Essays on Human Capital and Economic Development

by

Yasar Ersan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2023

Doctoral Committee:

Professor Hoyt Bleakley, Co-Chair
Professor Uday Rajan, Co-Chair
Professor John Bound
Professor Dean Yang

Yasar Ersan
yasare@umich.edu
ORCID iD: 0000-0001-6468-9204

ACKNOWLEDGEMENTS

This thesis is a product of collective efforts. It is not only a product of my individual effort. First, I thank my committee members, Hoyt Bleakley, John Bound, Uday Rajan, and Dean Yang for their constructive guidance to make my thesis come true. Without their guidance, everything was up in the air. All of my research ideas were crude and far from scientific. Throughout our meetings, I benefited from their experience and started constructing causal chains with arrows and circles. Their guidance helped me stop thinking like a high school student and I began to think scientifically, which has been increasing my understanding to assess the reliability and validity of my present and future studies. If my thesis has any analytical coherence and scientific value, although I cannot adequately express my gratitude, they deserve most thanks for their guidance. I also thank my co-authors for their cooperation and patience. I particularly thank Ilhan Can Ozen and Abdurrahman Aydemir.

I also thank Emir Murathanoglu, Guzhan Gulay, Mahmut Aydogmus, Luis Alberto Baldomero Quintana, Seven Agir, Murat Kirdar, Mark Dincecco, and Mohamed Saleh for their feedback and comments. I would like to thank the participants of Michigan's Economic Development Seminar (EDS), Michigan's Economic History Seminar (EHS), METU Economics Departmental Seminar, Spring 2023 WEFIDEV Seminar Series, Bogazici University Economics Seminar, and the 4^{th} LSE Workshop on Political Economy of Turkey for their helpful feedback and comments.

In the first chapter, I employ an administrative data set in my empirical analysis. I owe thanks to Adem Gun and Ugur Arslan for providing the overview of the policies and regulations of the individual retirement system in Turkiye and assembling the corresponding data. I also would like to express my gratitude to John Thiels for the technical support while running my STATA codes with high-speed computers. In the third chapter, similar to the first chapter, I exploit an administrative data set. I also thank Ozgur Uysal and Nagihan Altinisik for making the data ready to make estimations. Especially, Nagihan Altinisik spent hours and hours and was always reachable when there was an issue with my STATA codes. Sometimes we met through online meetings in which the internet connection was really bad to run my codes. Therefore, I will never forget her contribution to my thesis.

Working with administrative data requires multiple things at the same time. I do not know what he does or what sort of abilities he has, but Guzhan Gulay solved the technical or data access problems immediately and made everything run smoothly simultaneously. Furthermore, whenever I got hopeless and lost my motivation, he spent a significant amount of time reminding me of why I do Ph.D. In short, I thank Guzhan Gulay for what he did for me.

I owe debts to Laura Flak, Lauren Pulay, and Laura Howe for their assistance and help to overcome bureaucratic and administrative issues during my doctoral studies. Moreover, Pinar Ustel and Linroy Marshall opened their apartment for me while I was technically homeless. Their apartment was always a shelter for me. I also spent a really good time with their little and cute son, Leon. I will never forget their hospitality and generosity as well as memories.

I spent the first three years of my Ph.D. study in the US whereas in the last three years, I was mostly in Turkey due to COVID-19 and thanks to fellowships. During my Ph.D., I was away from my partner, Elif Burcu Saygin, and so was she away from me. Therefore, I can easily say that she spent as much as effort I spent during my study. There is no way to express my gratitude to her for her unconditional love and support. I am grateful to her and we did it together. Furthermore, I thank my mother Fatma Ersan, and my father Ismet Ersan. I also would like to thank my mother-in-law Guler Saygin and my father-in-law Mustafa Saygin for their support to me and Elif. Finally, I am certainly sure that I forgot to thank many people so I also express my gratitude to them.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

APPENDIX

ABSTRACT

This dissertation has three separate essays on human capital and economic development. In the first chapter, I examine how general education impacts savings for retirement and investment behavior in pension portfolios. In the second chapter, I study a historical human capital intervention with massive investment to shed light on to what extent it determines the long-run trajectory of local development. In the last chapter, I investigate how education determines stock market participation, the variation in investment behavior, and portfolio performance in stock market portfolios.

The first chapter studies the causal impacts of education on participation and wealth in defined contribution pension plans, using the 1997 Education Reform in Turkey that led to the substantial exogenous variation in schooling across birth cohorts. Employing a regression discontinuity design with an administrative data set spanning the universe of individual retirement accounts in Turkey, I find that the education reform increasing schooling by around half a year does not improve participation in defined contribution pension plans. Despite the strong positive correlation between education and the propensity to participate in defined contribution pension plans, I fail to find any causal evidence. However, I find that education reform improves pension wealth by around 3% for females but no improvement for males. I also examine financial channels through which education can potentially drive the wealth effects. Yet, I find no overall significant education impacts on equity participation and the share of wealth invested in equities. The increase in schooling does not impact financial sophistication, behavioral biases common in pension plans, also investment performance. Thus, education promotes wealth accumulation through the labor market channels rather than encouraging financial skills.

In the second chapter, co-authored with Ilhan Can Ozen, we explore a historical experiment. The American Board of Commissioners for Foreign Missions (ABCFM) had a significant foothold in Anatolia for over a century by investing heavily in human capital through a large number of schools, colleges, and various craft-skill training activities. Through extensive archival work, we study how ABCFM with its human capital intervention impacted long-run local economic development. Using the spatial variation in the built and functional mission stations, we find areas closer to ABCFM missions have presently higher general

development index by 0.07-0.12 standard deviation and higher income by 5%-17% in 10 km proximity. We identify the mission impacts by exploiting a placebo set from the group that was conceived but not carried out, and also an exogenous variation in the spatial distribution of ABCFM missions stemming from the re-partition of the working region. The mechanisms driving our results are labor productivity in the agriculture sector, which allows for greater skill differentiation and structural transformation. Gender roles are also significantly transformed through the transmission of cultural norms that put more emphasis on female education, and lower fertility rates.

In the last chapter, co-authored with Abdurrahman Aydemir, we examine the causal impacts of education on stock market participation, the variation in investment behavior, and portfolio performance, by leveraging the exogenous variation in schooling across birth cohorts stemming from the 1997 Education Reform in Turkey. Employing a regression discontinuity design with an administrative data set spanning the universe of stock market participants in Turkey, we find that the education reform increasing schooling by around half a year does not foster stock market participation. Despite the strong positive correlation between education and stock market participation, we fail to find any causal evidence. Moreover, our results indicate education does not significantly spur the likelihood of having a transaction account, and of holding stocks, bonds, funds, risky assets, and stocks that are counted to be safer and more liquid in the stock market portfolios. Consistently, results show that education boosts neither the share of wealth allocated to risky assets nor stock market wealth. Finally, we find no evidence that education significantly leads to higher portfolio returns. Consequently, general education is at best a minor factor in stock market participation and the variation in investors' stock market portfolios.

CHAPTER 1

Education and Savings for Retirement: Evidence from Pension Portfolios in Turkiye

1.1 Introduction

Rising life expectancy and falling birth rates worldwide accelerate population aging, which puts immense pressure to sustain adequate and financially sustainable levels of pensions. OECD (2019) documents that the number of people over 65 for every 10 workers in the OECD area was 2 in 1980, 3 in 2020, and projected to be 6 in 2060. Moreover, almost half of all working-age households in the US have zero retirement savings in their retirement accounts in 2016 (Morrissey (2016)). The undersaving for retirement is not only unique to developed countries but also a significant challenge for developing ones (Chetty (2015)). Consistently, the proportion of the elderly population with incomes below 50% of relative poverty thresholds is the largest among OECD countries in the US, Mexico, and Turkiye (OECD (2021)). Older people are the poorest in Turkiye as the old age average social pension salary was only 21% of the poverty line in 2017 (ILO (2021)). To cope with the inadequacy of savings for retirement and elderly poverty, many countries have introduced individual private pensions since the mid-1990s, which has led to the shift from defined benefit (DB) pension plans to defined contribution (DC) pension plans.[1] Consequently, the total asset in private pension plans in the OECD area was almost $42.5 trillion in 2018 (OECD (2019)). However, disintermediation goes hand in hand with the spread of DC pension plans, so individuals have to decide to join pension plans, conditional on participation, how much to save, and how to allocate. Most importantly, individuals bear the risks emerging in financial markets.

This article studies the causal impacts of schooling on the decision to participate, wealth accumulation in defined contribution pension plans, and the variation in investment behav-

[1] Defined benefit (DB) pension plans offer specified payment amounts in retirement whereas defined contribution (DC) pension plans allow individuals to contribute and invest on their own. For more details, see Arcanjo (2019).

ior in pension portfolios in a middle-income country, Turkiye. Schooling is a major force driving economic development and also its level is the primary indicator of economic development from the perspective of policymakers (Filmer et al. (2018)). Several studies have estimated the impact of human capital on labor, health, and various other outcomes.[2] Yet, less is known about the explanatory power of human capital in the decision to participate, wealth accumulation, and the variation in investors' portfolios in DC pension plans. Eliciting the causal relationship between education and the corresponding outcomes is a difficult task because it requires comprehensive and detailed data to track individual portfolios and a plausible exogenous variation in schooling. Exploiting a nationwide education reform extending compulsory schooling from 5 to 8 years for those born after January 1987 in Turkiye and a highly detailed and novel administrative data set, I study for the first time in the literature the causal effects of education on savings for retirement and investment decisions in pension portfolios.

Despite the OLS estimates documenting a strong positive correlation between years of schooling and participation in DC pension plans, yet, results indicate that schooling causally has no overall significant effect on participation, i.e. the propensity to have a positive balance in private retirement accounts, in DC pension plans. However, I find compelling evidence that schooling is a crucial input to accumulating more pension wealth in DC pension plans. The education reform increases wealth in DC pension plans by 3% for females, and null effects for males, indicating a strong heterogeneity by gender. Scaling those reduced-form estimates by the increase in years of schooling induced by the 1997 Education Reform reveals that an extra year of schooling improves pension wealth accumulation by 5-6% for females while having no effect for males.

Using various plausible proxies for financial literacy and sophistication, and also for behavioral biases and heuristics prevalent in DC pension plans, I examine the potential causal mechanisms driving the education effects on wealth accumulation in DC retirement accounts. Estimates reveal no overall economically significant evidence that schooling changes equity exposure in portfolios, particularly the likelihood of holding equity funds, and the share of wealth invested in either equity funds, stocks, or risky assets. Results also suggest that general education does not make a difference in avoiding financial mistakes, behavioral biases, and heuristics. Lastly, I find no supporting evidence that portfolio returns in DC pension plans significantly vary by education. Therefore, it is unlikely to argue that schooling increases financial skills such as financial literacy and sophistication. Overall, human capital does not appear to be a significant input to constructing financially desirable portfolios, at best it is a minor factor. Accordingly, the underlying channel of higher wealth accumulation

[2] Oreopoulos and Salvanes (2011).

in DC pension plans seems to be labor market skills gained through more education rather than financial skills.

To isolate the causal education effects, I exploit the substantial exogenous variation in schooling across month-year birth cohorts brought about by the 1997 Education Reform in Turkiye that extended compulsory schooling from 5 to 8 years. Taking advantage of the discontinuity at the month-year of birth of January 1987 arising from the education reform, I employ a regression discontinuity design. Those born after January 1987 serve as the treatment group, whereas those born before January 1987 form the control group in my quasi-experimental research design.

To uncover how the 1997 Education Reform has changed schooling outcomes, I first employ Household Labor Force Surveys in 2018 assembled by the Turkish Statistical Institute. Secondly, to quantify the impacts of schooling on the decision to participate, wealth accumulation, and the variation in investors' pension portfolios in DC pension plans, I use an unusually high-quality and detailed administrative data set spanning the universe of individual retirement accounts in Turkiye in December 2019 and 2020. The relevant data set is comprised of the month-end snapshots in December 2019 and 2020 with information on portfolio details such as fund choices, account balances, trading, and contribution amounts, and is provided by Borsa Istanbul Group which is the legal authority to keep the records of individual retirement accounts in DC pension plans in Turkiye.

1.1.1 Contribution to Literature

Prior studies have emphasized that demographics (Engström and Westerberg (2003), and Duflo et al. (2006)), behavioral problems (Benartzi and Thaler (2007)), peer effects (Duflo and Saez (2002, 2003)), the complexity of the pension plans (Iyengar and Kamenica (2010)), and financial education (Lusardi and Mitchell (2014)) are important drivers of participation and wealth accumulation in DC pension plans. On the one hand, by estimating the causal impacts of education in savings for retirement in DC pension plans this article speaks to the literature on savings for retirement. On the other hand, this article relates to the extensive literature on financial and non-financial returns of schooling (Oreopoulos and Salvanes (2011)). A large body of work documents the effects of schooling on labor income,[3] health and fertility,[4] crime,[5] financial behavior,[6] capital returns,[7] cognitive skills,[8] and domestic violence.[9] I build on the literature by studying the causal impacts of schooling on savings for retirement in DC pension plans. To the best of my knowledge, no study has estimated the causal effects of education on the corresponding outcomes. Results suggest that education promotes savings in DC pension plans, however, it does not impact the decision to join DC pension plans and the variation in investors' pension portfolios.

A growing body of literature hypothesizes that human capital is a close substitute for bonds rather than stocks if the correlation between labor income and stock returns is small.[10] Following this prediction, some empirical studies report a strong positive correlation between education and equity participation, both directly through stock holdings and indirectly through participation in DC pension plans (Gomes et al. (2021), and Egan et al. (2021)).[11] Black et al. (2018) and Cole et al. (2014) additionally find causal evidence that education promotes equity participation in Norway and the US, respectively. However, recent studies report that education has no causal impacts on financial outcomes such as capital returns (Fagereng et al. (2020a)), financial wealth, and investment behavior (Fagereng et al. (2021)). My study complements this literature in a unique setting with a credible identification strategy and is the first to provide causal estimates of education on equity exposure and risky share in DC pension plans which are the only financial portfolios for the vast majority of the population. Thus, I have access to a broader cross-section than in the studies with non-retirement account financial portfolios.[12] Despite the statistical precision, education effects are very small in magnitude, implying general education does not have overall significant impacts on equity exposure and risky share in DC pension plans.

An extensive literature documents that education is positively associated with financial

[3]Duflo (2001), Oreopoulos (2006), Acemoglu and Angrist (2000), Angrist and Keueger (1991), Card (1993), and Aydemir and Kirdar (2017).

[4]Lleras-Muney (2005), and Black et al. (2008).

[5]Lochner and Moretti (2004).

[6]Cole et al. (2014),Black et al. (2018), and Gray et al. (2021).

[7]Fagereng et al. (2020a).

[8]Carlsson et al. (2015).

[9]Erten and Keskin (2018).

[10]Cocco et al. (2005), Bodie et al. (1992), Fagereng et al. (2017), Guiso and Sodini (2013), Vissing-Jørgensen (2002), Catherine (2021), and Viceira (2001).

[11]The economic intuition is that, as the small correlation between labor income and stock returns implies the future labor income is certain, riskless labor income behaves like an extra endowment of the safe asset. Thus, an investor adjusts her financial portfolio to keep the share of total wealth invested in risky assets fixed. In short, certain future labor income is like a riskless bond and therefore is a close substitute for bonds.

[12]Survey of Consumer Finance in 2019 in the US documents that the share of households with direct stock holdings fell from 17% to 15% whereas households with either direct or indirect stock holdings increased from 52% to 53%. The increase in indirect stock holdings is partly due to the increasing role of 401(k) pension plans.

literacy and sophistication which avoids financial mistakes (Calvet et al. (2009b, 2007, 2009a), and Lusardi and Mitchell (2014)). Arguably, individuals gain more cognitive and numeracy skills with more education so are more financially literate and sophisticated. Additionally, a burgeoning literature emphasizes that investors are prone to behavioral biases and heuristics in DC pension plans that might result in welfare losses (Benartzi and Thaler (2007)).[13] The underlying reason for default effects —a prevalent behavioral problem in DC pension plans— is mainly the cognitive costs of evaluating different saving alternatives (Blumenstock et al. (2018)) and the limited computational capacity (Madrian (2014)). Given the causal contribution of schooling to cognitive skills (Carlsson et al. (2015)), this study adds to the literature by documenting the first causal evidence on the impacts of schooling on common financial mistakes, behavioral biases, and heuristics. However, I find no overall significant evidence that education causally lowers financial mistakes and behavioral problems, consequently, increasing financial literacy.

The rest of the article is organized as follows. In the next section, I briefly describe the Turkish Pension System and the 1997 Education Reform in Turkiye in Section 1.3. Section 1.4 introduces the data employed and renders the details of the research design with a particular emphasis on how I identify the causal impacts of education. Sections 1.5 and 1.6 present the findings and robustness checks, respectively. Section 1.7 concludes the article with a broad discussion of the findings.

1.2 Turkish Pension System

Turkish pension system consists of three pillars: i) mandatory pay-as-you-go public pension system, ii) occupational mostly defined benefit (DB) pension plans, and iii) the voluntary private pension system with fully funded defined contribution (DC) schemes. Even though anybody formally working is covered by the public pension scheme through social security pension, it has been far from being adequate for decades in Turkiye and the average pension salary was only 21% of the poverty line in Turkiye in 2017 as documented by a report of ILO (2021). To avoid the inadequacy of savings for retirement for post-retirement periods, the government introduced private pension accounts in 2003. There are three types of

[13]Most of the investors do not join in DC pension plans despite their advantages such as tax deduction and subsidy, however they opt-in after automatic enrollment nudges (Madrian and Shea (2001) and Thaler and Benartzi (2004)). Even after participation in pension plans, stickiness to default options (Cronqvist and Thaler (2004) Choi et al. (2005), Blumenstock et al. (2018), Beshears et al. (2009), Chetty et al. (2014), Brune et al. (2017), and Choi et al. (2004)), naive diversification strategies such as the conditional $\frac{1}{N}$ heuristic (Benartzi and Thaler (2001), Agnew (2006) and Huberman and Jiang (2006)) and inertia in trading i.e. lack of portfolio rebalancing or reshuffling (Agnew et al. (2003) and Sialm et al. (2015)) are commonly observed behavioral biases and heuristics in portfolios.

private pension plans. Voluntary and employer-sponsored pension plans have existed since 2003 whereas automatic enrollment pension plans have existed since 2017 after a nationwide automatic enrollment policy targeted most non-self-employed individuals working formally. After the automatic enrollment policy, opting out is a choice by investors as the government makes those working formally and younger than 45 enroll in the DC plans. The benefit of signing in pension plans or keeping enrolled in DC pension plans lies in the government's extra contribution. Put another way, for each Turkish Lira contributed to pension accounts by pension investors, the government also contributes 0.25 Turkish Lira for the voluntary and automatic enrollment pension plans, implying a matching rate of 25%.[14] Investors are entitled to the full amount of pension benefits once they are 56 years old given a minimum of 10 years of the coverage period. At the end of 2019, there were 18 licensed companies and 26 portfolio companies managing 404 different pension funds.[15] Almost 60% of 14 million active participants are 25-44 years old, which also rules out the concern that the population this article focuses on is too young to invest in DC pension plans as they are 28-38 years old.

Investors bear the responsibility of whether to participate or not in pension plans, how to invest, how much to invest, choosing the funds in their portfolios, and how to allocate their contributions unless they participate in an employer-sponsored pension plan.[16] In the voluntary pension plans, no limit exists on the amount of investment while the minimum contribution rate is 3% of monthly salary in the automatic enrollment pension plans, but possible to rise it if demanded by investors. Moreover, in the automatic enrollment plans, individuals choose whether to opt-out as opposed to opting in. I also present the share of different types of pension funds in terms of investment value in Figure A.1 in Appendix A, implying strong heterogeneity. Almost 30% of the total investment is in the default fund in line with the studies indicating a strong default effect in pension portfolios.[17] The concerning default fund must contain at least 50% bonds and bills, revenue sharing certificates, and rent certificates which are supposed to be less risky relative to stocks.[18] Figure A.2 in Appendix A illustrates an additional heterogeneity in pension funds with a particular emphasis on their beta coefficients and Sharpe ratios, suggesting a median fund fails to outperform the domestic benchmark equity index.[19] I calculate the beta coefficients using the monthly

[14] The matching rate has been 30% by January 22, 2022.

[15] For detailed information, see Peksevim and Akgiray (2019).

[16] In employer-sponsored pension plans, the authority responsible for managing the value in the pension accounts is the employer. Thus, in my estimations, I exclude those accounts that are only 1.5% of all pension accounts.

[17] Madrian and Shea (2001), Benartzi and Thaler (2007), Blumenstock et al. (2018), Carroll et al. (2009), Brune et al. (2017), Choi et al. (2004), Brown et al. (2016), and Beshears et al. (2009).

[18] For more details, see:
https://www.egm.org.tr/auto-enrollment-system-aes/auto-enrollment-funds/

[19] Beta coefficient shows how an individual fund moves (on average) when the overall stock market increases

returns of each pension fund and the domestic benchmark equity portfolio and also compute the Sharpe ratios using the monthly excess returns of each pension fund.

1.3 The 1997 Education Reform

Before 1997, the Turkish Education system consisted of a compulsory five years of primary school, a voluntary three years of junior high school, and a voluntary three years of high school. For the voluntary three years of junior high school, students were free to choose either a religious or secular school. During the 1990s in Turkiye, pro-Islamist parties gained sizeable popular support, which led a pro-Islamist party to form a government in 1996. With the help of the new government, the religious junior high schools and Quranic schools that were not regulated as heavily as regular education institutions increased their number of enrollments considerably. In this manner, the tension between the governing party and the military escalated. On February 28, 1997, the pro-Islamist party and many of its members were forced to resign from the government and banned from politics (Atılgan et al. (2015)).

The new governing coalition implemented a number of decisions to block the rise of the pro-Islamist social movement in Turkiye. One of the most radical consequences of these decisions appeared in education policy. On August 18, 1997, the Turkish parliament passed a law extending compulsory schooling from 5 to 8 years. By the new law, primary and junior high school was united under the new category of primary education. Moreover, the junior religious schools and Quranic schools were repealed from the education system so students could no longer attend them. Arguably, restricting religious education was the major motivation for the 1997 Education Reform. Accordingly, unlike some major schooling and health reforms in other contexts that went hand in hand with the economic developments and their requirements, the 1997 Education Reform was not driven by economic considerations. As a result, this reform did not coincide with a concurrent economic or any other policy that would have a third-factor impact on schooling.

The 1997 Turkish Education Reform became effective for the 1997-98 schooling year. It compelled the students that completed grade 5 or a lower grade at the end of the 1996-97 schooling year, and also those who did not hold a primary school degree at the beginning of the 1997-98 schooling year to attend mandatory 8 years of compulsory schooling. The education reform was nationwide and implemented strictly by the legal authorities. The

---

or decreases, and the Sharpe ratio measures the performance of pension funds compared to a risk-free asset, after adjusting for its risk. I use the BIST-100 index for the domestic benchmark equity portfolio. For more details, see https://www.borsaistanbul.com/en/sayfa/3542/bist-stock-indices The domestic benchmark equity index, BIST-100 is the primary indicator to quantify the performance of the 100 highest stocks that are publicly traded in Borsa Istanbul in Turkiye regarding the market value and trading volume.

school starting age in Turkiye was 6. Therefore, those born before January 1987 would not be affected by the education reform so they could either drop out or continue further schooling. However, those born after January 1987 and did not hold a primary school degree were mandated to continue at least three more years of schooling and so complete the 8 years of compulsory schooling. Even though, considering the imperfect compliance with the school starting age or grade repetition, there was a possibility that some cases might not have completely fitted this rule, the students born after January 1987 were more likely to comply with the new compulsory schooling than the older cohorts were.

Before 1997, the dropout rate after the completion of 5 years of compulsory schooling was quite higher. Taking into account enrollment ratios reported by national education statistics, one year before the 1997 Education Reform, almost 40% of students just dropped out of school after obtaining a primary school degree. In contrast, the 1997 Education Reform expanded compulsory schooling from 5 to 8 years in Turkiye, leading to a sharp increase in years of schooling. Strikingly, the number of students enrolled in grades 1-8 rose by almost 16% (Aydemir and Kirdar (2017)). The Ministry of Education responded to these increases with a set of measures to prevent the non-schooling of some students because of any shortage or a decline in the quality of schooling. Those measures comprise new school constructions, hiring of new teachers, transportation of students especially for rural areas, and boarding school constructions (Kirdar et al. (2016)).

The measures on the schooling infrastructure hereby raised significantly the student population. Whether the substantial increase in student population might have deteriorated the school quality might be a concern. The TIMSS 1999 and 2007, measuring the cognitive abilities of students across various countries in an internationally standardized way, conversely suggest that the mathematics score of Turkiye rose by 3 points while the average score among all participating countries dropped by 37 points from 1999 to 2007.[20] Similar patterns also arise in the domain of science. Therefore, it is quite hard to argue that the 1997 Education Reform deteriorated the quality of schooling while extending the duration of compulsory schooling.

[20] The scale is a random variable with a mean of 500 and a standard deviation of 100. For more details, see: https://timss.bc.edu/timss2007/pdf/timss2007_internationalmathematicsreport.pdf

1.4 Data and Research Design

1.4.1 Data

I benefit from a variety of data sources. First, I use the nationally representative 2018 Turkiye Household Labor Force Survey (HLFS) assembled by the Turkish Statistical Institute (TURKSTAT). HLFS data set has information about schooling, employment, the month-year of birth, and various demographic variables, which allows me to explore how the 1997 Education Reform altered the distribution of schooling by different month-year birth cohorts.

Employing the HLFS data, I generate various schooling outcomes. The primary schooling outcome of interest is the years of schooling.[21] Additionally, I produce four indicator variables each equal to one if an individual has at least completed the relevant schooling category from the set of primary, junior high school, high school, and college degrees. Yet, HLFS data has no information on pre-determined covariates such as the province of birth, and parental education since such variables are crucial for the verification of the research design. There exists a nationally representative data set with information on the related variables for the only female population, which is Domestic Violence against Women. Up to now, there are two of those surveys in 2008 and 2014, respectively.[22] The relevant schooling outcomes are described in Panel A in Table 1.1 in the bandwidth of those born 60 months before and after January 1987 since the analysis sample falls usually into the corresponding bandwidth.

The primary data to quantify the impacts of education on savings for retirement and the investment decisions in pension portfolios is an administrative data set provided by Borsa Istanbul Group, covering the universe of individual retirement accounts with month-end snapshots in Turkiye in December 2019 and 2020. The data is extraordinarily high quality and detailed as it has information on account balances, portfolio details, contribution amounts, trades, pension funds and their numbers, fund types, the composition of each pension fund by financial asset classes, the strategy of how an individual allocates her money into her funds, rate of return of each fund over time, month-year of birth, gender, and also the province of birth registration. I also note that Borsa Istanbul Group is the legal entity to keep the records of individual retirement accounts by law in Turkiye. Thus, the most striking feature of the concerning administrative data is that it is not prone to any reporting or measurement bias as it has the entire population of individual retirement accounts. Overall,

[21] The HLFS data has information on educational attainment but not the actual years of schooling. Therefore, I assigned 5 years of schooling for the primary school degree, 8 years of schooling for the junior high school degree, 11 years of schooling for the high school degree, 15 years of schooling for the college degree, and 17 years of schooling for the master's degree.

[22] For more information, see Erten and Keskin (2018).

it is quite unlikely to argue that selection bias emerges.

The administrative pension data set lacks information about individuals that have no pension account.[23] Both to measure the impact of schooling on participation i.e. the propensity to have a positive balance in individual retirement accounts and the propensity to hold equity and default funds, I resolve this issue by constructing month-year birth cohort level outcomes using the number of people in each specific month-year birth cohort provided by the TURKSTAT since the 1997 Education Reform is a treatment at the month-year birth cohort level. Then, I first sum the number of individual investors by each month-year birth cohort in the administrative pension data and later divide it by the number of people in each birth cohort in the whole of Turkiye's population, allowing me to have the ratios for each outcome on participation in pension plans and fund choices.

I first estimate the impact of the 1997 Education Reform on participation outcomes and choices of equity and default funds. For each outcome of interest, I compute the share of those participating in a set of pension plans in percentage terms in each month-year birth cohort. Therefore, in this part, the estimation is conducted at the month-year birth cohort level. The related outcomes are at the birth cohort level in months the percentage rates of those participating in any pension plans, participating in either voluntary, automatic, or employer-sponsored pension plans, and contributing monthly. I also note that participation in the related pension plans means that the balance in a specific account is greater than zero. The descriptive statistics for the related outcomes are presented in Panel B in Table 1.1. Moreover, I investigate the outcomes of the choices of funds. I particularly focus on the ratio of those holding equity, default funds, and those holding only default funds in pension portfolios in each month-year birth cohort in percentage terms. The concerning variables are documented in Panel C in Table 1.1.

I also explore individual investors' pension portfolios conditional on participation i.e. investors with positive balances in pension accounts. Thus, the unit of analysis is individual in that part of the estimations. There is strong heterogeneity in individual pension portfolios regarding portfolio size, contribution amounts, portfolio returns, equity exposure, portfolio beta coefficients, and Sharpe ratios as illustrated in Appendix A in Figure A.3. Furthermore, Figure A.4 in Appendix A exhibits a significant variation in portfolio risk measures.

Panel D in Table 1.1 presents the logarithm of pension wealth and monthly contribution amounts, annual portfolio returns in percentage terms, and indicator variable equal to one if an investor's annual portfolio return is greater than the annual return of the default pension

[23]The administrative data with individual retirement accounts has the information of around 24 million people at the end of 2019. The number of people who are 15-64 years old in the labor force was around 33 million. Thus, the administrative data set covers almost 75% of the working-age population in Turkiye in 2019. For more details, see https://data.tuik.gov.tr/Bulten/Index?p=Isgucu-Istatistikleri-Eylul-2019-30688

Table 1.1: Descriptive Statistics for Those Born 60 Months Before and After January 1987

	(1) Control	(2) Treatment	(3) Difference (2)-(1)
Panel A: Schooling Outcomes			
Primary School Degree	0.93	0.9	-0.03
	(0.26)	(0.30)	(0.00)
Junior High School Degree	0.64	0.86	0.23
	(0.48)	(0.34)	(0.00)
High School Degree	0.49	0.59	0.1
	(0.50)	(0.49)	(0.00)
College Degree	0.27	0.35	0.08
	(0.45)	(0.48)	(0.00)
Years of Schooling	9.2	10.31	1.15
	(4.72)	(4.64)	(0.04)
Observations	32781	29445	62226
Panel B: Participation in DC Pension Plans			
Pension Plans Participation Rate (%)	25.05	24.28	-0.77
	(0.63)	(0.44)	(0.10)
Voluntary Pension Plans Participation Rate (%)	24.07	19.61	-4.46
	(1.05)	(2.00)	(0.29)
Employer Sponsored Pension Participation Rate(%)	1.61	1.59	-0.02
	(0.08)	(0.17)	(0.02)
Automatic Pension Plans Participation Rate (%)	26.84	32.07	5.23
	(0.94)	(2.26)	(0.32)
Contributor Rate (%)	14.18	13.07	-1.11
	(0.58)	(0.47)	(0.10)
Panel C: Choices of Funds			
Equity Fund Ownership Rate (%)	2.21	2.91	-0.7
	(0.30)	(0.20)	(0.05)
Default Fund Ownership Rate (%)	14.5	12.94	1.57
	(0.93)	(0.21)	(0.12)
Only Default Fund Ownership Rate (%)	12.54	10.84	1.7
	(0.94)	(0.21)	(0.12)

Continued on next page

Table 1.1 – continued from previous page

Panel D: Pension Wealth, Contribution, and Performance			
Log of Pension Wealth	7.18	6.76	-0.42
	(2.35)	(2.26)	(0.00)
Log of Monthly Contribution	2.81	2.54	-0.27
	(2.71)	(2.62)	(0.00)
Annual Portfolio Rate of Return (%)	20.46	20.25	-0.21
	(2.29)	(2.39)	(0.00)
Outperforming Default	0.76	0.72	-0.04
	(0.45)	(0.43)	(0.00)
Panel E: Portfolio Shares of Funds, and Assets			
Equity Fund Share (%)	2.23	1.74	-0.49
	(8.32)	(7.24)	(0.01)
Stock Share (%)	10.19	9.13	-1.06
	(12.49)	(11.65)	(0.01)
Risky Assets Share (%)	78.25	75.88	-2.38
	(22.26)	(24.21)	(0.03)
Panel F: Behavioral Biases and Heuristics, and Financial Mistakes			
Default Fund Share (%)	46.4	54.86	8.46
	(48.76)	(48.64)	(0.06)
1/N Heruistic	0.65	0.67	0.01
	(0.47)	(0.48)	(0.00)
RSRL	0.44	0.45	0.01
	(1.17)	(1.16)	(0.00)
Disposition Effect	-0.28	-0.31	-0.03
	(0.43)	(0.45)	(0.00)
Portfolio Reshuffling	0.44	0.49	0.06
	(0.50)	(0.50)	(0.00)
Observations	1488304	1558031	3046335

Notes: The table displays the mean, standard deviations in parenthesis, and the difference between the treatment and control groups. The treatment group covers those born after January 1987 while the control group is based on those born before January 1987. Panel A uses data from the 2018 Household Labor Force Survey by TURKSTAT and presents descriptive statistics for individuals. Panel B and C use data at the cohort level. The remaining panels use the pension administrative data with the month-end snapshots of pension accounts at the investor level on December 31, 2019, provided by Borsa Istanbul Group. The variable definitions are provided in Appendix A.

fund. In Panel E in Table 1.1, I focus on the share of wealth invested in equity funds in percentage terms in pension portfolios. Moreover, the pension administrative data has information about the composition of each fund by asset classes, which allows me to generate the share of wealth invested in stocks, and risky assets in percentage terms.

Panel F in Table 1.1 displays the summary statistics of the variables emphasizing primary behavioral biases and heuristics, and financial mistakes. Following Benartzi and Thaler (2001, 2007), I construct the share of wealth invested in default funds in percentage terms in portfolios, and following Huberman and Jiang (2006) an indicator variable equal to one if an investor tends to follow the conditional $\frac{1}{N}$ heuristic that is allocating money evenly to all pension funds while contributing. Moreover, the inertia in portfolio choices and trading in pension plans for instance 401(k) plans in the US are prevalent (Agnew et al. (2003)). To examine whether inertia in portfolio reshuffling or rebalancing varies with education, I generate an indicator variable equal to one if an investor buys a pension fund other than existing funds in her portfolio over the year 2020.

For the financial mistakes variables which are also proxies for financial literacy and sophistication as suggested by Lusardi and Mitchell (2014), I pursue the strategy proposed by Calvet et al. (2009b). The first variable —relative Sharpe ratio loss (RSRL)— is a measure to quantify the loss due to under-diversification by comparing the Sharpe ratio of individual portfolios with the Sharpe ratio of domestic benchmark pension fund index provided by the Pension Monitoring Center.[24] The second variable is a measure to elicit the disposition effect which is the tendency of selling winning funds and keeping losing funds. Yet, the pension administrative data lack information on the initial buying prices of funds. Therefore, similar to Calvet et al. (2009b) I use the annual return of the domestic benchmark pension fund index and subsequently classify a fund winner if it outperforms the domestic pension fund index but I classify a fund loser if it fails to outperform the relevant index. Then, I compute the disposition effect concerning this classification.

The pension administrative data has no information about the educational attainment of investors. As the 1997 Education Reform induced a sharp increase in schooling for those born after January 1987, I mainly estimate the reduced-form impacts of the education reform. To show the association between years of schooling and participation in pension plans, I employ Turkiye Household Budget Surveys in 2018 and 2019. One more point to note is that in the HLFS data, around 5% of the observations are missing the month of birth. That might be a threat to the validity of estimates. To address that concern, I estimate the impacts of the 1997 Education Reform on attrition in the month of birth. Results show that the 1997 Education Reform has no significant effect on the concerning attrition.

[24] For more information, see: https://emeklilik.egm.org.tr/en/egm-bes-endeks

1.4.2 Research Design

1.4.2.1 Identification

The 1997 Education Reform and the school starting age of 6 mandated those born after January 1987 to complete junior high school or 8 years of schooling instead of 5 years prior to the reform. Using the cutoff of January 1987 in the birth cohorts in months, I adopt a regression discontinuity design (RD) with a running variable in the month-year of birth to establish the causal link between schooling and participation, wealth accumulation in DC pension plans, and the variation in investors' portfolios. While those born before January 1987 form the control group, those born after January 1987 stand for the treatment group in my research design. The identifying assumption is that there are no systematical differences other than being affected by the 1997 Education Reform or not between two cohorts born one month apart. As long as this assumption holds, the estimation strategy based on an RD design produces a treatment assignment that is as good as random. In section 1.4.2.2, I perform a set of validity checks to support the relevant assumption holds.

In line with the prior research (Oreopoulos (2006), Erten and Keskin (2018), and Aydemir et al. (2022)), I exploit the discontinuity in the month-year of birth to estimate the causal effects of education. The estimating equation on a sharp RD design is as follows:

$$\begin{gathered} y_i = \alpha + \beta T_i + f(x_i) + \epsilon_i \\ \forall x_i \in (c-h, c+h) \end{gathered} \tag{1.1}$$

where y_i is the specific outcome variable for either the month-year of birth cohort or the individual i. T_i stands for the treatment status and β is the main parameter of interest, x_i is the running variable in months which is re-centered around zero by subtracting the month-year of birth from January 1987 that is the cutoff value determining the treatment status, and h is the bandwidth around the cutoff point of c. The RD design enables the slope to vary on each side of the cutoff. $f(x_i)$ is the control function with a continuous n-order polynomial function of the running variable on each side of the cutoff point c. In all estimations, I use the local linear approach proposed by Cattaneo et al. (2019) and also provide the estimates relying on a quadratic control function in Appendix A. Since local linear RD estimates are often sensitive to the choice of bandwidth, it is necessary to choose it in a data-driven, and automatic way to avoid specification search and ad-hoc decisions. Thus, I implement the optimal bandwidth algorithm proposed by Calonico et al. (2014) which considers the conventional mean squared error optimality based on the fundamental bias-variance trade-off. For each outcome variable, I estimate the specific bandwidths separately using the concerning optimal bandwidth algorithm. I also report the local linear RD estimates in fixed bandwidth and the estimates with kink RD design to check whether the 1997 Education Reform changes the slope around the cutoff point of January 1987 in Appendix A. In all estimations, results are similar and robust to different RD designs.

As I explore the impact of the 1997 Education Reform on a large set of outcomes, I generate summary indexes for each field of outcomes following the procedure suggested by Kling et al. (2007) to avoid any complications arising from multiple hypothesis testing and specification searching. For each observation, I first subtract the control group mean of a certain outcome variable from the corresponding outcome variable and subsequently divide it by the standard deviation of the same outcome variable in the control group, leading to the standardized values of the outcomes in a certain field of outcomes. Then, I take the average of all the standardized values of the outcomes in a specific field of outcomes for each observation. Furthermore, following Lee and Card (2008), I adjust standard errors by clustering them at the month-year of birth to avoid any specification error concerns as the treatment is assigned at the month-year birth level and the running variable is discrete. When the unit of analysis is the month-year birth cohorts, regressions include controls for the month of birth. When the unit of analysis is the individual investors, I also control for the province of birth registration fixed effects in addition to the controls for the month of birth. Full sample regressions also control for gender.

I usually provide the reduced-form estimates of the 1997 Education Reform in estimations within a sharp RD design framework. The primary reason why I largely report the reduced-form estimates rather than also presenting instrumental variable estimates of an extra year of schooling is twofold. The first is data limitation as the pension administrative data lacks investor's education information. Following Angrist and Krueger (1992), it is nonetheless quite easy to calculate two-sample instrumental variable (TSIV) estimates, and in some parts of the article, I report them if the reduced-form estimates are significant. For all outcomes, TSIV estimates are available upon request. The last and most important reason is that the instrumental variable exploiting the 1997 Education Reform might not satisfy the exclusion restriction since the financial decisions are mainly determined at the household level. If the education reform changed the schooling of other household members, for instance, spousal education, then the 1997 Education Reform would clearly not only operate through the education of investors but also spousal education. Consequently, I mainly report the reduced-form estimates of the 1997 Education Reform.

1.4.2.2 Validity Checks

In many RD designs, the possibility of units i.e. individuals manipulating the value of the running variable is a threat to validity. Yet, it is hard to argue that individuals were able to manipulate their birth date as the 1997 Education Reform was executed when they were at the age of 11. To reinforce the relevant assertion, I provide three standard validity checks suggested by Cattaneo et al. (2019).

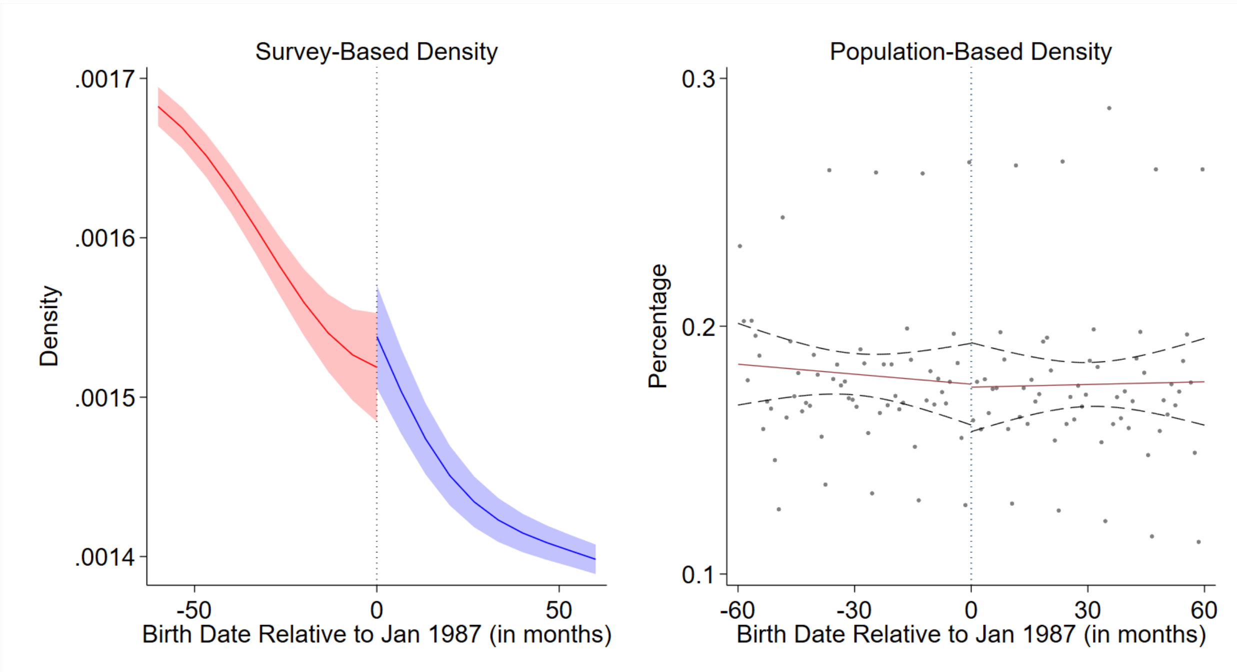

Figure 1.1: Density of the Running Variable (in months)

Notes: The Survey-Based Density graph uses data from the 2018 Household Labor Force Survey by TURKSTAT, which plots the density in monthly bins against the month-year of birth for those born before and after 60 months around the cutoff value of January 1987. The graph shows the results of Cattaneo et al. (2018) manipulation test of whether there is a discontinuity in the density of the running variable, the month-year of birth. Red and blue shaded areas present 95% confidence intervals for the point estimates fitted on each side. The Population-Based Density graph uses data covering the Turkish Population Share in each month-year of the birth cohort in 2019 assembled by TURKSTAT, which plots the population share in percentage terms in monthly bins against month-year of birth for those born before and after 60 months around the cutoff value of January 1987. The vertical line in each graph indicates the cutoff point. Black dashed lines indicate 95% confidence intervals around the mean of bins.

The first validity check tests whether around the cutoff the number of observations below the cutoff is considerably different than the number of observations above the cutoff or vice versa. Given that units can manipulate their birth date, then a jump in the density of the running variable at the cutoff is highly likely. I test this hypothesis by engaging the procedure suggested by Cattaneo et al. (2018). The corresponding procedure tests the null hypothesis of there is no discontinuity at the cutoff of January 1987. The left panel in Figure

1.1 documents that I fail to reject the null hypothesis with the p-value of 0.6267.

Using the share of the Turkish population in each month-year birth cohort in 2019 the right panel in Figure 1.1 implies the running variable is smooth around the cutoff and no sign of evidence of sorting in the full population exists. I also note that the numbers in each month-year birth cohort cover the whole population in Turkiye in 2019 provided by TURKSTAT, suggesting there is no issue of bias that might be led by misreporting or measurement errors. In addition, I provide local RD estimates on whether the 1997 Education Reform impacted the fraction of the population in the month-year birth cohorts in Appendix A. Results show that there is no evidence of birth date manipulation as all point estimates are small and indistinguishable from zero.

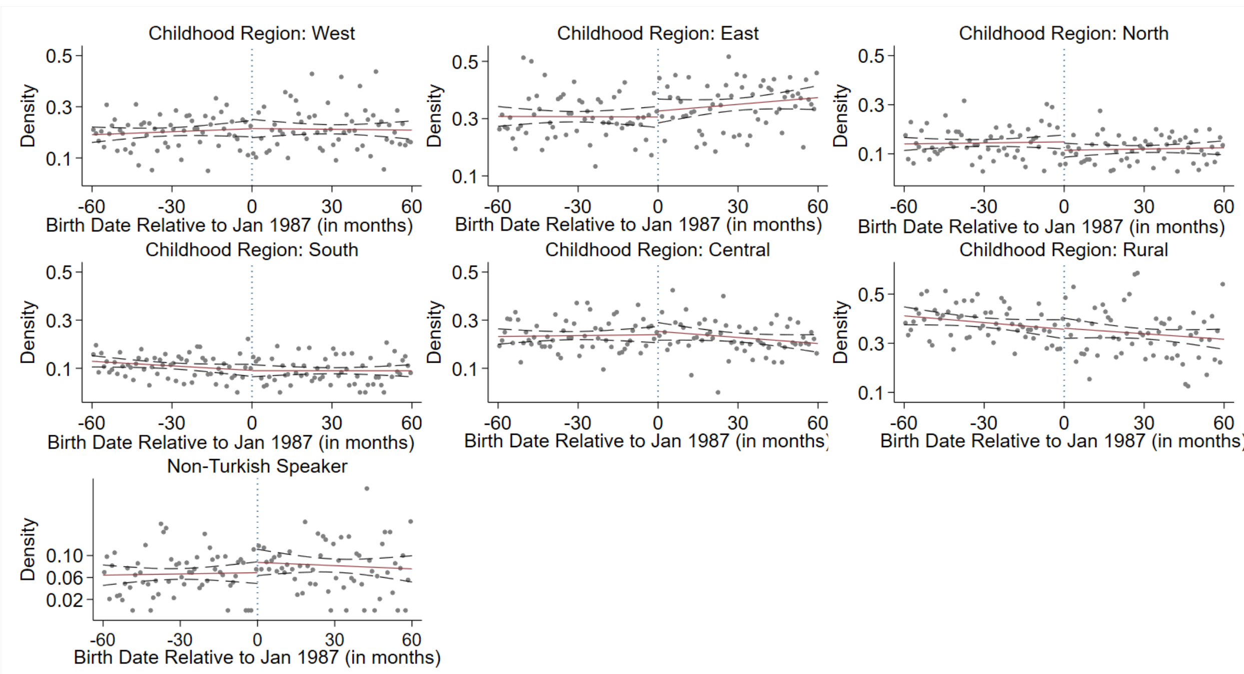

Figure 1.2: Balanced Covariates

Notes: All graphs use data from the 2008 and 2014 National Survey on Domestic Violence against Women in Turkiye by TURKSTAT. The sample includes females born before and after 60 months around the cutoff point, January 1987. The figures plot predetermined covariates in monthly bins against the month-year of birth of those born in January 1987. The vertical line in each graph indicates the cutoff point. Black dashed lines in each graph indicate 95% confidence intervals around the mean of bins.

The second validity check relies on the idea that predetermined covariates are continuous at the cutoff conditional on the fact that treatment is as good as random. However, data for predetermined covariates only exists for females. The Domestic Violence Against Women Surveys has information about the predetermined childhood regions and mother tongue. I generate five indicator variables for each geographical birth region, an indicator variable for whether a female was grown in rural areas, and an indicator variable for whether a female has a non-Turkish mother tongue. In Figure 1.2, I plot the binned means of the corresponding predetermined variables against the running variable in months. Graphical evidence indicates no significant differences in the predetermined covariates by the 1997 Education Reform. I also report local RD estimates consistent with the graphical evidence in Appendix A. I conclude that there is no overall significant evidence that those predetermined covariates are discontinuous at the cutoff as all point estimates are small and most are close to zero.

A further useful validity check works through artificial or placebo cutoffs in which in the absence of the treatment there would not be abrupt changes or jumps. I perform estimations with two placebo cutoffs which are January 1980 and January 1994. There are no economically and statistically significant treatment effects on schooling outcomes in both placebo cutoffs despite the remarkable increase at the true cutoff value of January 1987. In the result section, I will provide the graphical and regression evidence for the placebo cutoffs in more detail. Above all, it is unlikely to argue that there is any empirical evidence that the RD design is invalid in my research design.

1.5 Results

1.5.1 Schooling Outcomes

Employing the 2018 HLFS data, I begin the analysis by presenting visual evidence on the impacts of the 1997 Education Reform on schooling outcomes. In Figure 1.3, I plot the binned means of years of schooling in Panel A and junior high school completion in Panel B against the running variable in months. In all graphs, I restrict the sample to those born 60 months before and after January 1987. Moreover, I also present the graphical evidence by gender to uncover the heterogeneous effects of the 1997 Education Reform. Panel A in Figure 1.3 reveals years of schooling jump discontinuously at the cutoff for all samples. Panel B in Figure 1.3 depicts a marked increase in the propensity of holding at least a junior high school degree. Those graphs all together demonstrate that the 1997 Education Reform increased the propensity of holding at least a junior high school degree by 10-15% while increasing years of schooling by 0.5-1 year. Despite the graphical evidence with a clear and sharp increase in schooling outcomes induced by the 1997 Education Reform, I next document the local linear RD estimates.

I separately report local linear RD estimates for each schooling outcome in the full, male, and female samples in Table 1.2. For a specific schooling outcome, I estimate the optimal bandwidth according to the algorithm suggested by Calonico et al. (2014). I use the HLFS 2018 data in all estimations and the unit of analysis is individuals. All regressions include

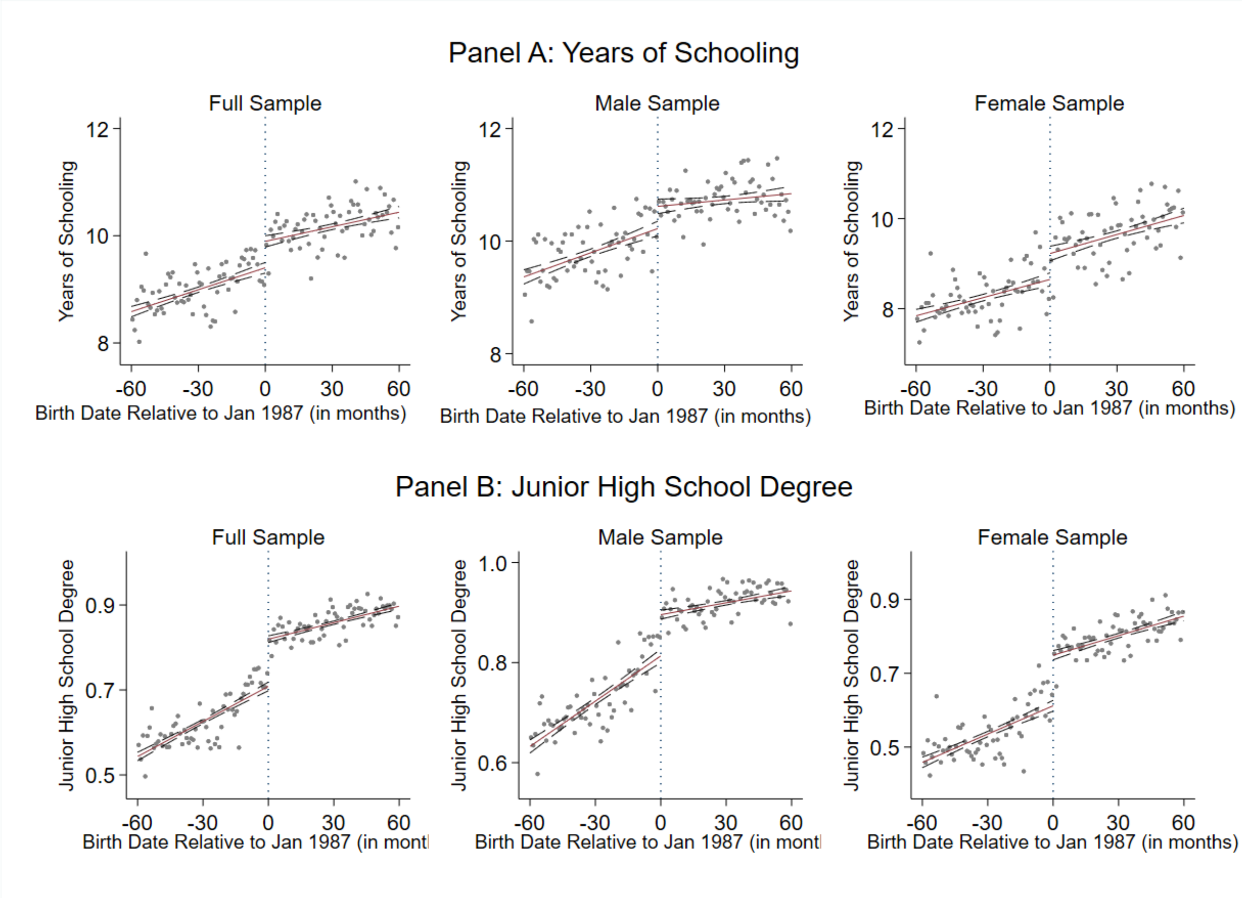

Figure 1.3: Education Reform vs Years of Schooling, and Junior High School Degree

Notes: All graphs use data from the 2018 Household Labor Force Survey assembled by TURKSTAT. The figures in Panel A plot years of schooling and the figures in Panel B plot the propensity to hold at least a junior high school degree in monthly bins against the month-year of birth of those born before and after 60 months around the cutoff point, January 1987. The vertical line in each graph indicates the cutoff point. Black dashed lines in each graph indicate 95% confidence intervals around the mean of bins. Full, male, and female sample figures are reported, respectively.

controls for months of birth. The regressions in the full sample also include an additional control for gender. All regressions use a linear polynomial function on each side of the cutoff value. Standard errors are clustered at the month-year birth level.

Column 1 in Table 1.2 presents that the reform-induced increase in years of schooling is 0.45, 0.38, and 0.44 for the full, male, and female samples, respectively. As the 1997 Education Reform made junior high school completion compulsory, column 3 in Table 1.2 reveals a large and precise increase in the fraction of those who completed junior high school education. Interestingly, the education reform displaced some females from schooling since it lowered the propensity to complete at least primary school for some part of the population. However, the point estimates are small in magnitude despite their precision.

Table 1.2: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	0.451***	-0.011*	0.055***	0.043***	0.031**	0.084***
	(0.105)	(0.007)	(0.010)	(0.012)	(0.014)	(0.019)
Control Mean	9.25	0.93	0.68	0.50	0.28	0.45
Control SD	4.73	0.26	0.47	0.50	0.45	0.84
Bandwidth	57.21	51.57	30.82	37.35	34.20	55.49
Observations	59347	53310	31744	39279	35837	57247
			Male			
Education Reform	0.381***	0.005	0.042***	0.038**	0.025	0.071***
	(0.148)	(0.007)	(0.015)	(0.019)	(0.017)	(0.026)
Control Mean	9.96	0.96	0.76	0.57	0.30	0.57
Control SD	4.33	0.19	0.43	0.50	0.46	0.77
Bandwidth	51.41	55.18	35.32	42.65	37.25	52.22
Observations	25664	27536	17678	21301	18902	26112
			Female			
Education Reform	0.448**	-0.032***	0.081***	0.049**	0.035*	0.085**
	(0.210)	(0.010)	(0.021)	(0.023)	(0.019)	(0.037)
Control Mean	8.56	0.89	0.57	0.44	0.26	0.33
Control SD	5.00	0.31	0.50	0.50	0.44	0.88
Bandwidth	49.82	65.53	40.32	39.94	38.87	49.24
Observations	26722	35135	21889	21311	20820	26722

Notes: Local linear RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

I present the local linear RD estimates for holding at least a high school degree and at least a college degree in columns 4 and 5 in Table 1.2. Thus, I argue that there are large spillover effects of education reform. The 1997 Education Reform encouraged those born after January 1987 to complete high school in all samples. For instance, females born after January 1987 have over 10% higher likelihood of completing high school relative to the mean of the control group. That pattern even is more amplified by the propensity to obtain a college degree. Relative to the control group mean, the 1997 Education Reform promoted having at least a college degree by more than 10% in the full sample, as well as around 14% in the female sample. Thus, the 1997 Education Reform with its spillover effects went beyond its primary purpose, which is of making those born after January 1987 complete at least junior high school.[25] The estimates are also in line with the previous studies examining the impact of the 1997 Education Reform on schooling outcomes.[26]

Point estimates imply that the propensity to complete at least high school is higher for those born after January 1987, even increasing the likelihood of holding a college degree. Thus, the 1997 Education Reform shifted right the entire distribution of years of schooling for those born after January 1987 as shown in Figure 1.4. The findings on high school and college education are noteworthy in at least one respect. One might claim that the education reform only fostered the propensity to hold at least a junior high school degree, which might not be enough for schooling to change the financial outcomes in DC pension plans that supposedly require a higher level of education. Yet, it is difficult to argue that this is a valid concern in my setting as the education reform shifted the whole distribution.

As the 1997 Education Reform made junior high school completion compulsory and so significantly changed the schooling outcomes, I assess whether there is an impact on the placebo cutoffs. Figure 1.5 reveals no overall significant impact in each placebo cutoff separately in contrast to the large and statistically significant point estimates at the true cutoff value of January 1987. I further report the local RD estimates for years of schooling and the propensity to have at least a junior high school degree in Appendix A in Table A.3. Estimates with placebo cutoffs are small and indistinguishable from zero, verifying that the 1997 Education Reform changed the schooling landscape in Turkiye.

I also perform some robustness checks. The first issue to touch upon is that around 5% of the individuals in the 2018 HLFS data have the missing month of birth information. To address the concern of whether the attrition in the month of birth is a threat to validity, I perform two robustness checks. First, I use the year of birth as the running variable. Yet, there are small numbers of data points and this might violate the requirements of the continuity-based RD approach. Thus, following Cattaneo et al. (2019) I employ a local randomization RD design to test whether the 1997 Education Reform has an impact on schooling outcomes in the closest window which is one year around the cutoff on each side.[27]

[25]Figure A.5 in Appendix A illustrates the jump in the high school and college education.

[26]Aydemir et al. (2022), Erten and Keskin (2018), and Gulesci et al. (2019).

[27]Local randomization RD design is based on the following procedure. It considers the closest window

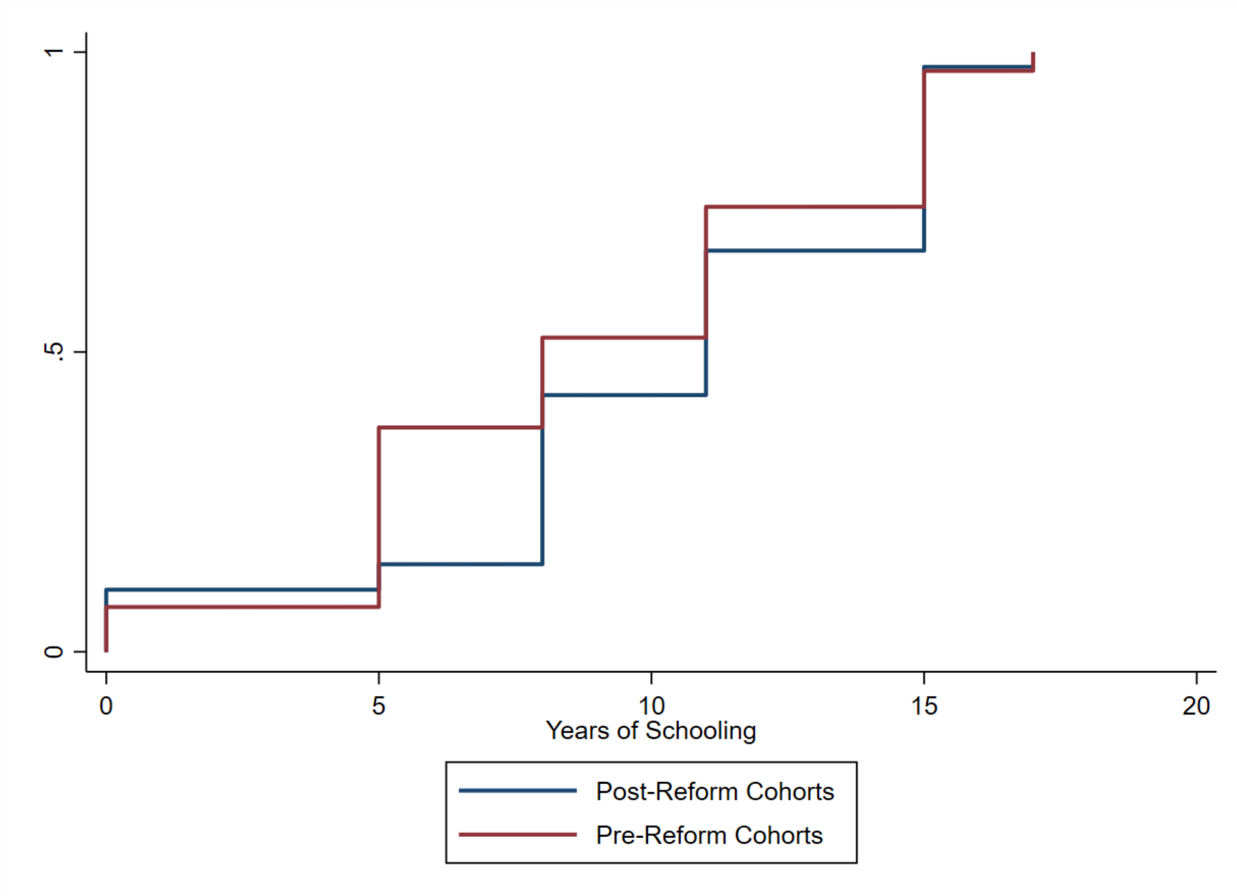

Figure 1.4: CDF of Years of Schooling by Education Reform across Cohorts

Notes: All graphs use data from the 2018 Household Labor Force Survey assembled by TURKSTAT. The sample producing this graph includes those born 57 months before or after January 1987 since the optimal bandwidth is 57 months. The figures plot the cumulative density of years of schooling by the cohorts of those born before and after January 1987, which is the determinant of treatment status arising from the 1997 Education Reform.

In Appendix A, Table A.4 reports very similar point estimates to the local RD estimates with a continuity-based approach.

As an additional robustness check for the attrition in the month of birth, using the local randomization RD design I assess whether the 1997 Education Reform causes the related attrition. For different window lengths from the closest one to the 6-year length, point estimates reveal that the attrition in the month of birth is orthogonal to the education reform. Table A.5 in Appendix A displays the local randomization RD estimates. Taking everything into account, I conclude that the 1997 Education Reform has significantly increased schooling in Turkiye, allowing me to have powerful first-stage estimates to examine how schooling

around the cutoff and then implements the Fisherian randomization by testing the null hypothesis of no treatment effect. For more details, see Cattaneo et al. (2019).

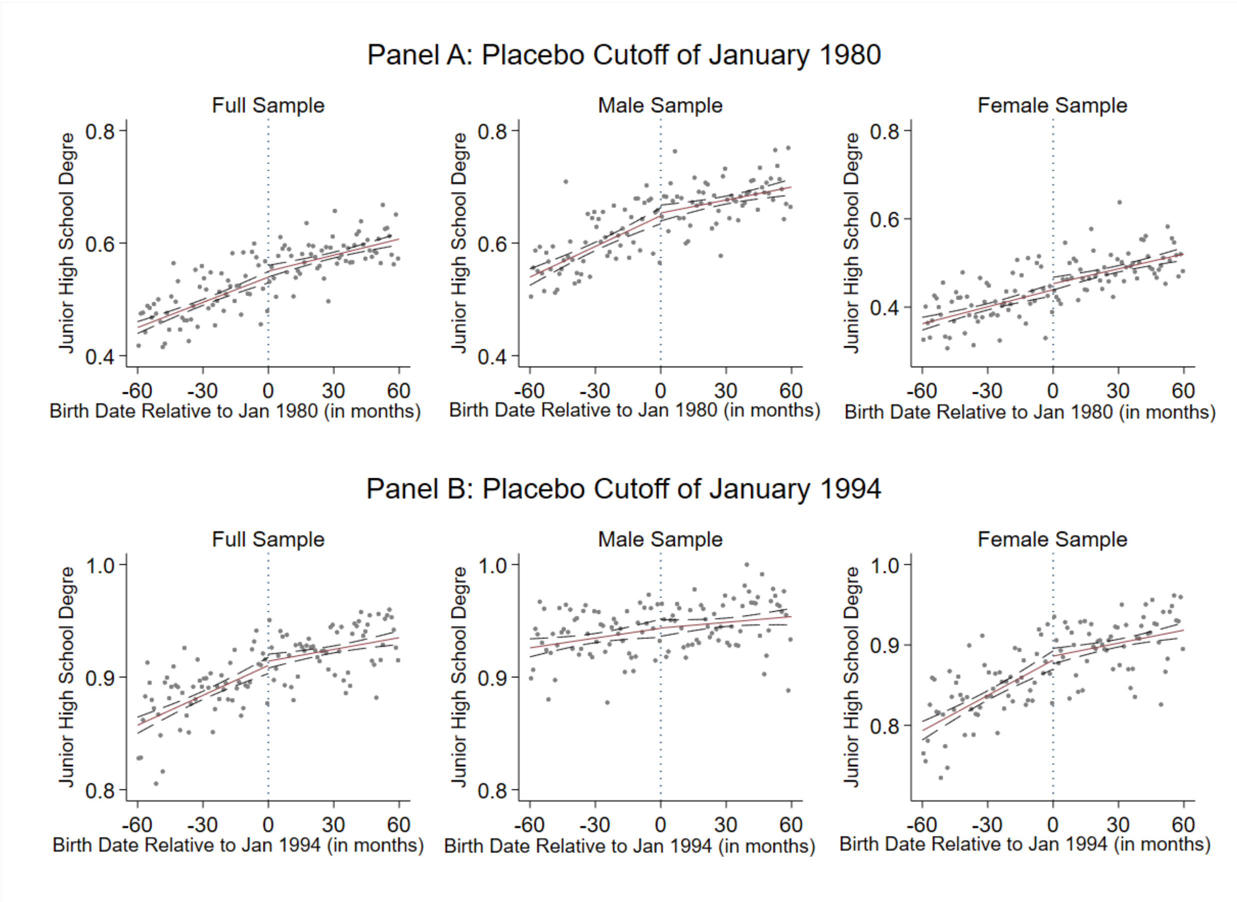

Figure 1.5: Placebo Cutoffs vs Junior High School Degree

Notes: All graphs use data from the 2018 Household Labor Force Survey by TURKSTAT. The figures plot the propensity of having at least a junior high school degree in monthly bins against the month-year of birth of those born before and after 60 months around the cutoff point, January 1980 in Panel A while the figures in Panel B plot for those born before and after 60 months around the cutoff point, January 1994. The vertical line in each graph represents the corresponding cutoff point. Black lines indicate 95% confidence intervals around the mean level. Full, male, and female sample figures are reported, respectively.

affects participation, wealth accumulation, and the variation in portfolios in DC pension plans.

1.5.2 Education and Savings for Retirement

1.5.2.1 Participation and Wealth Accumulation in DC Pension Plans

I now proceed with documenting the impacts of the 1997 Education Reform on outcomes related to participation —the propensity to have a positive balance in individual retirement accounts— and accumulated wealth in pension plans. Before presenting the causal effects of the 1997 Education Reform on pension outcomes, I document the positive association between years of schooling and pension plan participation. The pension administrative data has no information on education levels, but the 2018-19 Household Budget Surveys contain this information. Employing this data set, Figure 1.6 illustrates the relevant relationship in the sample of those born 35 months before or after January 1987 to show the concerning relationship around the RD bandwidth mostly falling into 35 months intervals in my estimations. The corresponding figure illustrates a robust positive relationship.

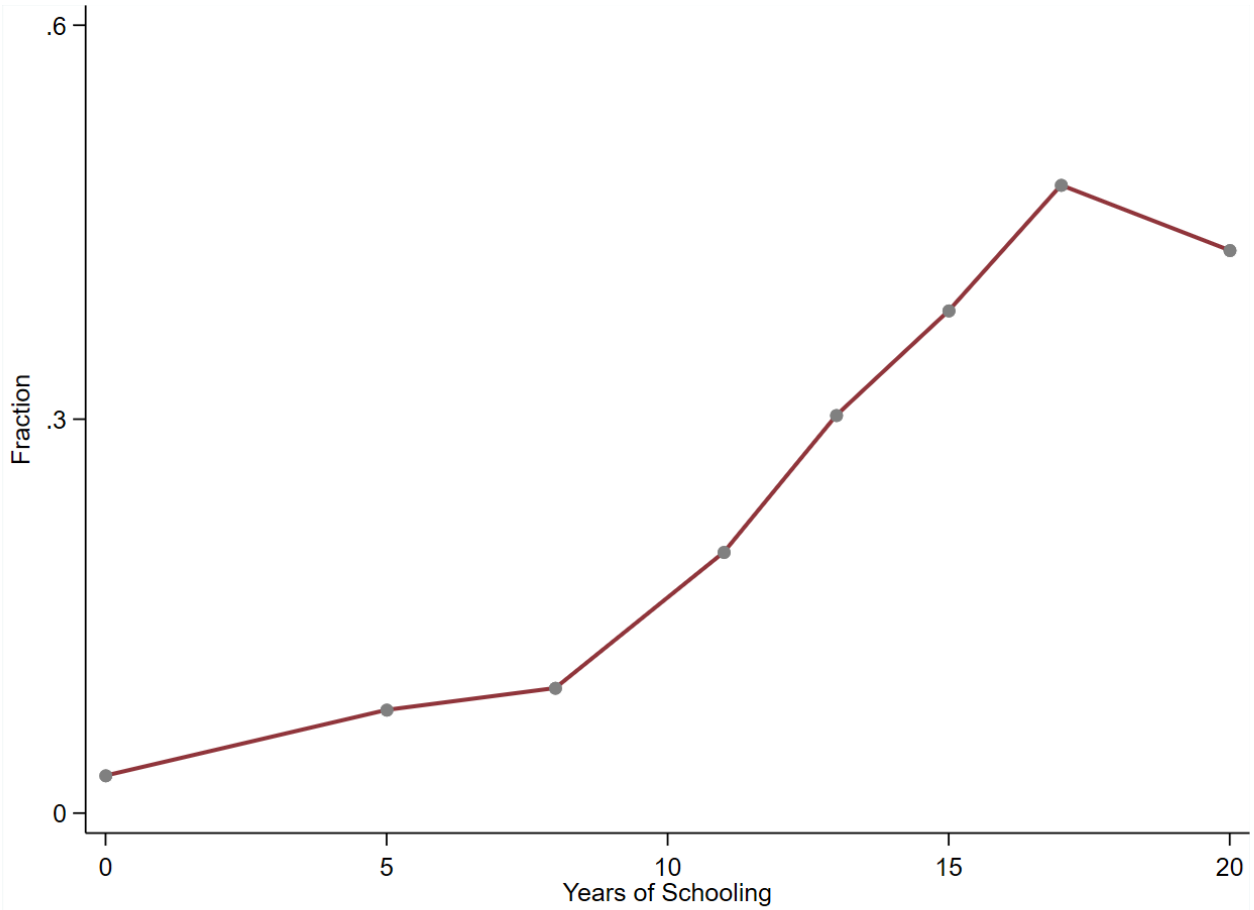

Figure 1.6: Schooling vs Participation in DC Pension Plans

Notes: The graph uses data from Household Budget Surveys of 2018-19 by TURKSTAT. The sample includes those born 35 months before or after January 1987 since the optimal bandwidth in the RD design is 35 months for the outcome of participation in DC pension plans. The figures plot the fraction of those participating in pension plans against the binned years of schooling.

I subsequently report the OLS estimates for the corresponding relationship in Table 1.3 which reveal that an extra year of schooling increases the participation in DC pension plans by 2.6% in the sample covering those born 35 months before or after January 1987, respectively. Nonetheless, as Figure 1.6 depicts, the participation rate is likely to be nonlinear regarding the years of schooling. Moreover, note that the education reform shifts the whole distribution of schooling in Turkiye right as the point estimates in Table 1.2 and the cumulative distribution function of years of schooling in Figure 1.4 reveal. Such a shift and non-linearities induced by the education reform make it harder to interpret the OLS coefficient stemming from regressing the propensity to participate in DC pension plans on years of schooling. Thus, to get a comparable benchmark OLS estimate relative to the causal estimate, I estimate an OLS coefficient with a different functional form mimicking the shift in the distribution of schooling induced by the education reform.

To do this, I first regress the indicator variables for educational attainment, and the reference category is those who completed primary school at most. Consistently, OLS estimates in column 2 show that the higher educational attainment the more participation in DC pension plans. Secondly, I multiply each OLS coefficient with the local RD point estimate measuring the corresponding increase in each educational degree led by the education reform. Then, I sum up all of the concerning products, which gives me a benchmark OLS estimate to compare with the causal estimate. Similarly, assuming the covariances are zero across those products, I also calculate the standard error. Column 2 in Table 1.3 shows that the benchmark OLS coefficient is 1.2%, implying that the shift in the distribution of educational attainment induced by the education reform is positively correlated with the propensity to participate in DC pension plans. However, these point estimates fail to represent a causal relationship so I investigate whether the OLS estimate truly represents the causal impact of schooling by using the 1997 Education Reform as an exogenous shock in schooling.

Having combined the number of people in each month-year birth cohort through TURKSTAT population data and the number of people participating in a specific pension plan via the pension administrative data, I generate the ratio of those participating in the corresponding pension plans in percentage terms. I also note that the 1997 Education Reform is a treatment assigned at the month-year birth cohort level. Panel A in Figure 1.7 plots the percentage of those participating in any DC pension plans against the running variable in months whereas panel B shows the logarithm of accumulated pension wealth in individual retirement accounts. As the data producing those graphs span the universe of pension plan participants, I underline that the confidence intervals are small in all graphs. Despite the graphical evidence illustrating jumps and kinks at the cutoff, such movements require a more refined analysis so I continue with local linear RD estimates.

I begin providing the local linear RD estimates in Table 1.4. I also note that the unit of analysis in Panel A in Table 1.4 is the birth cohorts in months. Considering the point estimates, for the full and male populations, the 1997 Education Reform does not change their decision to either participate in or contribute to DC pension plans as the point estimates are both small and indistinguishable from zero. Despite the precision, the impacts of the 1997 Education Reform on participation in DC pension plans and active contribution for the

Table 1.3: Education vs Participation in Pension Plans

	(1)	(2)
	Pension Participant	
Years of Schooling	0.027***	
	(0.001)	
Junior High School		0.043***
		(0.012)
High School		0.107***
		(0.015)
College		0.162***
		(0.018)
Weighted $\hat{\beta}$		0.012***
		0.03
Control Mean	0.20	0.20
Control SD	0.40	0.40
Observations	6113	6113

Notes: OLS estimates in all columns. All columns use data from Household Budget Surveys of 2018-19 by TURKSTAT. The sample includes those born 35 months before or after January 1987 since the optimal bandwidth in the RD design is 35 months for the outcome of participation in DC pension plans. The unit of analysis is individuals. In all columns, the outcome is a dummy variable equal to one if the individual participates in private pension plans. In the first column, the explanatory variable is years of schooling. In the second column, the explanatory variables are indicator variables for junior high school, high school, and college degrees, respectively. The reference category is the primary school degree at most. Weighted $\hat{\beta}$ is the weighted average of the point estimates of the indicator variables for the indicator variables of degrees in the second column regarding the shift induced by the Education Reform in degrees. All regressions include controls for gender with a dummy variable of being female and year of survey fixed effects for each survey year. The control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome in the control group. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

female cohorts are also economically very small.

The 1997 Education Reform substantially spurs only the participation in employer-sponsored DC pension plans for females. The effect size, which is calculated by dividing the corresponding point estimate by the standard deviation of the control group in the concerning outcome is 0.95 standard deviation.[28] The last column in Table 1.4 displays the impact of the education reform on the summary index for the participation field to avoid

[28] Most studies follow the rule developed by Cohen (2013) to interpret the impact of an intervention. The rule classifies the impact as trivial if the effect size is less than 0.1, small if less than 0.3 and greater than 0.1, moderate if less than 0.5 and greater than 0.3, and large if greater than 0.5.

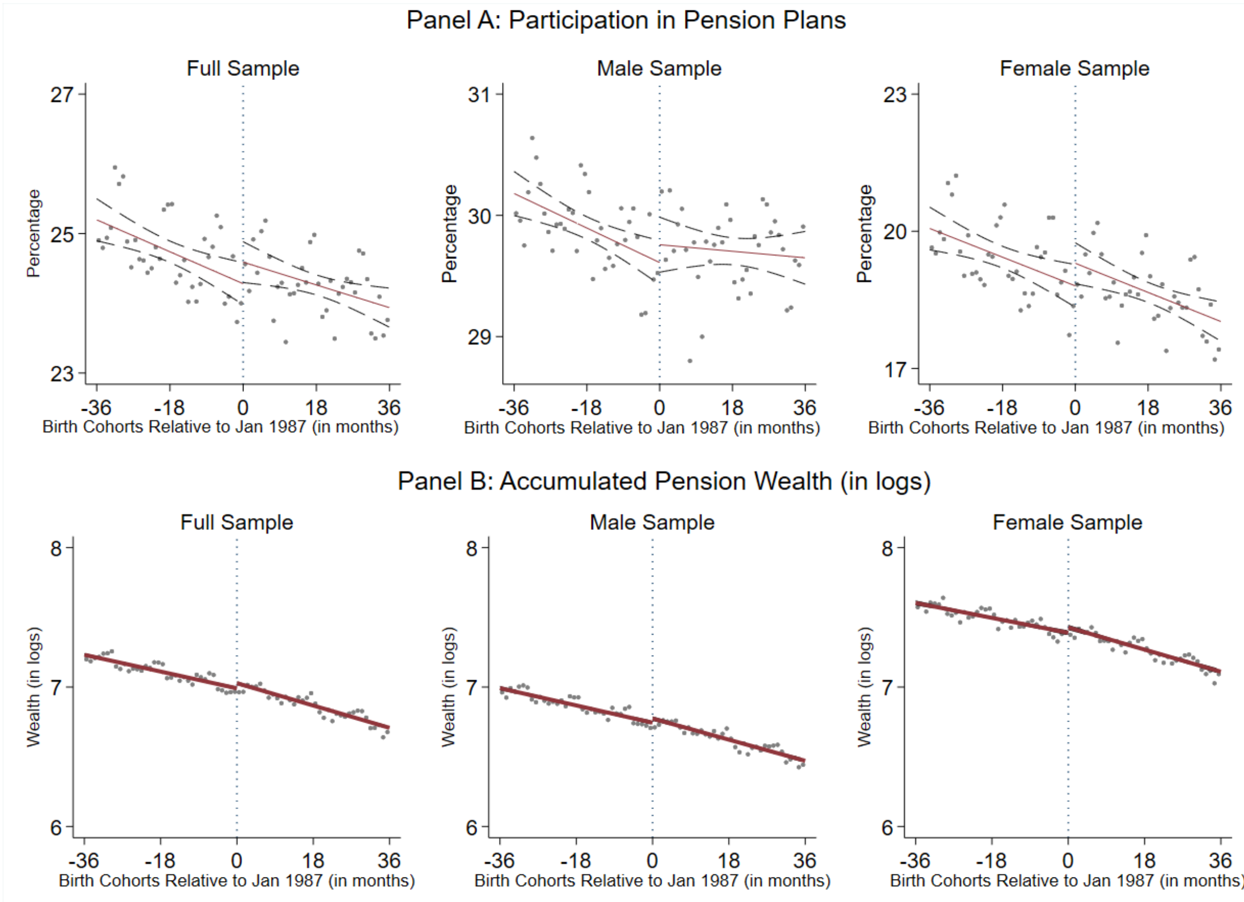

Figure 1.7: Education Reform vs Participation, and Wealth in Pension Plans

Notes: All graphs use administrative data assembled by Borsa Istanbul Group and TURKSTAT in 2019. The figures in Panel A plot the percentage of those participating in overall pension plans while the figures in Panel B plot the percentage of those participating in voluntary pension plans in monthly bins against the month-year of birth of those born before and after 36 months around the cutoff point, January 1987. The vertical line in each graph indicates the cutoff point. Black dashed lines in each graph indicate 95% confidence intervals around the mean of bins. Full, male, and female sample figures are reported, respectively.

the complications arising from multiple hypothesis testing. While estimates show that the 1997 Education Reform impacted the participation in pension plans for females, nonetheless, the effect size is small or moderate, to some extent negligible, except for the impacts on the participation in employer-sponsored pension plans. Above all, unlike the large OLS gradient of either basic years of schooling or the model accounting for the entire shift in the distribution of schooling induced by the education reform on the propensity to participate in DC pension plans, the local RD estimates of the 1997 Education Reform are small for females and also small and imprecise for males.

In Panel B in Table 1.4, I provide evidence on whether schooling differentiates wealth accumulation in DC pension plans. Since the savings in individual retirement accounts is

Table 1.4: Education Reform vs Participation and Wealth Accumulation in DC Pension Plans

	Panel A: Participation in DC Pension Plans						*Panel B: Savings*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participant	Voluntary	Employer Sponsored	Automatic	Contributor	Summary Index	(in logs)
				Full			
Education Reform	0.118	0.155**	0.042**	-0.059	0.120**	0.027***	0.015*
	(0.087)	(0.071)	(0.021)	(0.074)	(0.056)	(0.009)	(0.008)
Control Mean	24.75	23.71	1.64	27.58	13.88	1.60	7.12
Control SD	0.53	0.95	0.07	0.65	0.50	0.06	2.33
Bandwidth	35.05	42.60	27.82	29.00	34.05	35.19	37.76
Observations	71	85	55	59	69	71	1900641
				Male			
Education Reform	0.077	0.056	0.022	-0.174	0.043	0.015	0.007
	(0.119)	(0.069)	(0.037)	(0.131)	(0.071)	(0.010)	(0.010)
Control Mean	29.82	27.30	2.29	37.91	15.64	1.60	6.88
Control SD	0.30	0.63	0.10	0.76	0.39	0.04	2.33
Bandwidth	26.05	29.53	28.56	27.25	25.41	47.99	36.08
Observations	53	59	57	55	51	95	1146511
				Female			
Education Reform	0.205**	0.322***	0.066***	0.048	0.208***	0.064***	0.027***
	(0.105)	(0.084)	(0.017)	(0.105)	(0.075)	(0.014)	(0.010)
Control Mean	19.46	19.76	0.97	17.04	11.97	1.62	7.50
Control SD	0.78	1.14	0.07	0.70	0.59	0.11	2.26
Bandwidth	37.59	46.47	28.13	27.53	39.49	34.16	34.99
Observations	75	93	57	55	79	69	666926

Notes: Local linear RD estimates in all columns. Columns 1-6 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Column 7 uses administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-6 and individuals in column 7. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those having positive balances in DC retirement accounts in column 1, in DC voluntary retirement accounts in column 2, in DC employer-sponsorship retirement accounts in column 3, in DC automatic enrollment retirement accounts in column 4, and the percentage of those actively contributing to pension plans in column 5. In column 6, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variables. In column 7, the outcome is the log of the value of pension portfolios in Turkish Lira. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-6 and column 7 also includes controls for the birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

conditional on participation, I turn to employ the pension administrative data covering the universe of pension investors in Turkiye in 2019, so the unit of analysis is individual investors. Accordingly, column 6 in Table 1.4 reports the local linear RD estimates of the

1997 Education Reform, which reveal that those born after January 1987 accumulate more pension wealth, but a significant heterogeneity arises. In the full sample regressions, the point estimate is 0.015, implying that the 1997 Education Reform leads to higher wealth accumulation in DC pension plans by 1.5%. Despite the null effects of the 1997 Education Reform for the male population, the gradient of the education reform for the female population is 0.027, turning into an increase of 2.7%. Even though for brevity I mostly do not report the estimates relying on the two-sample instrumental variable (TSIV) strategy, a back-of-the-envelope calculation for TSIV estimates shows that an extra year of schooling improves the wealth in individual retirement accounts by 3% for the full population and 5.4% for females. Taking everything into account, the 1997 Education Reform led to higher pension wealth in DC pension plans, indicating education improves wealth accumulation.

1.5.2.2 Mechanisms

In this part, I provide evidence on the potential mechanisms of how schooling differentiates pension wealth accumulation using some plausible proxies suggested by the prior research for financial skills (Calvet et al. (2009b), and Benartzi and Thaler (2007)). To do this, I examine the impacts of the 1997 Education Reform on equity exposure i.e. the propensity to own equity and the share of wealth invested in equities or risky assets, financial sophistication, behavioral biases and heuristics, and portfolio returns in DC pension plans.

Do those with more education take advantage of the equity premium through holding equities and risky assets more in their pension portfolios and so have desirable financial behavior? Put another way, do they have more accumulated wealth in DC pension plans? Table 1.5 documents the local linear RD estimates of the education reform on equity exposure, risky share, and the propensity to hold default funds in the pension portfolios. The first four columns in Table 1.5 use the number of people in Turkiye through TURKSTAT population data and the number of pension investors in each month-year birth cohort owning the corresponding funds in their own DC retirement accounts. By doing so, I compute the percentage of those holding the related funds in each month-year birth cohort, thus the unit of analysis is the month-year birth cohorts. The point estimates in column 1 are precise but small. Lastly, columns 2 and 3 document the impacts of the 1997 Education Reform on the choice of default funds and the percentage of those having only default funds in their pension portfolios. Results also show that schooling is not a factor mitigating the default stickiness which is expressed in the propensity to own default pension funds or only default pension funds. In short, considering the summary index in column 4 in Table 1.5 the impacts are moderate in the equity fund ownership even in terms of effect size.

Table 1.5 displays the local linear RD estimates for the shares of wealth invested in equity funds, stocks, and risky assets. While column 5 in Table 1.5 documents the local linear RD estimates for the share of equity funds in pension portfolios, columns 6 and 7 present the local linear RD estimates for the portfolio shares of stocks and risky assets, respectively. I employ the administrative data with the individual retirement accounts. Therefore, the unit of analysis is individuals. Results show that 1997 Education changes the way how individuals allocate their investments precisely, but the point estimates in magnitude are small. Either comparing the point estimates with the control group mean or considering the effect size reveals no economically significant impacts of the education reform. Accordingly, education does not significantly change the level of equity exposure and the fraction of wealth invested in risky assets. Overall, I do not find support for the fact that general education promotes wealth accumulation through equity ownership that helps investors take advantage of the equity premium.

Does a higher level of general education allow investors to avoid behavioral biases and heuristics, and financial mistakes which possibly result in welfare losses? So, do investors earn higher returns and subsequently accumulate more wealth in DC retirement accounts? Point estimates do not provide clear evidence that general education helps investors overcome the concerning problems. In this part of the estimations, I keep employing the administrative data with the individual retirement accounts. I begin analyzing the impacts of the 1997 Education Reform on some prevalent behavioral biases and heuristics in DC pension plans documented by Benartzi and Thaler (2007) such as non-participation, inadequate contribution, stickiness to default fund or not exercising the right to choose, naive diversification strategies i.e. the conditional $\frac{1}{N}$ heuristic that evenly allocates money across funds while investing, and lack of portfolio reshuffling or rebalancing. As already shown, the participation impacts of schooling are minimal and it has no role on the propensity to hold default funds or only default funds in pension portfolios are not significantly changed by education in Table 1.5.

Column 1 in Table 1.6 documents that the 1997 Education Reform reduces the share of default funds in pension portfolios but the point estimate in magnitude is small. Columns 2 and 3 also reveal that schooling is not a factor in avoiding the conditional $\frac{1}{N}$ naive diversification strategy and the lack of portfolio reshuffling. For the female population, the 1997 Education Reform increases the contribution amount by 2.6%. Above all, I find no solid evidence that schooling is a factor mitigating most of the behavioral biases and heuristics common in retirement accounts.

Columns 5 and 6 in Table 1.6 display that schooling has no overall significant impact on lowering financial mistakes. Moreover, in the last two columns in Table 1.6 I estimate the impacts of the education reform on annual pension portfolio returns and whether education

Table 1.5: Education Reform vs Equity Exposure, and Default Option

	Panel A: Propensity to Own				*Panel B: Share of Wealth Invested in*			
	(1) Equity Fund	(2) Default Fund	(3) Only Default Fund	(4) Summary Index	(5) Equity Fund	(6) Stock	(7) Risky Assets	(8) Summary Index
				Full				
Education Reform	0.077***	-0.043	-0.067*	0.009	0.046**	0.057*	0.112*	0.005***
	(0.023)	(0.044)	(0.035)	(0.009)	(0.019)	(0.030)	(0.060)	(0.002)
Control Mean	2.81	13.01	10.98	1.52	2.20	10.08	77.95	-0.12
Control SD	0.15	0.20	0.18	0.03	8.27	12.39	22.52	0.75
Bandwidth	39.64	28.57	27.90	28.71	49.74	33.82	32.95	39.68
Observations	79	57	55	57	2506174	1690646	1640290	1997398
				Male				
Education Reform	0.089***	-0.051	-0.109	0.006	0.060***	0.116***	0.085	0.006***
	(0.034)	(0.089)	(0.082)	(0.013)	(0.021)	(0.031)	(0.059)	(0.002)
Control Mean	3.33	17.03	14.55	1.51	2.10	9.86	76.80	-0.15
Control SD	0.21	0.24	0.26	0.03	8.10	12.28	23.41	0.75
Bandwidth	45.12	22.07	26.24	31.52	46.87	40.68	31.67	38.49
Observations	91	45	53	63	1451167	1265014	978094	1202282
				Female				
Education Reform	0.071***	0.048	0.018	0.026	0.019	-0.034	0.169*	0.003
	(0.024)	(0.065)	(0.057)	(0.016)	(0.032)	(0.055)	(0.097)	(0.003)
Control Mean	2.36	9.02	7.37	1.54	2.34	10.46	79.77	-0.07
Control SD	0.17	0.34	0.26	0.09	8.49	12.57	20.92	0.73
Bandwidth	48.16	34.67	33.93	34.01	42.97	32.66	38.39	39.03
Observations	97	69	67	69	823369	629366	745603	764488

Notes: Local linear RD estimates in all columns. Columns 1-4 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Columns 5-8 use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-4 and individuals in the remaining columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the percentage of those holding equity funds in the total population. In column 2, the outcome is the percentage of those holding default pension funds in the total population. In column 3, the outcome is the percentage of those holding only default pension funds in the total population. In column 4, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variable. The outcome is the share of wealth invested in equity funds in column 5, in stocks in column 6, and in risky assets in percentage terms in column 7 in pension portfolios. In column 8, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-4 and the remaining columns also include controls for the birth registration certificate region fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

is a factor increasing the likelihood of earning higher returns than the returns of default funds. Despite the precision of both point estimates in the male sample, the impacts are small in magnitude as the education reform increases portfolio returns by 0.02 percent and the likelihood of outperforming the default funds by 0.002. All in all, no significant evidence that schooling increases financial literacy and sophistication exists, so higher returns as it is also verified by the summary index in column 9 of Table 1.6.

Then, what is the underlying channel of education for higher pension wealth, or savings for retirement? More schooling is clearly associated with higher wages in the labor market (Psacharopoulos and Patrinos (2018)). Indeed, Aydemir and Kirdar (2017) finds that the 1997 Education Reform causally increases wages by 3-4% and 1-2% for females and males, respectively. In line with this, the education reform leads to a 2-3% increase in accumulated pension wealth and contribution amounts to pension plans for females. Moreover, the participation rate in employer-sponsored pension plans for females born after January 1987 is substantially higher, indicating presumably those with more education are likely to work in firms offering better contracts.

Apart from that, education is likely to promote being formally employed, that is to say, the more education the less the informality as documented by Bleakley and Gupta (2023). Then, it is likely that those born after January 1987 might have participated through automatic enrollment plans more, so they save more in their individual retirement accounts. Yet, the point estimates of the 1997 Education Reform on participation through automatic enrollment are too small and imprecise as documented in Table 1.4, revealing that informality does not seem to be one of the possible labor market channels driving the results. Results hereby imply that the 1997 Education Reform operates through the labor market channels of greater earnings and a higher likelihood of being employed in the firms offering employer-sponsored pension plans rather than lowering informality, which boosts savings for retirement, and contribution amounts. Overall, the 1997 Education Reform does not seem to determine participation in pension plans and investment decisions through perhaps increasing financial literacy with cognitive skill gains.

Table 1.6: Education Reform vs Behavioral Biases and Heuristics, Financial Mistakes, and Performance

	(1) Default Fund Share	(2) 1/N Heuristic	(3) Portfolio Reshuffling	(4) Contribution (in logs)	(5) RSRL	(6) Disposition Effect	(7) Portfolio Return	(8) Outperforming Default	(9) Summary Index
					Full				
Education Reform	-0.356**	0.001	-0.004*	0.014**	-0.004*	0.002*	0.011*	0.003**	0.001
	(0.152)	(0.001)	(0.002)	(0.007)	(0.002)	(0.001)	(0.006)	(0.001)	(0.001)
Control Mean	46.73	0.37	0.44	2.75	0.44	-0.29	20.27	0.76	0.01
Control SD	48.78	0.48	0.50	2.69	1.16	0.43	2.39	0.43	0.34
Bandwidth	50.57	28.60	30.59	32.58	51.35	35.25	40.84	34.84	29.19
Observations	2551985	1436505	1535546	1640290	2597908	1780686	2049373	1738476	1485550
					Male				
Education Reform	-0.351**	-0.001	-0.002	0.005	0.001	0.001	0.017***	0.002*	0.000
	(0.167)	(0.002)	(0.002)	(0.008)	(0.003)	(0.001)	(0.007)	(0.001)	(0.001)
Control Mean	50.91	0.34	0.48	2.55	0.46	-0.31	20.23	0.74	0.00
Control SD	48.90	0.47	0.50	2.68	1.16	0.45	2.40	0.44	0.33
Bandwidth	50.66	30.41	34.88	32.06	40.59	31.39	36.14	28.61	30.43
Observations	1574751	946276	1071550	1010924	1263795	978094	1146511	886224	946276
					Female				
Education Reform	-0.274	0.006***	-0.007**	0.027***	-0.007**	0.003*	-0.000	0.005**	0.001
	(0.221)	(0.002)	(0.003)	(0.010)	(0.003)	(0.001)	(0.010)	(0.002)	(0.001)
Control Mean	40.45	0.41	0.38	3.07	0.41	-0.24	20.33	0.79	0.03
Control SD	47.91	0.49	0.49	2.68	1.16	0.41	2.36	0.41	0.35
Bandwidth	41.04	26.29	30.37	31.94	28.79	46.17	39.76	40.66	28.08
Observations	803354	511930	589270	609318	549943	901950	764488	784359	550281

Notes: Local linear RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the share of default funds in pension portfolios in percentage terms. In column 2, the outcome is a dummy variable equal to one if a pension investor follows the conditional $\frac{1}{N}$ heuristic while allocating money to pension funds. In column 3, the outcome is a dummy variable equal to one if a pension investor changed her funds in a year. In column 4, the outcome is the log of one plus the value of the monthly contribution in Turkish Lira. In column 5, the outcome is the relative Sharpe Ratio loss relative that indicates the loss from under-diversification to the domestic benchmark stock portfolio. In column 6, the outcome is the disposition effect measure. In column 7, the outcome is the annual net rate of return of pension portfolios after subtracting management fees of pension portfolios in percentage. In column 8, the outcome is a dummy variable equal to one if a pension investor outperforms the default fund in a year in terms of portfolio returns. In column 9, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

1.6 Robustness Checks

Participation bias extensively discussed in the literature (Lee (2009)) might still be a threat to the validity of estimates for portfolio outcomes. In words, portfolio outcomes are only observed for those who have positive balances in their individual retirement accounts. It is also important to figure out whether the 1997 Education Reform changes any portfolio outcomes through different channels, rather than solely by urging participation in pension plans. However, the 1997 Education Reform does not change the participation outcomes as its impacts are both small and imprecise for the full and male population and are small for the female population.

Results show that the participation rates for those born before and after January 1987 are similar. Moreover, the 1997 Education Reform does not impact the decision to participate in DC pension plans. As suggested by Lee (2009), if the treatment impacts of the 1997 Education Reform are promoting participant rather than non-participation and similar participation rates across control and treatment groups, estimates comparing control and treatment groups are valid. As in my setting, the point estimates are very small and imprecise overall, I can conclude that participation bias is not a concern.

I also follow the strategy proposed by Duflo (2001), testing the stability of estimates by adding the quadratic polynomials of participation rates in each month-year birth cohort to the regressions as controls. Moreover, the administrative data allows me to calculate the actual participation rates in each month-year birth cohort, which enables me to assess whether exact participation rates change estimates. Correspondingly, I add the participation rates in a quadratic polynomial form to each regression and later observe whether the estimates are robust to their inclusion. Regression with the additional controls of participation rates across month-year birth cohorts reproduces all tables related to the portfolio outcomes. For brevity, I present the concerning tables in Appendix A. Results show that point estimates are robust and not sensitive to the inclusion of participation rates in a quadratic polynomial form as controls. Overall, the participation bias does not seem to drive estimates.

1.7 Conclusion

This article examines the impacts of education on the decision to participate, wealth accumulation, and investment decisions in individual retirement accounts in DC plans. I identify the causal impacts of schooling by leveraging the exogenous variation stemming from the 1997 Education Reform in Turkiye which extended compulsory schooling from 5 to 8 years for those born after January 1987. The reform-induced increase in schooling is almost half a year for those born after January 1987 and a significant heterogeneity arises by gender with greater point estimates for females.

Employing an administrative data set spanning the universe of individual investors in DC pension plans in Turkiye in December 2019 and 2020, I do find compelling evidence that schooling is a crucial input to accumulating higher pension wealth in DC pension plans. Above all, a significant heterogeneity arises by gender, showing that the 1997 Education Reform causes females to have around 3% more wealth. In addition to the reduced-form estimates, two-sample instrumental estimates show that an extra year of schooling results in 5-6% more pension wealth in DC plans for females, but null effects for males. The OLS estimates show a strong correlation between years of schooling and the propensity to participate in DC pension plans. However, causal estimates demonstrate that schooling is not an important factor promoting substantially higher participation in DC pension plans.

I also examine the underlying channels of how education leads to more wealth accumulation in DC pension plans. To do this, I investigate the possible causal mechanisms related to financial skills, which are equity exposure i.e. the propensity to own equity and higher portfolio weights of equity or risky assets, financial literacy and sophistication, less exposure to behavioral biases and heuristics, and higher portfolio returns in DC pension plans. However, there is limited causal evidence that schooling causally promotes higher equity exposure in particular the propensity to hold equity funds and the share of wealth invested in equity funds or stocks or risky assets. My results also indicate that general education is not a causal factor in lowering financial mistakes, and behavioral problems prevalent in DC pension accounts. Thus, general education on its own is not sufficient for individuals to have adequate financial literacy. In line with the concerning estimates, general education does not lead to higher portfolio returns. Overall, I find supporting evidence that general education improves wealth accumulation in DC pension plans through labor market channels with greater labor income and a higher likelihood of being employed formally.

Existing literature points out that education and financial behavior are positively related and the relationship is robust. However, the strong positive correlation between education and financial behavior might not be causal and might reflect confounding factors such as genetic and family background. In short, education is possibly a factor in increasing abilities for the labor market rather than encouraging abilities related to financial markets. Therefore, combining general schooling with a curriculum aiming to foster financial skills deserves credit as a policy suggestion.

CHAPTER 2

Once Upon a Time in Anatolia: The Long Run Development Effects of American Missions in Anatolia

with Ilhan Can Ozen

2.1 Introduction

"Christianization had to come before civilization, he insisted, no matter the fact that the missionaries on the ground ... knew that it was secular education and foreign languages, that students demanded more than anything else."[1]

How does a historical intervention with a sizeable human capital investment affect long-run economic development? In this study, we explore a century-long historical adventure of the American Board of Commissioners for Foreign Missions (ABCFM), which planted educational seeds with its modern and secular education institutions in historical Ottoman Anatolia, contemporary Turkiye. Although the ultimate purpose was mainly related to religious ideals of Protestantism, the ABCFM massively invested in education, health, and various crafts and skill training activities with a particular emphasis on introducing schools and New England-type colleges in the 19^{th} and early 20^{th} century.[2]

Prior research studying mission effects is too focused on colonial settings, and especially in some of the most famous cases, in a pre-19^{th} century setup. This makes the result harder to interpret, as it involves multiple and sometimes conflicting channels which is hard to disentangle pure education effects on contemporary economic development and to generalize the findings of prior research. however, the missions in our context importantly focus on the 19^{th}

[1] ABC30, Rufus Anderson Papers, vol.10, memoranda of discussions in meetings of missionaries during Anderson's visits to the Levant,1843-44 cited in Makdisi (2011).

[2] **?**, **?**,Barton (1908), Kocabaşoğlu (2000), and Yücel (2005).

century time period and the case of education intervention that was designed in a modern, technical, and agriculture-oriented approach. Broadly, in our context, the Ottoman case was a striking exception, both because the colonizing dynamics never took place, and the domestic population converted to Protestantism at a very insignificant rate. This becomes critical, as it leaves the pure education effect undiminished, but cancels out, or limits the alternative hypothesis and channels, making the interpretation of our results more straightforward.

The Anatolian geography —the major part of the Middle East and Islamic societies— is critically affected during our study period by a host of growth and development processes. Put another way, it is a unique context with distinct cultural dynamics, particularly with the present religious homogeneity and with the ethnoreligious heterogeneity till the early 20^{th} century. According to existing research (Kuran (2018)), Islamic Law is admitted to be the primary reason for Middle East underdevelopment, through the weakened property rights, undergirded by the waqf system, depressing the capital accumulation. Regardless of its content, the policy implications of the corresponding research are limited. If history and historical institutions create immutable limitations for long-run development, the role of policy in economic development in the Middle East is limited and might be left non-examined. Our study offers a different explanation that opens a vital place for policy, where now the reasons for the underdevelopment of the Middle East are associated with the choices made in the crucial time periods. This article also reveals that the absence of historically well-designed policies such as human capital interventions might be one of the fundamental determinants of Middle East underdevelopment.

Using the development index provided by the State Planning Organization for each district in 2004, we figure out that a 10 km increase in the distance to ABCFM missions depresses the economic development by 0.07-0.12 standard deviation in Turkiye. Results also show that moving closer to an ABCFM mission by 10 km leads to a 5-17 percent increase in income per capita in 1996. We find empirical support for the underlying channels, chiefly the mechanisms of structural transformation, agricultural productivity, and human capital. In line with the established association between human capital and long-run growth, results show that locations close to ABCFM missions by 100 km turn out to have higher literacy by 2.5% and more female high school completion by around 8% in the long run.[3] Besides, moving closer to an ABCFM mission by 100 km lowers the child per woman by 0.2 child, which indicates that ABCFM missions change the gender norms in the long run by transmitting the pro-female cultural norms of Protestantism.

Prior research emphasizes the role of structural transformation and agricultural produc-

[3]Hanushek and Woessmann (2008, 2011, 2012).

tivity in the long-run economic development.[4] Our results are in line with those studies by pointing out that the historical human capital intervention by ABCFM ends up with the reallocation of economic activity from agriculture to industry and the increase in agricultural productivity by fostering skill generation and knowledge diffusion. Point estimates indicate a 100 km increase in proximity to the nearest ABCFM mission depresses the share of the population working in the industry by around 6% and also leads to an 11-13% increase in the share of the population working in agriculture. While getting closer to ABCFM missions lower, in the long run, the share of the population employed in agriculture, a 100 km decline in the proximity to the ABCFM missions promotes the share of agricultural production in the total national agricultural production by 2-7% and 2-5% in labor productivity in agriculture. We also examine several alternative mechanisms and subsequently rule out them.

To measure the long-term effects of ABCFM missions, similar to Caicedo (2018) we use the distance to the nearest ABCFM mission in km. Nonetheless, as Jedwab et al. (2022) points out, the non-random location choice of the ABCFM and the non-classical measurement arising from under-reporting of missions in remote and isolated areas in historical sources are threats to the validity. The latter problem simply arises from the non-identification of remote and underdeveloped areas missing in the mission atlases used in several studies. We address those challenges in two ways and subsequently establish the causal link.

Despite a rich set of controls for historical and geographical features in the regressions and detailed archival work exploiting all annual almanacs and reports of ABCFM over the period of 1810-1925, we first provide a placebo treatment analysis comparing the sites planned or suggested to have a mission facility but never equipped with the others. Placebo treatment analysis shows no economically and statistically significant effect on present-day development outcomes, implying that the mission activities are the determinant of the estimated impacts rather than location characteristics.

Secondly, in 1870 the ABCFM relinquished all its missions in the Syrian-Nestorian region to the Presbyterian Missionaries(ABCFM (1870)) and subsequently moved their activities completely into the geography of present Turkiye and so spatially relocated within the Ottoman Geography, which led to an exogenous variation in the distance to nearest ABCFM mission. This allows us to instrument the distance to the nearest ABCFM mission with the distance to the nearest Syrian-Nestorian mission station. Hence, the closest across-the-border sites in present Turkiye to the Syrian-Nestorian mission that were left by the ABCFM, are hypothesized to have a higher likelihood of ending up with an ABCFM mission. This is

[4] Caselli (2005), Bosworth and Collins (2008), Self and Grabowski (2007), Mellor (1999), Johnson (1997), Johnston and Mellor (1961), Hornbeck and Keskin (2015), Kögel and Prskawetz (2001), Gollin et al. (2002), Herrendorf et al. (2014), Restuccia et al. (2008).

because the movement cost would be minimized from the Syrian-Nestorian region to the geography of present Turkiye in this event. Estimates relying on instrumental variable strategy reveal the economically and statistically significant impact of ABCFM missions on present day-development with higher point estimates in terms of mean income and overall development index. Lastly, we further investigate the underlying channels driving the mission effects and examine the medium-run effects of the ABCFM missions in 1927 and the mission effects over time on population density which might be a proxy for economic development.

We joined several data sets into one by gathering information on the spatial distribution of the ABCFM missions and further details through in-depth archival work using primary historical sources either online or on-site whenever online investigation was not possible.[5] After detecting the locations of ABCFM missions, we utilize ArcGIS software to quantify spatial features. For the present outcomes, we mainly use two novel data sets, which are 1996 District GDP Data provided by the Turkish Statistical Institute (henceforth, TURKSTAT) and 2004 District Development Ranking Data provided by the State Planning Organization (henceforth, SPO). Finally, we exploit various historical censuses of the Ottoman Empire and Turkiye to control for initial conditions in the robustness checks and explore the mission effect in the medium run.

We perform several robustness checks. A critical robustness check to touch upon is the spatial autocorrelation in residuals. Kelly (2019) argues that spatial correlation might result in misleading statistical inferences with inflated t-statistics in regressions, particularly in settings where the observations are geographical units. We handle this issue with various approaches. We first correct the standard errors by implementing the procedure suggested by Conley (1999) for a range of different spatial correlation cutoffs. We later show that the inclusion of province-fixed effects removes the spatial autocorrelation in residuals. We further generate a new control variable to account for the extreme income disparities in the east-west axis and later add this variable into regressions, indicating that estimates are robust to its addition.

Last but not least, we run 1000 simulations by replacing the outcome and explanatory variable with the spatially correlated artificial noise. While replacing the outcome variable with the artificial spatial noise allows us to test the ability of our regressions to explain what it should not be able to capture, replacing the explanatory variable with the placebo noise variable allows us to figure out how often random noise outperforms our original variable. This practice can be treated as a randomization inference exercise, which also helps us re-

[5] We did the on-site archival inspection in Houghton Library at Harvard University to identify the locations of main stations and principal outstations. Our archival work covers chiefly the annual reports of the ABCFM in 1810-1920.

inforce that our results are causal. With this in mind, simulations manifest that the noise as either the explanatory variable or the outcome variable never outperforms the original explanatory and outcome variables. Therefore, simulations with placebo regressions stemming from artificial spatial noise reveal our estimates overall are unlikely to be driven by the spatial noise and unobservable factors.

2.1.1 Contribution to Literature

Prior research emphasizes that even quite historically remote missionary activities, still weigh in on current and important human capital and development outcomes (Jedwab et al. (2022)).[6] This is true not only at the aggregate level but in terms of ameliorating specific inequalities in the societies under question, especially certain types of missions targeted at increasing the knowledge and position of disadvantaged populations.[7] Our study follows this general literature as methodology, however, departs from it in some important respects. The problem with most of the existing literature is that the conversion channel might confound the pure human capital impacts. Additionally, colonial rule and the missions have moved hand in hand in many settings, leading to misinterpretation of findings. Our unique contribution lies in our non-colonial context without a channel of religious conversion where the demand for new education was more autogenous, and the education intervention was more modern, secular, and craft-based. This distinguishes our research and its chance of clean human capital effect identification. Similar points can be made about the fact that Ottoman society, with its Muslim dominant polity, and its well-established and well-divided multi-denominational society made the conversion channel non-functional.[8] Overall, this article builds on the literature, as the intervention is differentiated by increasing the skill and general knowledge of the population, rather than by transmitting religious values.

An extensive literature has related initial endowments to institutions to present economic outcomes (Sokoloff and Engerman (2000),[9] highlighting the geographical, cultural, and historical determinants of long-run economic development. Consistently, Kuran (2003, 2004, 2012) demonstrates that the institutions rooted in Islamic Law are the causes of the contemporary underdevelopment of the Middle East. However, this explanation limits the role of policy. Moreover, a major structural problem underdeveloped economies experience is

[6] Caicedo (2018), Waldinger (2017), Woodberry and Shah (2004), Calvi et al. (2020), Calvi and Mantovanelli (2018), Bai and Kung (2015), Koyama and Rubin (2022), and Cage and Rueda (2016).

[7] Nunn (2014), Okoye and Pongou (2014), Calvi et al. (2020), and Caicedo (2018).

[8] Makdisi (2011).

[9] Easterly and Levine (2003), Gallup et al. (1999), Acemoglu et al. (2001), Acemoglu et al. (2002), Guiso et al. (2006), Alesina and Giuliano (2015), Glaeser and Shleifer (2002), Spolaore and Wacziarg (2013) La Porta et al. (1998).

the low labor productivity in agriculture (Gollin et al. (2002)). Accordingly, a set of studies reports that productivity gaps in the agricultural sector are an important source of cross-country income variation (Caselli (2005), Bosworth and Collins (2008)), a key driver to reducing poverty (Self and Grabowski (2007), Mellor (1999)), a necessary condition for an economy to initiate the development process (Johnson (1997), Johnston and Mellor (1961), Hornbeck and Keskin (2015), Kögel and Prskawetz (2001)). We complement the literature by showing that historical human capital interventions with significant room for policy by paving the way for agricultural productivity subsequently leading to structural transformation play a major role in long-run development.

Recent studies establish that cultural values (Kuran (2018), and Guiso et al. (2006)), cultural festivals (Montero and Yang (2021)), and cultural transmission mechanisms (Becker and Woessmann (2011)) might be important drivers of the variation in the present-day economic development. Moreover, Turkish society is not free from the gender gap in human capital accumulation (Tansel and Güngör (2016)), and labor force participation (Dayıoğlu and Kırdar (2010)). We contribute to the literature on cultural values by presenting that even though the conversion channel is blocked in our setting, the cultural values of Protestantism that put more weight on female education were transmitted in the long run. Results indicate that the mission intervention promotes female education and fertility preferences. Thus, it adds to the empowerment of women despite of no impact on male education.

The rest of the article is organized as follows. In the next section, we briefly describe the historical intervention and setting. Section 2.3 introduces the data and renders the details of the research design with a particular emphasis on how we isolate the ABCFM mission impact. Sections 2.4, 2.5, and 2.6 report the findings for the primary development outcomes of the economic development index and the mean income, underlying mechanisms, and intermediate outcomes, respectively. In section 2.7, we perform several robustness checks. Section 2.8 concludes the article.

2.2 Historical Setting

"In the Ottoman Empire territories, a significant missionary education chain is known to have existed, from primary to higher education. In the year 1900, in Anatolia specifically, there were 400 different missionary schools in the different education stages. A considerable number of students, as high as 17500, attended these schools. For comparison, Yusuf Akcuraoglu estimates that these numbers represented 1/3 of the entire student body. In the same time period, the domestically financed elite education institutions consisted of 69 schools

distributed around Anatolia, with a total attendance of 6900 students."[10]

Ottoman Society and Polity at the start of 19^{th} century were in the midst of profound change and systemic challenges. The demands in society were beyond the capability and the plans of the government. The Tanzimat Decree in 1839 —a reform attempt to modernize Ottoman Society—, although a positive development in terms of societal reform, also made this demand-supply gap even larger, as the state defined all its subjects as having equal rights in its access to social and economic rights. The ethnic and religious competition between groups, and within groups created a positive impetus for increased education, and skills, in order to gain more advantageous positions in Ottoman society, post-Tanzimat.

The arrival of the first Protestant missionary to Ottoman Geography occurred in 1815 and was a British vicar. In 1820, two members of the ABCFM boarded in Izmir Harbor, with the goal of finding a place for the burgeoning overseas education activities, and more general activities associated with Christian conversion, and social transformation.[11] The first stage the missionaries had set for themselves was a kind of social mining, finding the open areas-populations that will be conducive to the missionary activity. Although they did not know about the Anatolian geography or the characteristics of the constituent populations, they would learn fast through the first excursions into the Anatolian territory and the first efforts to establish themselves in different localities. Moreover, after a century they had more than 70 mission stations but their rival the Catholic missions had only 10 mission stations that were mostly in Jerusalem and Istanbul.

After a certain time, the missionaries in Anatolia at the start of their excursions reported two regularities, the proselytizing of the Jewish, and Muslim populations to the Protestant cause faced serious opposition and harsh penalties arising from the apostasy laws prohibiting the conversion of Muslims, but there is some traction with the Armenian populations, with increased engagement with the Protestant missionaries if not conversion. Consequently, the missionaries updated the target areas of educational investment according to the preponderance of these populations, and the intensity of missionary work considerably increased around 1870 circa. Secondly, education (higher than the primary level) instruction in the foreign language and skill, craft, and printing activity allowed the missionaries to gain a significant foothold. The innovations that were introduced by the missionary activity to this geography, include both rudimentary industrial technologies, and many ways of increasing productivity in the agricultural areas. They were also instrumental in galvanizing, then dormant printing industry, which started to print Bibles but spread out for different kinds of

[10] Kocabaşoğlu (2000), translated from the original reference in Turkish.

[11] 'Turn the heathens to the true religion was their starting point, though their means to reach this goal was quite multidimensional.

publications.

Within less than a century, the ABCFM carried the know-how of the textile industry that took its source from New England, and the modern agriculture techniques of 19^{th} century US, which contributed to the generation of various skills and crafts for the different parts of Ottoman society.[12] As Finnie (2013) points out, the American missionaries boast for the first time to introduce the sewing machine, potato agriculture, oil lamp, and photo camera to Ottoman society.

The target group of enrollees was also significantly diverse, with many different denomination groups being enrolled in the ABCFM mission schools, including Muslim students. In line with the Protestant teachings, female literacy was given the same importance as males, which was also a very novel approach in Anatolia. In 1925, the Early Republican elites repealed the missionary activity to prioritize nation-building. However, four of the eight colleges have continued teaching with their programs and curriculum but under the surveillance of the Turkish Ministry of Education. One important point to note is that Ottomans had also schools. However, as the starting sentence we cite by Yusuf Akcuraoglu points out that the number of ABCFM's schools is not comparable with the relatively small number of elite schools of the Ottoman Empire. Thus, we conclude that state education seems to be an independent driver of long-run economic development with respect to its size.

Regarding the curricula that these educational institutions prepared and introduced to the Anatolian population, Mathematics, Algebra, Trigonometry, Natural Sciences, History, French, English, Turkish, and Geography were introduced into the curricula at all level missionary schools, subjects that were taught by the first time to these mass populations. It is likely to argue that this new curriculum is indeed in line with the demands of the 19^{th} century.[13] We single out the intervention in the 19^{th} century as significant in terms of education, skills, printing, and health intervention, which was hugely important for the host country in the 19^{th} century, in terms of its scale as 417 schools were opened in many different geographies, with a great variety of schools from theological seminary schools, girls boarding schools, high schools, and primary schools, headed by 8 New England type colleges which targeted giving the highest level English language education. Moreover, considering its nature as a multi-dimensional intervention, the secular curricula of these schools, both prioritized education in many different areas, but also prioritized agricultural and industrial work being undertaken by students.

[12]For instance, the industrial work in the Aintab mission in 1897 was narrated as *"the industrial department has afforded relief to great numbers, employing sixty-five girls in stocking making, thirty in felt embroidery, 300 in spinning and weaving, 200 in silk embroidery"* (ABCFM (1897), p.53).

[13]Somel (2001).

2.3 Data and Research Design

2.3.1 Data

We benefit from a variety of data sources as illustrated in Table 2.1 with a timeline chart displaying the chronology of the ABCFM's intervention and the data sets we employ. First, we use archival data to get the spatial distribution of ABCFM missions. Second, we employ TURKSTAT data on income and education levels, and the comprehensive data set assembled by SPO on present-day development measures.[14] Third, the Republican and the Ottoman population censuses allow us to measure the medium-run development impacts, and also to gather information on historical population counts to control the initial conditions, respectively. We utilize ArcGIS software to quantify the distance to the relevant mission facilities.

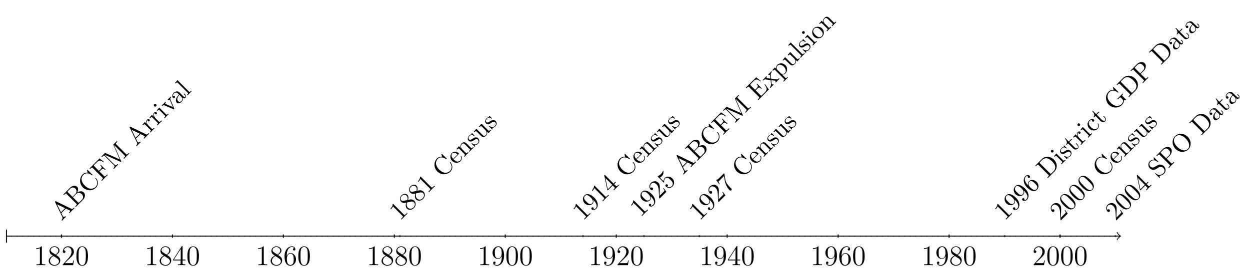

Table 2.1: Timeline of ABCFM Intervention and Data Sets

We conduct our empirical analysis at the district (sub-province) level, the second administrative unit following the province in modern Turkiye. For the development outcomes, we employ 2004 SPO District Development Ranking Data that has comprehensive information about a variety of standard development indicators such as urbanization, literacy, and occupational structure at the district level.[15]

For each district, SPO also notably constructs a standardized development index that is one of the main outcome variables.[16] It is estimated through principal component analysis of 32 socio-economic variables.[17] The corresponding index is advantageous in two respects.

[14]Dincer and Özaslan (2004).

[15]The SPO had a long tradition of portraying Turkish economic development from the 1960s till its repeal in 2011. Although the data set was published in 2004, it was, in general, put together through the Population Census of 2000 of the TURKSTAT.

[16]For detailed information, please refer to the link https://sbb.gov.tr/wp-content/uploads/2018/11/Ilcelerin_sosyo-ekonomik_gelismislik_si%C4%B1ralamasi_Arastarimasi-2004%E2%80%8B.pdf

[17]Briefly, they happen conventional development indicators lying on demography, agriculture and industry, urbanization, banking, health, transportation, human capital, and infrastructure characteristics.

First, it is free from external control or influence because already constructed in the SPO data set long before this article, helping to avoid the concerns related to specification search, data mining, and p-value hacking. Second, it avoids complications arising from the multiple hypothesis testing, as we assess the impact of the ABCFM missions on a large set of outcomes. SPO data set dismisses around 30 outlier districts in terms of economic development. Thus, it allows us to address the concern that outliers drive the results. We also do not include the districts in Istanbul in the sample because Istanbul has a distinct long-run development trajectory. Similarly, the SPO excludes almost all of the districts in Istanbul except five.

The second primary data set is 1996 District GDP Data assembled by the TURKSTAT on income levels, which provides a unique opportunity of having GDP per capita at the district (sub-province) level. It has more observations than the SPO data set, as it puts no restriction on highly-developed locations. Despite a smaller sample size of the SPO data with 867 units, the 1996 District GDP Data has all districts with 891 observations. We test whether the results are robust to the corresponding discrepancy regarding the sample size. Results in Appendix B in Table B.3 reveal no inconsistency.

Third, for schooling and fertility outcomes, we employ TURKSTAT 2000 Population Census. The information on land ownership distribution to calculate the GINI index and poverty rate is retrieved from the TURKSTAT 1997 Village Inventory Census. We utilize the TURKSTAT 1927 Population Census and 1927 Census of Agriculture to have the information on population features, the occupational distribution of the working population, the amount of cultivated area, the quantity of labor and total production in agriculture, and the prices of agricultural products in a specific district in 1927. The 1881 and 1914 Ottoman Population Censuses allow us to collect information on the features of historically settled communities in 19^{th} and early 20^{th} century as well as to control initial conditions.[18] Finally, to estimate the impacts of ABCFM missions on the population density over time we employ the 1927, 1965, and 2000 TURKSTAT Population Censuses together with 1881 and 1914 Ottoman Population Censuses.

To determine the spatial distribution of mission activity, we investigate the annual reports and the almanacs of the ABCFM starting in the 1810s and ending in the 1920s during our archival work.[19] Those annual reports and almanacs, the primary historical sources,

[18] Karpat (1985).

[19] With a highly hierarchically bureaucratic perspective, the ABCFM center in New England in the US obliged all missions abroad to report the outlook in their locations. The local reports sent by overseas mission stations were put together in the annual reports and later published by the ABCFM center in the US at their annual congresses. They document the annual activities of the ABCFM stations all over the world with information on the financial situation including accounting records, the number of staff in gender categories, the number and type of schools and health facilities, the number of students in those schools, the amount of donation, the newly established mission stations and its locations with geographic coordinates, and various

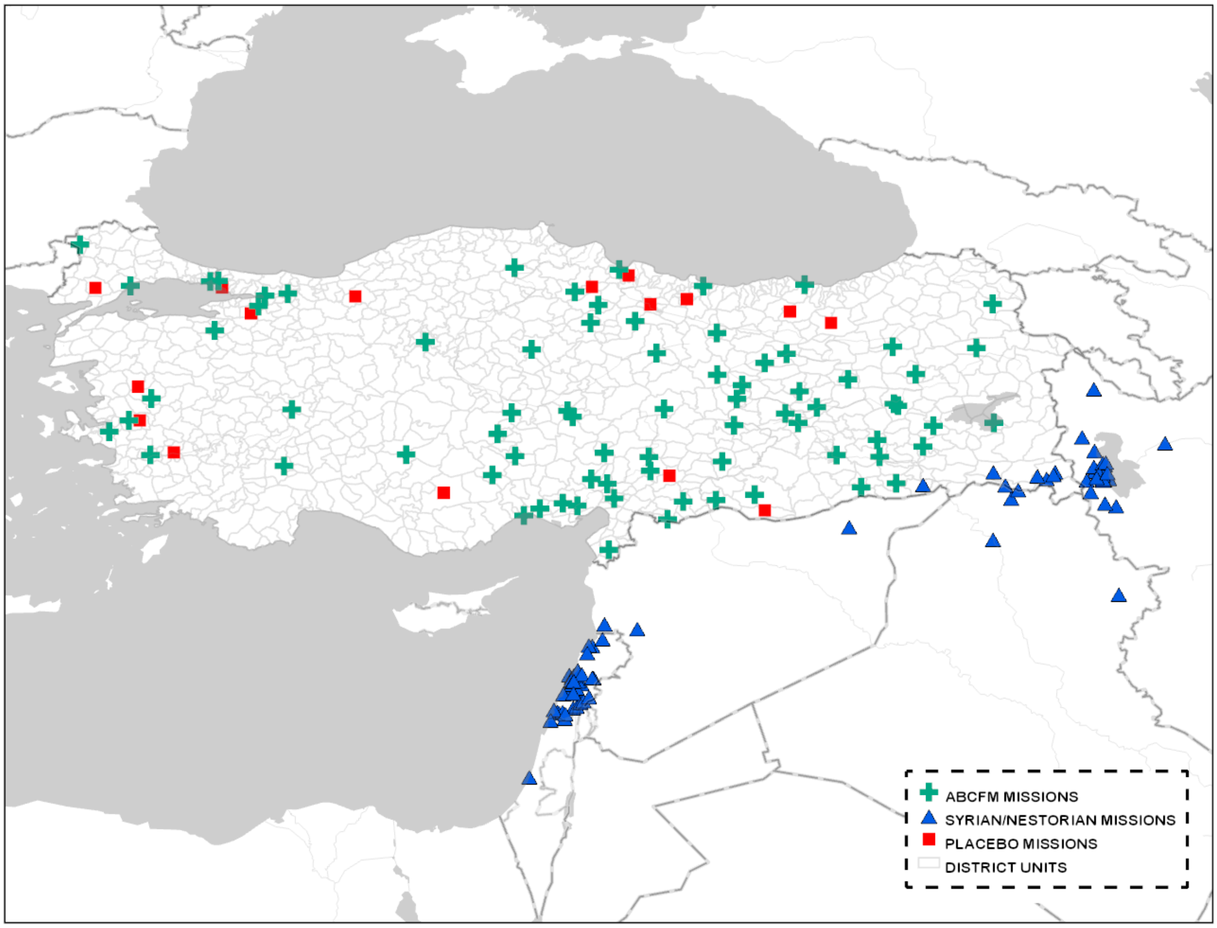

Figure 2.1: Spatial Distribution of Missions

Notes: The identification of any specific locations in terms of the mission facility it has is done through archival work. Archival work covers the annual reports and almanacs of ABCFM in the period of 1810-1920.

publish the location of main stations and principal out-stations. Even though there were also outstations other than the main stations and principal outstations that we measured the nearest distance from, most of the outstations were abandoned and also mainly directed from the main stations and principal outstations. As historical sources indicate that out-stations were not stable over time in terms of activity and in general were not equipped with modern education facilities.[20]

details.

[20]Kocabaşoğlu (2000) points out that out-stations were the units that took place in villages and were under the governance of the local community. The out-stations were far from having any influence on the decision process of ABCFM. Therefore, they existed for religious reasons rather than being allocated to education facilities.

Table 2.2: Descriptive Statistics

	More than 50 km					Less than 50 km				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Log of GDP per capita (in Dollars)	303	7.40	0.69	5.82	9.42	564	7.34	0.75	5.52	10.4
Development Index of the District	303	0.041	0.85	-1.63	3.99	564	-0.037	1.06	-2.01	7.95
Mission Distance	303	83.5	26.7	50.2	182.0	564	21.5	14.7	0.000030	49.8
Placebo Mission Distance	303	222.1	159.9	0.0028	893.1	564	190.7	140.0	0.0024	658.2
Distance to Syrian/Nestorian Missions	303	601.6	248.8	0	1134.2	564	469.5	282.1	0	1156.2
Average Ruggedness	303	14.4	7.57	0.85	35.2	564	12.3	6.22	0.88	31.5
Longitude	303	33.3	4.88	25.8	44.5	564	35.3	4.76	26.4	43.6
Latitude	303	39.5	1.64	36.2	42.0	564	39.1	1.38	36.0	42.0
Ann. Mean Temperature (.1 C degree)	303	108.4	27.1	42.9	179.8	564	112.6	31.5	42.9	189.4
Ann. Precipitation (mm)	303	654.1	149.5	369.4	1195.3	564	606.0	127.0	322.7	949.7
Elevation (m)	303	1013.6	497.8	64.3	2507.6	564	1034.2	532.9	49.2	2444.1
Forst-free period	303	199.5	33.9	127.2	337.7	564	203.9	35.9	127.2	333.8
Suitability index for wheat	303	32.9	15.0	2.88	69.5	564	35.6	13.8	5.83	84.5
Suitability index for barley	303	33.0	15.2	2.62	69.3	564	35.7	13.9	5.62	84.5
Suitability index for oat	303	32.5	15.0	2.81	69.1	564	35.4	13.8	4.97	84.5
Suitability index for cotton	303	3.92	6.19	0	29.5	564	6.74	8.44	0	41.0
Suitability index for olive	303	11.1	12.2	0	55.1	564	10.6	11.9	0	46.6
Dist. to Constantinople (travel days)	303	8.19	4.86	1.44	26.3	564	10.8	5.78	0.69	23.8
Dist. to Shore (travel days)	303	3.87	3.61	0.080	19.2	564	5.13	3.81	0.13	16.2
Length of primary rivers (km) per surface area (km2) in 2010	303	0.0064	0.018	0	0.14	564	0.010	0.020	0	0.13
Length of secondary rivers (km) per surface area (km2) in 2010	303	0.013	0.024	0	0.16	564	0.011	0.019	0	0.093
Distance to Nearest Custom Gate in km	303	321.2	159.0	4.82	619.1	564	255.0	143.1	4.59	582.2
Within Seljuk Sultanate	303	0.56	0.50	0	1	564	0.57	0.49	0	1
Within one and a half travel distance to 19th century ports	303	0	0	0	0	564	0.046	0.21	0	1

Notes: The table displays summary statistics of the main outcome variable and explanatory variable also control variables used in the regressions. For specific information on each variable about the source, definition, and construction, please refer to the Data Appendix in Appendix B.

Not only do the annual reports provide information on established missions but also information on either the planned or the recommended locations but not equipped with any mission facility. Put another way, those places never got full mission treatment. We refer to those as placebo missions and in the identification section, we discuss the possible reasons why they are taken as placebo. We also detect the missions in the Syrian/Nestorian region that were transferred to the Presbyterian Missionaries in 1870.[21] The geographical distribution of the areas having a mission facility and placebo missions, and also the missions left in 1870 in the Syrian/Nestorian region is depicted in Figure 2.1.

Benefiting from the ArcGIS software to calculate the distance in kilometers to the nearest actual, placebo, and transferred missions, we compute the distance from the centroid of a district to such points in kilometers (km), respectively. Additionally, we complement our data with some geographical variables such as ruggedness, elevation, the distance to the nearest custom gate, and the proximity to the coast and Istanbul. For some of the variables, we benefit from various sources, and further details for a specific variable are reported in Appendix B. Table 2.2 displays the descriptive statistics of key variables, separating sample districts into two groups for those closer than 50 km and those farther than 50 km.

2.3.2 Empirical Analysis

2.3.2.1 Econometric Model

To examine the impact of the ABCFM missions on present economic development, our econometric specification is as follows:

$$Y_{ij} = \alpha + \beta Dist_{ij} + \theta X_{ij} + \mu_j + \epsilon_{ij} \tag{2.1}$$

where Y_{ij} corresponds to the outcome of interest in district i in province j. The explanatory variable, the distance to the nearest ABCFM mission in kilometers, is denoted by $Dist_{ij}$. The main parameter of interest is β.

X_{ij} stands for the geographical and historical controls. The geographical controls include ruggedness, longitude, latitude, annual mean temperature, and precipitation, elevation, suitability indexes for wheat, barley, oat, cotton, olive, the distance to Istanbul in travel days, the distance to shore in travel days, the length of primary and secondary rivers per surface area in 2010 in kilometer square, and the distance to the nearest custom gate. The historical controls contain an indicator variable for whether the location is within the Seljuk Sultanate,

[21]For some specific archival documents not available online, we do on-site archival work in the Houghton Library at Harvard University.

likewise an indicator variable for whether the district is within one and a half-day travel distance to 19^{th} century major ports.[22] The variable μ_j captures the province-fixed effects accounting for the characteristics at the province level. ϵ_{ij} accounts for the idiosyncratic error term.

We re-estimate Equation 2.1 while shedding light on the underlying mechanisms and medium-run effects. In all specifications, standard errors are clustered at the province level.

2.3.2.2 Identification

The non-random settlement of the ABCFM missions raises the concern of endogeneity. As Jedwab et al. (2022) highlights several historical mission studies are likely to suffer from two sources of endogeneity driven by omitted variables and non-classical measurement error. As a result, a statistical association, OLS estimates in our case, might be far from causal as the higher likelihood of establishing the missions in safer, healthier, geographically more accessible, and economically more developed areas and the under-reporting of the missions in remote and isolated areas might lead to bias in estimates. We argue that a possible sample selection and reporting bias is unlikely in the ABCFM historical intervention.

The mission atlases, the primary historical source for several mission studies, are prone to non-classical measurement error emerging from significant under-reporting of the missions, predominantly in less developed locations (Jedwab et al. (2022)). Unlike the prior studies using the atlas published in 1912 as a single source of information, we benefit from each ABCFM annual report and almanac of the period of 1810-the 1920s which the local stations provided a detailed outlook of overseas missions in an organized manner. Overall, it is hard to argue that the non-classical measurement, reporting bias, is likely in our setting as our strategy to identify mission locations fundamentally differs from the prior studies.

The historical sources reveal that the ABCFM missionary activity mainly existed in remote, isolated, and underdeveloped areas in terms of 19^{th} century standards (Kocabaşoğlu (2000), Yücel (2005, 2011)). The sites where ABCFM planned or recommended to equip with a mission facility were usually chosen in advance before the physical presence of missionaries, resulting in sometimes ad-hoc and accidental choices.[23] Regressions also include a battery of controls to account for geographical and historical characteristics plus province-fixed effects that lead to the comparison of districts within the same province. Similarly, using Equation

[22] It is included in the set of controls for whether the specific district is in the hinterland of a port city and port. The size of a port's hinterland varies substantially and depends on the topography around it.

[23] For instance, the Aintab mission that became one of the most important main mission stations, after a while, was decided to be founded coincidentally. While going to Baghdad, an ABCFM missionary took a break in Aintab and suggested it be equipped with a mission facility even if he was not aware of Aintab before coming across it during his journey.

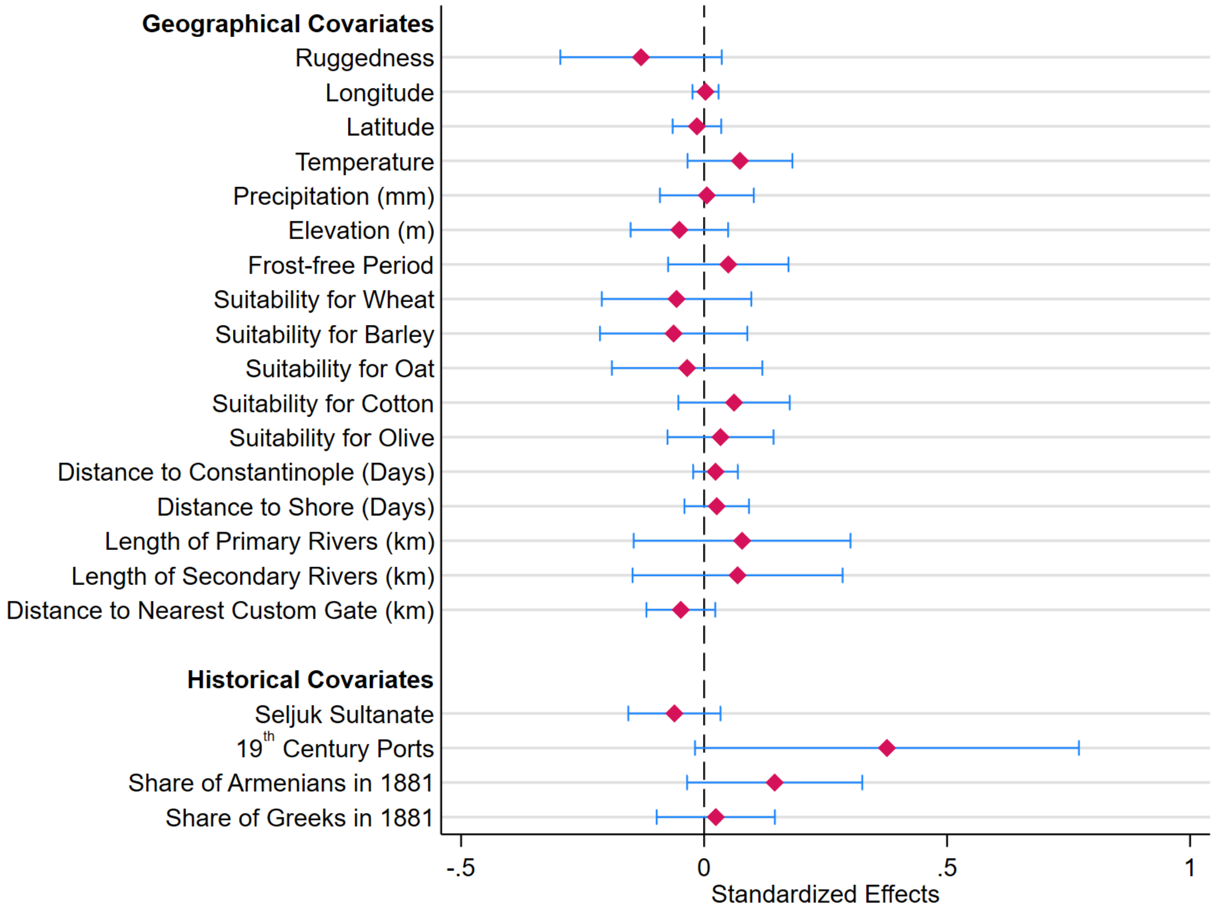

Figure 2.2: Balanced Covariates across Treatment and Control Sites

Notes: The graph presents the point estimates of treatment i.e. indicator variable equal to one if a site (district) has historically a mission facility and equal to zero otherwise. The pink dots plot the point estimates whereas the blue lines present the 95% confidence intervals for the relevant coefficient. The dark navy line indicates the value of zero.

2.1 conditional on only province-fixed effects we estimate the correlation between having a mission facility and the standardized predetermined covariates. Figure 2.2 indicates controlling for province-fixed effects makes the covariates of geographic and historical variables balanced across control and treatment sites at 95% confidence. With this in mind, in Table 2.3, we report the point estimates, revealing no systematic differences between the treatment and control sites.[24] We further adopt two distinct strategies to reinforce that the OLS estimates present the causal effects of ABCFM missions.

Our first strategy is to conduct a placebo treatment analysis. The rationale behind the placebo treatment analysis is that the mission impact is not driven by the location choices but by the true effect of mission activities. Hence, we design the placebo missions by ex-

[24]We generate a binary variable accounting for the mission presence in a district.

Table 2.3: Balanced Covariates across Treatment and Control Sites

	(1) Control		(2) Treatment		(1)-(2) Difference in Means	
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	N/Clusters	p-value
Average Ruggedness in the District	799	13.137	68	11.562	867	0.124
	80	(0.675)	45	(0.756)	80	
Longitude	799	34.428	68	36.858	867	0.832
	80	(0.585)	45	(0.641)	80	
Latitude	799	39.300	68	38.722	867	0.558
	80	(0.178)	45	(0.197)	80	
Ann. Mean Temperature (.1 C degree)	799	110.658	68	116.768	867	0.177
	80	(3.213)	45	(5.726)	80	
Ann. Precipitation (mm)	799	624.923	68	598.167	867	0.907
	80	(15.413)	45	(18.228)	80	
Elevation (m)	799	1020.802	68	1099.514	867	0.315
	80	(56.231)	45	(83.605)	80	
Forst-free period	799	201.958	68	207.024	867	0.426
	80	(3.795)	45	(5.994)	80	
Suitability index for wheat	799	34.727	68	34.142	867	0.465
	80	(1.518)	45	(1.818)	80	
Suitability index for barley	799	34.803	68	34.060	867	0.413
	80	(1.516)	45	(1.770)	80	
Suitability index for oat	799	34.400	68	34.121	867	0.654
	80	(1.484)	45	(1.945)	80	
Suitability index for cotton	799	5.510	68	8.644	867	0.288
	80	(0.811)	45	(1.407)	80	
Suitability index for olive	799	10.841	68	10.054	867	0.540
	80	(1.330)	45	(1.910)	80	
Dist. to Constantinople (travel days)	799	9.636	68	13.221	867	0.312
	80	(0.607)	45	(0.806)	80	
Dist. to Shore (travel days)	799	4.548	68	6.310	867	0.439
	80	(0.398)	45	(0.621)	80	
Length of primary rivers (km) per surface area (km2) in 2010	799	0.009	68	0.013	867	0.488
	80	(0.001)	45	(0.002)	80	
Length of secondary rivers (km) per surface area (km2) in 2010	799	0.011	68	0.011	867	0.528
	80	(0.001)	45	(0.002)	80	
Distance to Nearest Custom Gate in km	799	283.495	68	215.281	867	0.182
	80	(17.892)	45	(20.133)	80	
Within Seljuk Sultanate	799	0.571	68	0.574	867	0.203
	80	(0.056)	45	(0.083)	80	
Within one and a half travel distance to 19th century ports	799	0.025	68	0.088	867	0.062*
	80	(0.011)	45	(0.041)	80	
Sh. of Armenian pop. in 1881	764	7.200	65	14.839	829	0.113
	77	(1.168)	43	(2.091)	77	
Sh. of Greek pop. in 1881	764	6.597	65	4.693	829	0.695
	77	(1.164)	43	(1.349)	77	

Notes: This table compares the covariates between the districts with a mission facility and the districts without a mission facility. The p-value displays the test statistic of the hypothesis of difference in the means across the treatment and control groups is zero. Standard errors are clustered at the province level. Province-fixed effects are included in all regressions. ***p<0.01, **p<0.05, *p<0.1

ploiting the areas that were planned or suggested to have a mission facility by local staff but did not have. The corresponding locations and their circumstances were not evaluated through the physical presence of ABCFM missionaries as revealed by archival work.[25] More-

[25]For instance Malghara, Malkara in modern Turkiye, was mentioned in the 1852 Annual Report as fol-

over, historians state that ABCFM missionaries tried to spread missionary activity to the entire domain of the Ottoman Empire. However, the confusion in the early period and the circumstances that did not fit the plans in the field made the location choices to a large extent chaotic and coincidental.[26]

We hypothesize that the predetermined characteristics are balanced across the districts with an ABCFM mission and the districts without a mission facility. Figure 2.3 and 2.4 show that predetermined characteristics are balanced between the treatment and placebo sites, and between the control and placebo sites at the 95% confidence level, respectively. Table 2.4 reports no significant differences in standardized covariates between the districts with full treatment and the districts with placebo treatment. Table 2.5 documents that only the distance to the nearest custom gate differs between the placebo and control districts. Thus, we conclude that the distance to the nearest placebo mission allows us to conduct a placebo treatment analysis as the historical and geographical variables are balanced. Consequently, replacing the primary explanatory variable, the distance to the nearest mission, by the distance to the nearest placebo (non-established) mission instead, we re-estimate Equation 2.1.

Second, we adopt an instrumental variable approach (IV), two-stage least squares (2SLS), exploiting the exogenous variation in the spatial distribution of ABCFM missions brought about by the relinquishment of the ABCFM working field in the Syrian-Nestorian region. In 1870 the ABCFM transferred all its missions in the Syrian-Nestorian region to the Presbyterian Missionaries (ABCFM (1870)). They moved their operations into the area occupied by contemporary Turkiye. We argue that while making this spatial re-division within the Ottoman Geography, they considered the resettlement cost and sought to minimize it. Conditional on covariates, we hypothesize that the proximity to the Syrian-Nestorian missions predicts the likelihood of ending up with a mission facility for sites in contemporary Turkiye.

lows:
"Malghara has three hundred Armenian houses. Some of the people are hopeful inquirers after gospel truth."(ABCFM (1852), p.66)
However, Malghara had never been endowed with a mission facility according to the archival documents. Yagh-bassan, Yagbasan in modern Turkiye, is another placebo location in our research design and it was adverted in the 1882 Annual Report as follows:
"Yagh-bassan is a Greek village, on the spur of the same mountains, where is a good baud of brethren, and a promising opening for labor, but a pastor has not been found." (ABCFM (1882), p.32).
Moreover, Bazarjik, Pazarcik in modern Turkiye, was mentioned in 1913 Annual Report in the following passage:
"Bazaarjik, surrounded by 200 Moslem villages. .... The situation calls for the location and support of workers who are filled with desire and ability to reach the Moslems with the Gospel in each one of these centers and other centers not here named where the opportunity would be without limit. These are but illustrations of the opportunities opening, not only in the Marash field but in other parts of the Central Turkiye Mission as well as in similar regions in the missions to the north and west." (ABCFM (1913), p.78).

[26]Kocabaşoğlu (2000), Yücel (2005).

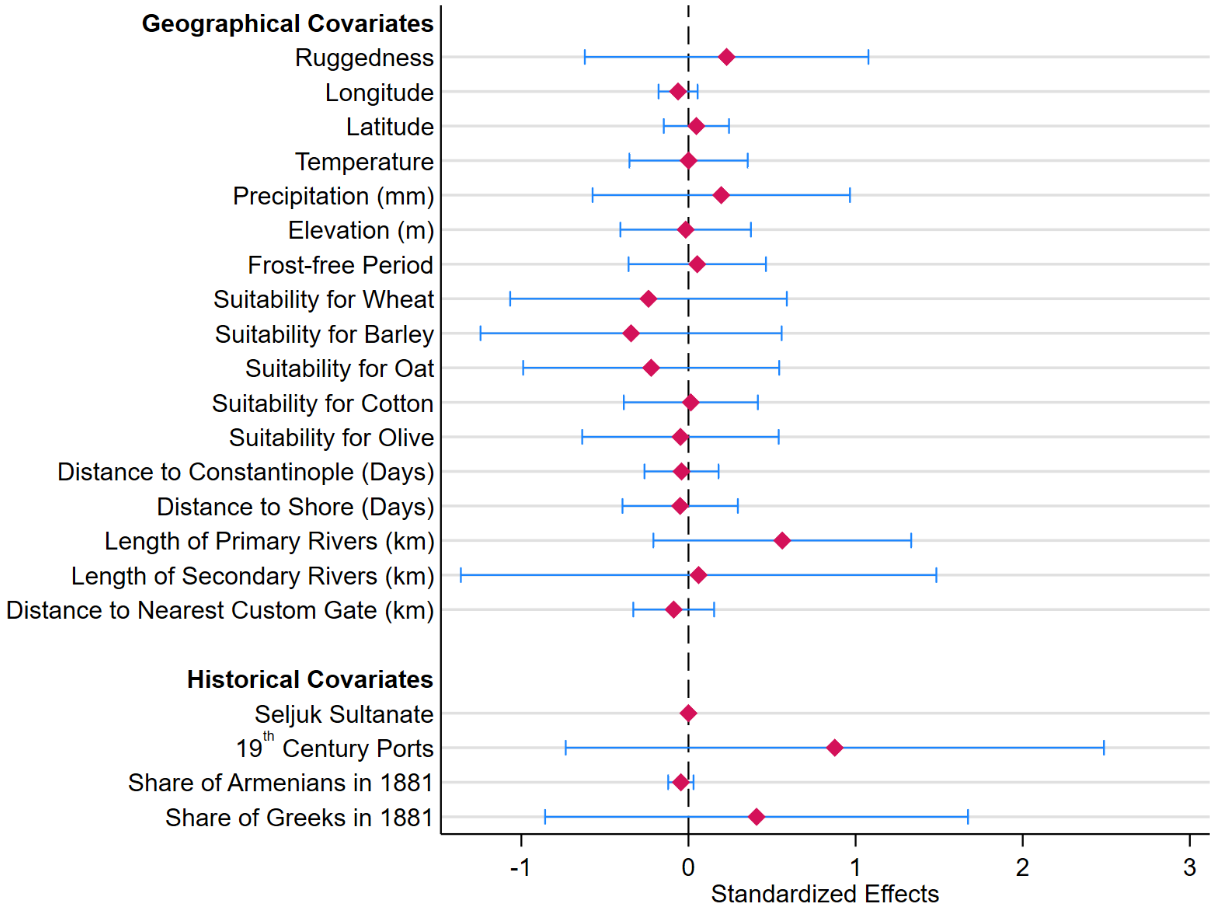

Figure 2.3: Balanced Covariates across Treatment and Placebo Sites

Notes: The graph presents the point estimates of placebo i.e. indicator variable equal to one if a site (district) has historically a placebo mission facility and equal to zero if a site has historically a mission facility. The pink dots plot the point estimates whereas the blue lines present the 95% confidence intervals for the relevant coefficient. The dark navy line indicates the value of zero.

Thus, the locations close to Syrian-Nestorian missions are likely to be close to an ABCFM mission. In short, this plausibly historical exogenous event allows us to instrument the main explanatory variable by the distance to the nearest Syrian-Nestorian mission.

The identifying assumption is that the distance to the nearest Syrian-Nestorian mission impacts the outcomes only through the explanatory variable conditional on the variables included in controls. A potential caveat is that distance to the present national border might have independent effects since the areas close to Syrian-Nestorian missions also tend to be close to the national borders, thus, violating the exclusion criteria of the instrument variable. To deal with this issue, we control the distance to the closest custom gate in regressions. Furthermore, in Figure 2.5 we report how geographical and historical characteristics differ by the instrument, i.e. the distance to the nearest Syrian-Nestorian mission. Estimates unfold that almost all of the geographical and historical features do not systematically vary by the

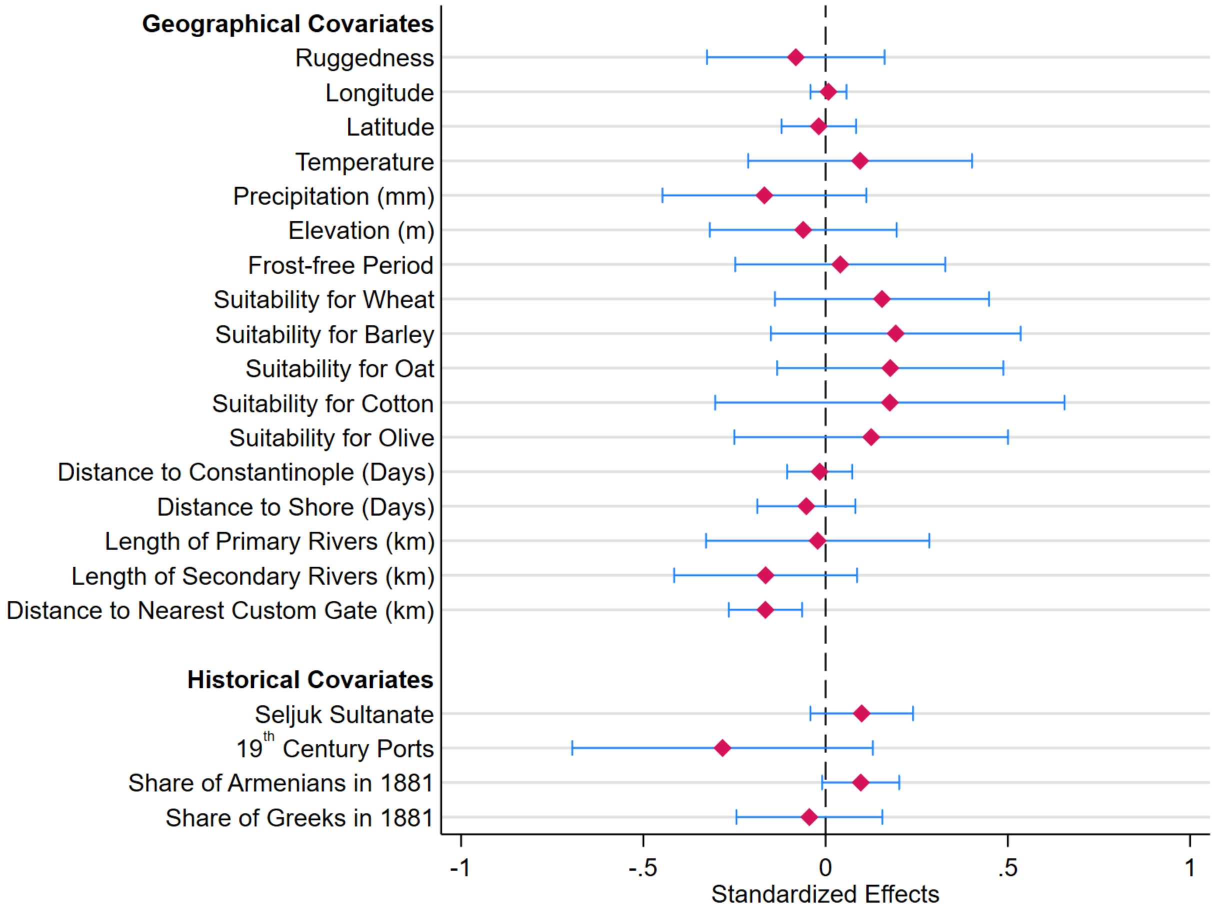

Figure 2.4: Balanced Covariates across Placebo and Control Sites

Notes: The graph presents the point estimates of placebo i.e. indicator variable equal to one if a site (district) has historically a placebo mission facility and equal to zero if a site has historically no mission facility. The pink dots plot the point estimates whereas the blue lines present the 95% confidence intervals for the relevant coefficient. The dark navy line indicates the value of zero.

instrumental variable. Despite the marginal significance of the point estimate, only elevation declines as a site gets farther from the Syrian-Nestorian missions. As a result, the covariates are not correlated with the distance to the nearest Syrian-Nestorian mission, reinforcing the exclusion criteria of our instrument likely holds.

The distance to the nearest relinquished mission has significant prediction power on the distance to the nearest ABCFM mission, leading to the first stage F-stat of 16. It also satisfies the conventional levels suggested by the literature against the problem of the weak instrument (Wang and Zivot (1998), and Bound et al. (1995)). Overall, the evidence of balanced covariates, the placebo analysis, and the instrumental variable approach plausibly establish the causal link between proximity to the ABCFM missions and present-day development outcomes in Turkiye. We follow the same instrumental variable and placebo treatment analysis strategies to examine the mechanisms and intermediate effects.

Table 2.4: Balanced Covariates across Treatment and Placebo Sites

	(1) Treatment		(2) Placebo		(1)-(2) Difference in Means	
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	N/Clusters	p-value
Average Ruggedness in the District	66	11.525	15	12.220	81	0.689
	43	(0.779)	13	(1.471)	50	
Longitude	66	36.973	15	33.928	81	0.107
	43	(0.650)	13	(1.405)	50	
Latitude	66	38.658	15	39.548	81	0.763
	43	(0.197)	13	(0.432)	50	
Ann. Mean Temperature (.1 C degree)	66	116.629	15	117.019	81	0.814
	43	(5.900)	13	(9.235)	50	
Ann. Precipitation (mm)	66	597.765	15	594.232	81	0.801
	43	(18.784)	13	(27.128)	50	
Elevation (m)	66	1117.686	15	859.418	81	0.756
	43	(84.813)	13	(151.349)	50	
Forst-free period	66	206.697	15	207.840	81	0.981
	43	(6.172)	13	(9.997)	50	
Suitability index for wheat	66	34.126	15	35.269	81	0.939
	43	(1.870)	13	(3.688)	50	
Suitability index for barley	66	34.033	15	35.771	81	0.700
	43	(1.821)	13	(3.917)	50	
Suitability index for oat	66	34.091	15	35.425	81	0.967
	43	(2.001)	13	(3.855)	50	
Suitability index for cotton	66	8.811	15	7.746	81	0.848
	43	(1.444)	13	(2.802)	50	
Suitability index for olive	66	9.696	15	14.544	81	0.752
	43	(1.935)	13	(4.298)	50	
Dist. to Constantinople (travel days)	66	13.510	15	7.809	81	0.867
	43	(0.797)	13	(1.201)	50	
Dist. to Shore (travel days)	66	6.470	15	2.668	81	0.933
	43	(0.626)	13	(0.440)	50	
Length of primary rivers (km) per surface area (km2) in 2010	66	0.014	15	0.007	81	0.141
	43	(0.003)	13	(0.003)	50	
Length of secondary rivers (km) per surface area (km2) in 2010	66	0.012	15	0.009	81	0.932
	43	(0.002)	13	(0.004)	50	
Distance to Nearest Custom Gate in km	66	211.273	15	275.875	81	0.322
	43	(20.317)	13	(43.343)	50	
Within one and a half travel distance to 19th century ports	66	0.076	15	0.067	81	0.458
	43	(0.039)	13	(0.063)	50	
Sh. of Armenian pop. in 1881	63	15.203	15	5.934	78	0.070*
	41	(2.132)	13	(1.490)	48	
Sh. of Greek pop. in 1881	63	3.982	15	11.968	78	0.880
	41	(1.176)	13	(4.053)	48	

Notes: This table compares the covariates between the districts with a mission facility and the districts with a placebo mission facility. The p-value displays the test statistic of the hypothesis of difference in the means across the treatment and placebo groups is zero. Standard errors are clustered at the province level. Province-fixed effects are included in all regressions. ***p<0.01, **p<0.05, *p<0.1

Table 2.5: Balanced Covariates across Placebo and Control Sites

	(1) Control		(2) Placebo		(1)-(2) Difference in Means	
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	N/Clusters	P-value
Average Ruggedness in the District	786	13.154	13	12.131	799	0.506
	80	(0.678)	12	(1.702)	80	
Longitude	786	34.434	13	34.057	799	0.749
	80	(0.586)	12	(1.563)	80	
Latitude	786	39.299	13	39.354	799	0.718
	80	(0.179)	12	(0.440)	80	
Ann. Mean Temperature (.1 C degree)	786	110.564	13	116.350	799	0.541
	80	(3.209)	12	(10.690)	80	
Ann. Precipitation (mm)	786	625.474	13	591.590	799	0.236
	80	(15.572)	12	(31.591)	80	
Elevation (m)	786	1022.556	13	914.737	799	0.635
	80	(56.298)	12	(171.860)	80	
Forst-free period	786	201.887	13	206.301	799	0.782
	80	(3.800)	12	(11.588)	80	
Suitability index for wheat	786	34.716	13	35.364	799	0.298
	80	(1.525)	12	(4.389)	80	
Suitability index for barley	786	34.785	13	35.899	799	0.267
	80	(1.522)	12	(4.676)	80	
Suitability index for oat	786	34.383	13	35.475	799	0.259
	80	(1.489)	12	(4.575)	80	
Suitability index for cotton	786	5.461	13	8.455	799	0.466
	80	(0.805)	12	(3.104)	80	
Suitability index for olive	786	10.799	13	13.414	799	0.509
	80	(1.327)	12	(4.800)	80	
Dist. to Constantinople (travel days)	786	9.656	13	8.441	799	0.717
	80	(0.610)	12	(1.287)	80	
Dist. to Shore (travel days)	786	4.575	13	2.921	799	0.436
	80	(0.402)	12	(0.458)	80	
Length of primary rivers (km) per surface area (km2) in 2010	786	0.009	13	0.008	799	0.888
	80	(0.001)	12	(0.003)	80	
Length of secondary rivers (km) per surface area (km2) in 2010	786	0.012	13	0.010	799	0.195
	80	(0.001)	12	(0.004)	80	
Distance to Nearest Custom Gate in km	786	283.803	13	264.845	799	0.002***
	80	(17.962)	12	(40.317)	80	
Within Seljuk Sultanate	786	0.570	13	0.615	799	0.165
	80	(0.056)	12	(0.157)	80	
Within one and a half travel distance to 19th century ports	786	0.025	13	0.000	799	0.176
	80	(0.011)	12	(0.000)	80	
Sh. of Armenian pop. in 1881	751	7.215	13	6.327	764	0.074*
	77	(1.184)	12	(1.613)	77	
Sh. of Greek pop. in 1881	751	6.544	13	9.640	764	0.660
	77	(1.160)	12	(3.116)	77	

Notes: This table compares the covariates between the districts without a mission facility and the districts with a placebo mission facility. The p-value displays the test statistic of the hypothesis of difference in the means across control and placebo groups is zero. Standard errors are clustered at the province level. Province-fixed effects are included in all regressions. ***$p<0.01$, **$p<0.05$, *$p<0.1$

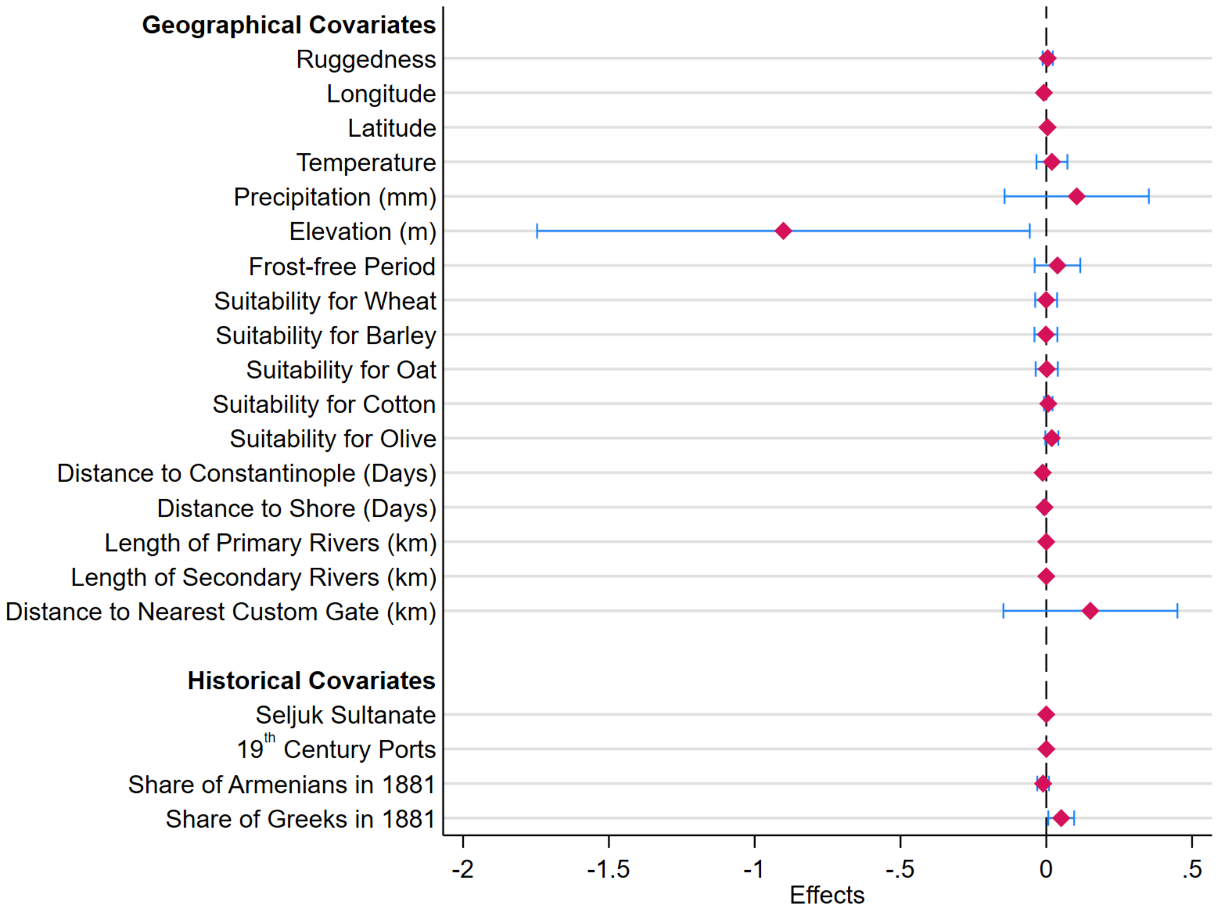

Figure 2.5: Distance to the Nearest Syrian/Nestorian Mission vs Covariates

Notes: The graph presents the point estimates of the distance to the nearest Syrian/Nestorian mission left to the Presbyterian missions in km. The pink dots plot the point estimates whereas the blue lines present the 95% confidence intervals for the relevant coefficient. The dark navy line indicates the value of zero.

2.4 Main Results

2.4.1 Development Index and Income per Capita

Employing 2004 SPO District Development Ranking Data, we begin our analysis by presenting visual evidence. Panel A in Figure 2.6 shows that districts close to ABCFM missions tend to be more developed presently. Similarly, binned scatter graph in Panel B indicates a similar pattern. Even though we find support for the positive mission effects visually, there is room for a more parsimonious analysis so we keep going with regression estimates.

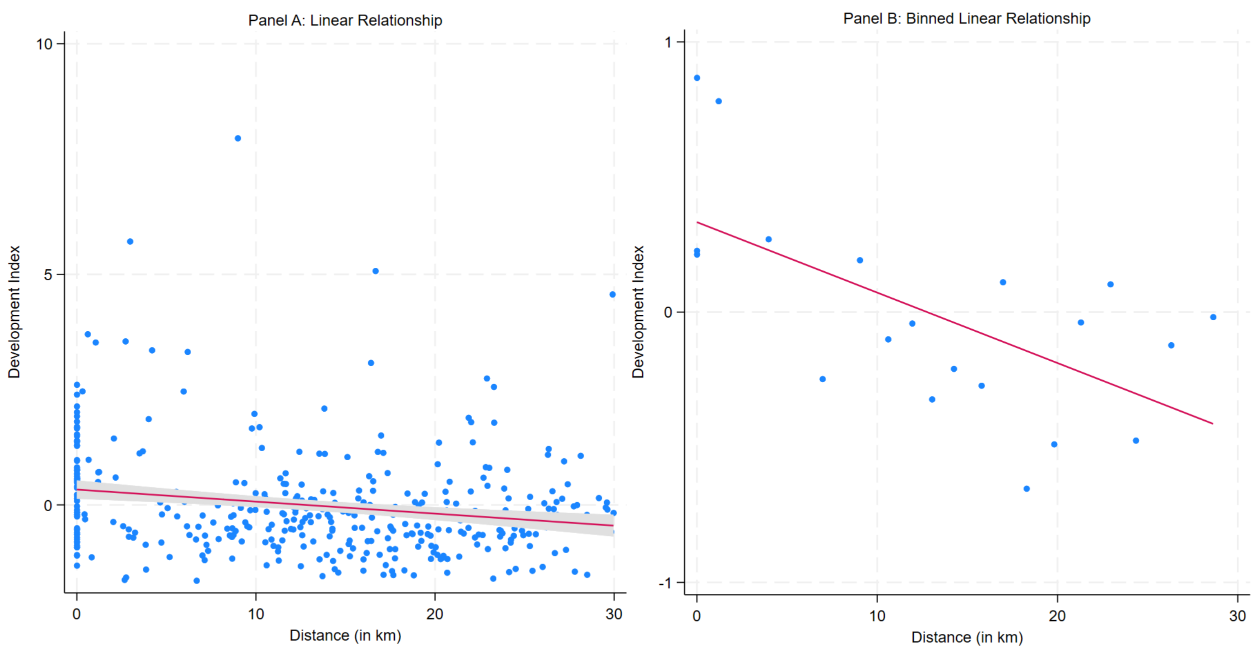

Figure 2.6: Distance to the Nearest ABCFM Missions vs Economic Development

Notes: Panel A depicts the unconditional graph of the development index in 2004 against the distance to the nearest ABCFM mission. The fitting curve is a linear line with a 95% confidence interval. Panel B depicts the unconditional binned scatter graph of the development index in 2004 against the distance to the nearest ABCFM mission. The data producing the graphs comes from the 2004 SPO District Development Data. The sample is restricted to districts closer than 30 km to the nearest mission.

To document the mission effects quantitatively on the development index in 2004, in Panel A in Table 2.6 we report the regression estimates. While the first column displays the estimates of the baseline specification with only province-fixed effects, the second column reports point estimates conditional on additional geographical and historical controls. Point estimates are firmly stable and robust to the inclusion of controls. None of the controls alter the coefficient of estimates. Column 2 also reveals that lowering the proximity to the nearest ABCFM mission by 10 km shifts the development index by 0.07 standard deviation.

To establish causality, we proceed with documenting the impacts of distance to the nearest placebo mission. Estimates reveal no statistically and economically significant impact in columns 3 and 4 as the gradient of the distance to the nearest placebo mission is small in magnitude and indistinguishable from zero. We lastly report IV-2SLS estimates. We also evoke that the districts close to the ABCFM missions in general are presently the underdeveloped part of modern Turkiye due to some other factors leading to a downward bias in the OLS estimates. Thus, we hypothesize that 2SLS estimates are greater than the OLS counterparts. Consistent with the above hypothesis, the corresponding 2SLS coefficient is larger than the OLS counterpart. Indeed, the IV estimates show that a 10 km increase in the proximity to the nearest mission lowers the development by 0.12 standard deviations in effect size.

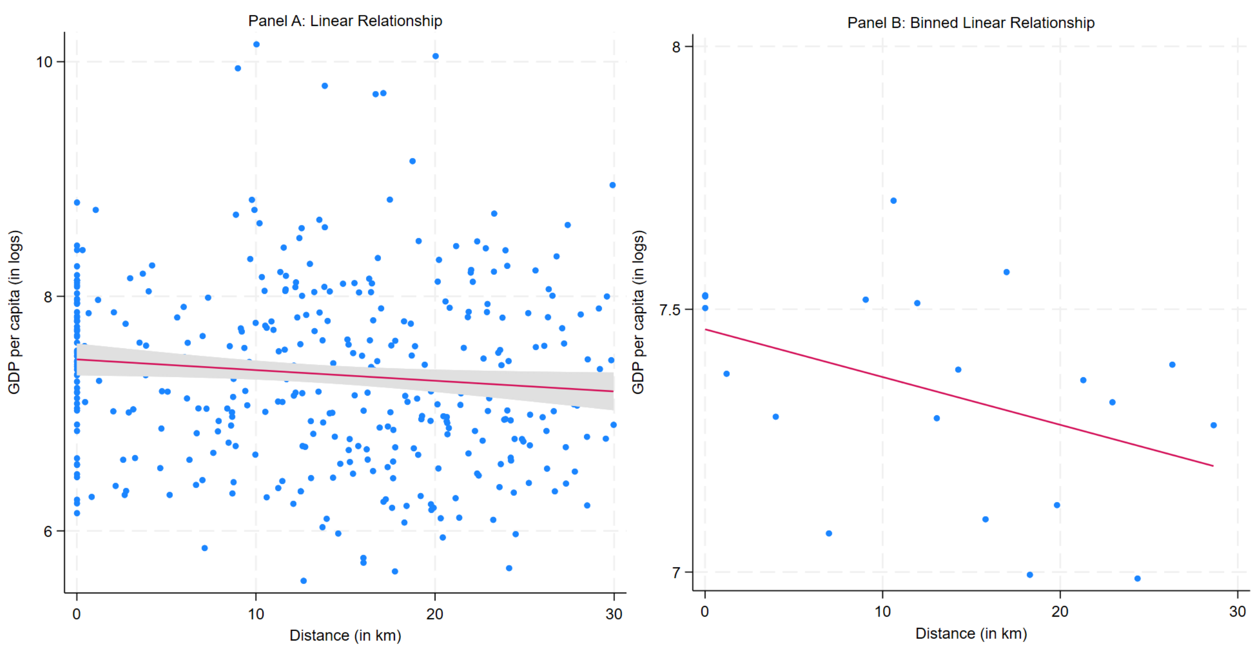

Figure 2.7: Distance to the Nearest ABCFM Missions vs GDP per Capita

Notes: Panel A depicts the unconditional graph of the log of GDP per capita in 1996 against the distance to the nearest ABCFM mission. The fitting curve is a linear line with a 95% confidence interval. Panel B depicts the unconditional binned scatter graph of the log of GDP per capita in 1996 against the distance to the nearest ABCFM mission. The data producing the graphs comes from the TURKSTAT 1996 District GDP Data. The sample is restricted to districts closer than 30 km to the nearest mission.

Employing TURKSTAT 1996 District GDP data, we continue our analysis by presenting graphical evidence. Panel A in Figure 2.7 plots the logarithm of income per capita against the distance to the nearest ABCFM mission, displaying the fact that districts close to ABCFM missions had a higher level of mean income, i.e. GDP per capita. Consistently, in Figure 2.7, binned scatter graph illustrates a similar pattern with more significant impacts of the ABCFM missions in Panel B. We keep reporting regression estimates. Column 1 in Panel B in Table 2.6 documents the baseline regression results with only province-fixed effects, while column 2 presents point estimates conditional on geographical and historical controls. The results are statistically significant and substantially stable, showing that the mission effect is not sensitive to including historical and geographical controls. In terms of magnitude, OLS estimates demonstrate that a 10 km increase in the distance to the nearest mission depresses the GDP per capita by five percentage points.

We further regress the distance to the nearest placebo mission on the logarithm of income per capita. Whereas being farther away from an ABCFM mission is associated with lower income per capita, distance to the nearest placebo mission has neither statistically nor economically significant impact since the coefficient is small in magnitude and indistinguishable from zero. Columns 3 and 4 of Panel B in Table 2.6 indicate the coefficients of the distance to the placebo missions is -0.001 and indistinguishable from zero, which is five-fold less than the gradient of distance to the nearest ABCFM mission in column 2 in Panel A. Finally, the 2SLS coefficient is three times greater than OLS estimates. In words, it is 0.017 in column 5 of Table 2.6, demonstrating that a 10 km increase in the distance to the nearest ABCFM mission lowers per capita income by 17 percent in contrast to the 5 percent OLS beta coefficient.[27] Above all, results suggest that the proximity to an ABCFM mission is associated with higher income per capita in the long run.

In short, it turns out that the historical intervention of ABCFM has a sizeable long-lasting contribution to economic development. Our point estimates are consistent with the existing literature yet contrast quite a bit in terms of magnitude. For instance, Caicedo (2018) finds evidence that a 10 km increase in the distance to a Jesuits mission in Paraguay leads to a 1.6 percent decline in income per capita. This contrast presumably arises from the fact that the intervention of ABCFM is heavy, multidimensional, and relatively new. In the next section, we investigate the chief mechanisms driving our results.

[27]Comparing our estimates with the rate of return of an extra year of schooling, which is 8% and 3% for women and men respectively, in Turkiye, the point estimates of the distance to the nearest ABCFM mission are great in magnitude (Aydemir and Kirdar (2017)).

Table 2.6: Mission Impact on Economic Development

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: Development Index in 2000				
Distance (in km)	-0.007***	-0.007***	0.000	-0.000	-0.012**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.36	0.45	0.34	0.43	0.44
Number of Clusters	80	80	80	80	80
First Stage F-stat					16.33
	Panel B: Log of GDP per Capita in 1996				
Distance (in km)	-0.005***	-0.005***	-0.001	-0.001	-0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.49	0.52	0.47	0.51	0.42
Number of Clusters	80	80	80	80	80
Sample Mean	7.36	7.36	7.36	7.36	7.36
First Stage F-stat					16.33

Notes: All columns in Panel A and B use data from 2004 SPO District Development Data, and TURKSTAT 1996 District GDP Survey respectively. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the standardized development index produced by SPO in all columns in Panel A and the logarithm of GDP per capita in Panel B. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap RK Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

2.5 Mechanisms

In this part, we explore the underlying mechanisms of how ABCFM contributed to long-term economic growth. The ABCFM historical intervention was massive and multi-dimensional, which carried know-how and provided substantial educational and health infrastructure, and several training courses in various crafts and skills. Hence, occupational specialization, agricultural productivity, human capital accumulation, and change in gender norms through the transmission of cultural values giving more importance to female education tend to be the main drivers of long-run economic development. One possible caveat is that some other channels are also likely to drive our estimates. To cope with such issues, we also assess some additional mechanisms and rule them out.

2.5.1 Structural Transformation and Agricultural Productivity

In this section, we focus on underlying mechanisms of structural transformation, in particular, the composition of the economic activity across the following sectors: i) agriculture, and ii) manufacturing, i.e. industry. In Panel A in Table 2.7, we report the regression estimates for the share of the population employed in the industry. Results show that being farther away from missions hinders the allocation of the labor force in the industry. The main coefficient of interest is quite insensitive to the inclusion controls and so stable. On top of that, IV estimates, in column 5, are consistent with the OLS estimates while placebo treatment analysis reveals a null effect. The 2SLS result demonstrates getting closer to an ABCFM mission by 10 km leads to an over 5 percent increase in the share of the population in the industry.

We, in addition, find significant evidence that the locations close to the ABCFM missions have a lower share of the population working in the agricultural sector. The point estimates in Panel B in Table 2.7 indicate a clear pattern of lower agricultural population share. In words, a 10 km increase in the nearest distance to the mission leads to a 1.15 percent increase in the population working in agriculture, which is more than a 1.5 percent increase relative to the sample mean. Estimations with the nearest distance to placebo mission locations in columns 3 and 4 present null estimates which are also much small relatively. IV estimate is consistent, and similar to the OLS coefficient even though it is imprecise. This result also implies that proximity to ABCFM missions brings about a structural transformation.

We explore the mission impact on the present-day agricultural production share in Panel C in Table 2.7.[28] The findings are striking, combined with the coefficient estimates on the population share in agriculture, and reveal that the proximity to the mission determines agricultural productivity. That is to say, districts close to ABCFM missions tend to have a larger agricultural production share. Furthermore, the point estimates are stable and statistically significant in Table 2.7. Similarly, the regressions with the distance to the nearest placebo mission locations produce a null effect. The estimated 2SLS coefficient is consistent with the OLS one. Yet, it has a greater magnitude as the IV strategy allows us to

[28]The share of agricultural production refers to the proportion of the national agriculture output.

Table 2.7: Mission Impact on Structural Transformation, and Agricultural Production and Productivity

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: Industrial Population Share in 2000				
Distance (in km)	-0.052***	-0.053***	-0.004	-0.003	-0.057*
	(0.009)	(0.010)	(0.004)	(0.005)	(0.030)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.38	0.41	0.35	0.39	0.41
Number of Clusters	80	80	80	80	80
Sample Mean	6.12	6.12	6.12	6.12	6.12
First Stage F-stat					16.33
	Panel B: Agricultural Population Share in 2000				
Distance (in km)	0.115***	0.129***	0.001	0.009	0.132
	(0.031)	(0.027)	(0.013)	(0.011)	(0.099)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.28	0.35	0.26	0.33	0.35
Number of Clusters	80	80	80	80	80
Sample Mean	69.46	69.46	69.46	69.46	69.46
First Stage F-stat					16.33
	Panel C: National Agricultural Production Share in 2000				
Distance (in km)	-0.002***	-0.001***	-0.000	-0.000*	-0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.32	0.39	0.29	0.38	-0.10
Number of Clusters	80	80	80	80	80
Sample Mean	0.12	0.12	0.12	0.12	0.12
First Stage F-stat					16.33
	Panel D: Log of Agricultural Labor Productivity in 2000				
Distance (in km)	-0.004**	-0.002	0.000	-0.000	-0.005
	(0.002)	(0.002)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.36	0.42	0.35	0.42	0.42
Number of Clusters	80	80	80	80	80
Sample Mean	0.51	0.51	0.51	0.51	0.51
First Stage F-stat					16.33

Notes: All columns in all panels use data from 2004 SPO District Development Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the percentage of the population employed in the industry sector in all columns in Panel A, the percentage of the population employed in agriculture in Panel B, the share of agricultural production of a district in Turkish total agricultural production in percentage terms in Panel C, and the logarithm of agricultural labor productivity in Panel D . Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentage terms. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

isolate the mission effect from confounding factors. Finally, we assess the impacts of ABCFM missions on agricultural labor productivity in Panel D. Point estimates show that moving farther away by 10 km from the nearest ABCFM mission depresses labor productivity in agriculture by 4-5%, however, the IV estimates are imprecisely estimated.

We argue that the proximity to ABCFM missions persistently has an important role in agricultural productivity, simultaneously lowering the population share in agriculture and promoting the proportion in the national agricultural output and labor productivity in agriculture. As the ABCFM intervened in Ottoman Anatolia through schools, health amenities, know-how and technology transfer, and vocational training in various skills and crafts, our findings are in line with the studies that emphasize the role of agricultural productivity in the long-run economic development (Caselli (2005)).

Overall, results suggest a reallocation of economic activity from the agricultural sector to the industry and service sectors. Also, we note that the agricultural production share in total national agricultural production and the labor productivity in agriculture is higher in the areas close to missions, suggesting that moving closer to ABCFM missions leads to higher agricultural productivity in the long run. Our results are in line with the literature that emphasizes progress in agricultural productivity and the shift of labor force from agriculture towards industry and service sectors is the main driver of long-run economic development (Caselli (2005), Herrendorf et al. (2014), and Restuccia et al. (2008)).

2.5.2 Human Capital and Shift in Gender Norms

We proceed with examining the effects of ABCFM mission activity on educational attainment. Analyzing the literacy and schooling rates and their relationship with the historical location of missions reveals a few important regularities. In Table 2.8, we assess the concerning ABCFM mission impacts by gender.

Comparing the distance to the nearest mission with full treatment sites with the distance to the nearest mission with placebo treatment sites documents a robust long-term link between historical exposure to ABCFM missions and literacy, as displayed in Panel A in Table 2.8. To make it more clear, moving farther away from the nearest ABCFM mission depresses the literacy rate by more than 2.5%. Despite the strong positive impact of ABCFM missions on literacy, the coefficients of distance to the nearest place mission are negligible and imprecise. Moreover, the IV-2SLS estimate in column 5 is very close to OLS estimates even though it is imprecisely estimated.

In Panel B and C in Table 2.8, we assess the long-run ABCFM mission effects on high school completion by gender. Panel B in Table 2.8 reveals a strong long-run ABCFM mission

Table 2.8: Mission Impact on Human Capital

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: Literacy Rate in 2000				
Distance (in km)	-0.026**	-0.026**	-0.004	0.001	-0.025
	(0.011)	(0.011)	(0.005)	(0.004)	(0.037)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.72	0.77	0.72	0.76	0.77
Number of Clusters	80	80	80	80	80
Sample Mean	83.41	83.41	83.41	83.41	83.41
First Stage F-stat					16.33
	Panel B: Female High School Completion Rate in 2000				
Distance (in km)	-0.036***	-0.040***	0.003	0.000	-0.084**
	(0.012)	(0.011)	(0.005)	(0.005)	(0.042)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.25	0.31	0.24	0.30	0.29
Number of Clusters	80	80	80	80	80
Sample Mean	9.28	9.28	9.28	9.28	9.28
First Stage F-stat					16.33
	Panel C: Male High School Completion Rate in 2000				
Distance (in km)	-0.004	-0.011	0.002	-0.003	-0.018
	(0.011)	(0.012)	(0.005)	(0.004)	(0.048)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.23	0.28	0.23	0.28	0.28
Number of Clusters	80	80	80	80	80
Sample Mean	15.23	15.23	15.23	15.23	15.23
First Stage F-stat					16.33

Notes: All columns use data from the TURKSTAT 2000 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the literacy rate in percentage terms in Panel A, the ratio of female high school graduates in percentage terms in Panel B, and the ratio of male high school graduates in percentage terms in Panel C. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentage terms. The first stage of the F-stat reports the Kleibergen-Paap RK Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

impact on female high school completion but we fail to find any causal impact on male counterpart. Being closer to an ABCFM mission by 100 km fosters female high school completion by 3.6%, which turns into a point estimate of over 35% relative to the sample mean. In addition, there is no significant impact stemming from being closer to the mission sites that did not get mission treatment. Considering the 2SLS estimate in column 5, we conclude that getting closer to the nearest ABCFM mission by 100 km promotes female high school completion by 8.4%, leading to a more than 90% increase relative to the sample mean. Our small and imprecise estimates in Panel C in Table 2.8 imply that ABCFM missions are not a factor in increasing male education, no matter what empirical specification we adapt. Consequently, the ABCFM mission effects are strongly heterogeneous by gender, which is in line with the Protestant ideals emphasizing female education. We interpret our findings as evidence of the transmission of cultural norms that favor female education, which leads to a shift in gender norms in the long run.

We also note that in Turkiye, the fundamental focus has switched from literacy promotion to prioritizing high school and higher education participation. Consistent with this, we also estimate the ABCFM mission impacts on primary school completion, which allows us a unique test for the validity of estimates on the domain of high school completion. Accordingly, primary school completion is compulsory then we expect that there is no impact of ABCFM missions on primary school completion. Table 2.9 indeed displays small and imprecise point estimates concerning the distance to the nearest ABCFM mission in the outcome of primary school completion in all estimation strategies we adopt. The null effects of ABCFM missions on primary school completion increase the credibility of our estimates regarding the high school completion and literacy rate, allowing us to conclude that what we estimate is not by chance or coincidental. All in all, the sites closer to ABCFM missions have substantially higher literacy rates and high school completion with a strong heterogeneity for the latter outcome favoring the females in the long run.

Table 2.9: Mission Impact on Primary School Education

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: Female Primary School Completion Rate in 2000				
Distance (in km)	0.011	0.011	0.001	-0.001	-0.064
	(0.019)	(0.020)	(0.009)	(0.007)	(0.077)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.34	0.35	0.34	0.35	0.34
Number of Clusters	80	80	80	80	80
Sample Mean	65.95	65.95	65.95	65.95	65.95
First Stage F-stat					16.33
	Panel B: Male Primary School Completion Rate in 2000				
Distance (in km)	-0.014	-0.023	0.001	0.003	-0.016
	(0.018)	(0.017)	(0.006)	(0.007)	(0.071)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.28	0.30	0.28	0.30	0.30
Number of Clusters	80	80	80	80	80
Sample Mean	69.70	69.70	69.70	69.70	69.70
First Stage F-stat					16.33

Notes: All columns use data from the TURKSTAT 2000 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the ratio of female primary school graduates in percentage terms in Panel A, and the ratio of male primary school graduates in percentage terms in Panel B. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentage terms. The first stage of the F-stat reports the Kleibergen-Paap RK Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

2.5.3 Additional Impacts

In the last part of the mechanisms section, we investigate alternative channels of the ABCFM missions on the long-run economic development. Our main objective is to provide complementary evidence that the channels hitherto discussed are the underlying ones driving the development effects of ABCFM missions. For brevity, we report the OLS estimates of full and placebo mission treatment effects in Table 2.10. All regressions include controls for historical and geographical features as well as province-fixed effects. The IV estimates are also reported in Appendix B.

Exploiting the 2000 TURKSTAT Population Census, we first investigate the impacts of ABCFM missions on fertility outcomes. Since ABCFM missions promote female education outcomes, we expect that it also contributes to empowering women. Column 1 in Panel A in Table 2.10 indicates that moving away 100 km from the nearest ABCFM mission increases child per woman by 2 children, which reinforces that ABCFM missions shift the gender norms through possibly the transmission of cultural values. In column 2, we present the estimates for the impacts of ABCFM missions on the infant mortality rate, which is indistinguishable from zero. Moreover, point estimates in column 3 reveal that the localities far away from ABCFM mission sites have a higher poverty rate. Despite the significant point estimates of full mission treatment, the point estimates of the placebo treatments are small and imprecise.

We continue with the mechanism of expropriation, which is crucial to single out this channel since ABCFM left modern Turkiye and its properties in 1925. As the investment by ABCFM was sizeable in terms of 19^{th} century circumstances, it is likely that the properties, in particular land and buildings, might have been captured by local elites. The ideal way of testing this hypothesis would be to compare the concentration of private assets in districts before and after the ABCFM left Ottoman Anatolia to capture how asset concentration was affected by the abandonment of ABCFM missions. However, no historical data set exists to conduct this analysis so we use a proxy variable for the land ownership concentration.

Table 2.10: Additional Mission Impacts

	(1) Fertility Rate	(2) Infant Mortality	(3) Poverty Rate	(4) Land Concentration	(5) Log Population Density	(6) Non-Resident Share	(7) Urbanization Rate	(8) Protestant Share
				Panel A: OLS Estimates of Treatment				
Distance (in km)	0.002**	0.027	0.032*	-0.000*	-0.004***	0.035**	-0.072**	-0.002**
	(0.001)	(0.027)	(0.018)	(0.000)	(0.001)	(0.014)	(0.034)	(0.001)
Observations	867	867	865	866	867	867	867	829
R-Squared	0.79	0.42	0.22	0.35	0.50	0.39	0.19	0.50
Number of Clusters	80	80	80	80	80	80	80	77
Sample Mean	3.73	41.25	6.22	0.69	3.96	1.51	44.05	0.32
				Panel B: OLS Estimates of Placebo				
Distance (in km)	0.000	-0.004	0.014	-0.000	0.000	0.010	-0.015	0.000
	(0.000)	(0.012)	(0.010)	(0.000)	(0.001)	(0.006)	(0.014)	(0.001)
Observations	867	867	865	866	867	867	867	829
R-Squared	0.79	0.41	0.22	0.34	0.49	0.39	0.18	0.50
Number of Clusters	80	80	80	80	80	80	80	77
Sample Mean	3.73	41.25	6.22	0.69	3.96	1.51	44.05	0.32

Notes: Columns 1, 5, 6, and 7 use data from the TURKSTAT 2000 Population Census. Columns 3, and 4 use data from 1997 Village Inventory Data. Column 2 uses data from 2004 SPO District Development Data. Column 8 uses data from the 1914 Ottoman Population Census. The explanatory variable in Panel A is the distance to the nearest ABCFM mission in km, and the distance to the nearest placebo mission location in km in Panel B. The outcome of interest is the fertility rate in column 1, the infant mortality rate in column 2, the poverty rate in column 3, the Gini index of land distribution in column 4, the logarithm of population density in column 5, the share of the non-resident population in column 6, the urbanization rate in percentage terms in column 7, and the share of the Protestant population in percentage terms in column 8. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap RK Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Using the 1997 Village Inventory Data, we construct a Gini index for land ownership distribution at the district level to see whether the proximity to ABCFM missions impacted it in 1997. We report regression estimates in Table 2.10 to explore the channel of expropriation, showing that the mechanism of expropriation is unlikely to drive the mission effect on long-run economic development. Yet, the IV estimates, in Appendix B, in terms of magnitude are small, reflecting that a 10 km decline in proximity to the nearest ABCFM mission leads to an increase of 2 percentage points in the concentration of land ownership relative to the sample mean. While interpreting such impact as either large or small is difficult, since our measure of asset concentration is far from perfect, we think that the channel of expropriation is not the primary one driving our results.

In columns 5 and 7 in Table 2.10, we document the ABCFM mission impacts on agglomeration we proxy by modern-day population density and urbanization rate. Results show that the areas closer to ABCFM missions have presently higher population density and urbanization rates despite the small and statistically insignificant impacts of placebo missions as reported in Panel B. An alternative channel might arise from the fact that people migrate to the districts close to ABCFM missions. Employing the TURKSTAT 2000 Population Census, we generate a proxy measure accounting for migration in locations. We subtract the total population number from the number of total residents in a district and then divide this value by the number of total residents. Note that positive values of the corresponding ratio imply out-migration. We report the regression estimates for this outcome in column 6 in Table 2.10. The gradient is positive, which supports that the human capital impacts of ABCFM missions do not seem to be driven through the sorting of people with more education to the districts historically closer to the ABCFM missions.

Lastly, we investigate the channel of conversion to Protestantism. Various studies figure out a relationship between Protestantism and long-run economic development.[29] In our setting, the conversion channel is not only closed today, and quite firmly has been closed since 1925, but the initial conversion was always quite negligible, meaning non-religious channels of effect need to be considered and prioritized in our setting. Unfortunately, no data set with information on the current denominations at the district level exists, and some surveys indicate that 99% of Turkish society is Muslim.[30] Hence, it is hard to say ABCFM missions had great success at inducing the conversion in the long term. To reinforce this hypothesis, we use the last census of the Ottomans in 1914 with information on denominations. We evaluate the impact of the ABCFM missions on the share of the Protestant population to show the short-run effects. Regression outputs in Table 2.10 point out that the ABCFM

[29] Weber and Kalberg (2013), Becker and Woessmann (2009), and Iyer (2016).

[30] For more details, see https://konda.com.tr/tr/konda-barometresi/

missions were not a significant factor in the share of the Protestant population.

2.6 Intermediate Outcomes

In this part, we report the medium-run impacts of ABCFM missions. To capture to what extent similar dynamics were available in the medium term, we employ various censuses of population, manufacturing, and agriculture.

We begin our analysis by presenting the mission effects over time on population density. As we acknowledge in the Historical Setting Section, the intensity of missionary activities tremendously accelerated in 1870 circa. The first reliable census in Anatolia was conducted in 1881 by the Ottomans and the second one was in 1914. Combining those censuses with various censuses of Modern Turkiye, we re-estimate the mission effects on population density by using Equation 2.1. Later, we plot the concerning coefficients to display the impacts of ABCFM missions on population density -a proxy of economic development- over time in Figure 2.8. Estimates show that the mission effects have gradually increased over time, but statistically significant point estimates have emerged after 1914.

We also hypothesize that in the medium run, the labor productivity in agriculture was higher in the sites closer to ABCFM missions whereas the land productivity was not affected as the land fertility was not able to be intervened by ABCFM missions. The underlying reason is that ABCFM introduced modern agricultural techniques through increasing human capital via schools and training in various skill crafts. In Table 2.11, we present the regression estimates for labor and land productivity, respectively. Our results confirm our hypothesis as the point estimates for labor productivity are both large and precise but the small and statistically insignificant estimates for land productivity. Moreover, we find no impact of placebo missions on labor productivity in agriculture as they are small and imprecise. Thus, we can conclude that the mission impacts we estimate are driven by the human capital investment in skills and education rather than selection on geographical characteristics. Put another way, if the corresponding estimates were driven by selection, then we would also observe ABCFM mission impacts on land productivity. Altogether, ABCFM's human capital investment in Anatolia also improves labor productivity in agriculture in the medium run.

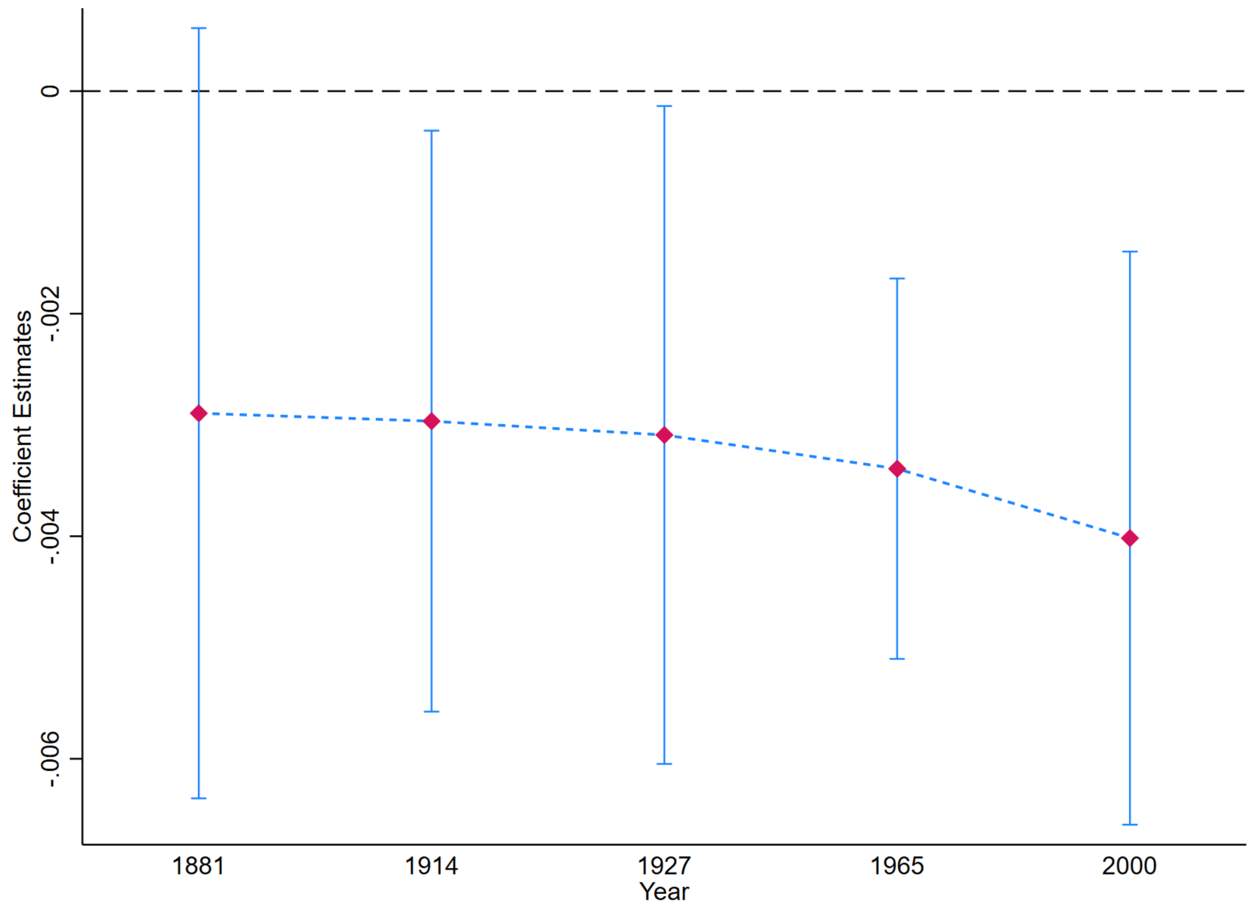

Figure 2.8: Distance to the Nearest ABCFM Missions vs Population Density over Time

Notes: The data producing the graph comes from the 1881 and 1914 Ottoman Population Censuses, and 1927, 1965, and 2000 Turkish Population Censuses. The graph displays the estimated coefficients and concerning 95 percent confidence intervals from estimating equation 2.1 for the log population density across time. The explanatory variable is the distance to the nearest ABCFM mission in km. The regressions include the full set of controls of province-fixed effects, and geographical, and historical characteristics.

The increase in labor productivity in agriculture in the medium run also led to the reallocation of economic resources from agriculture to industry and commercial sectors. We report our estimates in Table 2.12. For brevity, we report the IV estimates in Appendix B, so we only report the estimates relying on full and placebo mission treatments. The point estimates in columns 1 and 2 in Table 2.12 indicate that areas closer to ABCFM missions had a higher share of the population employed in the either industrial or commercial sector, implying the share of the population employed in agriculture declines. Despite the structural transformation led by ABCFM missions, treatment effects with placebo missions are small and imprecise. Above all, we conclude that ABCFM missions resulted in a structural transformation in the medium run.

Table 2.11: Mission Impact on Agricultural Productivity in 1927

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: Labor Productivity in Agriculture in 1927				
Distance (in km)	-41.306**	-31.608**	6.291	1.904	-118.901**
	(17.466)	(13.288)	(6.329)	(5.132)	(58.617)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	843	843	843	843	843
R-Squared	0.50	0.61	0.48	0.60	0.54
Number of Clusters	79	79	79	79	79
Sample Mean	7500.41	7500.41	7500.41	7500.41	7500.41
First Stage F-stat					17.69
	Panel B: Land Productivity in Agriculture in 1927				
Distance (in km)	-0.964	-0.715	-0.162	-0.290	0.405
	(1.038)	(1.067)	(0.402)	(0.436)	(3.482)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	843	843	843	843	843
R-Squared	0.46	0.49	0.46	0.49	0.49
Number of Clusters	79	79	79	79	79
Sample Mean	791.93	791.93	791.93	791.93	791.93
First Stage F-stat					17.69

Notes: All columns use data from the TURKSTAT 1927 Census of Agriculture. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the labor productivity in agriculture in 1927, i.e. agricultural output labor ratio in all columns in Panel A, and the land productivity in agriculture in 1927, i.e. agricultural output land ratio in all columns in Panel B. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Table 2.12: Mission Impact on Structural Transformation and Human Capital in 1927

	(1) Industrial Population Share	(2) Commerce Population Share	(3) Log Population Density	(4) Female Literacy Rate	(5) Male Literacy Rate	(6) Female Enrollment Rate	(7) Male Enrollment Rate
			Panel A: OLS Estimates of Treatment				
Distance (in km)	-0.024***	-0.018**	-0.003**	-0.015*	-0.020	-0.050***	0.008
	(0.007)	(0.007)	(0.001)	(0.008)	(0.015)	(0.018)	(0.050)
Observations	857	857	857	857	857	857	857
R-Squared	0.46	0.42	0.71	0.50	0.57	0.72	0.69
Number of Clusters	80	80	80	80	80	80	80
Sample Mean	4.51	3.38	2.89	2.03	13.08	8.93	43.01
			Panel B: OLS Estimates of Placebo				
Distance (in km)	-0.000	0.004*	0.001	0.004***	0.014***	0.009	0.014
	(0.004)	(0.002)	(0.001)	(0.001)	(0.005)	(0.007)	(0.021)
Observations	857	857	857	857	857	857	857
R-Squared	0.44	0.41	0.71	0.49	0.57	0.71	0.69
Number of Clusters	80	80	80	80	80	80	80
Sample Mean	4.51	3.38	2.89	2.03	13.08	8.93	43.01

Notes: All columns use data from the TURKSTAT 1927 Census of Population. The explanatory variable in Panel A is the distance to the nearest ABCFM mission in km, and the distance to the nearest placebo mission location in km in Panel B. The outcome of interest in 1927 is the percentage of the population employed in the industry in column 1, the percentage of the population employed in commerce in column 2, the logarithm of population density in column 3, the percentage of female literacy rate in column 4, the percentage of male literacy rate in column 5, the percentage of female enrolled in schooling in column 6, and the percentage of male enrolled in schooling in column 7. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentage terms. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***$p<0.01$, **$p<0.05$, *$p<0.1$

We also check whether the long-run impacts of ABCFM missions on human capital outcomes were apparent in the medium run. We highlight that the long-run human capital consequences of ABCFM are heterogeneous by gender, favoring female education but not male education. In columns 3,4,5, and 6 in Table 2.12, we document the ABCFM mission impacts on literacy rate and enrollment in schooling for each gender. The effects of ABCFM in the medium run are consistent with its effects in the long run. Results indicate that in the medium run, ABCFM human capital investment favored female education rather than male education. Despite the null effects on the male population, areas closer to ABCFM missions had higher female literacy and female enrollment in education. Besides, we underline that there is no robust impact of treatment with placebo missions. All in all, we conclude that the persistent pro-female human capital impacts were evident in the medium run, implying that the mechanism of transmission of cultural norms that shift the gender norms was at play in 1927.

2.7 Robustness

We now perform some robustness checks in this section. First, Kelly (2019) suggests that spatial correlation across geographical units might result in misleading inference in studies where the unit of observation is a geographical unit, and inflated t-statistics is likely a consequence of spatial noise rather than an accurate statistical relationship. We demonstrate that spatial noise does not induce our results. Second, we address sample discrepancies among the SPO District Development and TURKSTAT 1996 District GDP data. Lastly, we cope with the possible confounding impacts of World War I and its consequences such as the change in population composition.

2.7.1 Spatial Correlation

Adjacent units -districts- in our study setting, might result in spatial autocorrelation in residuals, leading to the fact that the long-term effect of ABCFM missions on economic development is the product of fitting spatial trends and downward biased standard errors failing to account for spatial autocorrelation in residuals. Consistently, Kelly (2019) points out that high statistically significant results are the artifact of uncorrected or arbitrary clustered standard errors for spatial autocorrelation in various studies. Whereas we adjust the standard errors by clustering at the province in all regressions, it is arbitrary. To avoid the misleading inferences emerging from arbitrary clustering and spatial autocorrelation in residuals, we propose various approaches.

A commonly proposed solution to account for spatial autocorrelation in residuals is to correct standard errors by implementing the procedure suggested by Conley (1999). We plot the Conley-adjusted t-values of the main explanatory variable in the regressions for the primary outcomes. Figure 2.9 by increasing the range of spatial autocorrelation with an increment of 5 km shows that t-statistics are below the value of 2 no matter the correlation range. Above all, the t-values of the explanatory variable are not sensitive to spatial autocorrelation.

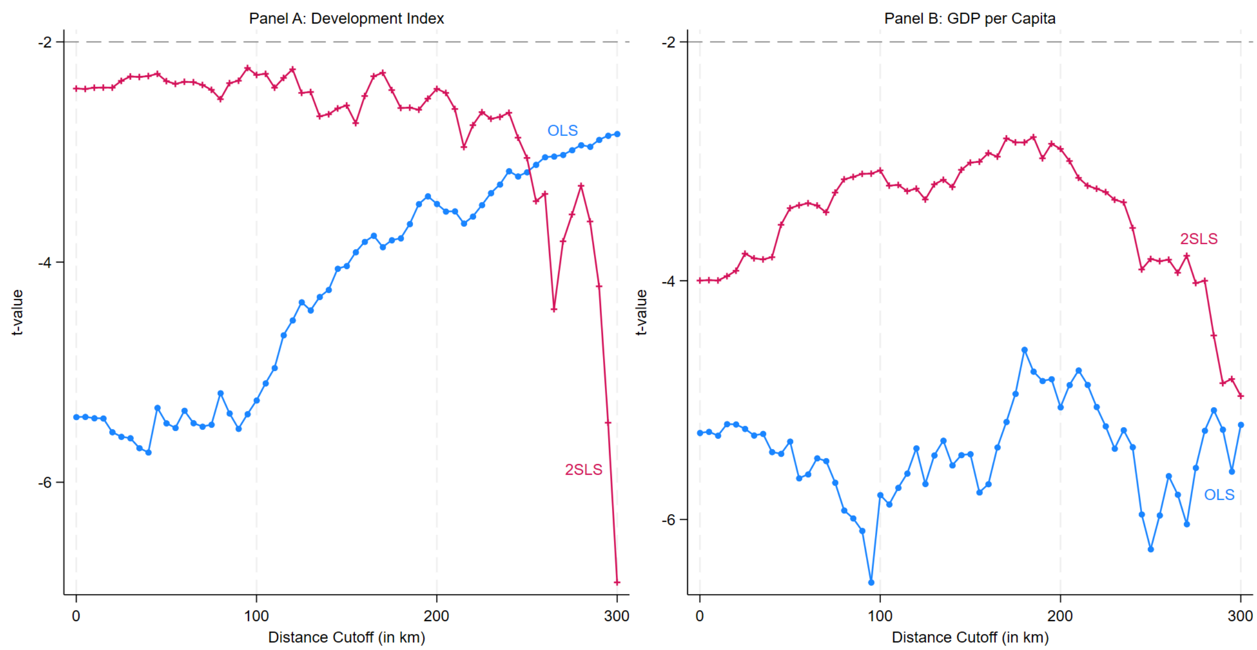

Figure 2.9: Conley (1999) Adjusted t-values

Notes: The data producing the graph in the left panel comes from the TURKSTAT 1996 District GDP Data. For the left panel, it comes from 2004 SPO District Development Data.

A substantial regional income disparity is present in Turkiye, in particular in the east-west geographical axis.[31] Accordingly, Figure 2.10 shows a strong eastward fall in the income per capita in 1996. Despite the inclusion of latitude and longitude in the controls of each regression, either being in the eastern part of modern Turkiye or the ABCFM mission presence might easily predict long-term economic development. To address this concern, we generate two variables equal to 1 if the neighborhood of a given district is further east than the 35-degree and 40-degree longitude, respectively. The inclusion of such control variables in estimations does not affect the point estimates. Table B.1 and B.2 report the estimates for

[31] Asik et al. (2020).

the logarithm of income per capita and the development index, respectively. The inclusion of control variables accounting for being in the eastern part does not affect the point estimates. Overall, the historical presence of ABCFM missions is still the predictor of long-term economic development even after controlling for being in the further east.

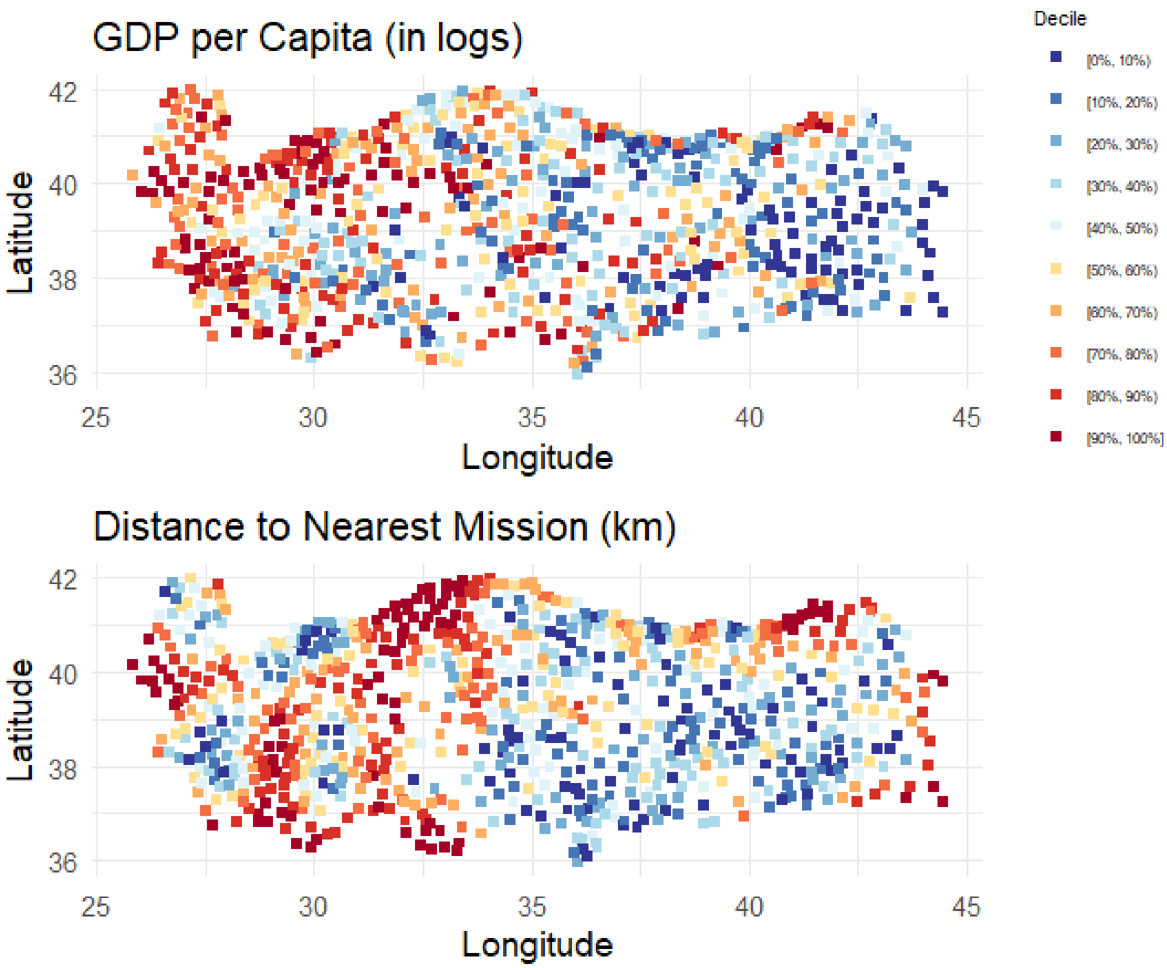

Figure 2.10: GDP per Capita in 1996 and Distance to the Nearest ABCFM Missions

Notes: The data producing the graphs comes from the TURKSTAT 1996 District GDP Data.

To investigate rigorously whether the spatial correlation exists among the residuals, we conduct a Moran test. As Table 2.13 presents the statistics of the Moran test, conditional on province-fixed effects, we fail to reject the hypothesis that no spatial correlation among the residuals. Note that regressions without province fixed-effects lead to a significant spatial autocorrelation among residuals. All regressions already include province-fixed effects, allowing us to rule out the spatial autocorrelation in residuals.

Following Kelly (2019) we perform simulations by generating a spatially correlated ar-

Table 2.13: Moran's z Values

	Log of GDP per Capita			Development Index		
	Moran's I	Moran's z	p-value	Moran's I	Moran's z	p-value
No Province FE	0.44	21.5	0.00	0.36	17.5	0.00
Province FE	0.0094	0.52	0.61	0.017	0.87	0.39

Notes: Table reports the Moran's I, Moran's z, and p-value of two main outcome variables: log of GDP per capita and standardized development index produced by SPO. Following(Kelly 2019), we assume that the weighting matrix gives equal weights to the five nearest neighbors of each district. The null hypothesis is no spatial correlation in the residuals of regressions either without or with province dummies.

tificial noise variable with the same spatial trend and correlation structure as the original dependent variables, the development index, and the logarithm of GDP per capita, across various correlation ranges.[32] We later compare the performance of our explanatory variable -the distance to the nearest ABCFM mission- and the noise variable. Spatial noise as an explanatory variable never outperforms the distance to the nearest ABCFM mission in explaining the variation in both the GDP per capita and the economic development index in various correlation ranges. In Appendix B, Table B.17 and Table B.18 reveal that the share of regressions where the noise variable outperforms the main explanatory variable is almost zero in any given correlation range. Besides, we replace the outcome variables with spatial noise in simulations, which indicates the distance to the nearest ABCFM mission does not have a significant ability to explain the placebo outcome arising from artificial noise. Thus, it is hard to say that the distance to the nearest ABCFM mission explains the spatial noise as it should not explain. With similarity to the randomization inference exercise, all in all, simulations with noise as either outcome or explanatory variable suggest that our estimates are unlikely to be driven by spatial autocorrelation as well as unobservable factors.

2.7.2 Extending Sample to Include All Districts

The SPO data set, as noted, has a lower number of districts than the TURKSTAT 1996 District GDP data since SPO excludes around 30 districts that are well-developed to avoid the outliers driving estimation. Nonetheless, the TURKSTAT data on the district's GDP covers all units in 1996. We report the estimates extending the study sample to all locations in Appendix B in Table B.3 for the primary outcome variable: the logarithm of income per

[32]We use a Matern function to estimate the spatial weight matrix with an exponential shape parameter for two different sites. The primary choice of parameter is the correlation range across sites. We include a set of correlation range from 25 km to 500 km, assuming the spatial correlation decays with the relevant distance.

capita. The point estimates are fully similar, showing that our findings are robust to the extension of the sample to all districts.

2.7.3 Accounting for Initial Conditions and Change in the Population Composition Post World War I

Existing studies report mixed results on how the legacy of minorities, particularly Armenians and Greeks, impacts the long-run economic development of Turkiye.[33] We acknowledge that ABCFM was inclined to settle among minorities with a particular emphasis on the Armenian community after 1840. Then, do places with a greater share of the Armenian population drive our results or do places have other characteristics that would interact with the distance to the nearest ABCFM mission in impacting the long-run economic development? If the answer is yes, then the estimates are biased. Even though the proposed strategies -placebo treatment analysis and IV-2SLS approach- to isolate the mission effect address this concern, we additionally provide some extra robustness checks.

The proper way is to control for the population share of minorities and population density and population numbers in the pre-ABCFM period in regressions, but the first population census of the Ottoman Empire was in 1881, approximately 50 years after ABCFM's first visit. On the one hand, it is the only available data set to study to what extent the ABCFM mission effect is responsive to the share of the minority populations and initial conditions. On the other hand, the share of the minority populations or the population density and population numbers seem to be outcomes or mediating factors in 1881 since the ABCFM missions that were built prior to 1881 might have positively impacted those outcomes.

Table B.4 and B.5 in Appendix B report the findings for main measures of interest, respectively, which unfold point estimates of the distance to the nearest ABCFM mission after controlling for the share of the Armenian and Greek populations, population density, and population are firmly stable and consistent. Note that the sample has fewer observations because three provinces were part of the Russian Empire in 1881. Estimated coefficients are quite similar to the point estimates in Panel A and B in Table 2.6 which documents our primary specifications that do not include the share of the minority populations and initial conditions as control variables. All in all, our results are robust to the inclusion of controls for minority presence and initial conditions.

We further include the interactions of the distance to the nearest ABCFM mission with the population shares of Armenians and Greeks separately in 1881 in Table B.6 and B.7 in Appendix B, respectively, allowing the mission effects to be heterogeneous with respect to the

[33] Arbatlı and Gokmen (2022), Asik et al. (2020), Sakalli (2017), Akarçay et al. (2021).

population shares of Armenians and Greeks in 1881 on the long-run development outcomes. By doing so, we additionally address the concern that the displacement of minorities or the change in population composition due to population replacement after World War I is a threat to the validity of estimates. The places with greater minority presence experienced a dramatic change in population composition post-World War I, which might have attenuated the mission impacts. Because, the Armenian and Greek minorities might have been the engine of transmission of human capital, knowledge, and various crafts. Suppose that it is true. Then, our estimates tend to be the lower bound of the mission impacts.

If the change in population composition attenuated the ABCFM mission impacts, then the places with a greater historical minority presence are likely to have lower mission effects. Consequently, the interaction terms of the distance to the nearest ABCFM mission with the population share of Armenians and Greeks in 1881 should be positive as the change in the population composition inhibits the mission impacts in the long run by blocking the transmission of historical mission effects. But in fact, the coefficient of the interaction term of distance to the nearest ABCFM mission with the Armenian population share in Table B.6 and B.7 in Appendix B are negative yet it is small in OLS specifications as well as both small and imprecise in the IV-2SLS specification in column 5. Moreover, the gradient of the interaction term of distance to the nearest ABCFM mission with the Greek population share is negative, small, and indistinguishable from zero. Prior studies also have figured out that the displacement of minorities depressed economic development in the short run.[34] To further show the short-run impact of the displacement of minorities, we use the population density in 1927 as a proxy for economic development. The corresponding coefficients are negative, small, and imprecise in Table B.8 in Appendix B. Overall, small, negative, and in most cases imprecise point estimates of the concerning interaction terms indicate that the change in population composition post World War I is not a valid concern, at worst implying that our estimates are the lower bound effects of ABCFM missions.

2.8 Conclusion

We examine the long-lasting impacts of a historical intervention focusing extensively on human capital investment on the long-run economic development in a setting with certain features. We explore a century-historical episode of the American Board of Commissioners for Foreign Missions (ABCFM) with modern secular education institutions. Results show compelling evidence that proximity to ABCFM missions led to a higher long-term income per capita and economic development, attributing this contribution to some important

[34]Sakalli (2017), Akarçay et al. (2021), Asik et al. (2020).

mechanisms. The ABCFM missions bring about a reallocation of economic activity from agriculture to industry by promoting higher agricultural productivity both in the mid-run and the long run. Moreover, we find causal evidence on the mechanisms of human capital, particularly in the domain of literacy for the full population and schooling for the female population. In addition to the positive impacts of the ABCFM missions on female education, fertility outcomes are significantly altered positively, leading to the empowerment of women and a shift in gender norms. We interpret this finding as evidence of the transmission of pro-female cultural norms of Protestantism in the long run.

We employ a data set by combining the archival work and present-day unique data sets. We isolate the mission effect by adopting two novel identification strategies. Comparing the sites planned or suggested to have a mission facility but never equipped, with the others, we rely on a placebo treatment analysis indicating no statistical and economic impacts of placebo missions on the long-run development outcomes. In addition, we employ a 2SLS strategy utilizing a historical episode resulting in the spatial re-partition of the working field of ABCFM in Ottoman geography. Leaving all its missions in the Syrian-Nestorian region to the Presbyterian Missionaries and fully putting weight on the geography of present Turkiye allows us to adopt an instrumental variable for the distance to the nearest ABCFM mission, which is our primary measure of interest. IV results are also consistent with the OLS estimates.

The geography this article focuses on is the historical Ottoman Anatolia, contemporary Turkiye. Turkish geography today is analyzed as the geography of education challenges and significant geographic heterogeneity. We provide evidence that more than 180 years ago, the missions that were established were able to increase mean income and economic development in the long term. This long-term effect, even in the presence of the stark discontinuity of the mission‘s operations is driven by productivity in agriculture leading to the reallocation of economic activity from agriculture to industry, the shift in gender norms through the transmission of cultural norms, and higher human capital accumulation.

A growing body of literature focuses on mission activity, and their long-term development effects, in detail and in specific research on the economic persistence of institutions.[35] There are certain characteristics about the Turkish case and the mission-population interaction in the Anatolian geography which makes us more confident that our estimates turn out to be the true effect of missionary activity, rather than any other channel confounding effect. I- The conversion channel is significantly closed or nonexistent for the entire duration of the missionary existence, and for the whole period after the disappearance of the mission. II- The nineteenth century is a crucial period in Ottoman history and economy, making the strength

[35] Cage and Rueda (2016), Nunn (2008), and Waldinger (2017).

of connections, and the presence of complementarities play a much more significant role than other mission geographies mostly relying on colonial settings. III- The mission effect is multidimensional, and significant effects are documented on the education dimension, the changing gender roles, and agricultural productivity. IV- The potential confounding variables of micro-development, mobility patterns, and historical emigration is controlled for, and either make our results a lower bound estimate of the true estimate or are not significantly affected by these additional variables/processes. Broadly, our study provides significant evidence to further support the external validity of empirical studies in the persistence and for the overall mission literature. Furthermore, we hope that our study is a good starting point to think about alternative channels and build on and rethink the already proposed reasons behind the underdevelopment of the Middle East.

CHAPTER 3

Education and Portfolio Investment: Evidence from Stock Market Portfolios in Turkiye

with Abdurrahman Aydemir

3.1 Introduction

Do those with more education participate in the stock market more? Does financial behavior vary by education? Does more education lead to better portfolio performance? Prior research has documented that general education and financial behavior are strongly correlated.[1] Particularly, more education is associated with a higher level of financial literacy and sophistication.[2] Moreover, Gomes et al. (2021) shows that those with more education are more likely to invest in risky assets, and the share of wealth allocated to risky assets -risky share- is substantially higher. Yet, less is known about to what extent this relationship is causal. In this study, we are interested in the causal effects of education on stock market participation, the variation in financial behavior in stock market portfolios, and portfolio performance. Yet, studying the causal education effects is a hard task as it requires a credible identification strategy and a comprehensive data set to track investors' portfolios. To do this, we adopt a regression discontinuity design (RD) by exploiting a discontinuity in birth cohorts in months arising from the 1997 Education Reform in Turkiye which exogenously extended compulsory schooling from five to eight years for those born after January 1987. Employing a novel data set spanning the universe of all stock market investors in Turkiye on December 31, 2021, and 2022, and using a quite large education intervention with high statistical power,

[1] Guiso and Sodini (2013), and Gomes et al. (2021).
[2] Lusardi and Mitchell (2014), and Calvet et al. (2007, 2009a,b).

we examine the causal impacts of general education on stock market participation, financial behavior, and portfolio performance.

The 1997 Education Reform increases years of schooling by half a year for those born after January 1987 with a strong heterogeneity by gender. In words, it favors female education more. Moreover, the education reform through its spillover effects substantially fosters high school and college completion though its primary purpose is just to promote junior high school completion. After quantifying the large and precise impacts of the education reform on schooling distribution, we turn to its causal impacts on stock market outcomes. Despite the OLS estimates revealing a strong positive correlation between years of schooling and stock market participation as well as the propensity to hold stocks, funds, and bonds, yet, we fail to find any causal evidence.

The point estimates for the overall stock market participation are statistically significant but small. OLS estimates show that an extra year of schooling is associated with a 1.6% increase in stock market participation whereas causal point estimates indicate approximately 0.35% increase by an extra year of schooling. Those point estimates are too small to argue that general education is a major input for stock market participation. We also find no significant evidence that education promotes the likelihood of having a transaction account, and of holding stocks, bonds, funds, risky assets, and stocks that are counted to be safer and more liquid in the stock market portfolios. In line with this, results show that education does not change the share of wealth allocated to those distinct assets. Our findings also show a null causal relationship between education and stock market wealth. Finally, we find no evidence that education significantly leads to higher portfolio returns. Overall, general education is at best a minor factor in stock market participation decisions and the variation in investors' stock market portfolios.

To isolate the causal effects of education, we leverage the exogenous variation stemming from the 1997 Education Reform in Turkiye which led to a substantial increase in schooling across birth cohorts in months. Utilizing the discontinuity in January 1987 in months, we adopt an RD design. In words, those born after January 1987 form the treatment group while those born before January 1987 stand for the control group in our quasi-experimental research design. The corresponding intervention is large and has high statistical power, allowing us to identify the causal education effects.

To elicit how the 1997 Education Reform has affected schooling outcomes, we use the Household Labor Force Surveys (HLFS) in 2018 assembled by the Turkish Statistical Institute. Later, to measure the causal effects of schooling on stock market participation and outcomes related to the stock market portfolios, we use a novel administrative data set spanning the universe of stock market investors in Turkiye in December 2021 and 2022. The data set is quite detailed to have the portfolio details such as which stock market assets are invested, the share of wealth invested in each asset, and the return of portfolios over time. The concerning data set is generated through the month-end snapshots on December 31, 2021, and 2022. The data owner is the Borsa Istanbul Group which is the legal authority to keep the records of individual transaction accounts in Turkiye. Therefore, the mismeasurement or reporting bias is not an issue in our research.

3.1.1 Contribution to Literature

We contribute to several branches of the literature. Several studies have documented the causal effects of education on various outcomes. Such outcomes vary from labor income,[3] health and fertility,[4] crime,[5] financial behavior,[6] capital returns,[7] cognitive skills,[8] and domestic violence.[9] Our study is most closely related to the work by Black et al. (2018) and Cole et al. (2014). Yet, our study starkly differs from those studies in some important respects. Both Black et al. (2018) and Cole et al. (2014) are likely to suffer the problem documented by de Chaisemartin and D'Haultfœuille (2022), Goodman-Bacon (2021), and Wooldridge (2021), as they adopt a difference in differences design (DID) with two-way fixed effects i.e. a staggered DID which relies on variation across groups of units receiving treatment at different times. Unless the assumption of constant treatment effects over time, then point estimates are biased. However, we exploit an RD design with an exogenous variation arising from a large compulsory schooling reform across birth cohorts, in addition, we have a richer set of outcomes through using detailed administrative data. Unlike the findings of those studies, we find no overall significant effect of education on stock market participation and the outcomes related to stock market portfolios. Consequently, we complement the literature on the returns of schooling.

Prior research has focused on explaining the reasons why a large proportion of the population does not own any stocks (Gomes et al. (2021)).[10] Furthermore, a growing body of literature hypothesizes that human capital is a close substitute for safe assets such as bonds so those with more education are more likely to invest in equities and to participate in the

[3]Duflo (2001), Oreopoulos (2006), Acemoglu and Angrist (2000), Angrist and Keueger (1991), Card (1993), and Aydemir and Kirdar (2017).

[4]Lleras-Muney (2005), and Black et al. (2008).

[5]Lochner and Moretti (2004).

[6]Cole et al. (2014),Black et al. (2018), and Gray et al. (2021).

[7]Fagereng et al. (2020a).

[8]Carlsson et al. (2015).

[9]Erten and Keskin (2018).

[10]Guiso and Sodini (2013), Haliassos and Bertaut (1995), Calvet et al. (2007), van Rooij et al. (2011), Grinblatt et al. (2011), Guiso et al. (2008), Malmendier and Nagel (2015), Fagereng et al. (2017), and Ameriks and Zeldes (2011).

stock market. Indeed, life cycle models predict that the optimal risky share -the wealth allocated to risky assets- is higher as human capital accumulates more (Gomes et al. (2021).[11] Consistently, Black et al. (2018) documents that those with more education tend to have higher risky share and are more likely to participate in the stock market. Yet, our article with a clean identification strategy and quite different findings departs from the existing literature. Results show that human capital does not seem to be a major factor driving stock market participation and the risky share.

Last but not least, our study speaks to the extensive literature on education and financial literacy, which reports that education and financial literacy is positively related.[12] Put another way, individuals are likely to have more cognitive and numeracy skills with more education, therefore, are more financially literate.[13] However, recently some studies document that financial outcomes such as capital returns (Fagereng et al. (2020a)), financial wealth, and investment behavior (Fagereng et al. (2021)) do not vary with education. Moreover, those recent studies emphasize the role of family background on financial wealth and financial behavior (Fagereng et al. (2020b)). Our study also builds on this literature by supporting the findings of those recent studies as our results show that education does not have a causal explanatory power in financial wealth and financial behavior.

The rest of the article is organized as follows. In the next section, we briefly discuss the 1997 Education Reform in Turkiye. Section 3.3 talks about the data and explains the details of the research design. Sections 3.4 and 3.5 present the findings and robustness checks, respectively. Section 3.6 concludes the article with a broad discussion of our results.

3.2 The 1997 Education Reform

See Section 1.3.

3.3 Data and Research Design

3.3.1 Data

We employ two main data sets to conduct our empirical analysis. The first main data set is the nationally representative 2018 Turkiye Household Labor Force Survey by the Turkish

[11] Cocco et al. (2005), Bodie et al. (1992), Fagereng et al. (2017), Guiso and Sodini (2013), Vissing-Jørgensen (2002), Catherine (2021), and Viceira (2001).

[12] Calvet et al. (2009b, 2007, 2009a), and Lusardi and Mitchell (2014).

[13] Indeed, Carlsson et al. (2015) with a reliable causal identification strategy finds significant evidence that those with more education have higher cognitive skills.

Statistical Insitute (TURKSTAT). It allows us to study how the 1997 Education Reform shifted schooling outcomes with respect to birth cohorts in months. Benefiting from the 2018 HLFS data, we generate a set of schooling outcomes covering years of schooling, and four indicator variables separately equal to one if an individual has at least a primary, junior high school, high school, and college degree.[14] However, HLFS data lacks information on pre-determined covariates which are important to validate our research design. To address this, we employ a nationally representative data set with information on the birthplace that was obviously determined prior to the 1997 Education Reform. The concerning data set is the Domestic Violence against Women which was conducted in 2008 and 2014.[15] The schooling outcomes we examine are presented in Panel A in Table 3.1 for those born 60 months before and after January 1987 as in most cases of our empirical analysis the sample falls into the concerning bandwidth of 60 months.

The primary data to quantify the causal impacts of education on stock market participation, and the variation in investment behavior is a comprehensive administrative data set spanning the universe of all individual stock market investors in Turkiye on December 31, 2021, and 2022. The data owner is the Borsa Istanbul Group, which is the legal entity to keep records of all investors with their portfolio details who have a transaction account. The relevant administrative data set is an unusually high-quality data set including portfolio choices, account balances, and demographic information such as gender, birth date, and birthplace. Such a feature of the administrative data set allows us to compute several outcome variables to examine how schooling impacts those portfolio outcomes. The most novel feature of the corresponding data is that it is not prone to any reporting bias and measurement errors since it covers the whole relevant population. In short, it is hard to argue that selection bias is an issue in our study.

A caveat of the administrative data is that it lacks information about those who do not have a transaction account. As we explore the causal impacts of schooling on stock market participation, direct stock, fund, bond, and risky asset ownership, we eliminate this problem by generating month-year birth cohort level outcomes by employing the data having information on the number of individuals in each month-year birth cohorts produced by TURKSTAT. The treatment, 1997 Education Reform, is a treatment at the month-year birth cohort level, allowing us to overcome the problem of people missing in the administrative data set. To generate the outcome variables at the cohort level, we first sum the number of

[14]The HLFS data has information on educational attainment but not the actual years of schooling. Therefore, we assigned 5 years of schooling for the primary school degree, 8 years of schooling for the junior high school degree, 11 years of schooling for the high school degree, 15 years of schooling for the college degree, and 17 years of schooling for the master's degree.

[15]For more information, see Erten and Keskin (2018).

Table 3.1: Descriptive Statistics for Those Born 60 Months Before and After January 1987

	(1) Control	(2) Treatment	(3) Difference (2)-(1)
Panel A: Schooling Outcomes			
Primary School Degree	0.93	0.9	-0.03
	(0.26)	(0.30)	(0.00)
Junior High School Degree	0.64	0.86	0.23
	(0.48)	(0.34)	(0.00)
High School Degree	0.49	0.59	0.1
	(0.50)	(0.49)	(0.00)
College Degree	0.27	0.35	0.08
	(0.45)	(0.48)	(0.00)
Years of Schooling	9.2	10.31	1.15
	(4.72)	(4.64)	(0.04)
Observations	32781	29445	62226
Panel B: Stock Market Participation Outcomes			
Stock Market Participation (%)	9.14	8.08	-1.05
	(0.51)	(0.44)	(0.10)
Direct Stock Ownership Rate (%)	5.25	5.04	-0.22
	(0.26)	(0.28)	(0.05)
Index Stock Ownership Rate (%)	3.51	3.16	-0.34
	(0.19)	(0.23)	(0.04)
Blue-Chip Stock Ownership Rate (%)	2.27	1.96	-0.32
	(0.15)	(0.17)	(0.03)
Risky Asset Ownership Rate (%)	6.18	5.82	-0.37
	(0.30)	(0.32)	(0.06)
Bond Ownership Rate (%)	0.10	0.07	-0.03
	(0.01)	(0.01)	(0.00)
Fund Ownership Rate (%)	3.06	2.40	-0.67
	(0.22)	(0.22)	(0.04)
Transaction Account Rate (%)	55.81	52.07	-3.73
	(1.36)	(2.91)	(0.41)

Continued on next page

Panel C: Portfolio Choices			
Share of Stocks	0.54	0.59	0.06
	(0.49)	(0.48)	(0.00)
Share of Risky Assets	0.65	0.7	0.05
	(0.47)	(0.45)	(0.00)
Share of Bonds	0.01	0.01	0.00
	(0.09)	(0.08)	(0.00)
Share of Blue-Chip Stocks	0.13	0.13	0.00
	(0.29)	(0.30)	(0.00)
Share of Index Stocks	0.24	0.26	0.02
	(0.39)	(0.40)	(0.00)
Share of Funds	0.33	0.30	-0.03
	(0.46)	(0.45)	(0.00)
Panel D: Stock Market Wealth, Risk, and Investment Strategies			
Risk Score (in logs)	3.68	3.69	0.01
	(0.84)	(0.82)	(0.00)
Single Stock Portfolio	0.18	0.22	0.05
	(0.38)	(0.42)	(0.00)
Risky Inertia	0.1	0.13	-0.04
	(0.77)	(0.64)	(0.00)
Stock Market Wealth (in logs)	6.69	6.65	-0.31
	(3.74)	(3.64)	(0.01)
Panel E: Portfolio Returns			
One Month Return (%)	-42.74	-43.88	-1.14
	(53.52)	(58.95)	(0.11)
Three Months Return (%)	-42.74	-43.95	-1.21
	(54.29)	(59.75)	(0.11)
Six Months Return (%)	-41.67	-43.15	-1.48
	(53.64)	(59.13)	(0.11)
Twelve Months Return (%)	104.13	112.83	8.7
	(120.45)	(126.04)	(0.24)
Observations	567758	506049	1073807

Notes: The table displays the mean, standard deviations in parenthesis, and the difference between the treatment and control groups. The treatment group covers those born after January 1987 while the control group is based on those born before January 1987. Panel A uses data from the 2018 Household Labor Force Survey by TURKSTAT and presents descriptive statistics for individuals. Panel B uses data at the cohort level. The remaining panels use the stock market administrative data on December 31, 2021, provided by Borsa Istanbul Group. The variable definitions are provided in the Data Appendix in Appendix C.

people having a positive balance in their transaction account and subsequently divide that number by the total number of individuals in each month-year birth cohort. Not only is this sort of strategy also helpful for us to compute the ratio in percentages of stock market participation at the birth cohort level, but also it is useful to calculate the percent ratio of direct stock ownership, fund, bond, and risky asset.

We first examine the causal impacts of the 1997 Education Reform on stock market participation, and asset choices such as direct stock ownership, bond, fund, risky asset ownership, and transaction account ownership. As we calculate the ratio of those having such assets at the month-year birth cohorts, we estimate the causal impacts of the 1997 Education Reform through regressions at the birth cohort level, so the unit of analysis is the birth cohorts. We also document the summary statistics of those concerning variables in Panel B in Table 3.1.

We also explore to what extent the education reform has explanatory power in explaining the variation in investors' investment behavior in stock market portfolios. Now, we have the opportunity to use the individual outcomes so the unit of analysis is individuals. In Panel C in Table 3.1, we report the summary statistics of portfolio choices with respect to some asset classes. To do this, we compute the share of wealth directly invested in stocks, risky assets, bonds, and funds. As the administrative individual portfolio data set has the information on which stocks are held by investors, we have a unique advantage to calculate the share of wealth invested in the stocks which are included in the BIST-100 benchmark index, market portfolio. This outcome allows us to measure how much an individual follows the market portfolio and refrains from risky and illiquid stocks. Similarly, we have the share of wealth invested in blue-chip stocks that are included in the BIST-30 index tracking the performance of the 30 most liquid and largest companies. Moreover, this outcome helps us unfold how education determines investment in the most liquid stocks. Overall, we can estimate the effects of the education reform on investment behavior through the outcomes related to the choices of assets in stock market portfolios at the extensive margin, and the outcomes related to the share of wealth invested in different assets at the intensive margin.

The unique features of the administrative data set allow us to assess whether those with more education have a higher level of risk appetite, the propensity to invest all of their money in a single stock, risky inertia, and portfolio size. The portfolio size and the number of stocks invested are already present in the administrative data set. Apart from this, we first use the logarithm of the risk score computed by Borsa Istanbul to track the risk appetite. Yet, the variable of risky inertia does not directly exist in the data set. To generate the concerning variable, we follow the measure suggested by Calvet et al. (2009b), which is a proxy for portfolio rebalancing. The descriptive statistics of those variables are reported in Panel D in Table 3.1.

Finally, we produce some outcome variables to investigate to what extent schooling affects portfolio performance i.e. the rate of return of portfolios. Since we do not have the realized returns of portfolios, we generate counterfactual portfolio returns by fixing the individual portfolios on December 31, 2021. Subsequently, we calculate the returns of portfolios after one, three, six months, and a year. This approach also enables us to examine how returns vary over time by education. Moreover, it provides us to understand whether those with more education outperform the portfolios of those with less education either in the short, medium, or long run. We display the summary statistics of return outcomes in the last panel in Table 3.1.

The administrative data set of individual stock market portfolios lacks information on the educational attainment of investors. The sharp increase in schooling for those born after January 1987 allows us directly to estimate the reduced-form impacts. Moreover, to show the simple correlation between years of schooling and stock market outcomes such as participation, direct stock, bond, and fund ownership, we employ Turkiye Household Budget Surveys in 2018 and 2019. A point to touch upon is that the HLFS data has missing information on the month of birth. If the relevant missing information is correlated with the education reform, this would be a threat to the validity of our estimates. To address this concern, we report the estimates of the 1997 Education Reform on the corresponding attrition. Point estimates are small and imprecise, implying that the attrition is orthogonal to the education reform.

3.3.2 Research Design

3.3.2.1 Identification

Considering the 1997 Education Reform and the school starting age of six, we highlight that those born after January 1987 had to complete junior high school or 8 years of schooling instead of 5 years prior to the reform, so a discontinuity emerges on January 1987 across birth cohorts. Using the cutoff of January 1987, we adopt a regression discontinuity design with a running variable in the month-year of birth to establish the causal link between schooling and stock market participation, and the variation in investors' portfolios. On the one hand, those born before January 1987 form the control group, on the other hand, those born after January 1987 form the treatment group in our quasi-experimental research design. The identifying assumption is that there are no systematical differences other than being affected by the 1997 Education Reform or not between two cohorts born one month apart. Given this assumption is satisfied, the estimation strategy based on an RD design produces a treatment assignment that is as good as random. In section 3.3.2.2, we perform a set of validity checks indicating that the identifying assumption holds.

In line with the existing research (Oreopoulos (2006), Erten and Keskin (2018), and Aydemir et al. (2022)), we exploit the discontinuity in the month-year of birth to estimate the causal effects of the education reform. The estimating equation is as follows:

$$\begin{aligned} y_i &= \alpha + \beta T_i + f(x_i) + \epsilon_i \\ \forall x_i &\in (c-h, c+h) \end{aligned} \tag{3.1}$$

where y_i is the specific outcome variable for either the month-year of birth cohort or the individual i. T_i is for the treatment status and β is the main parameter of interest, x_i is the running variable in months which is re-centered around zero by subtracting the month-year of birth from January 1987 that is the cutoff value determining the treatment status, and h is the bandwidth around the cutoff point of c. The RD design allows the slope to vary on each side of the cutoff. $f(x_i)$ is the control function with a continuous n-order polynomial function of the running variable on each side of the cutoff point c. In all estimations, we use the local linear approach proposed by Cattaneo et al. (2019) and also report the estimates with a quadratic control function in Appendix C. Since local linear RD estimates are often sensitive to the choice of bandwidth, it is crucial to choose it in a data-driven, and automatic way to avoid specification search and ad-hoc decisions. Thus, we use the optimal bandwidth algorithm proposed by Calonico et al. (2014) which considers the conventional mean squared error optimality based on the fundamental bias-variance trade-off. For any outcome variable, we estimate the specific bandwidths separately by using the optimal bandwidth algorithm. We also present the local linear RD estimates with a fixed bandwidth and the estimates with kink RD design to assess whether the 1997 Education Reform changes the slope around the cutoff point of January 1987 in Appendix C. In all estimations, results are not sensitive to different RD designs.

To avoid any possible complications arising from multiple hypotheses since we estimate the causal impacts of the education reform on multiple outcomes. and specification search or p-hacking, we compute summary indexes for each field of outcomes following the procedure proposed by Kling et al. (2007). To compute the concerning summary indexes, for each observation, we first subtract the mean of a certain outcome variable of the control group and then divide this by the control group standard deviation of that variable. So, for each outcome in a certain field of outcomes, we have a standardized value. Finally, we take the average of those standardized values in each field of outcomes. Furthermore, following Lee and Card (2008), we cluster standard errors at the month-year birth cohorts to avoid any specification error concerns as the treatment is assigned at the month-year birth level and the running variable is discrete. When the unit of analysis is the month-year birth cohorts, regressions include controls for the month of birth. When the unit of analysis is individual investors, we also include the controls for the province of birth registration fixed effects in addition to the controls for the month of birth. Full sample regressions also include a control variable for gender.

Most of the time, we report the reduced-form estimates of the 1997 Education Reform. The primary reason why we largely report the reduced-form estimates rather than also presenting instrumental variable estimates of an extra year of schooling is twofold. The first is data limitation as the stock market administrative data lacks investor's education information. Following Angrist and Krueger (1992), it is quite straightforward to calculate two-sample instrumental variable (TSIV) estimates, and we sometimes report them if the reduced-form estimates are significant. For all outcomes, TSIV estimates are available upon request. The last and most important reason is that the instrumental variable exploiting the 1997 Education Reform might not satisfy the exclusion restriction of the instrumental variable since the financial decisions are mainly determined at the household level. If the education reform also shifted the schooling of other household members, for instance, spousal education, then the 1997 Education Reform would clearly not only operate through the education of investors but also through spousal education. As a result, we mainly present the reduced-form effects of the 1997 Education Reform.

3.3.2.2 Validity Checks

See Section 1.4.2.2.

3.4 Results

3.4.1 Schooling Outcomes

See Section 1.5.1.

3.4.2 Education and Stock Market Outcomes

3.4.2.1 Stock Market Participation and Asset Ownership

We begin our analysis by reporting graphical evidence to display the statistical association between years of schooling and stock market outcomes. The administrative data set of transaction accounts has no information on the educational attainments of investors. Yet, the 2018-19 Household Budget Surveys have information on both educational attainment and stock market participation with a detail of which assets investors own. To present the correlation between years of schooling and overall stock market participation, we generate an indicator variable equal to one if a specific investor invests in any stock market asset such as stocks, bonds, and funds. Later, we plot the propensity to participate in the stock market and years of schooling in Figure 3.1. The concerning graphs exhibit a robust positive correlation in our study sample. The graph in Figure 3.1 covers those born 39 months before and after January 1987. The underlying reason for choosing a bandwidth of 39 months is that the optimal bandwidth usually falls into the corresponding bandwidth.

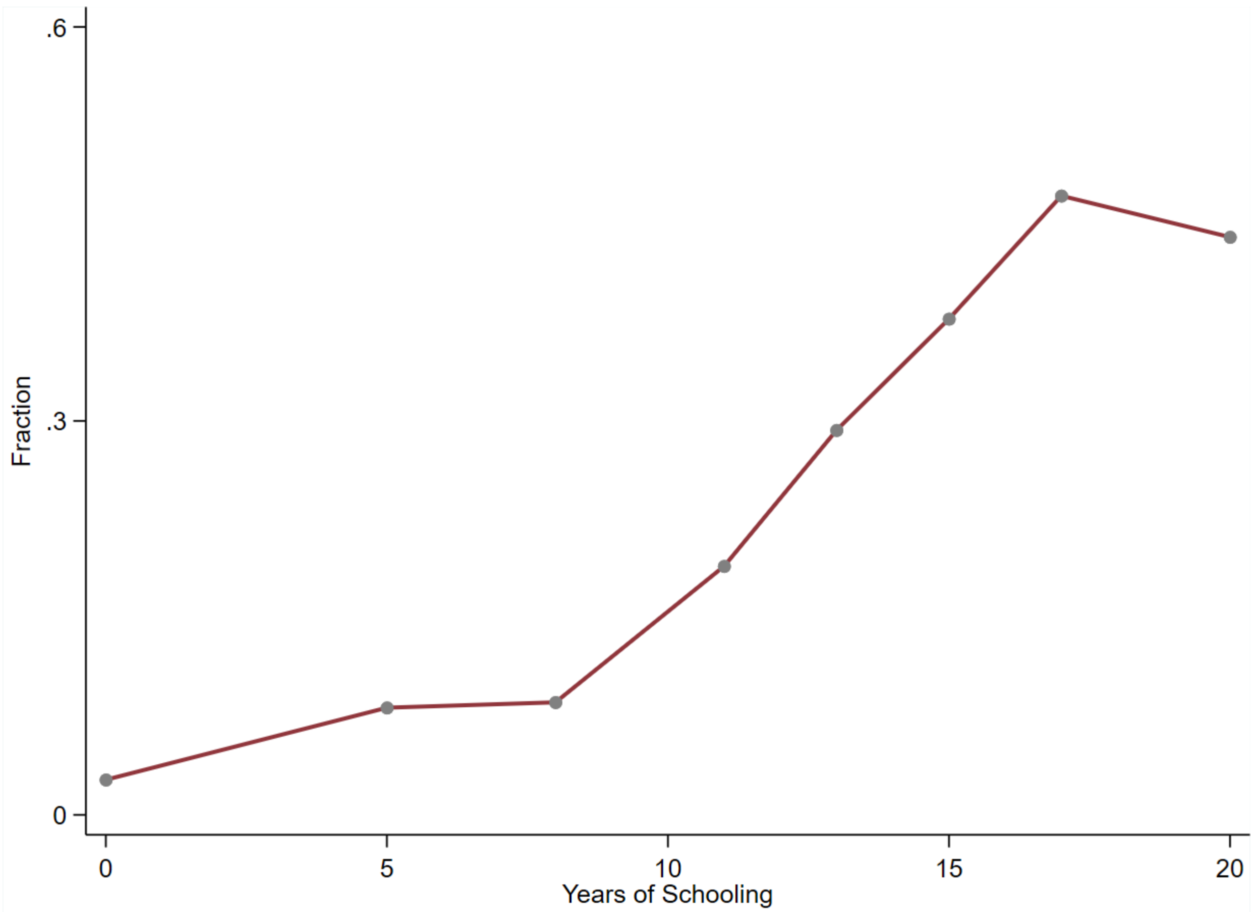

Figure 3.1: Schooling vs Participation in Stock Market

Notes: The graph uses data from Household Budget Surveys of 2018-19 by TURKSTAT. The sample includes those born 39 months before or after January 1987 since the optimal bandwidth in the RD design is 39 months for the outcome of participation in DC pension plans. The figures plot the fraction of those participating in the stock market against the binned years of schooling.

We subsequently report the OLS estimates in Table 3.2, revealing an extra year of schooling promotes stock market participation by 2.7% in the sample spanning those born 39 months before or after January 1987, respectively. Yet, as Figure 3.1 shows, the stock mar-

Table 3.2: Education vs Participation in Stock Market

	(1)	(2)
	Stock Market Participant	
Years of Schooling	0.027***	
	(0.001)	
Junior High School		0.032***
		(0.011)
High School		0.103***
		(0.013)
College		0.175***
		(0.016)
Weighted $\hat{\beta}$		0.012***
		(0.003)
Control Mean	0.19	0.19
Control SD	0.39	0.39
Observations	6905	6905

Notes: OLS estimates in all columns. All columns use data from Household Budget Surveys of 2018-19 by TURKSTAT. The sample includes those born 39 months before or after January 1987 since the optimal bandwidth in the RD design is 39 months for the outcome of stock market participation. The unit of analysis is individuals. In all columns, the outcome is a dummy variable equal to one if the individual participates in the stock market. In the first column, the explanatory variable is years of schooling. In the second column, the explanatory variables are indicator variables for junior high school, high school, and college degrees, respectively. The reference category is the primary school degree at most. Weighted $\hat{\beta}$ is the weighted average of the point estimates of the indicator variables for the indicator variables of degrees in the second column regarding the shift induced by the Education Reform in degrees. All regressions include controls for gender with a dummy variable of being female and year of survey fixed effects for each survey year. The control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome in the control group. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

ket participation rate is nonlinear with respect to the years of schooling. Moreover, the education reform shifts the whole distribution of schooling in Turkiye right as the point estimates in Table 1.2 and the cumulative distribution function of years of schooling in Figure 1.4 illustrate. The complete shift in educational attainment induced by the education reform makes it to compare the OLS coefficient stemming from regressing the propensity to participate in the stock market on years of schooling with the causal estimate difficult. Thus, to have a comparable benchmark OLS estimate relative to the causal estimate, we estimate an OLS coefficient adapting a different functional form mimicking the shift in the distribution of schooling induced by the education reform.

To do this, we first regress the indicator variables for each category of educational attainment, and the reference category is those who completed primary school at most. OLS estimates in column 2 show that the higher educational attainment the more participation in the stock market. In addition, we multiply each OLS coefficient with the local RD point estimate indicating the corresponding increase led by the education reform for each degree. Then, we sum up all of the concerning products, which allows us to have a benchmark OLS estimate to compare with the causal estimate. Similarly, assuming the covariances are zero across those products, we also calculate the standard error. Column 2 in Table 3.2 presents that the benchmark OLS coefficient is 1.2%, implying that the shift in the distribution of educational attainment induced by the education reform is positively correlated with the propensity to participate in the stock market. However, it is clear that those OLS estimates are biased, thus, we continue our analysis with causal estimates of the 1997 Education Reform. Figure 3.2 shows that the causal impacts are not as large as the OLS estimates. There seem to be small impacts but imprecise. Therefore, it is clear that there is room for more refined analysis so we proceed with the local linear RD estimates.

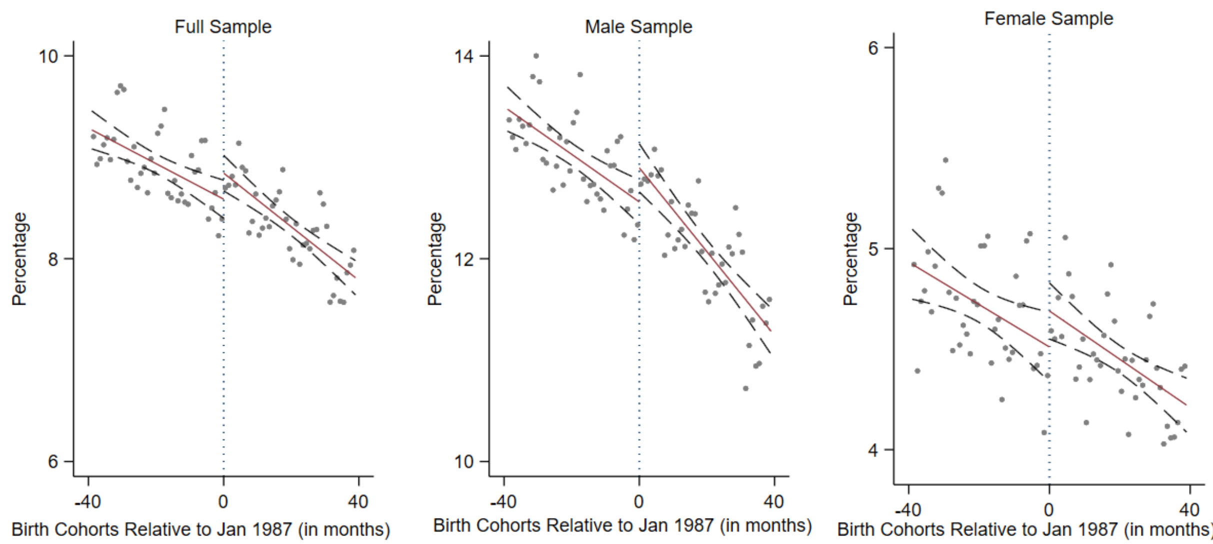

Figure 3.2: Education Reform vs Participation in Stock Market

Notes: All graphs use administrative data assembled by Borsa Istanbul Group and TURKSTAT in 2021. The figures plot the percentage of those participating in the stock market in monthly bins against the month-year of birth of those born before and after 39 months around the cutoff point, January 1987. The vertical line in each graph indicates the cutoff point. Black dashed lines in each graph indicate 95% confidence intervals around the mean of bins. Full, male, and female sample figures are reported, respectively.

We present the local linear RD estimates in Table 3.3. The point estimate in column 1 shows that the 1997 Education Reform increases the shares of stock market participants

Table 3.3: Education Reform vs Participation in Stock Market

	(1) Market Participant	(2) Direct Stock	(3) Index Stock	(4) Blue-Chip Stock	(5) Risky Asset	(6) Bond	(7) Fund	(8) Transaction Account	(9) Summary Index
					Full				
Education Reform	0.175***	0.131***	0.085***	0.060***	0.113***	0.004	0.040	0.445***	0.077***
	(0.054)	(0.036)	(0.029)	(0.022)	(0.041)	(0.003)	(0.027)	(0.109)	(0.024)
Control Mean	8.94	5.27	3.49	2.29	6.15	0.10	2.94	55.47	0.44
Control SD	0.35	0.24	0.19	0.15	0.30	0.01	0.15	1.25	0.18
Bandwidth	39.10	52.24	47.83	65.90	42.10	56.13	39.14	30.40	44.50
Observations	79	105	95	131	85	113	79	61	89
					Male				
Education Reform	0.233***	0.192***	0.129***	0.059	0.176***	0.013***	0.050	0.388***	0.090***
	(0.064)	(0.054)	(0.040)	(0.039)	(0.057)	(0.005)	(0.037)	(0.089)	(0.025)
Control Mean	13.07	8.02	5.35	3.46	9.22	0.14	4.18	72.12	0.45
Control SD	0.45	0.34	0.28	0.20	0.37	0.02	0.21	0.81	0.17
Bandwidth	41.62	54.85	55.85	46.59	47.95	71.28	40.38	27.60	43.59
Observations	83	109	111	93	95	143	81	55	87
					Female				
Education Reform	0.123**	0.064*	0.061**	0.040**	0.055	-0.013***	0.027	0.472***	0.069*
	(0.049)	(0.037)	(0.024)	(0.018)	(0.040)	(0.003)	(0.024)	(0.146)	(0.037)
Control Mean	4.80	2.43	1.62	0.99	3.02	0.06	1.69	38.19	0.13
Control SD	0.38	0.17	0.12	0.09	0.23	0.01	0.16	1.67	0.26
Bandwidth	43.66	41.94	45.65	48.85	42.58	42.53	42.34	30.73	42.95
Observations	87	83	91	97	85	85	85	61	85

Notes: Local linear RD estimates in all columns. All columns use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2021. The unit of analysis is the month-year birth cohorts in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those participating in the stock market in column 1, the percentage of those directly holding stocks in column 2, the percentage of those directly holding index stocks in column 3, the percentage of those directly holding blue-chip stocks in column 4, the percentage of those holding risky assets in column 5, the percentage of those directly holding bonds in column 6, the percentage of those directly holding funds in their stock market portfolios in column 7, the percentage of those having a transaction account in column 8. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

in percentages by 0.175 for the full sample, 0.233 for the male sample, and 0.123 for the female sample. All of those point estimates are statistically significant. However, comparing the local linear RD coefficients with the OLS coefficients reveal that the OLS estimates are upward biased and far from causal. Even if we consider the TSIV estimates by dividing the relevant point estimates by the reform-induced increase in years of schooling, the contribution of an extra year of schooling to the likelihood of stock market participation is smaller than half a percent. Equivalently, the causal impacts of the education reform are small on direct stock ownership either relative to the control group mean or in effect sizes. In columns 3 and 4, we document the causal effects of the education reform on the propensity to own index and blue-chip stocks. The impacts are small and precise consistent with the results on direct stock ownership. On the whole, despite the statistical significance point estimates show that education is a minor factor for stock ownership.

We now document the effects of the education reform on risky asset ownership, which denotes the ownership of any stock market assets except money market funds that are risk-free. In column 5, point estimates are precise for the full and male samples but imprecise for the female sample. Put another way, results demonstrate that those with more education are more likely to hold risky assets in their stock market portfolio but the magnitude is too small to say that they are substantial. We focus on the causal effects of the education reform on bond and fund ownership in columns 6 and 7, respectively. Consistently, the impacts are small and also indistinguishable from zero for the outcome of fund ownership. Moreover, in column 8 we report the causal impacts of the education reform on having a transaction account, a proxy for formal financial participation. The coefficients are close to half a percent, which implies that an extra year of schooling increases formal participation by around 1%. In the last column, we present the estimates for the summary index to address multiple hypothesis problems, showing that statistical power is not an issue in our estimations. In the last column, we present the estimates for the summary index to address multiple hypothesis problems, showing that statistical power is not an issue in our estimations. Rather, the issue is the too-small point estimates of the education reform. Overall, we find no significant evidence that education has an economically significant impact on stock market participation and the ownership of distinct financial assets despite of statistical significance of the concerning point estimates.

3.4.2.2 Education and the Variation in Portfolios

In this part, we proceed with the causal impacts of schooling on the variation in investors' behavior. We remind that the unit of analysis is individuals as we employ the individual level The first set of outcomes we are interested in is the share of wealth invested in distinct assets. We begin presenting the local linear RD estimates of the education reform on the share of wealth allocated to stocks in column 1 in Table 3.4. The coefficient of the education reform is neither significant nor large in all samples. Next, we report the estimates for the share of wealth invested in risky assets i.e. the risky share. Results show that education has no significant effect on the risky share in all samples. In the remaining columns, we document the causal impacts of education on the share of wealth invested in bonds, blue-chip stocks, index stocks, and funds. Yet, for all outcomes, results indicate no significant effects of the education reform. In the last column, we present the effects of the education reform on the summary index regarding the share of wealth invested in distinct assets. All in all, we fail to find any causal evidence that education is a factor driving the variation in investors' portfolios regarding the share of wealth invested in different assets.

We next present the local linear RD estimates of the 1997 Education Reform on investor's risk scores which shows the level of investor risk appetite. Column 1 in Table 3.5 documents the causal education effects on the logarithm of risk score, yet, in all samples the point estimates are imprecise and small. Moreover, we explore how single stock investment - investing all money into a single stock- varies by education. Point estimates in column 2 are small despite the statistical significance.

After quantifying the causal relationship between education and risk outcomes such as risk appetite and single-stock investment, we next present evidence of the relationship between schooling and portfolio inertia. To examine the corresponding causal association, we use the outcome variable, risky share inertia -a proxy for portfolio rebalancing- suggested by Calvet et al. (2009b). Nevertheless, point estimates indicate that there is no significant impact of the education reform on risky share inertia. We conclude that portfolio rebalancing does not change by education. Last but not least, we display the causal effects of the education reform on portfolio size or stock market wealth. Point estimates are substantially large but imprecise, which indicates an imprecise negative causal relationship between education and stock market wealth in all samples. The last column presents the point estimate for the summary index. In sum, it is hard to argue that education is not an explanatory factor in the variation of investment behavior in terms of risk appetite, rebalancing, single-stock investment, and stock market wealth.

Up to now, we investigate the causal impacts of education on various outcomes related to either stock market participation or the variation in investor's investment behavior. In

Table 3.4: Education Reform vs Variation in Portfolio Choices

	(1) Stock	(2) Risky Asset	(3) Bond	(4) Blue-Chip Stock	(5) Index Stock	(6) Fund	(7) Summary Index
				Full			
Education Reform	0.001	-0.000	-0.000	0.002	0.001	-0.001	0.001
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)
Control Mean	0.55	0.65	0.01	0.13	0.25	0.32	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	38.42	48.95	42.07	44.36	38.15	37.34	46.34
Observations	685035	864439	755220	791630	685035	668661	825909
				Male			
Education Reform	0.000	0.001	0.001***	0.002	0.002	-0.000	0.004**
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.003)	(0.002)
Control Mean	0.57	0.67	0.01	0.14	0.25	0.31	0.10
Control SD	0.49	0.46	0.08	0.30	0.39	0.45	0.46
Bandwidth	39.14	53.06	53.92	44.74	40.75	39.18	46.78
Observations	517965	694984	694984	582992	531309	517965	608200
				Female			
Education Reform	0.002	-0.002	-0.003***	0.000	-0.001	-0.005	-0.006*
	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
Control Mean	0.48	0.60	0.01	0.11	0.22	0.37	0.04
Control SD	0.49	0.48	0.10	0.28	0.38	0.47	0.46
Bandwidth	42.54	51.50	35.84	45.17	49.94	40.90	51.33
Observations	198736	240273	164691	213140	231400	188984	240273

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the share of wealth invested in stocks in column 1, in risky assets in column 2, in bonds in column 3, in blue-chip stocks in column 4, in index stocks in column 5, and in funds in column 6. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

addition, we document the causal impacts of education on investment performance by using monthly returns from one month to twelve months. As our results so far show no significant education impacts on stock market participation and investment strategies, it is quite plausible not to expect any variation in portfolio returns by education. As we have no information

Table 3.5: Education Reform vs Causal Channels

	(1) Risk Score	(2) Single Stock	(3) Risky Inertia	(4) Wealth (in logs)	(5) Summary Index
			Full		
Education Reform	-0.003	0.003*	-0.002	-0.017	-0.001
	(0.005)	(0.001)	(0.003)	(0.018)	(0.002)
Control Mean	3.68	0.19	0.14	6.94	-0.02
Control SD	0.84	0.39	0.79	3.72	0.53
Bandwidth	46.18	31.64	69.42	45.89	41.43
Observations	403890	561128	764908	808937	737564
			Male		
Education Reform	-0.005	0.004*	-0.002	-0.014	0.000
	(0.006)	(0.002)	(0.004)	(0.020)	(0.002)
Control Mean	3.70	0.19	0.13	7.00	-0.02
Control SD	0.82	0.39	0.78	3.79	0.53
Bandwidth	45.90	35.09	66.02	45.87	66.36
Observations	319435	463317	555448	595797	862541
			Female		
Education Reform	0.007	-0.001	-0.002	-0.027	-0.005
	(0.008)	(0.003)	(0.007)	(0.027)	(0.003)
Control Mean	3.57	0.18	0.15	6.77	-0.05
Control SD	0.90	0.38	0.78	3.52	0.54
Bandwidth	60.74	48.02	70.14	46.56	49.03
Observations	100792	227786	187330	217709	231400

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the logarithm of risk scores in column 1, an indicator variable equal to one if the investor invests all money into a single stock in column 2, the risk share inertia in column 3, and the logarithm of portfolio size in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

on the realized returns, we fix the existing portfolios on December 31, 2021, to calculate counterfactual returns. Later, we compute the returns of portfolios after one, three, six, and twelve months.

Table 3.6: Education Reform vs Portfolio Returns

	(1) One Month	(2) Three Months	(3) Six Months	(4) Twelve Months	(5) Summary Index
			Full		
Education Reform	-0.260	-0.451***	-0.353**	0.525	-0.003
	(0.173)	(0.160)	(0.148)	(0.428)	(0.003)
Control Mean	-42.74	-42.94	-41.98	105.40	-0.03
Control SD	53.52	54.70	54.31	121.31	0.93
Bandwidth	60.22	48.76	43.63	44.72	71.31
Observations	1058007	851769	762197	791630	1260754
			Male		
Education Reform	-0.361	-0.611***	-0.473**	0.779	-0.008**
	(0.224)	(0.206)	(0.203)	(0.525)	(0.004)
Control Mean	-42.17	-42.29	-41.49	107.25	-0.02
Control SD	56.00	57.30	56.25	120.98	0.98
Bandwidth	47.63	40.97	41.67	47.73	51.40
Observations	610844	524720	537070	618526	670196
			Female		
Education Reform	-0.129	-0.090	0.053	-0.138	-0.001
	(0.306)	(0.309)	(0.328)	(0.989)	(0.006)
Control Mean	-44.88	-44.64	-42.76	99.71	-0.07
Control SD	47.77	47.96	47.81	121.74	0.84
Bandwidth	58.20	59.50	68.88	44.82	54.93
Observations	268108	271848	314787	208638	254852

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the one-month return in column 1, the three-month return in column 2, the six-month return in column 3, and the annual return in percentages in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.6 reports the local linear RD estimates of the Education Reform on portfolio returns. The first column displays the coefficient for one-month returns. The education reform leads to a 0.26% lower one-month return for the full sample, but the coefficient is imprecise at the conventional levels of statistical significance. Furthermore, males with more education underperform males with less education, yet, the point estimate is marginally insignificant at the 10% significance level. The corresponding coefficient is small and imprecise in the female sample. Column 2 displays the estimates for the outcome of three-month returns. Heterogeneity by gender amplifies relative to the estimates for one-month returns. While males with more education underperform males with less education, no significant point estimates appear in the female sample. Similarly, it appears that the education reform leads to a worse portfolio performance for males and null effects for females in the outcome of six-month returns. However, males with more education outperform males with less education in terms of annual returns which we might refer to as long-run returns. In contrast, the 1997 Education Reform has no significant effect on annual portfolio returns in the female sample.

Independent of being statistically significant or not, those point estimates are too small relative to the control group mean. Even if we take into account TSIV estimates, despite being indistinguishable from zero, an extra year of schooling causally increases annual portfolio returns by at most 1% and 2% for the full and the male sample, respectively. Those TSIV estimates are too small relative to the control mean. Put another way, TSIV estimates turn to an increase in annual portfolio returns by less than 0.02%, which is simply a negligible number. Apart from this, it is much smaller in the female sample and imprecisely estimated. Overall, we conclude that education has no significant effect on portfolio returns in various periods.

3.5 Robustness Checks

We perform a robustness check to address the participation bias documented by Lee (2009). Put another way, we are able to observe the portfolio outcomes only for those participating in the stock market. Consequently, the participation bias is a threat to validity. By performing a robustness check, we also figure out to what extent participation in the stock market drives our results. Yet, considering the point estimates in column 1 in Table 3.3 reveals that the causal impacts of the 1997 Education Reform are statistically significant at conventional significance levels but small in magnitude. Even if we compare those coefficients with the control group mean, they are nonetheless small. Moreover, TSIV estimates for the full sample, the male sample, and the female sample indicate that an extra year of schooling fosters stock market participation by approximately 0.35%, 0.466%, and 0.246%, respectively. We note that all of those coefficients are less than 1%, which is small, at best implying a limited impact. As a result, it is unlikely to argue that participation bias confounds our estimates significantly.

We also adopt the strategy suggested by Duflo (2001). By doing so, we assess the stability of coefficients through the inclusion of the quadratic polynomials of participation rates in each month-year birth cohort as controls. A remarkable feature of the administrative data we utilize is that we are able to compute the actual participation rates in each month-year birth cohort. If participation bias drives our results, then the estimates would change significantly conditional on the ratio of stock market participants in each birth cohort. To do that, we include the participation rates in a quadratic polynomial form for each regression and subsequently evaluate whether the point estimates are robust. Results show that the point estimates are not sensitive to the additional controls of participation rates across month-year birth cohorts. For brevity, we report the relevant tables in Appendix C. Altogether, the participation bias is unlikely to drive our estimates.

3.6 Conclusion

We study the causal impacts of general education on stock market participation, the variation in investment behavior, and portfolio performance. To isolate the causal impacts of general education, we leverage the exogenous variation brought about by the Education Reform in Turkiye in 1997. The education reform extended compulsory schooling from five to eight years for those born after January 1987. With this in mind, in our research design, we treat those born before January 1987 as the control group and those born after January 1987 as the treatment group. Adopting a regression discontinuity design by exploiting the discontinuity in the birth date in months, we show that the reform-induced increase in schooling is almost half a year. The education reform is a large intervention as it shifts the whole schooling distribution right. In addition, it went beyond the policy target by significantly promoting higher degrees such as high school and college completion. A striking heterogeneity by gender also emerges favoring female education as the point estimates for the female sample are much higher.

Using an administrative data set covering the universe of all individuals with a transaction account in Turkiye in December 2021 and 2022, we find no significant evidence that general education is a factor determining stock market participation despite the positive robust correlation between education and stock market participation. Thus, we conclude that the corresponding association is far from causal. Then, we assess the role of general education in the ownership of stock market instruments such as stocks, funds, bonds, risky assets, and stocks that are included in either the market portfolio (index stocks) or the index of highly liquid and large market value companies (blue-chip stocks), and the propensity to have a transaction account. Yet, for all of those outcomes, the point estimates indicate no significant effects of general education. The small effects of the education reform on the propensity to have a transaction account particularly reveal that even for formal financial participation education is not ostensibly a significant input.

We also document the causal effects of general education on the share of wealth invested in risky assets, stocks, bonds, funds, and index or blue-chip stocks. The point estimates for those outcomes are imprecise and small. After documenting that, we proceed with the outcomes of risk appetite, extreme investment strategies such as investing all wealth into a single stock, risky share inertia as a proxy for portfolio rebalancing, and stock market wealth. Our findings imply that all corresponding outcomes do not vary with education except the stock market wealth. However, the coefficients are imprecise. Finally, we demonstrate the causal effects of education on portfolio returns over various sets of time spans. Either in the short-run -one-month or three-month returns- or in the long run -six-month or annual returns- the general education has no significant effect on portfolio performance. Overall, general education does not seem to be a causal determinant of stock market participation and the variation in investment behavior.

Prior research, on the one hand, documents that education and desirable financial behavior are positively related. On the other hand, existing studies with reliable causal identification strategies have reported mixed evidence on the impacts of education on stock market participation and investment behavior in stock market portfolios. Even though we figure out that general education and stock market participation, as well as stock, bond, and fund ownership, are positively associated, we fail to find any causal evidence. Such findings indicate that the corresponding correlation is an artifact of confounding variables perhaps family background or genetics. To sum up, it is clear that general education is at best a minor factor driving financial behavior. A certain policy implication of our study is that adding some components of financial education to the existing curriculum is likely to increase the efficiency of general education in terms of financial behavior.

APPENDIX A

Appendix to Chapter 1

A.1 Figures

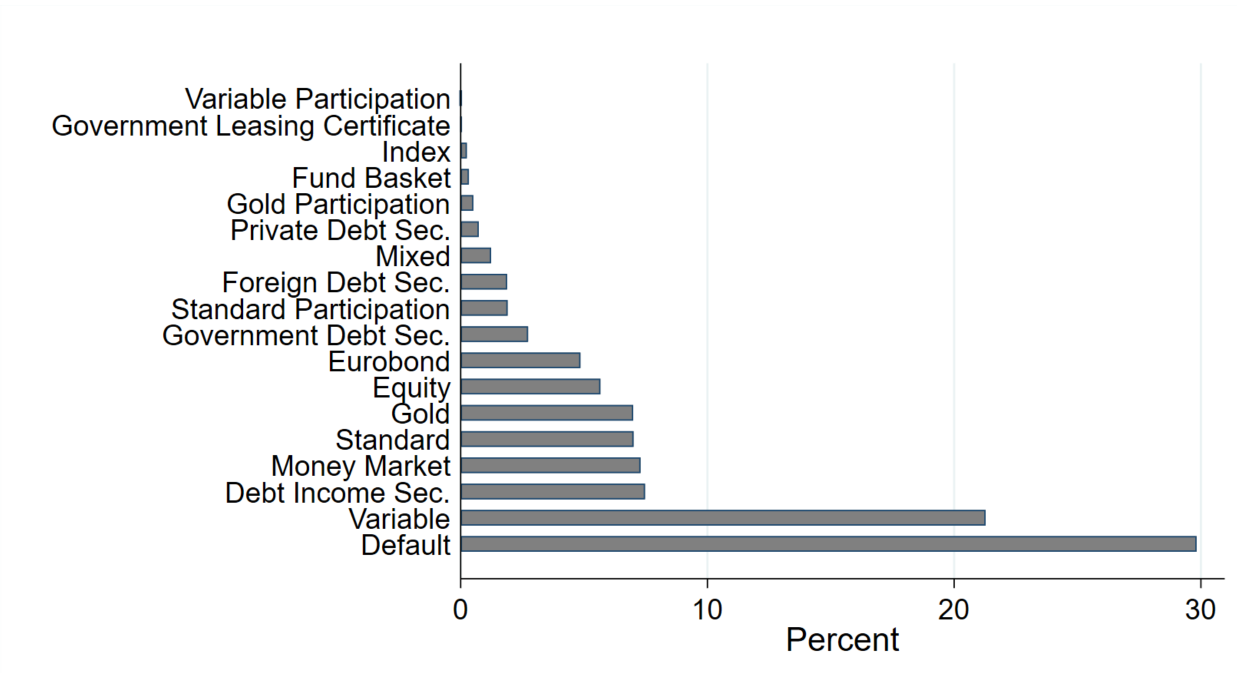

Figure A.1: Investment by Fund Types

Notes: The graph uses administrative data assembled by Borsa Istanbul Group. The figure plots the share of the total money invested in each fund type in percentage terms in 2019.

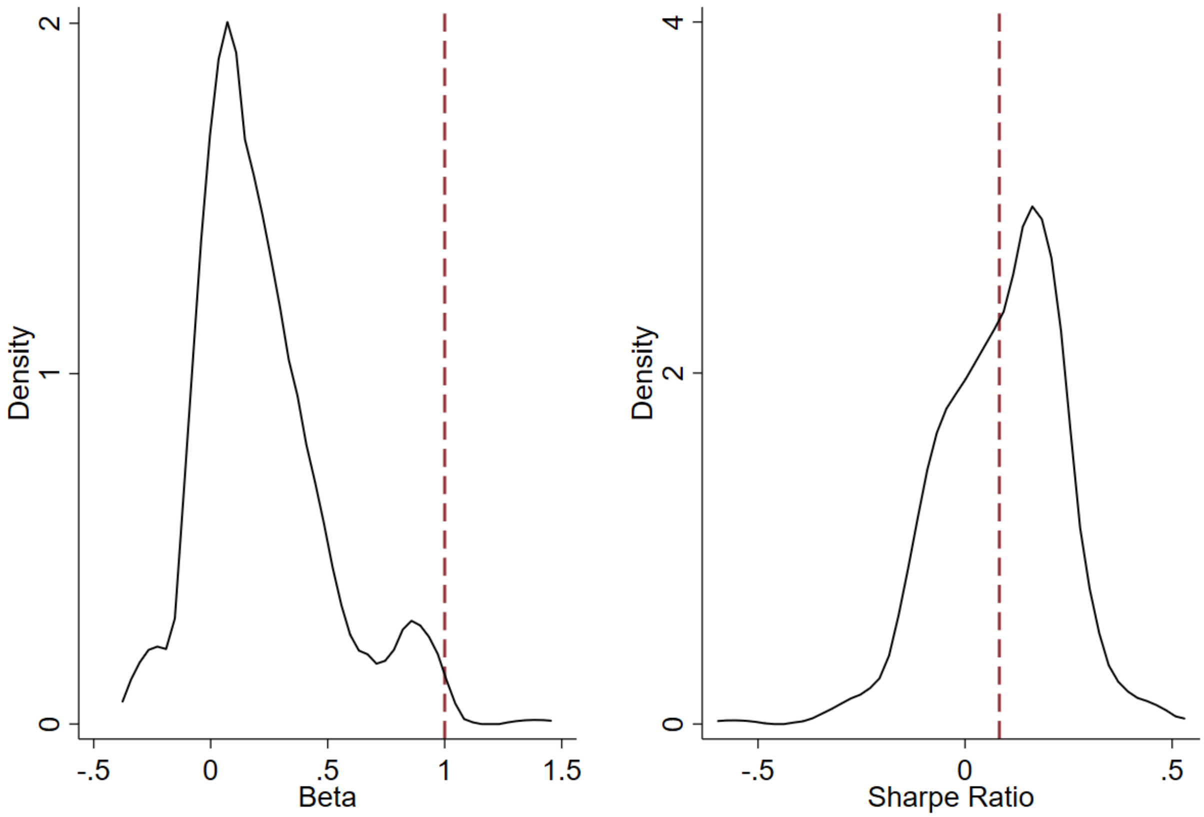

Figure A.2: Beta Coefficient and Sharpe Ratio of Funds

Notes: All graphs use administrative data assembled by Borsa Istanbul Group. The figures plot the beta coefficient and the Sharpe ratio distributions of all pension funds in 2019. The maroon color dashed line on the left panel represents the beta coefficient and on the right the Sharpe Ratio of the domestic benchmark index. To calculate the beta coefficients and Sharpe Ratios of pension funds, I use the monthly returns of pension funds. For the calculation of beta coefficients, I exploit the domestic benchmark equity index, BIST-100. For the risk-free rate in the calculations, I use the monthly return of BIST KYD Indices relying on 91 days to maturity T-bills of Turkiye. For more details, see the data appendix.

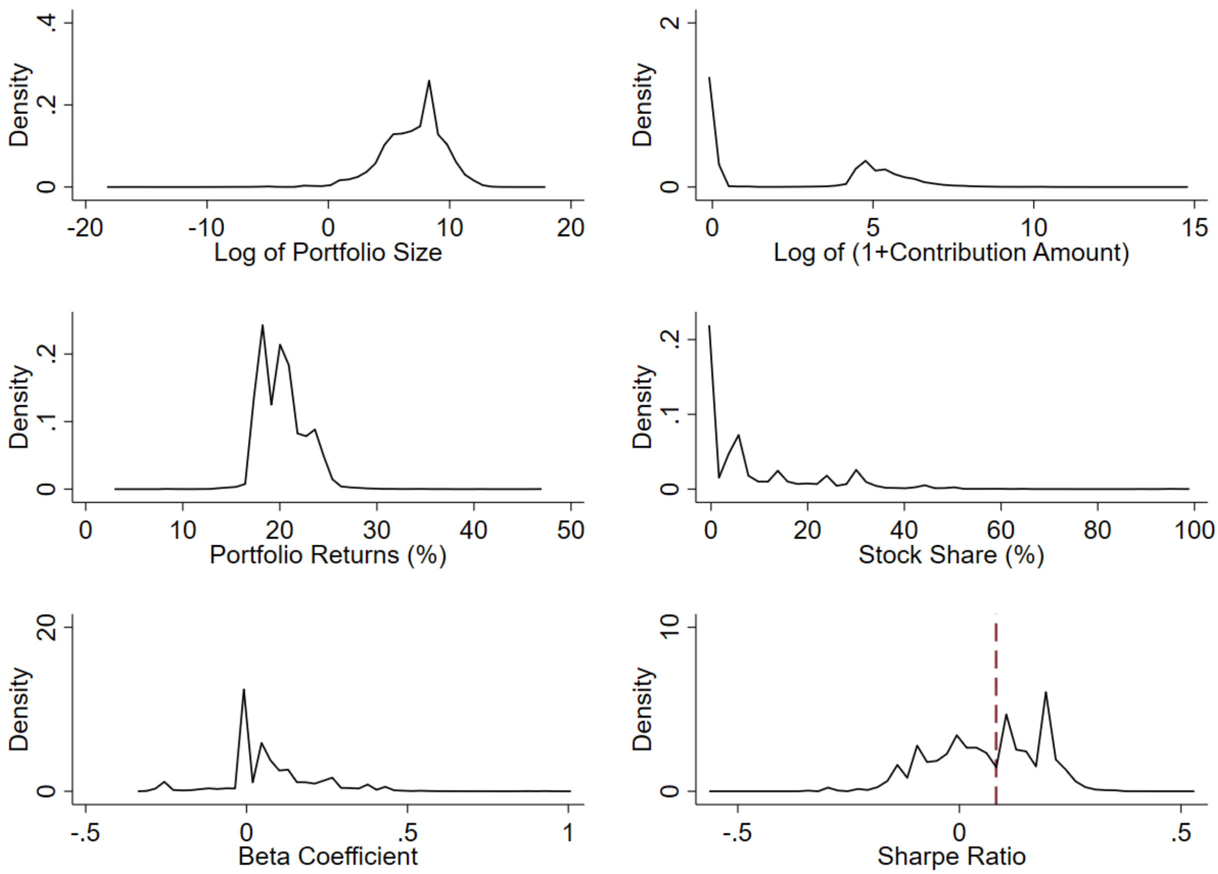

Figure A.3: Portfolio Characteristics

Notes: All graphs use administrative data assembled by Borsa Istanbul Group. The figures plot the distributions of various portfolio characteristics in 2019. The maroon color dashed line in the last graph represents the Sharpe Ratio of the domestic benchmark index. For the variable definitions, see the data appendix.

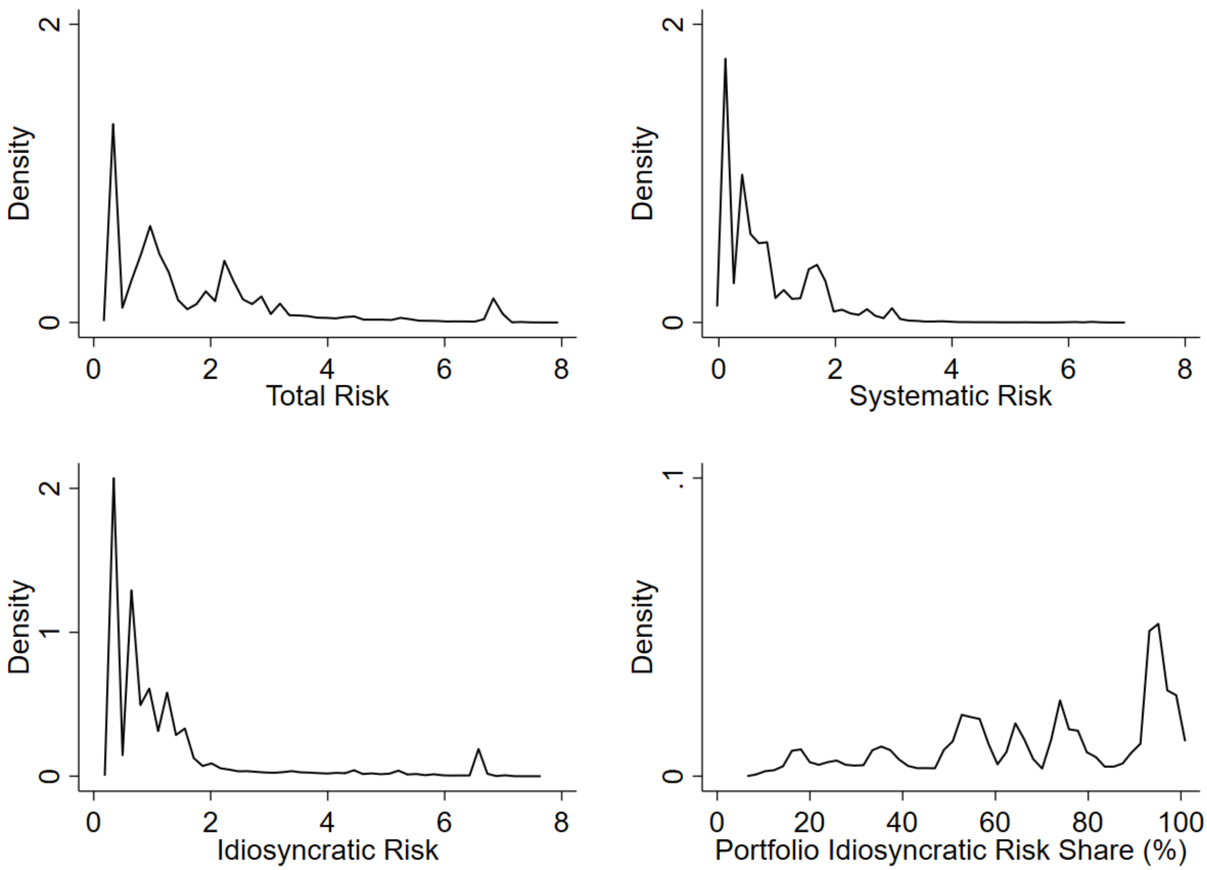

Figure A.4: Portfolio Risk Measures

Notes: All graphs use administrative data assembled by Borsa Istanbul Group. The figures plot the distributions of portfolio risk measures in 2019. For the variable definitions, see the data appendix.

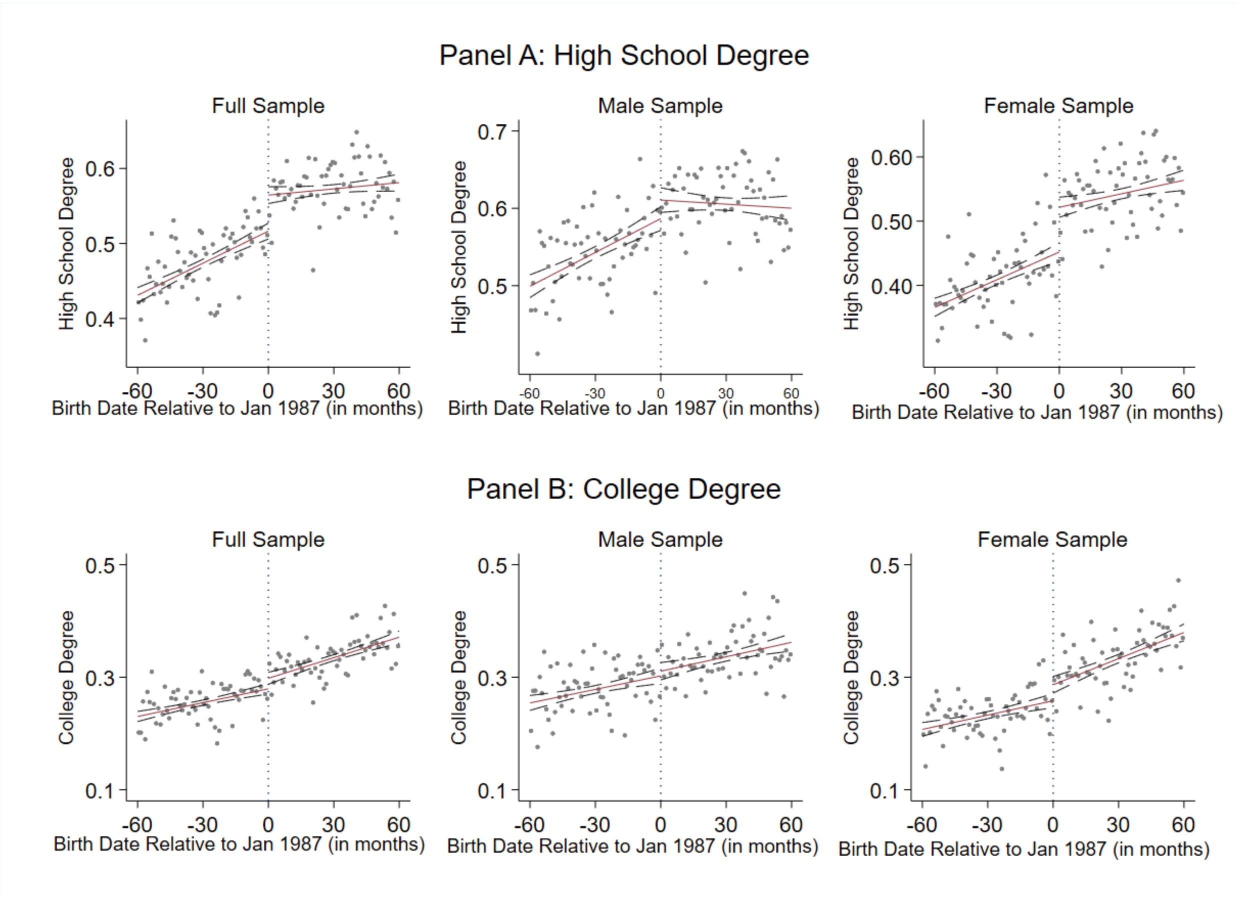

Figure A.5: Education Reform vs High School, and College Degree

Notes: All graphs use data from the 2018 Household Labor Force Survey assembled by TURKSTAT. The figures in Panel A the propensity to hold at least a high school degree and the figures in Panel B plot the propensity to hold at least a college degree in monthly bins against the month-year of birth of those born before and after 60 months around the cutoff point, January 1987. The vertical line in each graph indicates the cutoff point. Black dashed lines in each graph indicate 95% confidence intervals around the mean of bins. Full, male, and female sample figures are reported, respectively.

A.2 Regression Evidence for Validity Checks

Table A.1: Education Reform vs Population Shares of Month-Year Birth Cohorts

	(1) Linear RD h bandwidth	(2) Quadratic RD h bandwidth	(3) Linear RD h/2 bandwidth	(4) Linear RD 2h bandwidth
			Full	
Education Reform	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Control Mean	0.13	0.13	0.12	0.13
Control SD	0.02	0.02	0.02	0.02
Bandwidth	33.00	55.58	18.86	65.99
Observations	65	111	37	131
			Male	
Education Reform	-0.001	-0.001	-0.001*	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Control Mean	0.13	0.13	0.13	0.13
Control SD	0.02	0.02	0.02	0.02
Bandwidth	33.12	55.53	19.61	66.24
Observations	67	111	39	133
			Female	
Education Reform	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Control Mean	0.12	0.12	0.12	0.13
Control SD	0.02	0.02	0.02	0.02
Bandwidth	32.92	55.82	18.09	65.83
Observations	65	111	37	131

Notes: Local RD estimates in all columns. All columns use data covering the Turkish Population Share in each month of birth cohorts in 2019 assembled by TURKSTAT. The unit of analysis is the month-year birth cohorts. The main explanatory variable namely Education Reform is a dummy variable equal to one for the cohorts born after January 1, 1987. The outcome is the population share in each month of birth cohorts. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Column 1 reports RD estimates with the optimal bandwidth and a linear polynomial function on each side of the cutoff value. Column 2 reports RD estimates with a quadratic polynomial function on each side of the cutoff value. Columns 3 and 4 report RD estimates with a linear polynomial function on each side of the cutoff value in half and twice the optimal bandwidth estimated in column 1. All regressions include controls for the month of birth fixed effects with dummy variables for each month. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.2: Balanced Covariates

	Mother Tongue	Childhood Region					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Non-Turkish	Rural	West	East	North	South	Central
Education Reform	0.024	0.014	-0.037	0.075**	-0.040	0.009	0.000
	(0.016)	(0.030)	(0.034)	(0.036)	(0.027)	(0.024)	(0.028)
Control Mean	0.07	0.38	0.20	0.31	0.14	0.11	0.23
Control SD	0.25	0.48	0.40	0.46	0.35	0.31	0.42
Bandwidth	48.74	41.42	46.40	37.36	34.55	57.54	48.61
Observations	3534	3030	3370	2733	2508	4197	3510

Notes: Local linear RD estimates in all columns. All columns use data from 2008 and National Surveys on Domestic Violence against Women in Turkiye by TURKSTAT. The unit of analysis is females. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is a dummy variable equal to one if the specific female individual has a mother tongue other than Turkish. In column 2, the outcome is a dummy variable equal to one if the specific female individual grew up in rural Turkiye. In column 3, the outcome is a dummy variable equal to one if the specific female individual grew up in Western Turkiye. In column 4, the outcome is a dummy variable equal to one if the specific female individual grew up in Eastern Turkiye. In column 5, the outcome is a dummy variable equal to one if the specific female individual grew up in Northern Turkiye. In column 6, the outcome is a dummy variable equal to one if the specific female individual grew up in Southern Turkiye. In column 7, the outcome is a dummy variable equal to one if the specific female individual grew up in Central Turkiye. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects with dummy variables for each month and survey year fixed effects with dummy variables for each survey year. All regressions include controls for the month of birth and survey year dummies. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.3: Education Reform, Placebo Cutoffs vs Schooling Outcomes

	Years of Schooling			Junior High School Degree		
	(1) January 1987	(2) January 1980	(3) January 1994	(4) January 1987	(5) January 1980	(6) January 1994
				Full		
Treatment	0.451***	-0.047	-0.005	0.055***	0.003	-0.005
	(0.105)	(0.112)	(0.100)	(0.010)	(0.012)	(0.007)
Control Mean	9.25	8.27	10.48	0.68	0.52	0.90
Control SD	4.73	4.52	4.52	0.47	0.50	0.31
Bandwidth	57.21	49.38	36.52	30.82	46.55	49.70
Observations	59347	57174	34591	31744	53595	46336
				Male		
Treatment	0.381***	0.040	0.005	0.042***	-0.002	-0.001
	(0.148)	(0.121)	(0.129)	(0.015)	(0.016)	(0.007)
Control Mean	9.96	9.20	10.63	0.76	0.62	0.94
Control SD	4.33	4.22	3.97	0.43	0.49	0.24
Bandwidth	51.41	40.19	33.01	35.32	42.99	44.97
Observations	25664	22928	15247	17678	23792	19878
				Female		
Treatment	0.448**	0.029	0.032	0.081***	0.011	-0.010
	(0.210)	(0.144)	(0.153)	(0.021)	(0.015)	(0.011)
Control Mean	8.56	7.32	10.30	0.57	0.42	0.85
Control SD	5.00	4.61	5.03	0.50	0.49	0.36
Bandwidth	49.82	62.74	40.88	40.32	50.09	59.31
Observations	26722	36464	19778	21889	30009	28839

Notes: Local linear RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Treatment is a dummy variable equal to one for those born after January 1, 1987, in columns 1 and 4, or for those born after January 1, 1980, in columns 2 and 5, or for those born after January 1, 1994, in columns 3 and 6. In columns 1-3, the outcome is years of schooling. In columns 4-6, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects with dummy variables for each month and the regressions in the full sample additionally include controls for gender with a dummy variable of being female. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.3 Local Randomization RD Estimates with Schooling Outcomes

Table A.4: Local Randomization RD Estimates of Schooling Outcomes

Outcome	Obs Left	Obs Right	Estimate	Finite Sample p-value	Large Sample p-value	Control Mean
			Full			
Years of Schooling	6594	6660	0.412	0.000	0.000	9.367
Primary School Degree	6594	6660	-0.021	0.000	0.000	0.909
Junior High School Degree	6594	6660	0.097	0.000	0.000	0.710
High School Degree	6594	6660	0.041	0.000	0.000	0.508
College Degree	6594	6660	0.025	0.000	0.001	0.277
Summary Index	6594	6660	0.079	0.000	0.000	0.473
			Male			
Years of Schooling	3202	3178	0.341	0.000	0.001	10.116
Primary School Degree	3202	3178	-0.007	0.238	0.190	0.952
Junior High School Degree	3202	3178	0.079	0.000	0.000	0.807
High School Degree	3202	3178	0.029	0.022	0.021	0.568
College Degree	3202	3178	0.014	0.206	0.207	0.294
Summary Index	3202	3178	0.064	0.000	0.000	0.604
			Female			
Years of Schooling	3392	3482	0.500	0.000	0.000	8.660
Primary School Degree	3392	3482	-0.033	0.000	0.000	0.869
Junior High School Degree	3392	3482	0.117	0.000	0.000	0.618
High School Degree	3392	3482	0.054	0.000	0.000	0.451
College Degree	3392	3482	0.035	0.000	0.001	0.261
Summary Index	3392	3482	0.097	0.000	0.000	0.350

Notes: Local Randomization RD estimates. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The column of Outcome reports the corresponding schooling outcome. Column of Bandwidth reports the closest window length of one around the cutoff value of 1987. Columns of Obs Left and Obs Right report the number of observations in the right and left window of the cutoff value, 1987. Column of Estimate reports the local randomization RD estimates, that is the impact of Education Reform. Column of Finite Sample p-value displays the corresponding p-value for the relevant local randomization RD estimate in finite samples whereas the Column of Large Sample p-value documents the corresponding p-value for the relevant local randomization RD estimate asymptotically. Column of Control mean displays the mean of the corresponding outcome of those born before January 1, 1987. Full, male, and female sample estimates are reported, respectively.

A.4 Attrition

Table A.5: Local Randomization RD Estimates of Attrition

Bandwidth	Obs Left	Obs Right	Estimate	Finite Sample p-value	Large Sample p-value	Control Mean
			Full			
1	6594	6660	0.002	0.586	0.522	0.043
2	13374	13071	-0.000	0.920	0.931	0.043
3	20374	18944	0.000	0.930	0.976	0.044
4	27274	25103	0.001	0.638	0.606	0.045
5	34337	30904	0.002	0.246	0.250	0.045
6	42279	36871	0.004	0.004	0.005	0.045
			Male			
1	3202	3178	-0.000	1.000	0.956	0.045
2	6491	6276	-0.001	0.834	0.782	0.045
3	9835	9143	0.000	1.000	0.979	0.046
4	13185	12091	0.004	0.166	0.180	0.046
5	16616	14873	0.006	0.016	0.022	0.045
6	20501	17692	0.007	0.002	0.001	0.045
			Female			
1	3392	3482	0.005	0.394	0.335	0.040
2	6883	6795	0.001	0.928	0.874	0.042
3	10539	9801	0.000	1.000	0.987	0.043
4	14089	13012	-0.002	0.516	0.537	0.045
5	17721	16031	-0.001	0.522	0.510	0.046
6	21778	19179	0.001	0.504	0.481	0.046

Notes: Local Randomization RD estimates. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The outcome is a dummy variable equal to one if an individual lacks the month of birth information. Column of Bandwidth reports the window length around the cutoff value of 1987. Columns of Obs Left and Obs Right report the number of observations in the right and left window of the cutoff value, 1987. Column of Estimate reports the local randomization RD estimates, that is the impact of Education Reform. The column of Finite Sample p-value displays the corresponding p-value for the relevant local randomization RD estimate in finite samples whereas the column of Large Sample p-value documents the corresponding p-value for the relevant local randomization RD estimate asymptotically. Column of Control mean displays the mean of the corresponding outcome of those born before January 1, 1987. Full, male, and female sample estimates are reported, respectively.

A.5 Robustness Checks on Participation Bias

Table A.6: Education Reform vs Wealth Accumulation in DC Pension Plans

	(1) Full	(2) Male	(3) Female
Education Reform	0.013	0.005	0.028***
	(0.008)	(0.010)	(0.010)
Control Mean	7.12	6.88	7.50
Control SD	2.33	2.33	2.26
Bandwidth	37.76	36.08	34.99
Observations	1900641	1146511	666926

Notes: Local linear RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The dependent variable is the log of the value of pension portfolios in Turkish Lira in all columns. Columns 1,2, and 3 present the estimates for full, male, and female populations, respectively. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for quadratic polynomials of pension participation rate, the month of birth fixed effects, and birth registration certificate region fixed effects with dummy variables for each month and each relevant regions, and the regressions in the full sample additionally include controls for gender with a dummy variable of being female. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.7: Education Reform vs Equity Exposure

	(1) Equity Fund	(2) Stock	(3) Risky Assets	(4) Summary Index
		Full		
Education Reform	0.050***	0.053*	0.098	0.005***
	(0.019)	(0.028)	(0.060)	(0.002)
Control Mean	2.20	10.08	77.95	-0.12
Control SD	8.27	12.39	22.52	0.75
Bandwidth	49.74	33.82	32.95	39.68
Observations	2506174	1690646	1640290	1997398
		Male		
Education Reform	0.064***	0.123***	0.083	0.007***
	(0.020)	(0.031)	(0.061)	(0.002)
Control Mean	2.10	9.86	76.80	-0.15
Control SD	8.10	12.28	23.41	0.75
Bandwidth	46.87	40.68	31.67	38.49
Observations	1451167	1265014	978094	1202282
		Female		
Education Reform	0.027	-0.036	0.177*	0.004
	(0.032)	(0.055)	(0.099)	(0.003)
Control Mean	2.34	10.46	79.77	-0.07
Control SD	8.49	12.57	20.92	0.73
Bandwidth	42.97	32.66	38.39	39.03
Observations	823369	629366	745603	764488

Notes: Local linear RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The unit of analysis is the month-year birth cohorts. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the share of wealth invested in equity funds in column 5, in stocks in column 6, and in risky assets in percentage terms in column 7 in pension portfolios. In column 4, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for quadratic polynomials of pension participation rate, the month of birth fixed effects, and birth registration certificate region fixed effects with dummy variables for each month and each relevant regions, and the regressions in the full sample additionally include controls for gender with a dummy variable of being female. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.8: Education Reform vs Behavioral Biases and Heuristics, Financial Mistakes, and Performance

	(1) Default Fund Share	(2) 1/N Heuristic	(3) Portfolio Reshuffling	(4) Contribution (in logs)	(5) RSRL	(6) Disposition Effect	(7) Portfolio Return	(8) Outperforming Default	(9) Summary Index
					Full				
Education Reform	-0.386***	0.001	-0.004*	0.010	-0.004*	0.001*	0.012**	0.003***	0.001
	(0.148)	(0.001)	(0.002)	(0.007)	(0.002)	(0.001)	(0.006)	(0.001)	(0.001)
Control Mean	46.73	0.37	0.44	2.75	0.44	-0.29	20.27	0.76	0.01
Control SD	48.78	0.48	0.50	2.69	1.16	0.43	2.39	0.43	0.34
Bandwidth	50.57	28.60	30.59	32.58	51.35	35.25	40.84	34.84	29.19
Observations	2551985	1436505	1535546	1640290	2597908	1780686	2049373	1738476	1485550
					Male				
Education Reform	-0.360**	-0.002	-0.003	0.000	0.002	0.001	0.020***	0.002*	0.000
	(0.167)	(0.002)	(0.002)	(0.008)	(0.003)	(0.001)	(0.006)	(0.001)	(0.001)
Control Mean	50.91	0.34	0.48	2.55	0.46	-0.31	20.23	0.74	0.00
Control SD	48.90	0.47	0.50	2.68	1.16	0.45	2.40	0.44	0.33
Bandwidth	50.66	30.41	34.88	32.06	40.59	31.39	36.14	28.61	30.43
Observations	1574751	946276	1071550	1010924	1263795	978094	1146511	886224	946276
					Female				
Education Reform	-0.264	0.006***	-0.007**	0.026**	-0.008**	0.003*	-0.001	0.004**	0.001
	(0.220)	(0.002)	(0.003)	(0.010)	(0.004)	(0.001)	(0.010)	(0.002)	(0.001)
Control Mean	40.45	0.41	0.38	3.07	0.41	-0.24	20.33	0.79	0.03
Control SD	47.91	0.49	0.49	2.68	1.16	0.41	2.36	0.41	0.35
Bandwidth	41.04	26.29	30.37	31.94	28.79	46.17	39.76	40.66	28.08
Observations	803354	511930	589270	609318	549943	901950	764488	784359	550281

Notes: Local linear RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the share of default funds in pension portfolios in percentage terms. In column 2, the outcome is a dummy variable equal to one if a pension investor follows the conditional $\frac{1}{N}$ heuristic while allocating money to pension funds. In column 3, the outcome is a dummy variable equal to one if a pension investor changed her funds in a year. In column 4, the outcome is the log of one plus the value of the monthly contribution in Turkish Lira. In column 5, the outcome is the relative Sharpe Ratio loss relative that indicates the loss from under-diversification to the domestic benchmark stock portfolio. In column 6, the outcome is the disposition effect measure. In column 7, the outcome is the annual net rate of return of pension portfolios after subtracting management fees of pension portfolios in percentage. In column 8, the outcome is a dummy variable equal to one if a pension investor outperforms the default fund in a year in terms of portfolio returns. In column 9, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for quadratic polynomials of pension participation rate, the month of birth fixed effects, and birth registration certificate region fixed effects with dummy variables for each month and each relevant regions, and the regressions in the full sample additionally include controls for gender with a dummy variable of being female. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.6 Local Quadratic RD Estimates

Table A.9: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	0.403***	-0.009	0.049***	0.036**	0.032*	0.076***
	(0.128)	(0.008)	(0.013)	(0.016)	(0.017)	(0.023)
Control Mean	9.02	0.93	0.64	0.50	0.27	0.40
Control SD	4.69	0.26	0.48	0.50	0.45	0.83
Bandwidth	86.98	71.67	67.14	55.16	58.74	87.83
Observations	91154	74274	69926	57247	60380	92228
			Male			
Education Reform	0.356*	0.003	0.036*	0.034	0.031	0.066*
	(0.198)	(0.009)	(0.019)	(0.027)	(0.020)	(0.035)
Control Mean	9.75	0.96	0.72	0.55	0.29	0.53
Control SD	4.32	0.19	0.45	0.50	0.45	0.77
Bandwidth	76.24	70.74	70.97	59.08	58.10	76.40
Observations	38624	35368	35368	29523	29048	38624
			Female			
Education Reform	0.445*	-0.022	0.096***	0.041	0.033	0.086*
	(0.255)	(0.014)	(0.023)	(0.028)	(0.023)	(0.044)
Control Mean	8.37	0.89	0.53	0.43	0.26	0.29
Control SD	4.97	0.31	0.50	0.50	0.44	0.88
Bandwidth	83.75	80.66	95.05	64.15	60.34	84.91
Observations	45357	43713	52247	34567	32608	46293

Notes: Local quadratic RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.10: Education Reform vs Participation and Wealth Accumulation in DC Pension Plans

	Panel A: Participation in DC Pension Plans						*Panel B: Wealth*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participant	Voluntary	Employer Sponsored	Automatic	Contributor	Summary Index	(in logs)
				Full			
Education Reform	0.078	0.094	0.030	-0.068	0.131*	0.016	0.014
	(0.116)	(0.091)	(0.025)	(0.083)	(0.072)	(0.011)	(0.010)
Control Mean	24.90	23.86	1.62	27.10	14.17	1.61	7.17
Control SD	0.62	0.98	0.08	0.81	0.60	0.07	2.34
Bandwidth	43.05	48.69	57.96	50.74	57.97	55.32	54.42
Observations	87	97	115	101	115	111	2751779
				Male			
Education Reform	0.095	0.206**	0.009	-0.272	0.066	0.004	0.005
	(0.154)	(0.098)	(0.042)	(0.171)	(0.082)	(0.013)	(0.011)
Control Mean	30.02	27.42	2.29	37.29	15.95	1.60	6.94
Control SD	0.37	0.71	0.10	1.02	0.49	0.04	2.35
Bandwidth	47.45	32.53	57.42	45.86	48.77	57.98	55.67
Observations	95	65	115	91	97	115	1731225
				Female			
Education Reform	0.091	0.096	0.047**	0.129	0.281***	0.052***	0.029**
	(0.134)	(0.107)	(0.021)	(0.119)	(0.098)	(0.018)	(0.014)
Control Mean	19.72	19.83	0.92	16.69	12.30	1.63	7.51
Control SD	0.91	1.14	0.10	0.71	0.69	0.12	2.27
Bandwidth	47.83	49.78	63.83	55.46	67.76	57.68	42.92
Observations	95	99	127	111	135	115	823369

Notes: Local quadratic RD estimates in all columns. Columns 1-6 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Column 7 uses administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-6 and individuals in column 7. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those having positive balances in DC retirement accounts in column 1, in DC voluntary retirement accounts in column 2, in DC employer-sponsorship retirement accounts in column 3, in DC automatic enrollment retirement accounts in column 4, and the percentage of those actively contributing to pension plans in column 5. In column 6, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variables. In column 7, the outcome is the log of the value of pension portfolios in Turkish Lira. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-6 and column 7 also includes controls for the birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.11: Education Reform vs Equity Exposure, and Default Option

	Panel A: Propensity to Own				*Panel B: Share of Wealth Invested in*			
	(1) Equity Fund	(2) Default Fund	(3) Only Default Fund	(4) Summary Index	(5) Equity Fund	(6) Stock	(7) Risky Assets	(8) Summary Index
				Full				
Education Reform	0.054*	-0.028	-0.063	0.003	0.055**	0.006	0.076	0.001
	(0.028)	(0.059)	(0.048)	(0.013)	(0.024)	(0.037)	(0.074)	(0.002)
Control Mean	2.90	12.99	10.91	1.53	2.25	10.10	77.98	-0.12
Control SD	0.20	0.20	0.19	0.04	8.35	12.41	22.50	0.75
Bandwidth	57.09	45.42	43.81	40.60	76.87	39.03	35.48	34.92
Observations	115	91	87	81	3929283	1997398	1780686	1738476
				Male				
Education Reform	0.119***	-0.041	-0.117	0.004	0.052*	0.093**	0.070	0.005*
	(0.038)	(0.116)	(0.115)	(0.019)	(0.027)	(0.040)	(0.082)	(0.003)
Control Mean	3.52	16.82	14.39	1.52	2.13	9.86	76.93	-0.15
Control SD	0.26	0.31	0.32	0.03	8.16	12.28	23.32	0.75
Bandwidth	89.23	42.95	40.58	46.29	57.16	40.67	39.37	36.85
Observations	179	85	81	93	1798166	1265014	1232910	1146511
				Female				
Education Reform	0.056*	-0.014	-0.001	0.009	0.030	-0.104	0.095	-0.002
	(0.032)	(0.097)	(0.082)	(0.023)	(0.036)	(0.077)	(0.106)	(0.004)
Control Mean	2.40	9.05	7.38	1.55	2.39	10.50	79.76	-0.07
Control SD	0.18	0.35	0.26	0.10	8.54	12.61	20.92	0.73
Bandwidth	60.66	43.86	46.03	43.44	84.51	45.12	35.23	36.82
Observations	121	87	93	87	1668592	883063	682913	711844

Notes: Local quadratic RD estimates in all columns. Columns 1-4 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Columns 5-8 use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-4 and individuals in the remaining columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the percentage of those holding equity funds in the total population. In column 2, the outcome is the percentage of those holding default pension funds in the total population. In column 3, the outcome is the percentage of those holding only default pension funds in the total population. In column 4, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variable. The outcome is the share of wealth invested in equity funds in column 5, in stocks in column 6, and in risky assets in percentage terms in column 7 in pension portfolios. In column 8, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular-type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-4 and the remaining columns also include controls for the birth registration certificate region fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively.

Table A.12: Education Reform vs Behavioral Biases and Heuristics, Financial Mistakes, and Performance

	(1) Default Fund Share	(2) 1/N Heuristic	(3) Reshuffling	(4) Contribution (in logs)	(5) RSRL	(6) Disposition Effect	(7) Portfolio Return	(8) Outperforming Default	(9) Summary Index
					Full				
Education Reform	-0.265*	0.000	-0.005*	0.008	-0.001	0.001	0.008	0.002*	0.001
	(0.155)	(0.002)	(0.003)	(0.009)	(0.003)	(0.001)	(0.008)	(0.001)	(0.001)
Control Mean	46.63	0.37	0.44	2.78	0.44	-0.29	20.26	0.76	0.01
Control SD	48.77	0.48	0.50	2.70	1.16	0.43	2.39	0.43	0.34
Bandwidth	52.22	40.69	36.74	43.11	53.53	41.02	35.63	32.46	50.95
Observations	2650108	2049373	1858355	2199510	2696626	2097731	1780686	1640290	2551985
					Male				
Education Reform	-0.307	-0.002	-0.002	0.000	0.002	0.001	0.020**	0.003	0.000
	(0.200)	(0.002)	(0.003)	(0.010)	(0.004)	(0.002)	(0.009)	(0.002)	(0.001)
Control Mean	50.66	0.34	0.48	2.58	0.46	-0.31	20.22	0.74	0.01
Control SD	48.91	0.47	0.50	2.69	1.16	0.45	2.40	0.44	0.33
Bandwidth	55.96	42.89	37.66	43.71	50.85	40.57	35.23	36.31	56.18
Observations	1731225	1324678	1173036	1356106	1573221	1265014	1097773	1146511	1766510
					Female				
Education Reform	-0.195	0.005**	-0.008**	0.022	-0.005	0.000	-0.009	0.002	0.000
	(0.279)	(0.002)	(0.004)	(0.016)	(0.005)	(0.002)	(0.012)	(0.003)	(0.001)
Control Mean	40.35	0.41	0.38	3.10	0.41	-0.24	20.33	0.79	0.03
Control SD	47.89	0.49	0.49	2.69	1.16	0.41	2.36	0.41	0.35
Bandwidth	43.67	37.81	39.50	45.76	54.23	39.43	38.19	32.77	45.05
Observations	843404	727605	764488	883063	1053908	764488	745603	629366	883063

Notes: Local quadratic RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the share of default funds in pension portfolios in percentage terms. In column 2, the outcome is a dummy variable equal to one if a pension investor follows the conditional $\frac{1}{N}$ heuristic while allocating money to pension funds. In column 3, the outcome is a dummy variable equal to one if a pension investor changed her funds in a year. In column 4, the outcome is the log of one plus the value of the monthly contribution in Turkish Lira. In column 5, the outcome is the relative Sharpe Ratio loss relative that indicates the loss from under-diversification to the domestic benchmark stock portfolio. In column 6, the outcome is the disposition effect measure. In column 7, the outcome is the annual net rate of return of pension portfolios after subtracting management fees of pension portfolios in percentage. In column 8, the outcome is a dummy variable equal to one if a pension investor outperforms the default fund in a year in terms of portfolio returns. In column 9, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.7 Local Kink RD Estimates

Table A.13: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	-0.008	0.001	-0.005***	-0.003	-0.001	-0.002
	(0.011)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Control Mean	9.09	0.93	0.64	0.50	0.28	0.42
Control SD	4.70	0.26	0.48	0.50	0.45	0.84
Bandwidth	73.95	61.15	65.70	50.30	49.05	73.47
Observations	77043	63853	67860	52433	51460	77043
			Male			
Education Reform	-0.003	0.001	-0.005***	-0.004	0.001	-0.001
	(0.015)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Control Mean	9.80	0.96	0.72	0.56	0.29	0.54
Control SD	4.32	0.19	0.45	0.50	0.45	0.78
Bandwidth	71.40	67.01	69.32	53.96	60.43	72.05
Observations	35806	33721	34812	26578	30358	36728
			Female			
Education Reform	-0.020	0.002**	-0.002	-0.002	-0.002	-0.003
	(0.022)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)
Control Mean	8.48	0.89	0.52	0.44	0.26	0.31
Control SD	4.98	0.31	0.50	0.50	0.44	0.88
Bandwidth	68.46	81.89	96.20	59.20	50.43	68.63
Observations	36786	44273	53191	31788	27211	36786

Notes: Local kink RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.14: Education Reform vs Participation and Wealth Accumulation in DC Pension Plans

	Panel A: Participation in DC Pension Plans						*Panel B: Wealth*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participant	Voluntary	Employer Sponsored	Automatic	Contributor	Summary Index	(in logs)
				Full			
Education Reform	-0.016	-0.021	0.005**	0.010	0.009	0.000	0.002**
	(0.015)	(0.016)	(0.002)	(0.011)	(0.006)	(0.002)	(0.001)
Control Mean	24.80	23.49	1.63	27.23	14.20	1.60	7.16
Control SD	0.52	0.81	0.07	0.77	0.58	0.07	2.34
Bandwidth	40.13	35.16	52.79	44.72	63.27	42.61	50.82
Observations	81	71	105	89	127	85	2551985
				Male			
Education Reform	-0.013	-0.017	0.002	0.007	-0.004	-0.001	0.001
	(0.019)	(0.017)	(0.004)	(0.021)	(0.007)	(0.002)	(0.001)
Control Mean	29.97	27.49	2.29	37.38	16.05	1.60	6.93
Control SD	0.35	0.69	0.09	0.98	0.52	0.04	2.35
Bandwidth	42.31	37.99	55.93	42.43	57.11	44.73	53.10
Observations	85	75	111	85	115	89	1664757
				Female			
Education Reform	-0.024	-0.040***	0.011***	0.020	0.011	0.004**	0.004***
	(0.015)	(0.012)	(0.002)	(0.013)	(0.009)	(0.002)	(0.002)
Control Mean	19.70	19.93	0.94	16.77	12.26	1.63	7.52
Control SD	0.91	1.19	0.08	0.70	0.68	0.11	2.27
Bandwidth	46.47	52.65	55.84	49.58	63.74	49.94	45.90
Observations	93	105	111	99	127	99	883063

Notes: Local kink RD estimates in all columns. Columns 1-6 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Column 7 uses administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-6 and individuals in column 7. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those having positive balances in DC retirement accounts in column 1, in DC voluntary retirement accounts in column 2, in DC employer-sponsorship retirement accounts in column 3, in DC automatic enrollment retirement accounts in column 4, and the percentage of those actively contributing to pension plans in column 5. In column 6, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variables. In column 7, the outcome is the log of the value of pension portfolios in Turkish Lira. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-6 and column 7 also includes controls for the birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.15: Education Reform vs Equity Exposure, and Default Option

	Panel A: Propensity to Own				Panel B: Share of Wealth Invested in			
	(1) Equity Fund	(2) Default Fund	(3) Only Default Fund	(4) Summary Index	(5) Equity Fund	(6) Stock	(7) Risky Assets	(8) Summary Index
				Full				
Education Reform	-0.001	-0.014*	-0.008	-0.002	-0.002	-0.003	0.005	-0.000
	(0.002)	(0.008)	(0.006)	(0.002)	(0.002)	(0.006)	(0.013)	(0.000)
Control Mean	2.91	12.98	10.91	1.53	2.23	10.09	77.93	-0.12
Control SD	0.20	0.20	0.19	0.04	8.32	12.40	22.54	0.75
Bandwidth	60.80	41.55	45.94	41.21	63.84	36.38	31.31	33.34
Observations	121	83	91	83	3220574	1858355	1587412	1690646
				Male				
Education Reform	-0.000	-0.015	-0.003	-0.003	-0.003	-0.004	-0.028**	-0.001
	(0.002)	(0.014)	(0.015)	(0.002)	(0.003)	(0.007)	(0.013)	(0.001)
Control Mean	3.50	16.83	14.42	1.51	2.13	9.85	76.88	-0.15
Control SD	0.25	0.31	0.31	0.03	8.16	12.27	23.35	0.75
Bandwidth	83.26	41.13	37.79	41.54	57.46	38.10	36.29	33.56
Observations	167	83	75	83	1798166	1202282	1146511	1042058
				Female				
Education Reform	-0.000	-0.013	-0.016*	-0.002	-0.000	0.005	0.044**	0.001
	(0.003)	(0.011)	(0.009)	(0.002)	(0.003)	(0.009)	(0.020)	(0.001)
Control Mean	2.39	9.06	7.37	1.56	2.38	10.50	79.74	-0.07
Control SD	0.19	0.35	0.26	0.10	8.53	12.61	20.95	0.73
Bandwidth	55.45	44.30	44.40	45.89	66.28	43.33	32.42	36.51
Observations	111	89	89	91	1292849	843404	629366	711844

Notes: Local kink RD estimates in all columns. Columns 1-4 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Columns 5-8 use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-4 and individuals in the remaining columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the percentage of those holding equity funds in the total population. In column 2, the outcome is the percentage of those holding default pension funds in the total population. In column 3, the outcome is the percentage of those holding only default pension funds in the total population. In column 4, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variable. The outcome is the share of wealth invested in equity funds in column 5, in stocks in column 6, and in risky assets in percentage terms in column 7 in pension portfolios. In column 8, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-4 and the remaining columns also include controls for the birth registration certificate region fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.16: Education Reform vs Behavioral Biases and Heuristics, Financial Mistakes, and Performance

	(1) Default Fund Share	(2) 1/N Heuristic	(3) Reshuffling	(4) Contribution (in logs)	(5) RSRL	(6) Disposition Effect	(7) Portfolio Return	(8) Outperforming Default	(9) Summary Index
					Full				
Education Reform	-0.009	0.000	-0.001*	0.001	-0.001***	0.000	0.001	0.000	0.000
	(0.019)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Control Mean	46.94	0.37	0.44	2.78	0.44	-0.29	20.26	0.76	0.01
Control SD	48.79	0.48	0.50	2.70	1.16	0.43	2.39	0.43	0.34
Bandwidth	44.12	33.28	30.08	42.02	48.41	38.76	35.02	30.48	48.01
Observations	2253462	1690646	1535546	2148047	2463718	1947885	1780686	1535546	2465795
					Male				
Education Reform	0.012	0.000	-0.000	0.001	-0.002***	0.000	-0.000	-0.000	-0.000
	(0.022)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)
Control Mean	50.95	0.34	0.48	2.58	0.46	-0.31	20.22	0.74	0.01
Control SD	48.90	0.47	0.50	2.69	1.16	0.45	2.40	0.44	0.33
Bandwidth	49.88	33.74	35.88	42.32	44.53	40.06	35.16	33.64	56.72
Observations	1546455	1042058	1097773	1324678	1388200	1265014	1097773	1042058	1766510
					Female				
Education Reform	-0.020	0.000	-0.001**	0.002	-0.001	0.000	0.002	0.000	0.000
	(0.030)	(0.000)	(0.001)	(0.002)	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)
Control Mean	40.30	0.41	0.38	3.10	0.41	-0.24	20.33	0.79	0.03
Control SD	47.88	0.49	0.49	2.69	1.16	0.41	2.36	0.41	0.35
Bandwidth	46.99	37.01	34.43	44.14	51.50	39.18	36.93	34.09	46.44
Observations	901950	727605	666926	863887	994950	764488	711844	666926	901950

Notes: Local kink RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the share of default funds in pension portfolios in percentage terms. In column 2, the outcome is a dummy variable equal to one if a pension investor follows the conditional $\frac{1}{N}$ heuristic while allocating money to pension funds. In column 3, the outcome is a dummy variable equal to one if a pension investor changed her funds in a year. In column 4, the outcome is the log of one plus the value of the monthly contribution in Turkish Lira. In column 5, the outcome is the relative Sharpe Ratio loss relative that indicates the loss from under-diversification to the domestic benchmark stock portfolio. In column 6, the outcome is the disposition effect measure. In column 7, the outcome is the annual net rate of return of pension portfolios after subtracting management fees of pension portfolios in percentage. In column 8, the outcome is a dummy variable equal to one if a pension investor outperforms the default fund in a year in terms of portfolio returns. In column 9, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.8 Local Linear RD Estimates with Fixed Bandwidth

Table A.17: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	0.368***	-0.019***	0.088***	0.040***	0.018*	0.070***
	(0.102)	(0.006)	(0.010)	(0.010)	(0.009)	(0.018)
Control Mean	8.98	0.92	0.62	0.47	0.25	0.40
Control SD	4.69	0.26	0.49	0.50	0.44	0.83
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00
Observations	61311	61311	61311	61311	61311	61311
			Male			
Education Reform	0.273**	-0.001	0.065***	0.025**	0.002	0.050***
	(0.108)	(0.006)	(0.012)	(0.012)	(0.012)	(0.019)
Control Mean	9.78	0.96	0.72	0.54	0.28	0.54
Control SD	4.29	0.19	0.45	0.50	0.45	0.77
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00
Observations	29523	29523	29523	29523	29523	29523
			Female			
Education Reform	0.454***	-0.035***	0.110***	0.053***	0.032**	0.089***
	(0.172)	(0.009)	(0.017)	(0.017)	(0.013)	(0.030)
Control Mean	8.24	0.89	0.53	0.41	0.23	0.27
Control SD	4.91	0.32	0.50	0.49	0.42	0.87
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00
Observations	31788	31788	31788	31788	31788	31788

Notes: Local linear RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have a fixed bandwidth of 60 with a triangular-type kernel function in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.18: Education Reform vs Participation and Wealth Accumulation in DC Pension Plans

	Panel A: Participation in DC Pension Plans						*Panel B: Wealth*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participant	Voluntary	Employer Sponsored	Automatic	Contributor	Summary Index	(in logs)
				Full			
Education Reform	0.093	0.253***	0.105***	0.032	0.111**	0.049***	0.014*
	(0.072)	(0.070)	(0.017)	(0.056)	(0.051)	(0.008)	(0.008)
Control Mean	25.05	24.07	1.61	26.84	14.18	1.61	7.18
Control SD	0.63	1.05	0.08	0.94	0.58	0.07	2.35
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	119	119	3006699
				Male			
Education Reform	-0.045	0.080	0.087***	0.023	-0.002	0.022**	0.005
	(0.089)	(0.075)	(0.028)	(0.092)	(0.053)	(0.009)	(0.009)
Control Mean	30.11	27.99	2.28	36.81	16.06	1.60	6.94
Control SD	0.40	0.91	0.10	1.22	0.51	0.05	2.36
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	119	119	1856373
				Female			
Education Reform	0.241**	0.433***	0.124***	0.048	0.228***	0.107***	0.028***
	(0.095)	(0.082)	(0.015)	(0.082)	(0.071)	(0.013)	(0.010)
Control Mean	19.85	20.03	0.93	16.58	12.24	1.62	7.55
Control SD	0.93	1.22	0.09	0.78	0.69	0.12	2.28
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	119	119	1150326

Notes: Local linear RD estimates in all columns. Columns 1-6 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Column 7 uses administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-6 and individuals in column 7. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those having positive balances in DC retirement accounts in column 1, in DC voluntary retirement accounts in column 2, in DC employer-sponsorship retirement accounts in column 3, in DC automatic enrollment retirement accounts in column 4, and the percentage of those actively contributing to pension plans in column 5. In column 6, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variables. In column 7, the outcome is the log of the value of pension portfolios in Turkish Lira. RD estimates have a fixed bandwidth of 60 with a triangular-type kernel function in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-6 and column 7 also includes controls for the birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.19: Education Reform vs Equity Exposure, and Default Option

	Panel A: Propensity to Own				*Panel B: Share of Wealth Invested in*			
	(1) Equity Fund	(2) Default Fund	(3) Only Default Fund	(4) Summary Index	(5) Equity Fund	(6) Stock	(7) Risky Assets	(8) Summary Index
				Full				
Education Reform	0.089***	-0.096***	-0.102***	0.011	0.051***	0.110***	0.154***	0.007***
	(0.021)	(0.036)	(0.029)	(0.007)	(0.019)	(0.030)	(0.059)	(0.002)
Control Mean	2.91	12.94	10.84	1.55	2.23	10.19	78.25	-0.11
Control SD	0.20	0.21	0.21	0.05	8.32	12.49	22.26	0.75
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	3006699	3006699	3006699	3006699
				Male				
Education Reform	0.099***	-0.192***	-0.175***	0.004	0.062***	0.144***	0.094	0.007***
	(0.031)	(0.063)	(0.057)	(0.010)	(0.020)	(0.031)	(0.059)	(0.002)
Control Mean	3.41	16.70	14.23	1.52	2.14	9.96	77.22	-0.14
Control SD	0.23	0.34	0.36	0.03	8.18	12.37	23.10	0.76
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	1856373	1856373	1856373	1856373
				Female				
Education Reform	0.080***	0.007	-0.023	0.024*	0.033	0.052	0.243**	0.007**
	(0.024)	(0.054)	(0.046)	(0.014)	(0.030)	(0.051)	(0.098)	(0.003)
Control Mean	2.40	9.06	7.35	1.57	2.37	10.56	79.87	-0.07
Control SD	0.18	0.34	0.26	0.10	8.53	12.66	20.78	0.73
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	1150326	1150326	1150326	1150326

Notes: Local linear RD estimates in all columns. Columns 1-4 use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2019. Columns 5-8 use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is the month-year birth cohorts in columns 1-4 and individuals in the remaining columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the percentage of those holding equity funds in the total population. In column 2, the outcome is the percentage of those holding default pension funds in the total population. In column 3, the outcome is the percentage of those holding only default pension funds in the total population. In column 4, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007) for the former outcome variable. The outcome is the share of wealth invested in equity funds in column 5, in stocks in column 6, and risky assets in percentage terms in column 7 in pension portfolios. In column 8, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have a fixed bandwidth of 60 with a triangular-type kernel function in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns 1-4 and the remaining columns also include controls for the birth registration certificate region fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.20: Education Reform vs Behavioral Biases and Heuristics, Financial Mistakes, and Performance

	(1) Default Fund Share	(2) 1/N Heuristic	(3) Reshuffling	(4) Contribution (in logs)	(5) RSRL	(6) Disposition Effect	(7) Portfolio Return	(8) Outperforming Default	(9) Summary Index
					Full				
Education Reform	-0.454***	0.001	-0.004**	0.013**	-0.005**	0.002***	0.017***	0.003***	0.000
	(0.168)	(0.001)	(0.002)	(0.007)	(0.002)	(0.001)	(0.006)	(0.001)	(0.001)
Control Mean	46.40	0.37	0.44	2.81	0.44	-0.28	20.29	0.76	0.02
Control SD	48.76	0.48	0.50	2.71	1.16	0.43	2.40	0.43	0.34
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	3006699	3006699	3006699	3006699	3004145	3006699	3006699	3006699	3006699
					Male				
Education Reform	-0.369**	-0.001	-0.004*	0.002	-0.000	0.002*	0.022***	0.002	0.000
	(0.169)	(0.001)	(0.002)	(0.007)	(0.003)	(0.001)	(0.006)	(0.001)	(0.001)
Control Mean	50.57	0.34	0.47	2.61	0.46	-0.31	20.26	0.74	0.01
Control SD	48.91	0.47	0.50	2.70	1.16	0.45	2.42	0.44	0.33
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	1856373	1856373	1856373	1856373	1854567	1856373	1856373	1856373	1856373
					Female				
Education Reform	-0.581**	0.004**	-0.005*	0.030***	-0.013***	0.003*	0.009	0.005***	0.001
	(0.246)	(0.002)	(0.003)	(0.010)	(0.003)	(0.001)	(0.010)	(0.002)	(0.001)
Control Mean	39.87	0.41	0.38	3.12	0.41	-0.24	20.34	0.79	0.03
Control SD	47.80	0.49	0.49	2.70	1.16	0.41	2.37	0.41	0.35
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	1150326	1150326	1150326	1150326	1149578	1150326	1150326	1150326	1150326

Notes: Local linear RD estimates in all columns. All columns use administrative data covering the universe of all pension portfolios in 2019 assembled by Borsa Istanbul Group. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is the share of default funds in pension portfolios in percentage terms. In column 2, the outcome is a dummy variable equal to one if a pension investor follows the conditional $\frac{1}{N}$ heuristic while allocating money to pension funds. In column 3, the outcome is a dummy variable equal to one if a pension investor changed her funds in a year. In column 4, the outcome is the log of one plus the value of the monthly contribution in Turkish Lira. In column 5, the outcome is the relative Sharpe Ratio loss relative that indicates the loss from under-diversification to the domestic benchmark stock portfolio. In column 6, the outcome is the disposition effect measure. In column 7, the outcome is the annual net rate of return of pension portfolios after subtracting management fees of pension portfolios in percentage. In column 8, the outcome is a dummy variable equal to one if a pension investor outperforms the default fund in a year in terms of portfolio returns. In column 9, the outcome is a summary index of outcome variables explored constructed following Kling et al. (2007). RD estimates have a fixed bandwidth of 60 with a triangular-type kernel function in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.9 Data Appendix

A.9.1 Variables for Schooling Outcomes

Years of Schooling is generated by assigning 0, 5, 8, 11, 15, and 17 for those who have no degree, at least a primary school degree, at least junior high school degree, at least high school degree, at least college degree, and at least master degree, respectively.

Primary School is a dummy variable equal to 1 if a specific individual at least holds a primary school degree with 5 years of schooling.

Junior High School is a dummy variable equal to 1 if a specific individual at least holds a junior high school degree with 8 years of schooling.

High School is a dummy variable equal to 1 if a specific individual at least holds a high school degree with 11 years of schooling.

College is a dummy variable equal to 1 if a specific individual at least holds a college degree with 13-15 years of schooling.

A.9.2 Variables for the Participation in Pension Plans

Pension Plans Participation Rate is the percentage of those having a positive amount in their individual retirement accounts in any pension plans generated by dividing the number of individuals with positive balances by the number of individuals in the general population in the specific month-year birth cohort.

Voluntary Pension Plans Participation Rate is the percentage of those having a positive amount in their individual retirement accounts in the voluntary pension plans generated by dividing the number of individuals with positive balances in the voluntary pension plans by the number of individuals in the general population in the specific month-year birth cohort.

Employer Sponsored Pension Participation Rate is the percentage of those having a positive amount in their individual retirement accounts in the employer-sponsored pension plans generated by dividing the number of individuals with positive balances in the employer-sponsored pension plans by the number of individuals in the general population in the specific month-year birth cohort.

Automatic Pension Plans Participation Rate is the percentage of those having a positive amount in their individual retirement accounts in the automatic enrollment pension plans generated by dividing the number of individuals with positive balances in the automatic enrollment pension plans by the number of individuals in the general population in the specific month-year birth cohort.

Contributor Rate is the percentage of those contributing to any pension plans generated by dividing the number of individuals contributing to any pension plans by the number of individuals in the general population in the specific month-year birth cohort.

A.9.3 Variables for Choices of Funds

Equity Fund Ownership Rate is the percentage of those holding equity funds in their portfolios generated by dividing the number of individuals with equity funds in their portfolios by the number of individuals in the general population in the specific month-year birth cohort.

Default Fund Ownership Rate is the percentage of those holding default funds in their portfolios generated by dividing the number of individuals with default funds in their portfolios by the number of individuals in the general population in the specific month-year birth cohort. I treat a fund as the default fund if it is the starting fund or automatic enrollment standard fund.

Only Default Fund Ownership Rate is the percentage of those holding only default funds in their portfolios generated by dividing the number of individuals with only default funds in their portfolios by the number of individuals in the general population in the specific month-year birth cohort.

A.9.4 Variables for Pension Wealth, Contribution, and Performance

Logarithm of Pension Wealth is the log of the total wealth an investor has accumulated or simply the portfolio size in Turkish Liras.

Log of Monthly Contribution is the log of the total amount an investor contributes to pension plans monthly in Turkish Liras.

Annual Portfolio Rate of Return is the percentage returns of portfolios from December 2019 to December 2020.

Outperforming Default is a dummy variable equal to one if the investor's annual portfolio return is higher than the annual return of the default option (fund) in 2020.

Variables for Portfolio Shares of Funds, and Assets

Equity Fund Share is the percentage of wealth invested in equity funds in portfolios.

Stock Share is the percentage of wealth invested in stocks indirectly in portfolios. To calculate the total stock share in a portfolio, I use the stock share of each fund and calculate the total amount invested in stocks indirectly. Subsequently, I weigh the related amount by total wealth or portfolio size.

Risky Assets Share is the percentage of wealth invested in risky assets indirectly in portfolios. To calculate the total risky asset share in a portfolio, I use the share of each fund in risky assets that exclude money market instruments, deposits, reverse repo, and repo, and calculate the total amount invested in risky assets indirectly. Subsequently, I weigh the related amount by total wealth or portfolio size.

A.9.5 Variables for Behavioral Biases and Heuristics, and Financial Mistakes

Default Fund Share is the percentage of wealth invested in default funds in portfolios.

1/N Heuristic is a dummy variable equal to one if the condition I explain below is satisfied following Huberman and Jiang (2006):
Let s_{ij} be the share of investor i's contribution in fund j, and n_i is the total number of funds in i's portfolio, thus $\sum_{j=1}^{n_i} s_{ij} = 1$. Then, the Herfindahl index, defined for each investor i's portfolio as the sum of the squared fractions of contributions in each fund as follows:
$H_i = \sum_{j=1}^{n_i} s_j^2$
The value H_i is bounded between $\frac{1}{n_i}$ and 1 and it is equal to $\frac{1}{n_i}$ if investor i equally divides the contribution amount among n_i funds.

I treat an investor with the Herfindahl index close to $\frac{1}{n}$ as an investor prone to $\frac{1}{n}$ rule or naive diversification. Accordingly, I classify an investor as an investor with the $\frac{1}{n}$ rule if her Herfindahl index is bounded from above by the index that would lead to a portfolio in which the total deviation from an $\frac{1}{n}$ allocation is 20%.

RSRL (Loss from Underdiversification) is the measure quantifying the losses arising from insufficient diversification and it is called relative Sharpe ratio loss (RSRL) suggested by Calvet et al. (2007, 2009a,b). I construct it as follows:
$RSRL = 1 - \frac{S_i}{S_B}$, where S_i is the Sharpe Ratio of investor i's portfolio calculated by using excess monthly returns of the corresponding portfolio and S_B is the Sharpe Ratio of the domestic benchmark pension portfolio, calculated by using excess monthly returns. I exclude the portfolios with the funds that are not observable for at least 36 months and the number of portfolios excluded for this reason is around 7000, which is a negligible number.

Disposition Effect is the measure of the tendency of investors to sell the winning funds and to keep the losing ones. Since I am not able to observe the initial buying prices of any funds, to calculate it, I follow the strategy proposed by Calvet et al. (2009b). I compare the annual return of each fund in December 2020 with the annual return of the domestic benchmark pension portfolio. Subsequently, I classify a fund as the winner(loser) if its return is higher(lower) than the return of the corresponding pension portfolio. Accordingly, I subtract the proportion of funds held in the losing set from the proportions of funds sold in the winning set, which is the measure of the disposition effect.

Portfolio Reshuffling is a dummy variable equal to one if a pension investor buys a fund that has not been held in her portfolio over the year 2020.

Summary Index is a variable constructed through the procedure suggested by Kling et al. (2007). To make it clear, I first subtract the control group mean of the corresponding variable and subsequently divide it by the control group standard deviation of the concerning variable.

A.9.6 Control Variables

Birth Month Indicator Variables are 12 indicator variables for each month.

Birth Registration Certificate Region Indicator Variables are 26 indicator variables for each birth registration certificate region.

Gender is an indicator variable if a certain investor is female.

APPENDIX B

Appendix to Chapter 2

B.1 Historical Map

Figure B.1: Historical Map of the Spatial Distribution of ABCFM Missions

Source: American Board of Commissioners for Foreign Missions Archives, 1810-1961 (ABC 1-91) Houghton Library, Harvard University.

B.2 Estimates with Eastern Controls

Table B.1: Mission Impact on Development Index

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No East-West Control				
Distance (in km)	-0.007***	-0.007***	0.000	-0.000	-0.012**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.36	0.45	0.34	0.43	0.44
Number of Clusters	80	80	80	80	80
First Stage F-stat					16.33
	Panel B: East-West Control at the Longitude of 35				
Distance (in km)	-0.007***	-0.007***	-0.000	-0.000	-0.011**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.36	0.45	0.34	0.43	0.44
Number of Clusters	80	80	80	80	80
First Stage F-stat					15.14
	Panel C: East-West Control at the Longitude of 40				
Distance (in km)	-0.007***	-0.007***	0.000	-0.000	-0.012**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.36	0.45	0.34	0.43	0.44
Number of Clusters	80	80	80	80	80
First Stage F-stat					15.81

Notes: All columns use data from 2004 SPO District Development Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1, 2, and 3. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the standardized development index produced by SPO in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without Eastern control. Panel B reports the estimates with the Eastern control through a dummy variable equal to one if a district is in the east of 35-degree latitude. Panel C reports the estimates with the Eastern control through a dummy variable equal to one if a district is in the east of 40-degree latitude. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.2: Mission Impact on GDP per Capita

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No East-West Control				
Distance (in km)	-0.005***	-0.005***	-0.001	-0.001	-0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.49	0.52	0.47	0.51	0.42
Number of Clusters	80	80	80	80	80
Sample Mean	7.36	7.36	7.36	7.36	7.36
First Stage F-stat					16.33
	Panel B: East-West Control at the Longitude of 35				
Distance (in km)	-0.005***	-0.004***	-0.001	-0.001	-0.017***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.49	0.52	0.47	0.51	0.41
Number of Clusters	80	80	80	80	80
Sample Mean	7.36	7.36	7.36	7.36	7.36
First Stage F-stat					15.14
	Panel C: East-West Control at the Longitude of 40				
Distance (in km)	-0.005***	-0.004***	-0.001	-0.001	-0.016***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.50	0.53	0.48	0.51	0.44
Number of Clusters	80	80	80	80	80
Sample Mean	7.36	7.36	7.36	7.36	7.36
First Stage F-stat					15.81

Notes: All columns use data from TURKSTAT 1996 District GDP Survey. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1, 2, and 3. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the log of GDP per capita in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without Eastern control. Panel B reports the estimates with the Eastern control through a dummy variable equal to one if a district is in the east of 35-degree latitude. Panel C reports the estimates with the Eastern control through a dummy variable equal to one if a district is in the east of 40-degree latitude. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap RK Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

B.3 Estimates with the Extended Sample

Table B.3: Mission Impact on GDP per Capita

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	-0.005***	-0.005***	-0.001	-0.001	-0.017***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	891	891	891	891	891
R-Squared	0.50	0.53	0.48	0.51	0.42
Number of Clusters	80	80	80	80	80
Sample Mean	7.38	7.38	7.38	7.38	7.38
First Stage F-stat					15.71

Notes: All columns use data from TURKSTAT 1996 District GDP Survey. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the log of GDP per capita in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

B.4 Estimates with the Controls for Initial Conditions

Table B.4: Mission Impact on Development Index

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No Controls for Initial Conditions in 1881				
Distance (in km)	-0.008***	-0.008***	0.000	0.000	-0.011**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.36	0.46	0.34	0.44	0.46
Number of Clusters	77	77	77	77	77
First Stage F-stat					13.09
	Panel B: Controls for Initial Conditions in 1881				
Distance (in km)	-0.008***	-0.007***	-0.000	-0.000	-0.013**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.39	0.47	0.37	0.45	0.46
Number of Clusters	77	77	77	77	77
First Stage F-stat					15.42

Notes: All columns use data from 2004 SPO District Development Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the standardized development index produced by SPO in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without the controls for the minority population shares of the Armenians and Greeks, population density, and the logarithm of the population in 1881. Panel B reports the estimates with the controls for the minority population shares of the Armenians and Greeks, the logarithm of population density, and the logarithm of the population in 1881. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, and distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.5: Mission Impact on GDP per Capita

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No Controls for Initial Conditions in 1881				
Distance (in km)	-0.006***	-0.005***	-0.001	-0.001	-0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.49	0.52	0.46	0.50	0.40
Number of Clusters	77	77	77	77	77
Sample Mean	7.37	7.37	7.37	7.37	7.37
First Stage F-stat					13.09
	Panel B: Controls for Initial Conditions in 1881				
Distance (in km)	-0.005***	-0.005***	-0.001	-0.001	-0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.49	0.52	0.47	0.51	0.39
Number of Clusters	77	77	77	77	77
Sample Mean	7.37	7.37	7.37	7.37	7.37
First Stage F-stat					15.42

Notes: All columns use data from TURKSTAT 1996 District GDP Survey. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the log of GDP per capita in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without the controls for the minority population shares of the Armenians and Greeks, population density, and the logarithm of the population in 1881. Panel B reports the estimates with the controls for the minority population shares of the Armenians and Greeks, the logarithm of population density, and the logarithm of the population in 1881. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, and distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

B.5 Estimates with the Controls for Population Replacement

Table B.6: Mission Impact on Development Index

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No Controls for Population Replacement				
Distance (in km)	-0.007***	-0.007***	0.000	-0.000	-0.012**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.36	0.45	0.34	0.43	0.44
Number of Clusters	80	80	80	80	80
First Stage F-stat					16.33
	Panel B: Controls for Population Replacement				
Distance (in km)	-0.006***	-0.005***	-0.000	-0.000	-0.013
	(0.001)	(0.002)	(0.001)	(0.001)	(0.009)
Share of Armenians in 1881	0.006	0.009**	0.013*	0.015**	0.001
	(0.006)	(0.004)	(0.007)	(0.006)	(0.009)
Share of Greeks in 1881	0.020***	0.014**	0.012*	0.005	0.008
	(0.005)	(0.005)	(0.007)	(0.006)	(0.008)
Share of Armenians in 1881*Distance (in km)	-0.000***	-0.000***	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share of Greeks in 1881*Distance (in km)	-0.000	-0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.40	0.48	0.38	0.46	0.46
Number of Clusters	77	77	77	77	77
First Stage F-stat					9.63

Notes: All columns use data from 2004 SPO District Development Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the standardized development index produced by SPO in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without the controls for the change in population composition after World War I. Panel B reports the estimates with the controls for the change in population composition after World War I by including the share of Armenians and Greeks in 1881 and their interaction terms with the distance to the nearest mission, the logarithm of population density, and the logarithm of the population in 1881. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Standard errors are clustered at the province level. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.7: Mission Impact on GDP per Capita

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No Controls for Population Replacement				
Distance (in km)	-0.005***	-0.005***	-0.001	-0.001	-0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.49	0.52	0.47	0.51	0.42
Number of Clusters	80	80	80	80	80
Sample Mean	7.36	7.36	7.36	7.36	7.36
First Stage F-stat					16.33
	Panel B: Controls for Population Replacement				
Distance (in km)	-0.005***	-0.004***	-0.001	-0.001	-0.025**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.010)
Share of Armenians in 1881	-0.004	-0.001	0.004	0.004	-0.020*
	(0.004)	(0.004)	(0.005)	(0.005)	(0.012)
Share of Greeks in 1881	0.002	-0.002	-0.001	-0.006*	-0.016*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)
Share of Armenians in 1881*Distance (in km)	-0.000*	-0.000**	-0.000*	-0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share of Greeks in 1881*Distance (in km)	0.000	0.000	0.000	0.000*	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.49	0.52	0.48	0.51	0.33
Number of Clusters	77	77	77	77	77
Sample Mean	7.37	7.37	7.37	7.37	7.37
First Stage F-stat					9.63

Notes: All columns use data from TURKSTAT 1996 District GDP Survey. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the log of GDP per capita in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without the controls for the population replacement after the First World War. Panel B reports the estimates with the controls for the change in population composition after World War I by including the share of Armenians and Greeks in 1881 and their interaction terms with the distance to the nearest mission, the logarithm of population density, and the logarithm of the population in 1881. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. Standard errors are clustered at the province level. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.8: Mission Impact on Population Density in 1927

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
	Panel A: No Controls for Population Replacement				
Distance (in km)	-0.003**	-0.003**	0.000	0.001	0.009
	(0.002)	(0.001)	(0.000)	(0.001)	(0.006)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	857	857	857	857	857
R-Squared	0.66	0.71	0.65	0.71	0.61
Number of Clusters	80	80	80	80	80
Sample Mean	2.89	2.89	2.89	2.89	2.89
First Stage F-stat					15.63
	Panel B: Controls for Population Replacement				
Distance (in km)	-0.003**	-0.003**	0.000	0.001*	0.006
	(0.001)	(0.001)	(0.000)	(0.000)	(0.008)
Share of Armenians in 1881	-0.005	-0.003	0.009	0.009	0.005
	(0.005)	(0.004)	(0.006)	(0.005)	(0.007)
Share of Greeks in 1881	-0.003	-0.003	-0.003	-0.004	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)
Share of Armenians in 1881*Distance (in km)	-0.000	0.000	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share of Greeks in 1881*Distance (in km)	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.76	0.79	0.76	0.79	0.76
Number of Clusters	77	77	77	77	77
Sample Mean	2.89	2.89	2.89	2.89	2.89
First Stage F-stat					9.63

Notes: All columns use data from the TURKSTAT 1927 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the logarithm of population density in 1927 in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. Panel A reports the estimates without the controls for the population replacement after the First World War. Panel B reports the estimates with the controls for the change in population composition after World War I by including the share of Armenians and Greeks in 1881 and their interaction terms with the distance to the nearest mission, the logarithm of population density, and the logarithm of the population in 1881. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Sample Mean presents the sample mean of the outcome variable. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

B.6 Additional Outcomes and IV-2SLS Estimates

Table B.9: Mission Impact on Infant Mortality

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	0.016	0.027	-0.001	-0.004	-0.024
	(0.028)	(0.027)	(0.011)	(0.012)	(0.099)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.40	0.42	0.40	0.41	0.41
Number of Clusters	80	80	80	80	80
Sample Mean	41.25	41.25	41.25	41.25	41.25
First Stage F-stat					16.33

Notes: All columns use data from 2004 SPO District Development Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the infant mortality rate per thousand in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable per thousand. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.10: Mission Impact on Fertility Rate

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	0.002	0.002**	-0.000	0.000	0.003
	(0.001)	(0.001)	(0.000)	(0.000)	(0.003)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.74	0.77	0.74	0.77	0.77
Number of Clusters	80	80	80	80	80
Sample Mean	3.76	3.76	3.76	3.76	3.76
First Stage F-stat					16.33

Notes: All columns use data from the TURKSTAT 2000 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the fertility rate in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Table B.11: Mission Impact on Poverty Rate

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	0.043*	0.032*	0.010	0.014	0.101
	(0.025)	(0.018)	(0.009)	(0.010)	(0.064)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	865	865	865	865	865
R-Squared	0.18	0.22	0.18	0.22	0.20
Number of Clusters	80	80	80	80	80
Sample Mean	6.22	6.22	6.22	6.22	6.22
First Stage F-stat					15.79

Notes: All columns use data from 1997 Village Inventory Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the poverty rate in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentages. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.12: Mission Impact on the Concentration in Land Ownership

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	-0.000**	-0.000*	-0.000	-0.000	-0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	866	866	866	866	866
R-Squared	0.29	0.35	0.29	0.34	0.26
Number of Clusters	80	80	80	80	80
Sample Mean	0.69	0.69	0.69	0.69	0.69
First Stage F-stat					15.89

Notes: All columns use data from 1997 Village Inventory Data. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the Gini index of land distribution in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***$p<0.01$, **$p<0.05$, *$p<0.1$

Table B.13: Mission Impact on Population Density

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	-0.004***	-0.004***	0.000	0.000	0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.41	0.50	0.40	0.49	0.46
Number of Clusters	80	80	80	80	80
Sample Mean	3.96	3.96	3.96	3.96	3.96
First Stage F-stat					16.33

Notes: All columns use data from the TURKSTAT 2000 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the logarithm of population density in 2000 in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.14: Mission Impact on Urbanization Rate

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	-0.072*	-0.072**	-0.009	-0.015	0.001
	(0.038)	(0.034)	(0.014)	(0.014)	(0.132)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.12	0.19	0.11	0.18	0.18
Number of Clusters	80	80	80	80	80
Sample Mean	44.05	44.05	44.05	44.05	44.05
First Stage F-stat					16.33

Notes: All columns use data from the TURKSTAT 2000 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the share of the urban population in percentage terms in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentages. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

Table B.15: Mission Impact on Internal Migration

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	0.021	0.035**	-0.002	0.010	-0.007
	(0.015)	(0.014)	(0.008)	(0.006)	(0.063)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	867	867	867	867	867
R-Squared	0.30	0.39	0.30	0.39	0.39
Number of Clusters	80	80	80	80	80
Sample Mean	1.51	1.51	1.51	1.51	1.51
First Stage F-stat					16.33

Notes: All columns use data from the TURKSTAT 2000 Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the share of the non-resident population in percentage terms in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Table B.16: Mission Impact on Religious Conversion

	(1) OLS Treated	(2) OLS Treated	(3) OLS Placebo	(4) OLS Placebo	(5) IV Treated
Distance (in km)	-0.003***	-0.002**	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)
Geo and Hist Controls	NO	YES	NO	YES	YES
Observations	829	829	829	829	829
R-Squared	0.48	0.50	0.47	0.50	0.50
Number of Clusters	77	77	77	77	77
Sample Mean	0.32	0.32	0.32	0.32	0.32
First Stage F-stat					13.09

Notes: All columns use data from the 1914 Ottoman Population Census. The explanatory variable is the distance to the nearest ABCFM mission in km in columns 1,2 and 5. The remaining columns have the explanatory variable of the distance to the nearest placebo mission location. The outcome of interest is the share of the population in the Protestant Sect in all columns. Columns 1-4 report the OLS estimates whereas column 5 presents the 2SLS estimates. All regressions include controls for province-fixed effects with dummy variables for each province. Geographic controls are ruggedness, longitude, latitude, annual mean temperature, annual mean precipitation, elevation, frost-free period, suitability indexes for wheat, barley, oat, cotton, and olive, distance to Constantinople, distance to shore, length of primary and secondary rivers per surface area (km2) in 2010, distance to the nearest custom gate. Historical controls are whether the district is within the Seljuk Sultanate and whether it is within one and a half travel distance to 19^{th} century major ports. Sample Mean presents the sample mean of the outcome variable in percentages. The first stage of the F-stat reports the Kleibergen-Paap rk Wald F statistic of the 2SLS estimation for the first stage. Standard errors are clustered at the province level. The number of clusters indicates the total number of provinces in the standard error adjustments. ***p<0.01, **p<0.05, *p<0.1

B.7 Spatial Correlation Simulations

Table B.17: Development Index Simulation Results

		Explanatory Noise					Dependent Noise			
Correlation Range (km)	Mean z-value of Moran's I	Outperforming Share	p=0.05	p=0.01	p=0.001	p=0.0001	p=0.05	p=0.01	p=0.001	p=0.0001
25	15.14	0.00	0.04	0.01	0.00	0.00	0.08	0.02	0.00	0.00
50	26.25	0.00	0.05	0.01	0.00	0.00	0.08	0.03	0.01	0.00
75	31.56	0.00	0.06	0.02	0.00	0.00	0.14	0.05	0.01	0.00
100	34.79	0.00	0.05	0.01	0.00	0.00	0.12	0.05	0.02	0.00
125	36.78	0.00	0.06	0.02	0.00	0.00	0.14	0.05	0.01	0.00
150	38.10	0.00	0.06	0.02	0.00	0.00	0.13	0.04	0.01	0.01
175	39.23	0.00	0.06	0.01	0.00	0.00	0.13	0.05	0.01	0.01
200	39.89	0.00	0.06	0.01	0.00	0.00	0.15	0.06	0.02	0.01
225	40.54	0.00	0.06	0.01	0.00	0.00	0.13	0.04	0.02	0.00
250	40.84	0.00	0.07	0.01	0.00	0.00	0.11	0.05	0.01	0.01
275	41.40	0.00	0.06	0.02	0.00	0.00	0.14	0.05	0.02	0.01
300	41.57	0.00	0.07	0.01	0.00	0.00	0.13	0.06	0.02	0.01
325	41.98	0.00	0.05	0.01	0.00	0.00	0.15	0.05	0.02	0.00
350	42.07	0.00	0.05	0.01	0.00	0.00	0.15	0.07	0.02	0.01
375	42.38	0.00	0.06	0.01	0.00	0.00	0.14	0.05	0.02	0.00
400	42.58	0.00	0.08	0.01	0.00	0.00	0.15	0.06	0.01	0.00
425	42.63	0.00	0.07	0.02	0.00	0.00	0.15	0.06	0.01	0.01
450	42.87	0.00	0.06	0.01	0.00	0.00	0.15	0.06	0.01	0.00
475	42.81	0.00	0.06	0.01	0.00	0.00	0.15	0.07	0.01	0.01
500	42.89	0.00	0.06	0.02	0.00	0.00	0.17	0.06	0.02	0.01

Notes: We generate 1000 times spatial noise for each correlation range in kilometers. We regress the development index on spatial noise and report the results under the column of explanatory noise and spatial noise on the distance to the nearest ABCFM mission and report the results under the column of dependent noise. Outperforming share presents the share of regressions where the spatial noise used as the explanatory variable outperforms our original explanatory variable, distance to the nearest ABCFM mission.

Table B.18: GDP per Capita Simulation Results

		Explanatory Noise					Dependent Noise			
Correlation Range (km)	Mean z-value of Moran's I	Outperforming Share	p=0.05	p=0.01	p=0.001	p=0.0001	p=0.05	p=0.01	p=0.001	p=0.0001
25	15.14	0.00	0.04	0.01	0.00	0.00	0.08	0.02	0.00	0.00
50	26.25	0.00	0.05	0.01	0.00	0.00	0.08	0.03	0.01	0.00
75	31.56	0.00	0.07	0.01	0.00	0.00	0.14	0.05	0.01	0.00
100	34.79	0.00	0.07	0.02	0.00	0.00	0.12	0.05	0.02	0.00
125	36.78	0.00	0.06	0.02	0.00	0.00	0.14	0.05	0.01	0.00
150	38.10	0.00	0.06	0.01	0.00	0.00	0.13	0.04	0.01	0.01
175	39.23	0.00	0.06	0.02	0.00	0.00	0.13	0.05	0.01	0.01
200	39.89	0.00	0.08	0.02	0.01	0.01	0.15	0.06	0.02	0.01
225	40.54	0.00	0.06	0.02	0.00	0.00	0.13	0.04	0.02	0.00
250	40.84	0.00	0.07	0.03	0.00	0.00	0.11	0.05	0.01	0.01
275	41.40	0.00	0.06	0.02	0.00	0.00	0.14	0.05	0.02	0.01
300	41.57	0.00	0.08	0.02	0.00	0.00	0.13	0.06	0.02	0.01
325	41.98	0.00	0.07	0.02	0.01	0.01	0.15	0.05	0.02	0.00
350	42.07	0.00	0.06	0.02	0.00	0.00	0.15	0.07	0.02	0.01
375	42.38	0.00	0.08	0.02	0.01	0.01	0.14	0.05	0.02	0.00
400	42.58	0.00	0.07	0.02	0.00	0.00	0.15	0.06	0.01	0.00
425	42.63	0.00	0.07	0.02	0.00	0.00	0.15	0.06	0.01	0.01
450	42.87	0.00	0.08	0.02	0.00	0.00	0.15	0.06	0.01	0.00
475	42.81	0.00	0.07	0.02	0.00	0.00	0.15	0.07	0.01	0.01
500	42.89	0.00	0.08	0.02	0.00	0.00	0.17	0.06	0.02	0.01

Notes: We generate 1000 times spatial noise for each correlation range in kilometers. We regress the log of GDP per capita on spatial noise and report the results under the column of explanatory noise and spatial noise on the distance to the nearest ABCFM mission and report the results under the column of dependent noise. Outperforming share presents the share of regressions where the spatial noise used as the explanatory variable outperforms our original explanatory variable, distance to the nearest ABCFM mission.

B.8 Data Appendix

We obtain a substantial portion of our data from Seyhun Orcan Sakalli, which he already used in his study Sakalli (2017). By the way, we owe thanks to him for sharing the data with us. His data is very helpful chiefly for constructing our control variables and the population variables in 1881 and 1914 which come from the Ottoman Population Censuses. The variable definitions are also available online at
https://sites.google.com/site/sosakalli/ on June 1, 2023.
Although he explains the construction of variables, we also explain how we construct the variables we use in our estimations in the next part. Before going over the details, there are two important points we touch upon. The first issue is that the borders of districts have changed over time. At first, there were 369 districts in 1914 while there were 872 in 2000. To match the districts, we follow the strategy by Sakalli (2017) as he exploits various official documents to track the border changes over time. For more details, see the link provided above.

The second issue is the historical population counts in 1881 and 1914, which are the years of the two Ottoman Population Censuses. We had access to the digitized version of those censuses from the data set of Sakalli (2017). For more, details see the link above.

The remaining data sets are obtained from several sources and we document them in the next section.

B.8.1 District Level Data Sources

The district-level data comes from various sources.

1) District Development Ranking Data in 2004 by State Planning Organization

Source location: The original data can be downloaded at
https://www.sanayi.gov.tr/merkez-birimi/b94224510b7b/sege on June 1, 2023.

2) District Level GDP Data in 1996 by the Turkish Statistical Institute

Source location: The original data can be downloaded at
https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

3) District Level Fertility, and Educational Attainment Data in 2000 by the Turkish

Statistical Institute

Source location: The original data can be downloaded at https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

4) District Level Village Inventory Census in 1997 by the Turkish Statistical Institute

Source location: The original data can be downloaded at https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

5) District Level Population Census in 1965 by the Turkish Statistical Institute

Source location: The original data can be downloaded at https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

6) District Level Agriculture Census in 1927 by the Turkish Statistical Institute

Source location: The original data can be downloaded at https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

7) District Level Profession Census in 1927 by the Turkish Statistical Institute

Source location: The original data can be downloaded at https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

8) District Level Population Census in 1927 by the Turkish Statistical Institute

Source location: The original data can be downloaded at https://kutuphane.tuik.gov.tr/yordambt/yordam.php# on June 1, 2023.

B.8.2 Geographical Distribution of ABCFM Missions

To identify the mission locations, we benefit from three main sources. We scanned all annual reports of ABCFM from 1810 to 1930 and all almanacs. The annual reports and almanacs can be obtained online from the HathiTrust Digital Library. For some documents, we also visited the Houghton Library at Harvard University. In each annual report, the locations of main stations and principal out-stations are documented with various details such as the geographic coordinates, number of people working, number of students, etc. After identifying the mission locations, we categorize those sites as treated and placebo. The treated sites got the full treatment through either the presence of a main station or a principal outstation. The placebo sites had never got the treatment but they were planned to have a mission facility. Besides, we identify the missions in the Syrian/Nestorian regions through the annual reports. Those missions were left to the Presbyterian Church in 1870.

We use ArcGIS software to calculate the distance to ABCFM missions, and placebo missions in kilometers from the centroid of a certain district. The spatial distribution of the ABCFM missions is presented in Figure 2.1.

Source location: The original annual reports and almanacs can be downloaded at https://www.hathitrust.org/ on June 1, 2023.

B.8.3 Outcome Variables

Development Index in 2000 is provided by the State Planning Organization (SPO) in 2004 District Development Ranking Data. It is a standardized variable with a mean of 0 and a standard deviation of 0. The SPO computes the development index by employing a principal component analysis using 32 socioeconomic variables.

Logarithm of GDP per Capita in 1996 is provided by Turkish Statistical Institute (TURKSTAT) in 1996 District GDP Data. We take the logarithm of the GDP per capita.

Industrial Population Share in 2000 is provided by the SPO in 2004 District Development Ranking Data. It is the percentage of the population employed in the industrial sector.

Agricultural Population Share in 2000 is provided by the SPO in 2004 District Development Ranking Data. It is the percentage of the population employed in the agriculture sector.

National Agricultural Production Share in 2000 is provided by the SPO in 2004 District Development Ranking Data. It is the percentage share of national agricultural production in a certain district.

Log of Agricultural Labor Productivity in 2000 is the logarithm of agricultural value divided by the number of people working in the agriculture sector. The relevant data comes from the 2004 District Development Ranking Data by the SPO and the 2000 Population Census by the TURKSTAT.

Literacy Rate in 2000 is provided by the SPO in 2004 District Development Ranking Data. It is the percentage of the literate population older than 6.

Female High School Completion Rate in 2000 is provided by the TURKSTAT in the 2000 Population Census. It is the percentage of females who completed high school.

Male High School Completion Rate in 2000 is provided by the TURKSTAT in the 2000 Population Census. It is the percentage of males who completed high school.

Female Primary School Completion Rate in 2000 is provided by the TURKSTAT in the 2000 Population Census. It is the percentage of females who completed primary school.

Male Primary School Completion Rate in 2000 is provided by the TURKSTAT in the 2000 Population Census. It is the percentage of males who completed primary school.

Fertility Rate is provided by the TURKSTAT in the 2000 Population Census. It is the ratio of total children divided by the total number of women.

Infant Mortality is provided by the SPO in 2004 District Development Ranking Data.

Poverty Rate is provided by the TURKSTAT in the 2000 Population Census.

Land Concentration is provided by the TURKSTAT in the 1997 Village Inventory Census Population. The variable is taken from Sakalli (2017) and is a GINI index representing land concentration.

Log Population Density is provided by the SPO in 2004 District Development Ranking Data.

Non-Resident Share is provided by the TURKSTAT in the 2000 Population Census. It is the percentage of the non-resident population.

Urbanization Rate is provided by the SPO in 2004 District Development Ranking Data. It is the percentage of the population living in urban areas.

Protestant Share is provided by **?**. It is the percentage of the Protestant population in 1914. The original data comes from Karpat (1985).

Labor Productivity in Agriculture in 1927 is provided by the TURKSTAT in the 1927 Census of Agriculture. It is the total agricultural value in 1927 divided by the number of people working in the agriculture sector.

Land Productivity in Agriculture in 1927 is provided by the TURKSTAT in the 1927 Census of Agriculture. It is the total agricultural value in 1927 divided by the hectare cultivated.

Industrial Population Share in 1927 is provided by the TURKSTAT in the 1927 District Level of Profession Census. It is the percentage of the population employed in the industrial sector in 1927.

Commerce Population Share in 1927 is provided by the TURKSTAT in the 1927 District Level of Profession Census. It is the percentage of the population employed in the commerce sector in 1927.

Log Population Density in 1927 is provided by the TURKSTAT in the 1927 District Level of Profession Census.

Female Literacy in 1927 is provided by Sakalli (2017). The original data comes from the TURKSTAT in the 1927 District Level Population Census. It is the percentage of females older than 6 who are literate.

Male Literacy in 1927 is provided by Sakalli (2017). The original data comes from the TURKSTAT in the 1927 District Level Population Census. It is the percentage of males older than 6 who are literate.

Female Enrollment Rate in 1927 is provided by Sakalli (2017). The original data comes from the TURKSTAT in the 1927 District Level Population Census. It is the percentage of females between 6-14 years old who are enrolled in a school.

Male Enrollment Rate in 1927 is provided by Sakalli (2017). The original data comes from the TURKSTAT in the 1927 District Level Population Census. It is the percentage of males between 6-14 years old who are enrolled in a school.

Log Population Density in 1881 is provided by Sakalli (2017). The original data comes from Karpat (1985).

Log Population Density in 1914 is provided by Sakalli (2017). The original data comes from Karpat (1985).

Log Population Density in 1927 is provided by Sakalli (2017). The original data comes from the TURKSTAT in the 1927 District Level Population Census.

Log Population Density in 1965 is provided by Sakalli (2017). The original data comes from the TURKSTAT in the 1965 Population Census.

B.8.4 Control Variables

Average Ruggedness is calculated by using ArcGIS software.

Longitude is the value of longitude crossing in the district center.

Latitude is the value of latitude crossing in the district center.

Temperature is provided by Sakalli (2017). The mean annual temperature over the period 1960-1990 comes originally from GAEZ v3.0 data set (Iiasa (2012)). It is calculated at the district level using the zonal statistics tool of the ArcGIS® software in Celsius.

Temperature is provided by Sakalli (2017). The mean annual precipitation over the period 1960-1990 comes originally from GAEZ v3.0 data set (Iiasa (2012)). It is calculated at the district level using the zonal statistics tool of the ArcGIS® software in millimeters.

Elevation is provided by Sakalli (2017). It originally comes from NASA's Shuttle Radar Topography Mission (SRTM) data set (Jarvis et al. (2008)). We calculate the mean elevation of a certain district by exploiting the zonal statistics tool of ArcGIS® software in meters.

Frost-free Period is provided by Sakalli (2017). The mean annual frost-free period over the period 1960-1990 comes originally from GAEZ v3.0 data set (Iiasa (2012)). It is calculated at the district level using the zonal statistics tool of the ArcGIS® software.

Suitability Index for Wheat is provided by Sakalli (2017).

Suitability Index for Barley is provided by Sakalli (2017).

Suitability Index for Oat is provided by Sakalli (2017).

Suitability Index for Cotton is provided by Sakalli (2017).

Suitability Index for Olive is provided by Sakalli (2017).

Suitability indices originally come from GAEZ v3.0 data set (Iiasa (2012)). The concerning indices reflect the suitability of land for cultivating crops with low inputs and without irrigation (rain-fed) and take values between 0 and 100. We calculate the mean suitability to cultivate wheat, cotton, and olive using the zonal statistics tool of ArcGIS® software.

Distance to Constantinople (Days) is provided by Sakalli (2017). It is originally computed using ArcGIS® software in travel days.

Distance to Shore (Days) is provided by Sakalli (2017). It is originally computed using ArcGIS® software in travel days.

Length of Primary Rivers (km) is provided by Sakalli (2017). Information on rivers comes from the USGS HydroSHEDS data set (Lehner et al. (2008)). It is calculated using ArcGIS® software and normalized by the surface area of districts.

Length of Secondary Rivers (km) is provided by Sakalli (2017). Information on rivers comes from the USGS HydroSHEDS data set (Lehner et al. (2008)). It is calculated using ArcGIS® software and normalized by the surface area of districts.

Distance to Nearest Custom Gate (km) the distance to the nearest custom gate from the centroid of a certain district. It is calculated using ArcGIS® software.

Seljuk Sultanate is provided by Sakalli (2017). It is an indicator variable equal to one if a certain district is within the territories of the Seljuk Sultanate of Rum.

19^{th} **Century Ports** is provided by Sakalli (2017). It is the travel distance to the major nineteenth-century ports, namely the ports of Constantinople (Istanbul), Trabzon (Trebizond), Mersin, Iskenderun (Alexandretta), Samsun, and Izmir (Smyrna) using ArcGIS®software. It is an indicator variable equal to one if a certain district within one and a half travel days from one of the major nineteenth-century ports.

Share of Armenians in 1881 is provided by Sakalli (2017). It is the percentage of the Armenian population in a certain district in 1881. The original source is Karpat (1985).

Share of Greeks in 1881 is provided by Sakalli (2017). It is the percentage of the Greek population in a certain district in 1881. The original source is Karpat (1985).

Share of Armenians in 1881*Distance (in km) is the interaction variable of the share of Armenians in 1881 with the distance to the nearest ABCFM missions or distance to the nearest placebo mission.

Share of Greeks in 1881*Distance (in km) is the interaction variable of the share of Greeks in 1881 with the distance to the nearest ABCFM missions or distance to the nearest placebo mission.

East-West Control at the Longitude of 35 is an indicator variable equal to one if the center of a certain district is further east than the longitude of 35 degrees.

East-West Control at the Longitude of 40 is an indicator variable equal to one if the center of a certain district is further east than the longitude of 40 degrees.

Province Fixed Effects is a set of indicator variables for the Turkish provinces.

B.8.5 Simulations

We follow the procedure suggested by Kelly (2019). We run placebo regressions. To run those regressions, we generate artificial dependent and explanatory variables by simulating the data 1000 times. The noise variable has the same correlation with the residuals of our original regressions. We use a Matern function with an exponential smoothing parameter. In words, the spatial correlation decays exponentially. The choice parameter is the correlation range and we estimate the spatial correlation in various ranges. To generate the noise variable, we use the longitude and latitude of a certain district.

After generating the noise variable, we first run regressions where the noise variable is the explanatory variable. Then, we compute the number of specifications in which the noise variable outperforms our original explanatory regarding the t-values. Put another way, we count the specifications in which the absolute value of the t-statistic is greater than the absolute value of the t-statistics of our original independent variable, distance to the nearest ABCFM mission in kilometers. Moreover, we also report the percentage of the specifications in which the coefficient of the noise variable is significant at the conventional levels of significance from 0.05 to 0.0001.

We also replace the dependent variable with the noise variable. Similarly, we count the number of specifications in which our original explanatory variable explains the noise variable at the conventional levels of significance from 0.05 to 0.0001. In addition, we present the mean z-value of Moran's I in our simulation tables.

APPENDIX C

Appendix to Chapter 3

C.1 Robustness Checks on Participation Bias

Table C.1: Education Reform vs Variation in Portfolio Choices

	(1) Stock	(2) Risky Asset	(3) Bond	(4) Blue-Chip Stock	(5) Index Stock	(6) Fund	(7) Summary Index
				Full			
Education Reform	-0.001	-0.002	-0.000	0.002	0.001	-0.002	0.000
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)
Control Mean	0.55	0.65	0.01	0.13	0.25	0.32	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	36.81	48.78	42.36	48.14	43.56	34.57	50.90
Observations	654324	864439	755220	864439	773454	613254	893811
				Male			
Education Reform	-0.002	-0.001	0.001***	0.002	0.002	0.000	0.003*
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.003)	(0.002)
Control Mean	0.57	0.67	0.01	0.14	0.25	0.31	0.10
Control SD	0.49	0.46	0.08	0.30	0.39	0.45	0.46
Bandwidth	38.34	47.53	57.96	48.36	52.95	39.18	55.22
Observations	505503	618526	748167	636653	682633	517965	722107
				Female			
Education Reform	0.003	-0.003	-0.003***	-0.000	-0.001	-0.006	-0.006**
	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
Control Mean	0.48	0.60	0.01	0.11	0.22	0.37	0.04
Control SD	0.49	0.48	0.10	0.28	0.38	0.47	0.46
Bandwidth	42.77	50.37	37.08	48.29	58.30	41.00	53.57
Observations	198736	235673	175150	227786	273839	193815	249795

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the share of wealth invested in stocks in column 1, in risky assets in column 2, in bonds in column 3, in blue-chip stocks in column 4, in index stocks in column 5, and in funds in column 6. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for quadratic polynomials of stock market participation rate, the month of birth fixed effects, and birth registration certificate region fixed effects with dummy variables for each month and each relevant regions, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.2: Education Reform vs Causal Channels

	(1) Risk Score	(2) Single Stock	(3) Risky Inertia	(4) Wealth (in logs)	(5) Summary Index
			Full		
Education Reform	0.000	0.002	-0.002	-0.028	-0.002
	(0.005)	(0.002)	(0.004)	(0.017)	(0.002)
Control Mean	3.68	0.19	0.13	6.95	-0.02
Control SD	0.84	0.39	0.77	3.73	0.53
Bandwidth	54.66	32.41	61.83	50.48	47.93
Observations	470289	579265	678620	893811	839920
			Male		
Education Reform	-0.002	0.003	-0.002	-0.021	-0.000
	(0.005)	(0.002)	(0.004)	(0.019)	(0.002)
Control Mean	3.70	0.19	0.14	7.00	-0.01
Control SD	0.82	0.39	0.79	3.79	0.52
Bandwidth	53.77	37.83	74.22	48.09	48.97
Observations	372189	493511	620707	636653	636653
			Female		
Education Reform	0.010	-0.001	-0.012	-0.049*	-0.007**
	(0.009)	(0.003)	(0.007)	(0.025)	(0.003)
Control Mean	3.57	0.18	0.13	6.77	-0.05
Control SD	0.90	0.38	0.75	3.52	0.54
Bandwidth	56.23	48.56	47.63	49.79	54.70
Observations	94156	227786	126024	231400	254852

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the logarithm of risk scores in column 1, an indicator variable equal to one if the investor invests all money into a single stock in column 2, the risk share inertia in column 3, and the logarithm of portfolio size in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for quadratic polynomials of stock market participation rate, the month of birth fixed effects, and birth registration certificate region fixed effects with dummy variables for each month and each relevant regions, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.3: Education Reform vs Portfolio Returns

	(1) One Month	(2) Three Months	(3) Six Months	(4) Twelve Months	(5) Summary Index
			Full		
Education Reform	-0.269	-0.480***	-0.351**	0.508	-0.006**
	(0.175)	(0.157)	(0.144)	(0.427)	(0.003)
Control Mean	-42.88	-42.97	-41.98	105.31	-0.04
Control SD	53.72	54.99	54.31	121.21	0.94
Bandwidth	54.42	44.26	43.67	45.59	54.99
Observations	949208	779998	762197	808937	963363
			Male		
Education Reform	-0.281	-0.573***	-0.390*	0.532	-0.008**
	(0.227)	(0.208)	(0.206)	(0.533)	(0.004)
Control Mean	-42.17	-42.29	-41.50	107.33	-0.02
Control SD	56.00	57.30	56.33	121.03	0.99
Bandwidth	47.11	40.82	40.68	46.31	42.78
Observations	610844	524720	524791	608200	556484
			Female		
Education Reform	0.007	0.101	0.169	-0.338	0.001
	(0.295)	(0.290)	(0.331)	(0.907)	(0.005)
Control Mean	-44.90	-44.50	-42.86	98.08	-0.07
Control SD	47.82	47.85	48.04	120.47	0.82
Bandwidth	57.80	65.65	64.52	58.09	71.80
Observations	263546	299823	295012	273839	335135

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the one-month return in column 1, the three-month return in column 2, the six-month return in column 3, and the annual return in percentages in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for quadratic polynomials of stock market participation rate, the month of birth fixed effects, and birth registration certificate region fixed effects with dummy variables for each month and each relevant regions, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C.2 Local Linear RD Estimates with Fixed Bandwidth

Table C.4: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	0.368***	-0.019***	0.088***	0.040***	0.018*	0.070***
	(0.102)	(0.006)	(0.010)	(0.010)	(0.009)	(0.018)
Control Mean	8.98	0.92	0.62	0.47	0.25	0.40
Control SD	4.69	0.26	0.49	0.50	0.44	0.83
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00
Observations	61311	61311	61311	61311	61311	61311
			Male			
Education Reform	0.273**	-0.001	0.065***	0.025**	0.002	0.050***
	(0.108)	(0.006)	(0.012)	(0.012)	(0.012)	(0.019)
Control Mean	9.78	0.96	0.72	0.54	0.28	0.54
Control SD	4.29	0.19	0.45	0.50	0.45	0.77
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00
Observations	29523	29523	29523	29523	29523	29523
			Female			
Education Reform	0.454***	-0.035***	0.110***	0.053***	0.032**	0.089***
	(0.172)	(0.009)	(0.017)	(0.017)	(0.013)	(0.030)
Control Mean	8.24	0.89	0.53	0.41	0.23	0.27
Control SD	4.91	0.32	0.50	0.49	0.42	0.87
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00
Observations	31788	31788	31788	31788	31788	31788

Notes: Local linear RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have a fixed bandwidth of 60 with a triangular-type kernel function in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.5: Education Reform vs Participation in Stock Market

	(1) Market Participant	(2) Direct Stock	(3) Index Stock	(4) Blue-Chip Stock	(5) Risky Asset	(6) Bond	(7) Fund	(8) Transaction Account	(9) Summary Index
					Full				
Education Reform	0.204***	0.142***	0.099***	0.059***	0.142***	0.004	0.041*	1.101***	0.095***
	(0.045)	(0.034)	(0.027)	(0.022)	(0.037)	(0.003)	(0.023)	(0.110)	(0.022)
Control Mean	9.14	5.25	3.51	2.27	6.18	0.10	3.06	55.81	0.48
Control SD	0.44	0.26	0.19	0.15	0.30	0.01	0.22	1.36	0.19
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	119	119	119	119	119
					Male				
Education Reform	0.250***	0.196***	0.129***	0.075**	0.199***	0.015***	0.042	1.018***	0.103***
	(0.058)	(0.053)	(0.039)	(0.035)	(0.054)	(0.005)	(0.034)	(0.107)	(0.023)
Control Mean	13.26	7.98	5.33	3.50	9.23	0.13	4.32	72.77	0.49
Control SD	0.51	0.35	0.27	0.21	0.38	0.02	0.27	1.13	0.17
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	119	119	119	119	119
					Female				
Education Reform	0.157***	0.088***	0.069***	0.042**	0.085**	-0.009***	0.040*	1.192***	0.107***
	(0.043)	(0.032)	(0.022)	(0.017)	(0.035)	(0.003)	(0.021)	(0.143)	(0.033)
Control Mean	4.89	2.45	1.63	1.01	3.05	0.06	1.77	38.34	0.18
Control SD	0.39	0.17	0.13	0.09	0.22	0.01	0.19	1.72	0.27
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	119	119	119	119	119	119	119	119	119

Notes: Local linear RD estimates in all columns. All columns use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2021. The unit of analysis is the month-year birth cohorts in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those participating in the stock market in column 1, the percentage of those directly holding stocks in column 2, the percentage of those directly holding index stocks in column 3, the percentage of those directly holding blue-chip stocks in column 4, the percentage of those holding risky assets in column 5, the percentage of those directly holding bonds in column 6, the percentage of those directly holding funds in their stock market portfolios in column 7, the percentage of those having a transaction account in column 8. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.6: Education Reform vs Variation in Portfolio Choices

	(1) Stock	(2) Risky Asset	(3) Bond	(4) Blue-Chip Stock	(5) Index Stock	(6) Fund	(7) Summary Index
				Full			
Education Reform	0.002	0.001	0.000	0.002*	0.001	-0.003	0.002
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)
Control Mean	0.54	0.65	0.01	0.13	0.24	0.33	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	1049097	1049097	1049097	1049097	1049097	1049097	1049097
				Male			
Education Reform	0.003	0.001	0.001***	0.003**	0.002	-0.002	0.005**
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)
Control Mean	0.54	0.65	0.01	0.13	0.24	0.33	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	771442	771442	771442	771442	771442	771442	771442
				Female			
Education Reform	0.002	-0.002	-0.002***	0.000	-0.001	-0.005	-0.005*
	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
Control Mean	0.54	0.65	0.01	0.13	0.24	0.33	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	60.00	60.00	60.00	60.00	60.00	60.00	60.00
Observations	277655	277655	277655	277655	277655	277655	277655

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the share of wealth invested in stocks in column 1, in risky assets in column 2, in bonds in column 3, in blue-chip stocks in column 4, in index stocks in column 5, and in funds in column 6. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.7: Education Reform vs Causal Channels

	(1) Risk Score	(2) Single Stock	(3) Risky Inertia	(4) Wealth (in logs)	(5) Summary Index
			Full		
Education Reform	-0.001	0.001	-0.003	-0.016	-0.001
	(0.005)	(0.001)	(0.004)	(0.016)	(0.002)
Control Mean	3.68	0.18	0.13	6.96	-0.02
Control SD	0.84	0.38	0.77	3.74	0.53
Bandwidth	60.00	60.00	60.00	60.00	60.00
Observations	511035	1049097	654916	1049097	1049097
			Male		
Education Reform	-0.003	0.001	-0.003	-0.010	0.000
	(0.005)	(0.002)	(0.004)	(0.018)	(0.002)
Control Mean	3.68	0.18	0.13	6.96	-0.02
Control SD	0.84	0.38	0.77	3.74	0.53
Bandwidth	60.00	60.00	60.00	60.00	60.00
Observations	412434	771442	497259	771442	771442
			Female		
Education Reform	0.007	-0.001	-0.005	-0.032	-0.006*
	(0.008)	(0.003)	(0.007)	(0.026)	(0.003)
Control Mean	3.68	0.18	0.13	6.96	-0.02
Control SD	0.84	0.38	0.77	3.74	0.53
Bandwidth	60.00	60.00	60.00	60.00	60.00
Observations	98601	277655	157657	277655	277655

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the logarithm of risk scores in column 1, an indicator variable equal to one if the investor invests all money into a single stock in column 2, the risk share inertia in column 3, and the logarithm of portfolio size in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.8: Education Reform vs Portfolio Returns

	(1) One Month	(2) Three Months	(3) Six Months	(4) Twelve Months	(5) Summary Index
			Full		
Education Reform	-0.261	-0.358**	-0.262*	0.620	-0.005*
	(0.173)	(0.157)	(0.148)	(0.394)	(0.003)
Control Mean	-42.74	-42.74	-41.67	104.13	-0.03
Control SD	53.52	54.29	53.64	120.45	0.94
Bandwidth	60.00	60.00	60.00	60.00	60.00
Observations	1033659	1033659	1033786	1049097	1049097
			Male		
Education Reform	-0.310	-0.456**	-0.374**	0.885*	-0.006*
	(0.210)	(0.191)	(0.188)	(0.484)	(0.003)
Control Mean	-42.74	-42.74	-41.67	104.13	-0.03
Control SD	53.52	54.29	53.64	120.45	0.94
Bandwidth	60.00	60.00	60.00	60.00	60.00
Observations	761811	761811	761903	771442	771442
			Female		
Education Reform	-0.134	-0.091	0.050	-0.150	-0.001
	(0.305)	(0.310)	(0.347)	(0.903)	(0.006)
Control Mean	-42.74	-42.74	-41.67	104.13	-0.03
Control SD	53.52	54.29	53.64	120.45	0.94
Bandwidth	60.00	60.00	60.00	60.00	60.00
Observations	271848	271848	271883	277655	277655

Notes: Local linear RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the one-month return in column 1, the three-month return in column 2, the six-month return in column 3, and the annual return in percentages in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a linear polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C.3 Local Quadratic RD Estimates

Table C.9: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	0.403***	-0.009	0.049***	0.036**	0.032*	0.076***
	(0.128)	(0.008)	(0.013)	(0.016)	(0.017)	(0.023)
Control Mean	9.02	0.93	0.64	0.50	0.27	0.40
Control SD	4.69	0.26	0.48	0.50	0.45	0.83
Bandwidth	86.98	71.67	67.14	55.16	58.74	87.83
Observations	91154	74274	69926	57247	60380	92228
			Male			
Education Reform	0.356*	0.003	0.036*	0.034	0.031	0.066*
	(0.198)	(0.009)	(0.019)	(0.027)	(0.020)	(0.035)
Control Mean	9.75	0.96	0.72	0.55	0.29	0.53
Control SD	4.32	0.19	0.45	0.50	0.45	0.77
Bandwidth	76.24	70.74	70.97	59.08	58.10	76.40
Observations	38624	35368	35368	29523	29048	38624
			Female			
Education Reform	0.445*	-0.022	0.096***	0.041	0.033	0.086*
	(0.255)	(0.014)	(0.023)	(0.028)	(0.023)	(0.044)
Control Mean	8.37	0.89	0.53	0.43	0.26	0.29
Control SD	4.97	0.31	0.50	0.50	0.44	0.88
Bandwidth	83.75	80.66	95.05	64.15	60.34	84.91
Observations	45357	43713	52247	34567	32608	46293

Notes: Local quadratic RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.10: Education Reform vs Participation in Stock Market

	(1) Market Participant	(2) Direct Stock	(3) Index Stock	(4) Blue-Chip Stock	(5) Risky Asset	(6) Bond	(7) Fund	(8) Transaction Account	(9) Summary Index
					Full				
Education Reform	0.194***	0.070	0.085**	0.055*	0.107**	0.003	0.056*	0.410***	0.053*
	(0.062)	(0.049)	(0.036)	(0.031)	(0.050)	(0.004)	(0.032)	(0.154)	(0.031)
Control Mean	9.17	5.28	3.52	2.29	6.19	0.10	3.10	55.54	0.48
Control SD	0.44	0.25	0.19	0.15	0.29	0.01	0.23	1.30	0.19
Bandwidth	66.05	53.95	67.73	70.68	63.28	69.78	68.77	31.44	57.89
Observations	133	107	135	141	127	139	137	63	115
					Male				
Education Reform	0.246***	0.107	0.085	0.056	0.151**	0.019***	0.078*	0.510***	0.070**
	(0.078)	(0.075)	(0.055)	(0.050)	(0.071)	(0.006)	(0.044)	(0.140)	(0.033)
Control Mean	13.28	8.02	5.34	3.52	9.23	0.14	4.41	72.27	0.49
Control SD	0.52	0.34	0.27	0.22	0.38	0.02	0.30	0.94	0.17
Bandwidth	63.04	55.84	62.06	65.34	60.55	91.55	76.97	30.27	55.98
Observations	127	111	125	131	121	183	153	61	111
					Female				
Education Reform	0.158***	0.080*	0.074***	0.048**	0.087*	-0.016***	0.038	0.312*	0.058
	(0.057)	(0.042)	(0.028)	(0.020)	(0.046)	(0.004)	(0.031)	(0.180)	(0.044)
Control Mean	4.93	2.43	1.63	1.03	3.05	0.06	1.80	38.21	0.21
Control SD	0.37	0.18	0.14	0.10	0.23	0.01	0.20	1.61	0.27
Bandwidth	74.74	74.29	77.19	83.83	80.07	65.49	69.27	35.17	67.75
Observations	149	149	155	167	161	131	139	71	135

Notes: Local quadratic RD estimates in all columns. All columns use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2021. The unit of analysis is the month-year birth cohorts in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those participating in the stock market in column 1, the percentage of those directly holding stocks in column 2, the percentage of those directly holding index stocks in column 3, the percentage of those directly holding blue-chip stocks in column 4, the percentage of those holding risky assets in column 5, the percentage of those directly holding bonds in column 6, the percentage of those directly holding funds in their stock market portfolios in column 7, the percentage of those having a transaction account in column 8. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth for columns. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.11: Education Reform vs Variation in Portfolio Choices

	(1) Stock	(2) Risky Asset	(3) Bond	(4) Blue-Chip Stock	(5) Index Stock	(6) Fund	(7) Summary Index
				Full			
Education Reform	-0.000	-0.003	-0.000	0.001	0.001	-0.002	-0.001
	(0.003)	(0.003)	(0.000)	(0.001)	(0.002)	(0.003)	(0.002)
Control Mean	0.54	0.65	0.01	0.13	0.24	0.33	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	50.46	55.33	68.04	57.60	53.37	49.23	51.33
Observations	893811	982059	1210679	1017341	944779	878229	910469
				Male			
Education Reform	-0.001	-0.002	0.001**	0.001	0.002	-0.000	0.001
	(0.003)	(0.003)	(0.000)	(0.002)	(0.002)	(0.003)	(0.002)
Control Mean	0.57	0.67	0.01	0.14	0.25	0.31	0.10
Control SD	0.49	0.46	0.09	0.30	0.39	0.45	0.46
Bandwidth	51.12	56.21	69.31	50.41	53.06	51.73	48.63
Observations	670196	735823	901753	658138	694984	670196	636653
				Female			
Education Reform	0.001	-0.003	-0.003***	-0.001	-0.001	-0.005	-0.007**
	(0.006)	(0.005)	(0.001)	(0.002)	(0.003)	(0.006)	(0.003)
Control Mean	0.47	0.59	0.01	0.11	0.22	0.37	0.03
Control SD	0.49	0.48	0.11	0.28	0.38	0.47	0.47
Bandwidth	55.56	67.44	70.62	65.78	66.33	58.63	77.02
Observations	259952	316453	331113	306270	311376	273839	364007

Notes: Local quadratic RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the share of wealth invested in stocks in column 1, in risky assets in column 2, in bonds in column 3, in blue-chip stocks in column 4, in index stocks in column 5, and in funds in column 6. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.12: Education Reform vs Causal Channels

	(1) Risk Score	(2) Single Stock	(3) Risky Inertia	(4) Wealth (in logs)	(5) Summary Index
			Full		
Education Reform	-0.004	0.004**	-0.003	-0.032	-0.000
	(0.007)	(0.002)	(0.004)	(0.026)	(0.002)
Control Mean	3.68	0.18	0.15	6.95	-0.02
Control SD	0.84	0.39	0.82	3.73	0.53
Bandwidth	59.14	48.18	98.00	52.42	51.17
Observations	511035	864439	1061389	927699	910469
			Male		
Education Reform	-0.006	0.005**	-0.003	-0.024	0.002
	(0.008)	(0.002)	(0.006)	(0.030)	(0.003)
Control Mean	3.70	0.19	0.14	7.01	-0.01
Control SD	0.83	0.39	0.81	3.80	0.53
Bandwidth	58.34	50.34	84.71	52.40	52.68
Observations	406913	658138	701791	682633	682633
			Female		
Education Reform	0.006	-0.000	-0.021**	-0.031	-0.005
	(0.011)	(0.004)	(0.009)	(0.032)	(0.004)
Control Mean	3.57	0.17	0.15	6.79	-0.05
Control SD	0.90	0.38	0.78	3.54	0.54
Bandwidth	84.65	59.14	67.12	62.39	58.69
Observations	138627	277655	179374	292028	273839

Notes: Local quadratic RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the logarithm of risk scores in column 1, an indicator variable equal to one if the investor invests all money into a single stock in column 2, the risk share inertia in column 3, and the logarithm of portfolio size in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.13: Education Reform vs Portfolio Returns

	(1) One Month	(2) Three Months	(3) Six Months	(4) Twelve Months	(5) Summary Index
			Full		
Education Reform	-0.400*	-0.640***	-0.472**	0.283	-0.009**
	(0.233)	(0.214)	(0.186)	(0.626)	(0.004)
Control Mean	-42.85	-42.88	-41.84	104.55	-0.04
Control SD	53.67	54.48	53.81	120.68	0.94
Bandwidth	55.64	54.02	54.46	56.97	55.38
Observations	967607	949208	949328	1000546	982059
			Male		
Education Reform	-0.473*	-0.770***	-0.592**	0.655	-0.012**
	(0.285)	(0.260)	(0.252)	(0.727)	(0.005)
Control Mean	-42.02	-42.18	-41.31	106.96	-0.02
Control SD	55.46	56.50	55.53	120.65	0.98
Bandwidth	59.05	55.96	57.37	53.45	60.84
Observations	761811	713085	738935	694984	789631
			Female		
Education Reform	-0.242	-0.270	-0.141	-0.597	-0.003
	(0.434)	(0.463)	(0.518)	(1.298)	(0.008)
Control Mean	-44.88	-44.64	-42.96	97.52	-0.07
Control SD	47.77	47.96	48.07	120.11	0.83
Bandwidth	58.73	59.68	60.25	63.18	60.56
Observations	268108	271848	278266	296339	284176

Notes: Local quadratic RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the one-month return in column 1, the three-month return in column 2, the six-month return in column 3, and the annual return in percentages in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with a quadratic polynomial function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C.4 Local Kink RD Estimates

Table C.14: Education Reform vs Schooling Outcomes

	(1) Years of Schooling	(2) Primary School	(3) Junior High School	(4) High School	(5) College	(6) Summary Index
			Full			
Education Reform	-0.008	0.001	-0.005***	-0.003	-0.001	-0.002
	(0.011)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Control Mean	9.09	0.93	0.64	0.50	0.28	0.42
Control SD	4.70	0.26	0.48	0.50	0.45	0.84
Bandwidth	73.95	61.15	65.70	50.30	49.05	73.47
Observations	77043	63853	67860	52433	51460	77043
			Male			
Education Reform	-0.003	0.001	-0.005***	-0.004	0.001	-0.001
	(0.015)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Control Mean	9.80	0.96	0.72	0.56	0.29	0.54
Control SD	4.32	0.19	0.45	0.50	0.45	0.78
Bandwidth	71.40	67.01	69.32	53.96	60.43	72.05
Observations	35806	33721	34812	26578	30358	36728
			Female			
Education Reform	-0.020	0.002**	-0.002	-0.002	-0.002	-0.003
	(0.022)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)
Control Mean	8.48	0.89	0.52	0.44	0.26	0.31
Control SD	4.98	0.31	0.50	0.50	0.44	0.88
Bandwidth	68.46	81.89	96.20	59.20	50.43	68.63
Observations	36786	44273	53191	31788	27211	36786

Notes: Local kink RD estimates in all columns. All columns use data from the 2018 Household Labor Force Survey by TURKSTAT. The unit of analysis is individuals. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. In column 1, the outcome is years of schooling. In column 2, the outcome is a dummy variable equal to one if the specific individual has at least a primary school degree. In column 3, the outcome is a dummy variable equal to one if the specific individual has at least a junior high school degree. In column 4, the outcome is a dummy variable equal to one if the specific individual has at least a high school degree. In column 5, the outcome is a dummy variable equal to one if the specific individual has at least a college degree. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and the regressions in the full sample additionally include controls for gender. Control mean and standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.15: Education Reform vs Participation in Stock Market

	(1) Market Participant	(2) Direct Stock	(3) Index Stock	(4) Blue-Chip Stock	(5) Risky Asset	(6) Bond	(7) Fund	(8) Transaction Account	(9) Summary Index
					Full				
Education Reform	0.001	-0.011*	-0.000	0.001	-0.007	0.000	-0.000	-0.025	-0.003
	(0.005)	(0.006)	(0.003)	(0.002)	(0.005)	(0.000)	(0.003)	(0.032)	(0.003)
Control Mean	9.16	5.25	3.51	2.29	6.19	0.10	3.09	55.54	0.44
Control SD	0.44	0.26	0.19	0.15	0.29	0.01	0.23	1.30	0.18
Bandwidth	63.29	42.36	61.96	66.95	51.96	62.23	65.16	31.03	47.60
Observations	127	85	123	133	103	125	131	63	95
					Male				
Education Reform	-0.010	-0.017*	-0.009*	-0.004	-0.009	0.000	0.001	0.021	-0.003
	(0.007)	(0.010)	(0.005)	(0.004)	(0.008)	(0.000)	(0.003)	(0.029)	(0.004)
Control Mean	13.26	7.98	5.34	3.51	9.23	0.14	4.39	72.21	0.46
Control SD	0.52	0.35	0.28	0.21	0.37	0.02	0.29	0.90	0.16
Bandwidth	57.72	44.13	56.12	59.00	49.68	81.27	73.78	29.35	48.33
Observations	115	89	113	117	99	163	147	59	97
					Female				
Education Reform	0.009**	0.003	0.004*	0.002	0.007**	0.000	-0.004	-0.031	0.000
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.000)	(0.003)	(0.034)	(0.004)
Control Mean	4.94	2.45	1.64	1.02	3.04	0.07	1.77	38.21	0.18
Control SD	0.38	0.17	0.13	0.10	0.22	0.01	0.19	1.61	0.27
Bandwidth	70.59	69.20	67.02	72.52	74.05	66.82	58.79	35.90	60.50
Observations	141	139	135	145	149	133	117	71	121

Notes: Local kink RD estimates in all columns. All columns use data from population numbers assembled by TURKSTAT and the relevant investor numbers assembled by Borsa Istanbul Group in 2021. The unit of analysis is the month-year birth cohorts in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the percentage of those participating in the stock market in column 1, the percentage of those directly holding stocks in column 2, the percentage of those directly holding index stocks in column 3, the percentage of those directly holding blue-chip stocks in column 4, the percentage of those holding risky assets in column 5, the percentage of those directly holding bonds in column 6, the percentage of those directly holding funds in their stock market portfolios in column 7, the percentage of those having a transaction account in column 8. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth for columns. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.16: Education Reform vs Variation in Portfolio Choices

	(1) Stock	(2) Risky Asset	(3) Bond	(4) Blue-Chip Stock	(5) Index Stock	(6) Fund	(7) Summary Index
				Full			
Education Reform	-0.000	0.000	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control Mean	0.55	0.65	0.01	0.13	0.24	0.33	0.08
Control SD	0.49	0.47	0.09	0.29	0.39	0.46	0.46
Bandwidth	45.69	54.16	65.64	50.47	49.77	47.76	51.37
Observations	808937	963363	1155637	893811	878229	839920	910469
				Male			
Education Reform	-0.001**	-0.000	0.000	-0.001***	-0.000	0.000	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control Mean	0.57	0.67	0.01	0.14	0.25	0.31	0.10
Control SD	0.49	0.46	0.09	0.30	0.39	0.45	0.46
Bandwidth	42.91	52.16	64.50	42.60	48.96	46.58	46.48
Observations	556484	682633	836061	556484	636653	608200	608200
				Female			
Education Reform	0.000	0.000	0.000	0.000	0.000	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Control Mean	0.47	0.60	0.01	0.11	0.22	0.37	0.04
Control SD	0.49	0.48	0.11	0.28	0.38	0.47	0.46
Bandwidth	47.89	57.00	76.38	60.14	55.88	53.34	57.75
Observations	221394	264723	359211	284176	259952	249795	269174

Notes: Local kink RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the share of wealth invested in stocks in column 1, in risky assets in column 2, in bonds in column 3, in blue-chip stocks in column 4, in index stocks in column 5, and in funds in column 6. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.17: Education Reform vs Causal Channels

	(1) Risk Score	(2) Single Stock	(3) Risky Inertia	(4) Wealth (in logs)	(5) Summary Index
			Full		
Education Reform	0.002*	-0.000	0.000	-0.002	-0.000
	(0.001)	(0.000)	(0.000)	(0.002)	(0.000)
Control Mean	3.68	0.18	0.15	6.95	-0.02
Control SD	0.84	0.39	0.82	3.73	0.53
Bandwidth	47.51	50.40	90.17	54.76	47.98
Observations	410635	893811	984169	963363	839920
			Male		
Education Reform	0.002*	0.000	-0.000	-0.002	-0.000
	(0.001)	(0.000)	(0.000)	(0.003)	(0.000)
Control Mean	3.71	0.18	0.14	7.01	-0.01
Control SD	0.82	0.39	0.79	3.80	0.52
Bandwidth	48.41	51.59	74.30	50.53	48.52
Observations	341008	670196	620707	658138	636653
			Female		
Education Reform	-0.000	-0.000	-0.001	0.001	-0.000
	(0.001)	(0.000)	(0.001)	(0.003)	(0.000)
Control Mean	3.57	0.17	0.15	6.79	-0.05
Control SD	0.90	0.38	0.78	3.54	0.54
Bandwidth	79.23	59.09	68.06	58.10	54.75
Observations	130587	277655	182099	273839	254852

Notes: Local kink RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the logarithm of risk scores in column 1, an indicator variable equal to one if the investor invests all money into a single stock in column 2, the risk share inertia in column 3, and the logarithm of portfolio size in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.18: Education Reform vs Portfolio Returns

	(1) One Month	(2) Three Months	(3) Six Months	(4) Twelve Months	(5) Summary Index
			Full		
Education Reform	0.030	0.017	0.018	0.001	0.000
	(0.026)	(0.024)	(0.023)	(0.064)	(0.000)
Control Mean	-42.90	-42.90	-41.86	104.73	-0.04
Control SD	53.71	54.47	53.80	120.74	0.94
Bandwidth	53.00	53.19	53.23	54.11	53.07
Observations	930893	930893	931013	963363	944779
			Male		
Education Reform	-0.003	-0.013	-0.007	-0.020	-0.000
	(0.029)	(0.027)	(0.028)	(0.070)	(0.000)
Control Mean	-42.10	-42.18	-41.34	106.96	-0.02
Control SD	55.52	56.50	55.64	120.65	0.98
Bandwidth	56.54	55.58	55.84	53.08	58.74
Observations	726637	713085	713172	694984	760755
			Female		
Education Reform	0.109**	0.098**	0.071	0.078	0.002**
	(0.051)	(0.049)	(0.051)	(0.127)	(0.001)
Control Mean	-45.06	-44.78	-43.09	98.24	-0.07
Control SD	47.83	48.07	48.13	120.62	0.84
Bandwidth	52.47	54.90	56.07	57.87	54.42
Observations	239957	249536	259226	269174	254852

Notes: Local kink RD estimates in all columns. All columns use the administrative data by Borsa Istanbul Group in 2021. The unit of analysis is individuals in all columns. The main explanatory variable namely Education Reform is a dummy variable equal to one for those born after January 1, 1987. The outcome is the one-month return in column 1, the three-month return in column 2, the six-month return in column 3, and the annual return in percentages in column 4. The outcome is a summary index of outcome variables following Kling et al. (2007) for the former outcome variables in the last column. RD estimates have the optimal bandwidth with a triangular type kernel function calculated through the algorithm by Calonico et al. (2014) in all columns. Each column reports RD estimates with the first derivative of a linear function on each side of the cutoff value. All regressions include controls for the month of birth fixed effects and birth registration certificate region fixed effects, and the regressions in the full sample additionally include controls for gender. Control mean and control standard deviation (SD) display the mean and standard deviation of the corresponding outcome of those born before January 1, 1987. Standard errors are clustered at the month-year birth cohort level. Robust standard errors are in parentheses. Full, male, and female sample estimates are reported, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C.5 Data Appendix

Variables for Schooling Outcomes

Years of Schooling is generated by assigning 0, 5, 8, 11, 15, and 17 for those who have no degree, at least a primary school degree, at least junior high school degree, at least high school degree, at least college degree, and at least master degree, respectively.

Primary School is a dummy variable equal to 1 if a specific individual at least holds a primary school degree with 5 years of schooling.

Junior High School is a dummy variable equal to 1 if a specific individual at least holds a junior high school degree with 8 years of schooling.

High School is a dummy variable equal to 1 if a specific individual at least holds a high school degree with 11 years of schooling.

College is a dummy variable equal to 1 if a specific individual at least holds a college degree with 13-15 years of schooling.

C.5.1 Variables for the Stock Market Participation and Asset Ownership

Stock Market Participation Rate is the percentage of those participating in the stock market i.e. those having a positive amount of wealth in the transaction account generated by dividing the number of individuals with positive balances by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021.

Stock Ownership Rate is the percentage of those holding stocks in their stock market portfolio generated by dividing the number of individuals owning stocks by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021.

Index Stock Ownership Rate is the percentage of those holding index stocks in their stock market portfolio generated by dividing the number of individuals owning index stocks by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021. The index stocks are the stocks that are included in the BIST-100 Index, which tracks the market benchmark portfolio.

Blue-chip Stock Ownership Rateis the percentage of those holding blue-chip stocks in their stock market portfolio generated by dividing the number of individuals owning blue-chip stocks by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021. The blue-chip stocks are the stocks that are included in the BIST-30 Index, which tracks the most liquid and largest market capitalization companies.

Risky Asset Ownership Rate is the percentage of those holding risky assets in their stock market portfolio generated by dividing the number of individuals owning risky assets by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021. Risky assets exclude money market funds.

Bond Ownership Rate is the percentage of those holding bonds in their stock market portfolio generated by dividing the number of individuals owning bonds by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021.

Fund Ownership Rate is the percentage of those holding funds in their stock market portfolio generated by dividing the number of individuals owning funds by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021.

Transaction Account Ownership Rate is the percentage of those having transaction accounts generated by dividing the number of individuals with a transaction account by the number of individuals in the general population in the specific month-year birth cohort on December 31, 2021.

C.5.2 Variables for the Variation in Portfolio Choices

Stock Share is the share of wealth invested in stocks directly in stock market portfolios.

Risky Assets Share is the share of wealth invested in risky assets in stock market portfolios.

Bond Share is the share of wealth invested in bonds in stock market portfolios.

Blue-chip Stock Share is the share of wealth invested in blue-chip stocks in stock market portfolios.

Index Stock Share is the share of wealth invested in index stocks in stock market portfolios.

Fund Share is the share of wealth invested in funds in stock market portfolios.

C.5.3 Variables for Causal Channels

Risk Score is the logarithm of one plus the risk score of a certain investor computed by Borsa Istanbul Group.

Single Stock is an indicator variable equal to one if a certain investor invests all of her money in a single stock.

Risky Inertia is the absolute value of the difference between the logarithm of the risky share on December 31, 2022, and the logarithm of the risky share on December 31, 2021. We follow the strategy proposed by Calvet et al. (2009b).

Wealth (in logs) is the logarithm of the portfolio size.

C.5.4 Variables for Portfolio Returns

One Month is the one-month return of portfolios on December 31, 2021.

Three Months is the three-month return of portfolios on December 31, 2021.

Six Months is the six-month return of portfolios on December 31, 2021.

Twelve Months is the twelve-month (annual) return of portfolios on December 31, 2021.

Summary Index is a variable constructed through the procedure suggested by Kling et al. (2007). To make it clear, I first subtract the control group mean of the corresponding variable and subsequently divide it by the control group standard deviation of the concerning variable.

C.5.5 Control Variables

Birth Month Indicator Variables are 12 indicator variables for each month.

Birth Registration Certificate Region Indicator Variables are 26 indicator variables for each birth registration certificate region.

Gender is an indicator variable if a certain investor is female.

BIBLIOGRAPHY

ABCFM (1852). *Annual report - American Board of Commissioners for Foreign Missions.* Boston: Press of T. R. Marvin.

ABCFM (1870). *Annual report - American Board of Commissioners for Foreign Missions.* Riverside Press.

ABCFM (1882). *Annual report - American Board of Commissioners for Foreign Missions.* Riverside Press.

ABCFM (1897). *Annual report - American Board of Commissioners for Foreign Missions.* THF Board Press: Boston.

ABCFM (1913). *Annual report - American Board of Commissioners for Foreign Missions.* The Board Congregational House.

Acemoglu, D. and Angrist, J. (2000). How large are human-capital externalities? evidence from compulsory schooling laws. *NBER Macroeconomics Annual*, 15:9–59.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*. *The Quarterly Journal of Economics*, 117(4):1231–1294.

Agnew, J., Balduzzi, P., and Sundén, A. (2003). Portfolio choice and trading in a large 401(k) plan. *American Economic Review*, 93(1):193–215.

Agnew, J. R. (2006). Do behavioral biases vary across individuals? evidence from individual level 401(k) data. *Journal of Financial and Quantitative Analysis*, 41(4):939–962.

Akarçay, A., Davutyan, N., and Polat, S. (2021). Economic consequences of demographic engineering: Turkey and wwi. *Available at SSRN 3828518.*

Alesina, A. and Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4):898–944.

Ameriks, J. and Zeldes, S. P. (2011). *How do household portfolio shares vary with age?* SSRN.

Angrist, J. D. and Keueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, 106(4):979–1014.

Angrist, J. D. and Krueger, A. B. (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association*, 87(418):328–336.

Arbatlı, C. E. and Gokmen, G. (2022). Human capital transfers and sub-national development: Armenian and greek legacy in post-expulsion turkey. *Journal of Economic Growth*, pages 1–43.

Arcanjo, M. (2019). Retirement pension reforms in six european social insurance schemes between 2000 and 2017: More financial sustainability and more gender inequality? *Social Policy and Society*, 18(4):501–515.

Asik, G., Karakoç, U., and Pamuk, S. (2020). Regional inequalities and the west-east divide in turkey since 1880.

Atılgan, G., Aytekin, E. A., Ozan, E. D., Saraçoğlu, C., Şener, M., Uslu, A., and Yeşilbağ, M. (2015). *Osmanlı'dan günümüze Türkiye'de siyasal hayat*. Yordam Kitap.

Aydemir, A. and Kirdar, M. G. (2017). Low wage returns to schooling in a developing country: Evidence from a major policy reform in turkey. *Oxford Bulletin of Economics and Statistics*, 79(6):1046–1086.

Aydemir, A. B., Kırdar, M. G., and Torun, H. (2022). The effect of education on internal migration of young men and women: incidence, timing, and type of migration. *Labour Economics*, 74:102098.

Bai, Y. and Kung, J. K.-s. (2015). Diffusing Knowledge While Spreading God's Message: Protestantism and Economic Prosperity in China, 1840–1920. *Journal of the European Economic Association*, 13(4):669–698.

Barton, J. L. (1908). *Daybreak in Turkey*. Pilgrim Press.

Becker, S. O. and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History*. *The Quarterly Journal of Economics*, 124(2):531–596.

Becker, S. O. and Woessmann, L. (2011). The effects of the protestant reformation on human capital. In *The Oxford Handbook of the Economics of Religion*, pages 93–110. Oxford University Press.

Benartzi, S. and Thaler, R. (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives*, 21(3):81–104.

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review*, 91(1):79–98.

Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2009). The importance of default options for retirement saving outcomes: Evidence from the united states. In *Social security policy in a changing environment*, pages 167–195. University of Chicago Press.

Black, S. E., Devereux, P. J., Lundborg, P., and Majlesi, K. (2018). Learning to Take Risks? The Effect of Education on Risk-Taking in Financial Markets. *Review of Finance*, 22(3):951–975.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2008). Staying in the Classroom and Out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Births. *The Economic Journal*, 118(530):1025–1054.

Bleakley, H. and Gupta, B. (2023). Mind the Gap: Schooling, Informality, and Fiscal Externalities in Nepal. *The World Bank Economic Review*, page lhad013.

Blumenstock, J., Callen, M., and Ghani, T. (2018). Why do defaults affect behavior? experimental evidence from afghanistan. *American Economic Review*, 108(10):2868–2901.

Bodie, Z., Merton, R. C., and Samuelson, W. F. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control*, 16(3):427–449.

Bosworth, B. and Collins, S. M. (2008). Accounting for growth: Comparing china and india. *Journal of Economic Perspectives*, 22(1):45–66.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.

Brown, J. R., Farrell, A. M., and Weisbenner, S. J. (2016). Decision-making approaches and the propensity to default: Evidence and implications. *Journal of Financial Economics*, 121(3):477–495.

Brune, L., Giné, X., Goldberg, J., and Yang, D. (2017). Savings defaults and payment delays for cash transfers: Field experimental evidence from malawi. *Journal of Development Economics*, 129:1–13.

Cage, J. and Rueda, V. (2016). The long-term effects of the printing press in sub-saharan africa. *American Economic Journal: Applied Economics*, 8(3):69–99.

Caicedo, F. V. (2018). The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America*. *The Quarterly Journal of Economics*, 134(1):507–556.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Calvet, L., Campbell, J., and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009a). Fight or Flight? Portfolio Rebalancing by Individual Investors. *The Quarterly Journal of Economics*, 124(1):301–348.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009b). Measuring the financial sophistication of households. *American Economic Review*, 99(2):393–98.

Calvi, R., Hoehn-Velasco, L., and Mantovanelli, F. G. (2020). The protestant legacy: missions, gender, and human capital in india. *Journal of Human Resources*, pages 0919–10437R2.

Calvi, R. and Mantovanelli, F. G. (2018). Long-term effects of access to health care: Medical missions in colonial india. *Journal of Development Economics*, 135:285–303.

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. Working Paper 4483, National Bureau of Economic Research.

Carlsson, M., Dahl, G. B., Öckert, B., and Rooth, D.-O. (2015). The Effect of Schooling on Cognitive Skills. *The Review of Economics and Statistics*, 97(3):533–547.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2009). Optimal Defaults and Active Decisions*. *The Quarterly Journal of Economics*, 124(4):1639–1674.

Caselli, F. (2005). Chapter 9 accounting for cross-country income differences. volume 1 of *Handbook of Economic Growth*, pages 679–741. Elsevier.

Catherine, S. (2021). Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle. *The Review of Financial Studies*. hhab136.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, 18(1):234–261.

Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review*, 105(5):1–33.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., and Olsen, T. (2014). Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark. *The Quarterly Journal of Economics*, 129(3):1141–1219.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2004). For better or for worse: Default effects and 401 (k) savings behavior. In *Perspectives on the Economics of Aging*, pages 81–126. University of Chicago Press.

Choi, J. J., Laibson, D. I., and Madrian, B. C. (2005). Are empowerment and education enough? underdiversification in 401 (k) plans. *Brookings Papers on Economic Activity*, 2005(2):151–213.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies*, 18(2):491–533.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Cole, S., Paulson, A., and Shastry, G. K. (2014). Smart Money? The Effect of Education on Financial Outcomes. *The Review of Financial Studies*, 27(7):2022–2051.

Conley, T. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Cronqvist, H. and Thaler, R. H. (2004). Design choices in privatized social-security systems: Learning from the swedish experience. *American Economic Review*, 94(2):424–428.

Dayıoğlu, M. and Kırdar, M. G. (2010). Determinants of and trends in labor force participation of women in turkey.

de Chaisemartin, C. and D'Haultfœuille, X. (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey. *The Econometrics Journal*, page utac017.

Dincer, B. and Özaslan, M. (2004). *İlçelerin sosyo-ekonomik gelişmişlik sıralaması araştırması (2004)*. Ankara: Devlet Planlama Teşkilatı Bölgesel Gelişme ve Yapısal Uyum Genel Müdürlüğü.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Duflo, E., Gale, W., Liebman, J., Orszag, P., and Saez, E. (2006). Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H amp;R Block. *The Quarterly Journal of Economics*, 121(4):1311–1346.

Duflo, E. and Saez, E. (2002). Participation and investment decisions in a retirement plan: the influence of colleagues' choices. *Journal of Public Economics*, 85(1):121–148.

Duflo, E. and Saez, E. (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment*. *The Quarterly Journal of Economics*, 118(3):815–842.

Easterly, W. and Levine, R. (2003). Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics*, 50(1):3–39.

Egan, M. L., MacKay, A., and Yang, H. (2021). What drives variation in investor portfolios? evidence from retirement plans. Working Paper 29604, National Bureau of Economic Research.

Engström, S. and Westerberg, A. (2003). Which individuals make active investment decisions in the new swedish pension system? *Journal of Pension Economics and Finance*, 2(3):225–245.

Erten, B. and Keskin, P. (2018). For better or for worse?: Education and the prevalence of domestic violence in turkey. *American Economic Journal: Applied Economics*, 10(1):64–105.

Fagereng, A., Gottlieb, C., and Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance*, 72(2):705–750.

Fagereng, A., Guiso, L., Holm, M. B., and Pistaferri, L. (2020a). K-returns to education. *Available at SSRN 3526034*.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020b). Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170.

Fagereng, A., Mogstad, M., and Rønning, M. (2021). Why do wealthy parents have wealthy children? *Journal of Political Economy*, 129(3):703–756.

Filmer, D., Langthaler, M., Stehrer, R., and Vogel, T. (2018). Learning to realize education's promise. *World Development Report. The World Bank*.

Finnie, D. H. (2013). *Pioneers east*. Harvard University Press.

Gallup, J. L., Sachs, J. D., and Mellinger, A. D. (1999). Geography and economic development. *International Regional Science Review*, 22(2):179–232.

Glaeser, E. L. and Shleifer, A. (2002). Legal Origins*. *The Quarterly Journal of Economics*, 117(4):1193–1229.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *The American Economic Review*, 92(2):160–164.

Gomes, F., Haliassos, M., and Ramadorai, T. (2021). Household finance. *Journal of Economic Literature*, 59(3):919–1000.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. Themed Issue: Treatment Effect 1.

Gray, D., Montagnoli, A., and Moro, M. (2021). Does education improve financial behaviors? quasi-experimental evidence from britain. *Journal of Economic Behavior Organization*, 183:481–507.

Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). Iq and stock market participation. *The Journal of Finance*, 66(6):2121–2164.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2):23–48.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Trusting the stock market. *The Journal of Finance*, 63(6):2557–2600.

Guiso, L. and Sodini, P. (2013). Household finance: An emerging field. In *Handbook of the Economics of Finance*, volume 2, pages 1397–1532. Elsevier.

Gulesci, S., Meyersson, E., and Trommlerová, S. K. (2019). The Effect of Compulsory Schooling Expansion on Mothers' Attitudes Toward Domestic Violence in Turkey. *The World Bank Economic Review*, 34(2):464–484.

Haliassos, M. and Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, 105(432):1110–1129.

Hanushek, E. A. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–68.

Hanushek, E. A. and Woessmann, L. (2011). Chapter 2 - the economics of international differences in educational achievement. volume 3 of *Handbook of the Economics of Education*, pages 89–200. Elsevier.

Hanushek, E. A. and Woessmann, L. (2012). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4):267–321.

Herrendorf, B., Rogerson, R., and Ákos Valentinyi (2014). Chapter 6 - growth and structural transformation. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2 of *Handbook of Economic Growth*, pages 855–941. Elsevier.

Hornbeck, R. and Keskin, P. (2015). Does agriculture generate local economic spillovers? short-run and long-run evidence from the ogallala aquifer. *American Economic Journal: Economic Policy*, 7(2):192–213.

Huberman, G. and Jiang, W. (2006). Offering versus choice in 401(k) plans: Equity exposure and number of funds. *The Journal of Finance*, 61(2):763–801.

Iiasa, F. (2012). Global agro-ecological zones (gaez v3. 0). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*.

ILO (2021). World social protection report 2020–22: Social protection at the crossroads – in pursuit of a better future. Technical report, ILO.

Iyengar, S. S. and Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, 94(7):530–539.

Iyer, S. (2016). The new economics of religion. *Journal of Economic Literature*, 54(2):395–441.

Jarvis, A., Reuter, H. I., Nelson, A., Guevara, E., et al. (2008). Hole-filled srtm for the globe version 4. *available from the CGIAR-CSI SRTM 90m Database (http://srtm. csi. cgiar. org)*, 15(25-54):5.

Jedwab, R., Meier zu Selhausen, F., and Moradi, A. (2022). The economics of missionary expansion: evidence from africa and implications for development. *Journal of Economic Growth*, 27(2):149–192.

Johnson, D. G. (1997). Agriculture and the wealth of nations. *The American Economic Review*, 87(2):1–12.

Johnston, B. F. and Mellor, J. W. (1961). The role of agriculture in economic development. *The American Economic Review*, 51(4):566–593.

Karpat, K. (1985). *Ottoman Population, 1830—1914: Demographic and Social Characteristics*. The University of Wisconsin Press.

Kelly, M. (2019). The standard errors of persistence.

Kirdar, M. G., Dayioğlu, M., and Koç, (2016). Does longer compulsory education equalize schooling by gender and rural/urban residence? *World Bank Economic Review*, 30(3):549–579.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kocabaşoğlu, U. (2000). *Anadolu'daki Amerika: Kendi belgeleriyle 19. yüzyılda Osmanlı İmparatorluğu'ndaki Amerikan misyoner okulları*. İmge Kitabevi.

Koyama, M. and Rubin, J. (2022). *How the World Became Rich: The Historical Origins of Economic Growth*. John Wiley & Sons.

Kuran, T. (2003). The islamic commercial crisis: Institutional roots of economic underdevelopment in the middle east. *The Journal of Economic History*, 63(2):414–446.

Kuran, T. (2004). Why the middle east is economically underdeveloped: Historical mechanisms of institutional stagnation. *Journal of Economic Perspectives*, 18(3):71–90.

Kuran, T. (2012). *The long divergence*. Princeton University Press.

Kuran, T. (2018). Islam and economic performance: Historical and contemporary links. *Journal of Economic Literature*, 56(4):1292–1359.

Kögel, T. and Prskawetz, A. (2001). Agricultural productivity growth and escape from the malthusian trap. *Journal of Economic Growth*, 6(4):337–357.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. (1998). Law and finance. *Journal of Political Economy*, 106(6):1113–1155.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies*, 76(3):1071–1102.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674. The regression discontinuity design: Theory and applications.

Lehner, B., Verdin, K., and Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10):93–94.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *The Review of Economic Studies*, 72(1):189–221.

Lochner, L. and Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, 94(1):155–189.

Lusardi, A. and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1):5–44.

Madrian, B. C. (2014). Applying insights from behavioral economics to policy design. *Annual Review of Economics*, 6(1):663–688.

Madrian, B. C. and Shea, D. F. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics*, 116(4):1149–1187.

Makdisi, U. (2011). *4. The Artillery of Heaven*. Cornell University Press.

Malmendier, U. and Nagel, S. (2015). Learning from Inflation Experiences *. *The Quarterly Journal of Economics*, 131(1):53–87.

Mellor, J. (1999). Faster, more equitable growth–the relation between growth in agriculture and poverty reduction agricultural policy development project (research report no. 4). *Washington, DC: United States Agency for International Development*.

Montero, E. and Yang, D. (2021). Religious festivals and economic development: Evidence from the timing of mexican saint day festivals. Working Paper 28821, National Bureau of Economic Research.

Morrissey, M. (2016). The state of american retirement. *Economic Policy Institute, Washington, DC*.

Nunn, N. (2008). The Long-term Effects of Africa's Slave Trades*. *The Quarterly Journal of Economics*, 123(1):139–176.

Nunn, N. (2014). *Gender and Missionary Influence in Colonial Africa*, page 489–512. Cambridge University Press.

OECD (2019). Pensions at a glance 2019: Oecd and g20 indicators. Technical report, OECD.

OECD (2021). Pensions at a glance 2021: Oecd and g20 indicators. Technical report, OECD.

Okoye, D. and Pongou, R. (2014). Historical Missionary Activity, Schooling, and the Reversal of Fortunes: Evidence from Nigeria. MPRA Paper 58052, University Library of Munich, Germany.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, 96(1):152–175.

Oreopoulos, P. and Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives*, 25(1):159–84.

Peksevim, S. and Akgiray, V. (2019). Reforming the pension system in turkey: Comparison of mandatory and auto-enrolment pension systems in selected oecd countries. Technical report, OECD.

Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, 26(5):445–458.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250.

Sakalli, S. O. (2017). Secularization and religious backlash: Evidence from turkey. Technical report, Technical report, mimeo.

Self, S. and Grabowski, R. (2007). Economic development and the role of agricultural technology. *Agricultural Economics*, 36(3):395–404.

Sialm, C., Starks, L. T., and Zhang, H. (2015). Defined contribution pension plans: Sticky or discerning money? *The Journal of Finance*, 70(2):805–838.

Sokoloff, K. L. and Engerman, S. L. (2000). Institutions, factor endowments, and paths of development in the new world. *Journal of Economic Perspectives*, 14(3):217–232.

Somel, S. A. (2001). *The modernization of public education in the Ottoman Empire, 1839-1908: Islamization, autocracy, and discipline*, volume 22. Brill.

Spolaore, E. and Wacziarg, R. (2013). How deep are the roots of economic development? *Journal of Economic Literature*, 51(2):325–69.

Tansel, A. and Güngör, N. D. (2016). Gender effects of education on economic development in turkey. In *Women, Work and Welfare in the Middle East and North Africa: The Role of Socio-demographics, Entrepreneurship and Public Policies*, pages 57–86. World Scientific.

Thaler, R. and Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1):S164–S187.

van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2):449–472.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance*, 56(2):433–470.

Vissing-Jørgensen, A. (2002). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy*, 110(4):825–853.

Waldinger, M. (2017). The long-run effects of missionary orders in mexico. *Journal of Development Economics*, 127:355–378.

Wang, J. and Zivot, E. (1998). Inference on structural parameters in instrumental variables regression with weak instruments. *Econometrica*, 66(6):1389–1404.

Weber, M. and Kalberg, S. (2013). *The Protestant ethic and the spirit of capitalism*. Routledge.

Woodberry, R. D. and Shah, T. S. (2004). Christianity and democracy: The pioneering protestants. *Journal of democracy*, 15(2):47–61.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*.

Yücel, İ. (2005). Kendi belgeleri ışığında amerikan boardın osmanlı ülkesindeki teşkilatlanması.

Yücel, İ. (2011). Yakın şarkta batı nüfuzu üzerine bir değerlendirme: Anadolu'daki amerikan misyoner hastaneleri (1880-1930). *HÜTAD*, (15).