GPA measurement. Results suggest that losing a sibling 2 years before the 9th grade has a negative impact on the 9th-grade GPA, and this impact is more pronounced and prevalent across different ages at the time of sibling loss for children with lower socioeconomic backgrounds. Moreover, there is a decrease in the probability of general track choice in the upper-secondary school following a sibling loss. Delving into the potential mechanisms, I find significant increases in the probability of antidepressant prescriptions for the surviving children and their parents. Furthermore, mothers' probability of taking sick leave increases, and their employment probability decreases after a child's loss, while fathers' labor market outcomes remain mostly unaffected except for a decrease in earnings 3 years after death.

This chapter mainly contributes to the large literature on the spillover effects of health shocks, providing evidence on the understudied causal impact of child death on the surviving siblings' outcomes, as well as examining the mechanisms of mental well-being and parents' labor market trajectories following the loss. The impact of such an experience extends beyond the immediate grief, influencing both short and long-term life outcomes for all surviving family members. As this chapter reveals, the education of surviving siblings, the labor market outcomes of parents - especially mothers - and the mental health of all family members are adversely affected. Severe social and economic implications could potentially follow these effects, which requires dedicated policy attention. With the help of the appropriate policies regarding mental health support and labor market institutions, the social and economic long-term costs of such events could be minimized on a broader scale, along with the potential resulting inequalities for future generations. Pointing out the large negative consequences of such a devastating life experience, this chapter highlights the importance of future research exploring the long-term and intergenerational effects of such losses. Although these extreme cases are not as likely to occur as mild to severe health shocks of children, their consequences are very important to be explored considering the severe effects documented in this chapter.

References

Aaskoven, M. S., Kjær, T., & Gyrd-Hansen, D. (2022). Effects of parental health shocks on children's school achievements: A register-based population study. Journal of Health Economics, 81, 102573.

Adhvaryu, A., Daysal, N. M., Gunnsteinsson, S., Molina, T., & Steingrimsdottir, H. (2022). Impacts of Child Health on Families: Evidence from Childhood Cancers.

Agamile, P., & Lawson, D. (2021). Rainfall shocks and children's school attendance: evidence from Uganda. Oxford Development Studies, 49(3), 291-309.

Alam, S. A. (2015). Parental health shocks, child labor and educational outcomes: Evidence from Tanzania. Journal of health economics, 44, 161-175.

Anaya, L., Iriberri, N., Rey-Biel, P., & Zamarro, G. (2022). Understanding performance in test-taking: The role of question difficulty order. Economics of Education Review, 90, 102293.

Andrew, A., Cattan, S., Costa Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., ... & Sevilla, A. (2020). Inequalities in children's experiences of home learning during the COVID-19 lockdown in England. Fiscal Studies, 41(3), 653-683.

Baker, M., & Milligan, K. (2016). Boy-girl differences in parental time investments: Evidence from three countries. Journal of Human Capital, 10(4), 399-441.

Balart, P., Oosterveen, M., & Webbink, D. (2018). Test scores, noncognitive skills and economic growth. Economics of Education Review, 63, 134-153.

Banks, J., & Mazzonna, F. (2012). The effect of education on old age cognitive abilities: evidence from a regression discontinuity design. The Economic Journal, 122(560), 418-448.

Bargain, O., Boutin, D., & Champeaux, H. (2019). Women's political participation and intrahousehold empowerment: Evidence from the Egyptian Arab Spring. Journal of Development Economics, 141, 102379.

Batut, C., Coly, C., & Schneider-Strawczynski, S. (2021). It's a man's world: culture of abuse, # MeToo and worker flows. Università Bocconi, The Dondena Centre.

Becker, G. S. (1992). Human capital and the economy. Proceedings of the American philosophical society, 136(1), 85-92.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical

Society: series B (Methodological), 57(1), 289-300.

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., ... & Simonsen, M. (2021). Sibling spillovers. The Economic Journal, 131(633), 101-128.

Bolton, J. M., Au, W., Chateau, D., Walld, R., Leslie, W. D., Enns, J., ... & Sareen, J. (2016). Bereavement after sibling death: a population-based longitudinal case-control study. World Psychiatry, 15(1), 59-66.

Bratti, M., & Mendola, M. (2014). Parental health and child schooling. Journal of health economics, 35, 94-108

Breining, S. N. (2014). The presence of ADHD: Spillovers between siblings. Economics Letters, 124(3), 469-473.

Breivik, A. L., & Costa-Ramón, A. (2021). The career costs of children's health shocks. University of Zurich, Department of Economics, Working Paper, (399).

Burton, P., Chen, K., Lethbridge, L., & Phipps, S. (2017). Child health and parental paid work. Review of Economics of the Household, 15(2), 597-620.

Chetty, R., Friedman, J., Hendren, N., & Stepner, M. (2020). "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data" Opportunity Insights Working Paper

Cheng, I. H., & Hsiaw, A. (2022). Reporting sexual misconduct in the # MeToo era. American Economic Journal: Microeconomics, 14(4), 761-803.

Conti, G., Heckman, J., & Urzua, S. (2010). The education-health gradient. American economic review, 100(2), 234-238.

Currie, J., & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. The Quarterly journal of economics, 118(4), 1495-1532.

Del Boca, D., Oggero, N., Profeta, P., & Rossi, M. (2020). Women's and men's work, housework and childcare, before and during COVID-19. Review of Economics of the Household, 18(4), 1001-1017.

DeMars, C. E., Bashkov, B. M., & Socha, A. B. (2013). The Role of Gender in Test-Taking Motivation under Low-Stakes Conditions. Research & Practice in Assessment, 8, 69-82.

Dhanaraj, S. (2016). Effects of parental health shocks on children's schooling: Evidence from Andhra Pradesh, India. International Journal of Educational Development, 49, 115-125.

Di Pietro, G. (2018). The academic impact of natural disasters: evidence from L'Aquila earthquake. Education Economics, 26(1), 62-77.

Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. Proceedings of the National Academy of Sciences, 108(19), 7716-7720.

El-Mallakh, N., Maurel, M., & Speciale, B. (2018). Arab Spring protests and women's labor market outcomes: Evidence from the Egyptian revolution. Journal of Comparative Economics, 46(2), 656-682.

Eriksen, T. L. M., Gaulke, A., Skipper, N., & Svensson, J. (2021). The Impact of Childhood Health Shocks on Parental Labor Supply. Journal of Health Economics, 102486.

Ferhat, S., Hotte, R., & Verwimp, P. (2022). Entry into Marriage, Motherhood and the Arab Spring: Evidence from Egypt. ECARES.

Fernández, R. (2011). Does culture matter?. Handbook of social economics, 1, 481-510.

Fletcher, J., Hair, N. L., & Wolfe, B. L. (2012). Am I my brother's keeper? Sibling spillover effects: The case of developmental disabilities and externalizing behavior (No. w18279). National Bureau of Economic Research.

Fletcher, J., Mailick, M., Song, J., & Wolfe, B. (2013). A sibling death in the family: Common and consequential. Demography, 50(3), 803-826.

Fletcher, J., Vidal-Fernandez, M., & Wolfe, B. (2018). Dynamic and heterogeneous effects of sibling death on children's outcomes. Proceedings of the National Academy of Sciences, 115(1), 115-120.

Gautier, T. (2021). Essays in development economics (Doctoral dissertation).

Gerhardt, C. A., Fairclough, D. L., Grossenbacher, J. C., Barrera, M., Jo Gilmer, M., Foster, T. L., ... & Vannatta, K. (2012). Peer relationships of bereaved siblings and comparison classmates after a child's death from cancer. Journal of pediatric psychology, 37(2), 209-219.

Ghazalian, P. L. (2022). The effects of the Arab Spring on female labour force participation in the MENA region. Economics of Transition and Institutional Change, 30(4), 869-900.

González de San Román, A., & De La Rica, S. (2012). Gender gaps in PISA test scores: The impact of social norms and the mother's transmission of role attitudes.

Gunnsteinsson, S., & Steingrimsdottir, H. (2019). The Long-Term Impact of Children's Disabilities on Families (No. 6-2019). Working paper.

Gevrek, Z. E., Gevrek, D., & Neumeier, C. (2020). Explaining the gender gaps in mathematics achievement and attitudes: The role of societal gender equality. Economics of Education Review, 76, 101978.

Golin, M. (2021). Gender Gaps in Home Schooling Time. mimeo.

Ikeda, M., & Yamaguchi, S. (2021). Online learning during school closure due to COVID-19. The Japanese Economic Review, 72(3), 471-507.

Johnson, E., & Reynolds, C. L. (2013). The effect of household hospitalizations on the educational attainment of youth. Economics of Education Review, 37, 165-182.

Johnson, R. C. (2010). The health returns of education policies from preschool to high school and beyond. American Economic Review, 100(2), 188-194.

Khamis, S. (2011). The Arab" feminist" spring?. Feminist Studies, 37(3), 692-695.

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and gender inequality: Evidence from Denmark. American Economic Journal: Applied Economics, 11(4), 181-209.

Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer?. Review of Income and Wealth, 55(1), 128-165.

Kristiansen, I. L. (2021). Consequences of serious parental health events on child mental health and educational outcomes. Health Economics, 30(8), 1772-1817.

Kvist, A. P., Nielsen, H. S., & Simonsen, M. (2013). The importance of children's ADHD for parents' relationship stability and labor supply. Social Science & Medicine, 88, 30-38.

Le, H. T., & Nguyen, H. T. (2017). Parental health and children's cognitive and noncognitive development: New evidence from the longitudinal survey of Australian children. Health Economics, 26(12), 1767-1788.

Levy, R. E., & Mattsson, M. (2023). The effects of social movements: Evidence from# MeToo. Available at SSRN 3496903.

Liu, J., Modrek, S., & Sieverding, M. (2019). The effects of political protests on youth human capital and well-being in Egypt. Social Science & Medicine, 243, 112602.

Luca, D. L., & Bloom, D. E. (2018). The Returns to Parental Health: Evidence from Indonesia (No. w25304). National Bureau of Economic Research.

Lundberg, S., Pabilonia, S. W., & Ward-Batts, J. (2007). Time allocation of parents and investments in sons and daughters. Unpublished paper.

Mammen, K. (2011). Fathers' time investments in children: do sons get more?. Journal of Population Economics, 24(3), 839-871.

Mendolia, S., Nguyen, N., & Yerokhin, O. (2019). The impact of parental illness on children's schooling and labour force participation: evidence from Vietnam. Review of Economics of the Household, 17(2), 469-492.

Montolio, D., & Taberner, P. A. (2021). Gender differences under test pressure and their impact on academic performance: a quasi-experimental design. Journal of Economic Behavior & Organization, 191, 1065-1090.

Morefield, B. (2010). Parental health events and children's skill development. University of North Carolina at Greensboro-Department of Economics Working Papers, 10-11.

Neuman, E. (2018). Source country culture and labor market assimilation of immigrant women in Sweden: evidence from longitudinal data. Review of Economics of the Household, 16(3), 585-627.

Nobel Prize Outreach AB. (2024). Tawakkol Karman - Facts. NobelPrize.org. https://www.nobelprize.org/prizes/peace/2011/karman/facts/

Nollenberger, N., Rodríguez-Planas, N., & Sevilla, A. (2016). The math gender gap: The role of culture. American Economic Review, 106(5), 257-61.

OECD (2019), PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris, https://doi.org/10.1787/5f07c754-en.

Parman, J. (2013). Childhood health and sibling outcomes: The shared burden and benefit of the 1918 influenza pandemic (No. w19505). National Bureau of Economic Research.

Price, J. (2008). Parent-child quality time does birth order matter?. Journal of human resources, 43(1), 240-265.

Rodríguez-Planas, N., & Nollenberger, N. (2018). Let the girls learn! It is not only about math… it's about gender social norms. Economics of Education Review, 62, 230-253.

Shihada, I. (2011). Women and the Arab Spring: Expectations and Concerns. VOLUME 8 ISSUE, 245.

Singerman, D. (2013). Youth, gender, and dignity in the Egyptian uprising. Journal of Middle East Women's Studies, 9(3), 1-27.

Stans, R. (2020). A Setback Set Right? Unfortunate Timing of Family Distress and Educational Outcomes (No. crctr224_2020_206). University of Bonn and University of Mannheim, Germany.

Sun, A., & Yao, Y. (2010). Health shocks and children's school attainments in rural China. Economics of Education Review, 29(3), 375-382.

Thamarapani, D., Rockmore, M., & Friedman, W. (2020). The Educational and Fertility Effects of Sibling Deaths. Available at SSRN 3172537.

Thomas, D., Beegle, K., Frankenberg, E., Sikoki, B., Strauss, J., & Teruel, G. (2004). Education in a Crisis. Journal of Development economics, 74(1), 53-85.

Vaalavuo, M., Salokangas, H., & Tahvonen, O. (2023). Gender inequality reinforced: the impact of a child's health shock on parents' labor market trajectories. Demography, 60(4), 1005-1029.

Van den Berg, G. J., Lundborg, P., & Vikström, J. (2017). The economics of grief. The Economic Journal, 127(604), 1794-1832.

Wantchekon, L., Klašnja, M., & Novta, N. (2015). Education and human capital externalities: evidence from colonial Benin. The quarterly journal of economics, 130(2), 703-757.

Weldeegzie, S. G. (2017). Growing-up unfortunate: War and human capital in Ethiopia. World Development, 96, 474-489.

Yi, J., Heckman, J. J., Zhang, J., & Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcomes. The Economic Journal, 125(588), F347-F371.

Yu, Y., Liew, Z., Cnattingius, S., Olsen, J., Vestergaard, M., Fu, B., ... & Li, J. (2017). Association of mortality with the death of a sibling in childhood. JAMA pediatrics, 171(6), 538-545.