WHEN SCHOOLS TAKE THE LEAD, STUDENTS DON'T REPEAT

EVIDENCE FROM EGYPT'S 2008-2010 FINANCIAL DECENTRALIZATION PILOT

By
Aly Abdou

Submitted to
Central European University
Department of Economics and Business

In partial fulfillment of the requirements for the degree of Master of Art in Economic Policy in Global Markets

Supervisor: Professor Sergey Lychagin

Budapest, Hungary
Year
2019

Abstract

In this paper, I evaluate the impact of Egypt's financial decentralization pilot program (2008-2010) on the experience of grade retention for primary school students. The program decentralized school funding from the central government to local schools using an enrollment and needs-based formula and provided training to school staff to help them better manage the funds. Using household survey data on 28,331 respondents from the 1996/1997 cohorts and 26,881 respondents from the 1994/1995 cohorts, I estimate that the pilot program significantly decreases grade retention by 5.2 percentage points for primary school students. However, the program had a larger impact on females and students from above average income households compared to males and students coming from below average income households. Additionally, the program also had positive spill-over effects on untreated education stages. I also observe heterogenous effects in different states that are correlated with the state's development indicators. The results suggest that the financial decentralization interventions can have a positive on a student's grade retention and a smaller impact if the student is from a disadvantaged socio-economic background.

CEU eTD Collection

Acknowledgments

I would like to express my deepest appreciation to my thesis supervisor Professor Sergey Lychagin (Central European University) for his guidance and support. Our discussions and his insightful comments were a major contribution to the success of this work. I also extend my gratitude to Professor Gabor Bekes (Central European University) for his enthusiasm and knowledge. The foundational work in the Data Analysis modules set me on the right path since the first year. I also thank Professor Julius Horvath (Central European University) for making me feel at home and for keeping his door always open when help was needed. I also thank Professor Thomas Rooney for reviewing and editing the drafts of this study.

Finally, I would like to thank my mother and my partner. I would have not been able to complete this work without their unconditional love and support.

CEU eTD Collection

Table of Contents

CEU eTD Collection

List of tables & figures

CEU eTD Collection

List of abbreviations

BOT – Boards of Trustees, Parents, and Teachers
ERF – Economic Research Forum
LFS – Labor Force Survey
MOE – Ministry of Education
MOF – Ministry of Finance
NDP – National Democratic Party
SBM – School-based Management
SDS – Sustainable Development Strategy
TEU – Training and Evaluation Unit

CEU eTD Collection

Introduction

In 2016, Egypt embarked on an ambitious all-inclusive reform program called the Sustainable Development Strategy (SDS): Egypt Vision 2030. The SDS aspires to improve the economic and social quality of life for Egyptians. Moreover, it seeks to reduce the social intergenerational and gender gaps and achieve a balanced geographical distribution of services. The strategy sees the education sector as one of the core contributors to Egypt's social transformation by 2030. The improvement of education quality is one of the main adopted specific objectives. The adopted framework relies on the principles of "inclusive sustainable development" and "balanced regional development" to ensure that "no one is left behind" (Ministry of Planning, Monitoring and Administrative Reform 2016).

The 2014 Egyptian constitution obliges the government with a minimum public expenditure on education of 4 percent of GDP. Failing to achieve the constitutional commitment , the government relied on international organizations and donors to help support the 2016 reform package (World Bank 2018). According to the SDS: Egypt Vision 2030, the government is committed to increase the public expenditure on pre-university education from the current 3 % to 5% by 2030 (Ministry of Planning, Monitoring and Administrative Reform 2016).

Compared to neighboring low-middle income countries, Egypt has successfully improved its key indicators for access to education. Nonetheless, the quality of the education received by the Egyptian population has failed to show similar trends of improvement. The education system is not delivering the expected learning outcomes. Although more students are enrolled in schools, fewer are actually learning. One in every 5 students in grade 3 lack the crucial reading skills and proceed to grade 4 practically illiterate. Furthermore, graduates from public schools are realizing that their accumulated skills are insufficient to compete on the modern markets of the 21st century (World Bank 2018).

For the past two decades, Egypt has been experimenting and cautiously promoting school-based management and decentralization interventions to face the key challenges impeding its centralized

CEU eTD Collection

education system. In 2000, Egypt introduced both the Parent-Teacher councils and Boards of Trustees, Parents and Teachers councils in various schools in Egypt with the objective of including the wider local community in the education process. An empirical assessment of those interventions found that the interventions had not significantly affected education outcomes measured at the state level (Nasser-Ghodsi 2006).

Following the worldwide trend of the decentralization of education in developing countries (P. J. Gertler, Premand, and Rawlings 2011; Jimenez and Sawada 2013; Pradhan et al. 2011; Skoufias and Shapiro 2006), it is predictable that Egypt continue with its effort in achieving a fully decentralized education system. Nonetheless, the lack of empirical evidence of the impact of decentralization interventions on student outcomes in Egypt is one of the impediments for the advancement of decentralization policies in Egypt (El Baradei 2015).

In this paper, I attempt to clear the ambiguity associated with the impact of decentralization on one aspect of education quality – grade retention – on different sub-groups of the population (Carr-Hill et al. 2015). Due to the high possibility for students who experience grade retention to drop-out and have adverse labor outcomes, it can potentially serve as a good proxy for education quality (Eide and Showalter 2001; Manacorda 2012; Roderick 1994). I use household survey data to perform an impact evaluation of a financial decentralization pilot program implemented in the fiscal years 2008/2009 and 2009/2010. The program's main objective was to arrive at an equitable distribution of funds from the central government to local schools. The program developed a needs and enrollment-based funding formulae to account for variations in regional poverty and enrollment rates. Funds were previously characterized by a top-down hierarchal allocation that generated huge disparities between states (F. H. Healey, Crouch, and Hanna 2014; H. Healey, Hanna, and Attalla 2016).

I find that the fiscal decentralization intervention had an overall positive impact on grade retention. However, my results also indicate that the intervention had a higher impact on female students

CEU eTD Collection

compared to male students. Additionally, the intervention was found to have a higher impact on students from more advantaged households compared to student from less advantaged households. Finally, I also show that the interventions had heterogenous effects on different states that are possibly explained by the state's socio-economic indicators.

This paper can be of relevance to policymakers in Egypt and elsewhere as it sheds light on the overall impact of a specific form of decentralization implemented in Egypt and highlights asymmetric effects for gender and income groups. As policymakers in Egypt are currently amidst an ambitious education reform effort, this paper can further enrich the debate by providing factual evidence on the implications of financial decentralization that achieves an equitable school funding on grade retention.

The paper is organized as follows: Section 1 is split to three sub-sections where I begin by providing an overview of the theory of SBM and decentralization, review impact assessments from experiences around the world, and describe Egypt's institutional framework while focusing on the 2008/2009 financial decentralization pilot program. In section 2, I present the data sources, elaborate on the construction of my dataset and outcome variable, and present the difference-in-difference empirical strategy and the assumptions that should be made to reach an unbiased difference-in-difference estimator. I conclude section 2 by presenting my results for the average impact of the pilot program and a further decomposition of the main effects by gender, income groups, and regions. Finally, in Section 3 I discuss the results in light of the available literature and highlight relevant policy implications.

CEU eTD Collection

Section 1 Literature Review

1.1 An overview of School-based Management

For decades, international agencies such as the World Bank and the U.S. Agency for International Development (USAID) have been promoting and financing decentralization reforms in low and middle income countries (Smoke 2001). In fact, one of the major proposals for educational improvement in the World Bank's World Development Report (WDR) 2018 is School-based management (SBM) (World Bank 2017). In their survey of the literature on SBM, (Malen, Ogawa, and Kranz 1990) define SBM as a formal change in the government structures which identifies individual schools as the initial unit of improvement and allows for the decentralization of decision-making authority to be the main means through which the improvement is motivated and sustained.

The extent to which a decision-making authority is devolved to lower levels determines the intensity of the decentralization program. A wide selection of SBM programs are implemented across the world due to the variety of the combinations available in terms of who receives the decision-making authority and in terms of the degree of decision-making actually decentralized (Santibañez 2006). In general, SBM reforms transfers authority over one or more of the following activities: curriculum development; staff management; budgeting; maintenance and infrastructure; and monitoring and evaluation (Patrinos, Barrera-Osorio, and Fasih 2009). Consequently, due to the uniqueness of both the implemented SBM program and the pre-intervention conditions of each country and the unavailability of a standardized measure of decentralization, most topical surveys have found it tough to reach a generalizable impact for SBM (Carr-Hill et al. 2015; Westhorp et al. 2014).

The complexity of modern education systems heavily strain the technical, managerial, and financial capacity of governments (King and Guerra 2005). In the past, centralized systems in general and especially in low-income countries, have relied heavily on physical inputs – such as classrooms, teachers, and textbooks – to ensure a successful education system (Jimenez and Paqueo 1996). On the

CEU eTD Collection

contrary, the primary incentives that impact learning outcomes are of an institutional nature such as school autonomy, accountability, and exposure to choice and competition (Hanushek and Wößmann 2007). In theory, parents who are concerned with maximizing their children's learning outcomes are able to choose the most productive school, in terms of academic results for example. Subsequently, this demand side pressure can potentially incentivize schools to perform better in competition for students. In a similar manner, allowing schools to become autonomous can improve learning outcomes if they were also held accountable.

As explained above, it is presumed that devolving decision-making authority to the level of individual schools can have an ultimate positive impact on student education outcomes (Carr-Hill et al. 2015). In fact, there are several causal pathways by which decentralized decision-making can eventually have a positive impact on student's learning outcomes. First, increases in accountability combined with higher responsiveness to local needs is assumed to strengthen the stakeholder's perceptions of the education provisions, which is in turn expected to have a positive impact on enrollment, attendance, and repetition rates along with decreasing corruption within schools. Second, higher accountability can potentially make the school's recruitment decisions be based on teacher's performance rather than recruitment just on the basis of qualifications. Eventually, when the teacher's performance is monitored and factored in during the recruitment process, it can lead to reduced absenteeism, increases teacher motivation and ultimately, improvements in the quality of teaching within schools. Third, decentralization of funding, which is aimed at making school spending become more efficient, combined with the strengthening of community participation channels can lead to more resources being available to schools, which is not only an important factor for an improvement in education quality but also crucial for decreasing the cost of education provision. Therefore, school-based decision-making models can impact final education outcomes through the abovementioned intermediate outcomes. It is assumed that an effective high intensity SBM project can positively impact

CEU eTD Collection

enrollment, attendance and repetition rates, in addition to improving the equality of access to education.

1.2 School-based Management around the world

In this study, I am primarily concerned with the impact of the funding decentralization pilot on individual level grade retention, possibly capturing one aspect of education quality. In this section, I review the literature that investigate the effects of devolving decision-making to individual schools. As noted earlier, since SBM programs greatly differ when implemented due to the specific objectives of the reforming government in each case, attempts to generalize the results from each study should be questioned. I rely on four review studies performed on the topic of school-based management and education interventions in order to identify the trends of the impact decentralization reforms on grade repetition (Carr-Hill et al. 2015; Patrinos, Barrera-Osorio, and Fasih 2009; Santibañez 2006; Westhorp et al. 2014). The four review studies applied rigorous methods to qualify the papers included in the surveys. This allows me to only focus on credible sources that estimate a causal impact of various SBM interventions on student retention.

(Skoufias and Shapiro 2006) employ a panel data regression analysis and propensity score matching to evaluate Mexico's Quality Schools Program (PEC). The PEC program is an intervention program which offers grants to schools with the aim of decentralizing decision-making within schools. They find that the PEC program significantly decreases repetition by 0.31 percentage points. However, they don't find any significant effects of the intervention for schools in indigenous communities.

(P. Gertler, Patrinos, and Rodríguez-Oreggia 2012) evaluate a different SBM intervention in Mexico, AGE. The AGE program is a parental empowerment program that provide resources and training to parents to improve rural schools in Mexico. The authors used a three-year randomized difference-in-differences design to show that the program significantly reduced failure and repetition rates.

CEU eTD Collection

In El Salvador, (Jimenez and Sawada 2013) use observational data to investigate if the SBM intervention affects educational outcomes measured by retention rates. The SBM intervention they study is the EDUCO program which is designed to enhance community participation in the education process through the contribution of meals, time, and labor to improve schools. After controlling for selection bias and other issues of endogeneity they find that the EDUCO program had a significantly positive impact on retention rates.

In Brazil, Paes de Barros and Mendonca, in an article featured in (Savedoff 1998), use census data, household surveys and evaluation data to empirically study the impact of three SBM interventions on retention rates along with other education outcomes. They find that SBM interventions positively impacted retention rates in states where the principal was elected, school councils were established, and had financial autonomy. They included a series of control variables to deal with biases resulting from omitted variables. After the inclusion of the control variables, they found that the impact on retention rates was the only effect that remained robust and highlighted financial autonomy as being more important than other interventions.

In Indonesia, (Pradhan et al. 2011) evaluate the effect of four randomized interventions with the aim to strengthen school committees in primary schools. Schools were randomly allocated to a control group or a treatment group that receives a grant in addition to a combination of three interventions: training for committee members, election of committee members, or a collaboration between committee and village council. In a two-year period after the implementation, the study finds no significant effects of any of the interventions on student retention rates.

Generally, the literature shows a significantly positive impact of the various SBM interventions on retention rates in different contexts. However, the values of the impact vary among studies primarily due to the uniqueness of each intervention and the unavailability of a standardized measure to allow for a comparison. Also, most studies don't conduct any sub-group analysis relating to individual

CEU eTD Collection

characteristics, such as gender or socioeconomic status, which can be important to understand if the impact of those interventions affect sub-populations differently. However, there is some general evidence that SBM intervention have a stronger impact on wealthier students at a younger age and relatively weaker impact on disadvantaged communities (Carr-Hill et al. 2015). Later I show that the impact of the financial decentralization pilot program in Egypt didn't differ from the global literature in terms of its impact on grade retention. Moreover, through performing sub-group analysis I show that decentralization program has different effects on communities that vary based on their gender, income levels, and/or regional characteristics.

1.3 The case of Egypt

1.3.1 A centralized framework

The Egyptian education system has a long tradition of severe centralized control over both the administrative and financial aspects. This is a well-documented observation that was concluded in many previous studies (ElRashidy 2018; Ginsburg et al. 2010). It is characterized by a hierarchal authority structure where the Ministry of Education (MOE) is at the top of the pyramid with the majority of the decision-making power. Additionally, the MOE holds the power to override any decisions made at the subsequent lower levels. Although regional educational directorates were devolved some powers by the Education Law 139/1981, the devolved authority turned out to be more rhetorical than substantive (Hammad 2013). Subsequently, the regional directorates were unable to change the education system to fit the specific needs of their regional environments (Khaleel 2003). The structure is further broken down to constitute of district offices (idara) directly responsible for schools in their geographical area within each state (governorate). This bureaucratic structure restrains and limits local initiative as each lower administrative level must receive approval from a higher office reaching eventually to the MOE which controls the education processes in the country (Hammad and Norris 2009).

CEU eTD Collection

The problem becomes even worse at the local school level where schools have minimal decision-making power over issues such as selection of textbooks, evaluation of teachers, and allocation of budgets (Hanson 1990). The role of principals and head teachers, as noted by (Shehata 1997) found in (Hammad and Norris 2009), are only limited to being the MOE's representatives within schools to oversee the implementation of the directives issued by the central authorities or coordinate between the school teachers and refer school problems to an authority further up for solutions.

Furthermore, the national budget structure reflects the centralized configuration described above. Spending requests must pass from the village to the district to the state (governorate) level which are then compiled to form the budget request by each ministry and submitted to the Ministry of Finance (MOF). Eventually, budget ceilings are determined in a negotiation process between the state (governorate) directorate and the Ministry of Finance before being approved by the parliament as part of the annual budget process (Boex 2011). This mode for budgeting puts schools in disadvantaged areas at risk of not receiving adequate funding as it allows for favoritism and nepotism to dominate the budgeting process (Dixon, Bhuiyan, and Üstüner 2018).

This heavy centralization had a negative impact on the quality of the Egyptian education system as it allowed no room for local schools to get rewarded for demonstrating personal initiative. Moreover, Egypt's population growth adds further stress on the government's limited resources. The number of children enrolled in elementary school increased from 9.5 million in 2005 to 12.2 million in 2017. However, it is estimated that 30 percent of school children lack the basic skills for reading and writing (Education in Egypt 2019). Also, the prevalence of repetition, absenteeism, and dropping-out is negatively impacting the efficiency of the education system (Krafft 2012). To further complicate the problem, since school resources are not distributed in an equitable manner, the education system further amplifies the existing socio-economic inequalities as students from disadvantaged socio-economic communities report more infrastructure problems (Krafft 2012). Therefore, the increase in

CEU eTD Collection

enrollment figures although is an improvement towards access indicators do not necessarily reflect the true situation regarding the quality of education.

Grade repetition can be regarded as a form of delay in the education process practiced by many countries to allow students experiencing difficulties to catch-up with their peers. Many researchers – in different contexts – concluded that it is correlated with dropping out of school and has adverse labor implications (Eide and Showalter 2001; Manacorda 2012; Roderick 1994). Therefore, it can serve – for the purpose of this study – as a proxy for educational quality.

The "*Survey of young people in Egypt*" (Roushdy et al. 2010) specifically reports on the situation of grade repetition in Egypt. Table 1.1(see section Tables & Figures) reports that about 6.3 percent of the students currently attending primary school experience grade retention while 27.9 percent of the children who previously attended school experienced grade repetition in primary school. Repetition rates are relatively higher for the secondary and the post-secondary education stages compared to the primary and preparatory stages. Among the sample of people who previously attended school, about 39.4 percent of those who attended preparatory and stopped during or after this level repeated a year at least once. This shows that experiencing grade repetition at an early stage is critical for the progression to higher education stages.

The survey results by (Roushdy et al. 2010) are further broken-down by sub-samples: First, they find that male students are more likely to experience grade repetition than females. The gender gap is smaller in the basic education stages and widens at the higher education stages where males are twice as likely to experience grade repetition; Second, they find that repetition rates were found to be higher among rural students relative to urban students. Specifically, 20.4 percent of students in rural upper Egypt experience repetition making them the highest impacted group; Third, they find that wealth is strongly correlated with repetition. About 23.4 percent of people coming from the 20% poorest

CEU eTD Collection

households experience grade repetition relative to 9.2 percent of the wealthiest quintile. The gap between wealth groups is smaller at higher education stages compared to basic education.

Therefore, based on their results from the survey, it is clear that the experience of repeating a grade is non-uniform across different sub-groups and implementing policies that doesn't factor in these structural differences is more than likely to impact the different sub-groups in an heterogenous manner.

1.3.2 Recent SBM/Decentralization initiatives in Egypt

For the past two decades, the Egyptian government – led by the National Democratic Party (NDP) at the time – with support and funding from international organizations began to promote and experiment with school-based management interventions and education decentralization reforms as a way to improve the education system. As mentioned earlier, the prime objective of SBM and decentralization reforms is to strengthen the local community and/or officials at the lower levels of the hierarchy. Unfortunately, the government only managed to make half steps towards achieving decentralization. The political climate at the time was characterized by tensions between the government and the public. Among the reasons that impeded the progress of decentralization and SBM reforms was; the fear of the Islamist groups from dominating the local councils, the public officials doubted the efficiency of the local bodies, or resistance from public officials that benefitted from enjoying greater authority and power(Ginsburg et al. 2010; Ibrahim 2010).

On the other hand, international organizations pushed forward with promoting decentralization and SBM as a tool for improving the education in the Arab world (El Baradei 2015; Ginsburg et al. 2010). Some examples of those initiatives are the 2001 experimental decentralization effort of the education system in Alexandria, the 2003 National Standards of Education issued by the MOE, in which they encourage stakeholders within schools and the community to get involved in the annual standardization of education to receive accreditation. Other initiatives included the establishment of

CEU eTD Collection

the Training and Evaluation units (TEU) in 2000. The objective of the TEUs was to provide local schools greater discretion over self-assessment and identifying their own training needs. Also in 2005, the Boards of Trustees, Parents, and Teachers (BOT) was established in all schools with main aim of strengthening the local community's involvement within the school and locate other sources of funding for the school (Hammad and Norris 2009).

In this paper, I focus on a specific SBM intervention that managed to use a funding formula to decentralize school funding directly from the Ministry of Finance (MOF) to individual schools.

1.3.3 The 2008/2009 financial decentralization pilot program

In 2008/2009, the Egyptian government piloted a financial decentralization program in three pilot governorates; Fayoum, Luxor, and Ismailia. This section is exclusively based on the work of (H. Healey, Hanna, and Attalla 2016) where they report on the technical effort undertaken to implement the aforementioned pilot program. After the pilot program, the Egyptian government continued to decentralize a fraction of the education budget until 2013 when the program was ended due to the volatile political environment at the time. Therefore, in this paper, I only assess the impact of the pilot program in the fiscal years 2008/2009 and 2009/2010 on student's grade retention and exclude the following years from the analysis.

A total of 8.4M EGP and 10.1M EGP from the Government of Egypt's budget were decentralized to all primary schools in three states (governorates) in the fiscal years 2008/2009 and 2009/2010 respectively. The main objectives of the initiative were achieving; equity in funding, local control over finances, accountability, transparency, and capacity building. These intermediate outcomes as theorized in the literature on SBM and decentralized education finance should eventually improve student's education outcomes.

To achieve equity in funding, an enrollment and needs-based funding formula was developed where the poverty level of each state (governorate) was factored-in to determine the funding that will

CEU eTD Collection

eventually reach each school. As noted earlier, the funding to schools was previously wrongly channeled to schools that had relatively less need for the funding. These disparities in funding can negatively impact the quality of education in the most disadvantaged areas. (H. Healey, Hanna, and Attalla 2016) examine the district's non-personnel recurrent budgets and find that the money they received from the Ministry of Finance (MOF) varied in per-student terms by ratios of as much as 100 to 1.

Along with providing the local schools with control over finances comes a lot of freedom on the ways this money can be used. When the intervention was evaluated at the end of the pilot program, it was found that schools actually effectively used the decentralized funds in areas that required additional funding and then proceeded to collect the additional funds from the community. The evaluation at the end of the first year showed that 73 percent of the schools participating in the pilot programs displayed their expenditures to the public – achieving higher transparency – and a 51 percent increase in community funding was achieved, demonstrating higher accountability.

Additionally, continued trainings and implementation support was given to all stakeholders involved in the process to increase their individual and institutional capacity. An important element in this program was the institutionalization of the initiative in a ministerial decree which established a decentralization support unit within the MOE's organizational structure. Moreover, the pilot program was monitored by a committee formed from high level officials and technical experts from the Ministry of Finance (MOF), Ministry of Education (MOE), the Ministry of Local Development, the Ministry of Economic Development, and the Central Agency for Organization and Administration. So, it is safe to claim that there was a high level of government support and political will during the pilot phase of the intervention.

In the following section, I proceed to evaluate the impact of this financial decentralization pilot program on grade retention to capture one aspect of education quality. As far as I am aware, an impact

CEU eTD Collection

evaluation of this decentralization pilot program is important as the program wasn't empirically evaluated so far. In general, there is a lack of empirical evidence to support education decentralization interventions in Egypt (El Baradei 2015). However, one paper stands out where (Nasser-Ghodsi 2006) used state-level data to evaluate the impact of the Boards of Trustees, Parents, and Teachers initiative and finds no significant impact on student education outcomes. In this paper I use household survey data which allows me to perform sub-sample analysis in order to understand the impact of the decentralization intervention on different groups. Due to the lack of empirical papers on the Egyptian pilot, I believe that this evaluation can shed some light on relevant issues that can advance the debate on SBM and financial decentralization in Egypt and the Arab world.

CEU eTD Collection

Section 2 Impact Assessment

2.1 Data Overview

An ideal dataset for the identification of the true causal effect of the funding decentralization pilot program on grade retention should meet two criteria. Individual level information on grade retention should be available and selection into the treatment group should be effectively randomized (Lechner 2010). On the other hand, uncovering causal effects from a repeated cross-section of observational data can be problematic because we are unsure if the treatment and control groups are effectively random groups. An experimental setting is one where the Ministry of Education randomly chooses several primary schools to enter the pilot where the retention experience for all students was observed for an adequate period of time before and after the implementation of the pilot program.

In this paper, I use labor force surveys (LFS) for the years 2006-2011 produced by the Economic Research forum (ERF) and the Central Agency for Public Mobilization and Statistics in Egypt (CAPMAS). The Egypt Labor Force Survey is an annual representative population survey that includes the results of quarterly survey rounds that represent the labor force characteristics during the year. One of the aims of the survey is to collect information on the employment status and distribution by geographical area, gender, age, and education status. The responses used to construct the LFS are collected across the year over 4 rounds.[1] Since the LFS surveys households and its members, it also includes information on the education status of children in households. One of the variables of interest collected in the LFS is the number of successfully completed years of school. Because the academic year in Egypt begins in September and ends in June, the last two rounds in each survey year include information on the previous successfully completed academic year.

CEU eTD Collection

[1] The duration of each round is as follows; round 1 is from 1st of January to the 31st of March, round 2 is from the 1st of April to the 30th of June, round 3 is from the 1st of July to the 30th of September, round 4 is from the 1st of October to the 31st of December.

I compiled all years of the survey to produce a cross sectional dataset of individuals that spans the period of the decentralization pilot program. The dataset includes educational responses from the Academic years 2005/2006, 2006/2007, and 2009/2010 when the pilot program ended.[2]

I use the academic year as the time period of reference. The 1st pre-treatment period is found using the first and second rounds of the 2007 LFS.[3] The students report their successful years of education based on their performance in the 2005/2006 academic year. The 2nd pre-treatment period is found using the 3rd and 4th rounds of the 2007 LFS and the 1st and 2nd rounds of the 2008 LFS. The students report their successful years of education based on their performance in the 2006/2007 academic year. Finally, the post-treatment period is found using the 3rd and 4th rounds of the 2010 LFS and the 1st and 2nd rounds of the 2011 LFS. The students report their successful years of education based on their performance in the 2009/2010 academic year. Table 2.1 shows a visual representation of this structure.

I supplement the dataset with subnational data from the Global Data Lab[4] to include subnational covariates in my regression models. The subnational data includes information on the subnational human development index, income index, and health index (Smits and Permanyer 2019). The subnational covariates are of relevance because the treatment occurred at the level of the state and controlling for different subnational characteristics can capture endogenous effects.

Test scores are also often used as a proxy for education quality, especially in education decentralization studies (Faguet and Sánchez 2008; Galiani and Schargrodsky 2002). I was unable to use test-score data on Egypt as the available survey of young people in Egypt produced by the ERF, or the Trends in International Mathematics and Science Study (TIMSS) and Progress in international

CEU eTD Collection

2 Since, the pilot program was implemented nationwide for the academic year 2010/2011, I was unable to include responses from the 2012 survey and further in the analysis.
3 I was unable to use the 3rd and 4th round of the 2006 LFS because the successful years of education variable was coded differently.
4 https://globaldatalab.org

Reading Literacy Study (PIRLS), is not compatible with the decentralization pilot under study for Egypt mainly due to the political unrest in 2011 which prevented the collection of data during the period or participation in the assessments. The only available option was to construct the grade retention variable from the responses of the education outcomes in the Labor Force Survey.

The data I use includes information on 28,331 individuals from the 1996/1997 cohort observed for 3 time periods. The analysis includes two time periods before the implementation of the decentralization pilot program and only one time period after the implementation. Treatment status is well defined at the state level. There are 1,612 observations from the treatment states and 26,719 observations from the control states. In addition to observing the successfully completed years of education, enrollment status, and individual characteristics, we also observe the parent's education and labor features and household characteristics.

I restrict my analysis to non-migrants since research has found that parental migration adversely impacts their children's education and health outcomes (Meng and Yamauchi 2015). I only retain individuals living with both their biological parents as evidence shows that children in other family structures (single parent, step parent) are more likely to achieve lower education outcomes (Parke 2003). I drop observations with parents who are less than 25 years old or more than 80 years old. I only retain individuals who have "student" as their main activity status to eliminate the impact of child-work which can be influential especially in a rural setting. I also limit the size of households to 12 members to arrive at a comparable sample. Since compulsory schooling in Egypt starts at age 6 and pre-school permits children starting from the age of 4, I eliminated all observations where the difference between their years of education and age is 3 or less.

The experience of grade retention has been used by researchers as a proxy for education quality. (Cockx, Picchio, and Baert 2019) find that grade retention has adverse effects on schooling outcomes where the less able students are impacted the most. However (Psacharopoulos and Velez 1993)

CEU eTD Collection

discredit the idea that grade retention may have a substantial effect on future labor outcomes. But since I am concerned about the quality of the education process, grade retention can serve as a good proxy for education quality as it is documented that repeating a grade is closely associated with an increase in the odds of dropping out (Roderick 1994).

To construct the main outcome variable of this study (grade retention), I calculate the average successful years of schooling for each cohort, state, and academic year. I consider an individual to have experienced grade retention if their successfully completed years of school is less than the floor value of the average years of schooling for their age, state, and academic year.[5] Therefore, the estimates I compute can be considered as lower bound estimates.

Table 2.2 shows the proportions and standard errors of key variables in the 1996/1997 cohort population by treatment status.[6] Panel A reports the proportions of individual characteristics, B parent's characteristics, C reports household characteristics, and D regional characteristics. Looking at Primary Enrollment and Lower Secondary enrollment in A, we can see that the larger share of students was promoted from the primary stage to the lower secondary stage during the implementation of the decentralization pilot, with a small fraction left behind. Based on observed individual and parent characteristics, we can see that both groups are not substantially different from each other. However, looking at household characteristics, we can see that treatment observations demonstrate higher number of members and a different dwelling structure than control groups. I deal with this variation in the sample by implementing the k-nn nearest neighbor matching to arrive at a relatively more comparable sample as part of the robustness checks. All group-specific and time-fixed effects will be controlled for using the model discussed below.

CEU eTD Collection

[5] If the average successfully completed years of school for a certain cohort, in a specific state, in period t is 4.5, I consider an individual to have experienced retention if their successful years of schooling is less than 4. I also ran the models using the mean value instead of the floor and the results were the same (see section 2.3.6)

[6] Table A.1.1 and A.1.2 in appendix I shows the proportions and standard errors of the 1994/1995 cohort.

Due to data restrictions I implement a standard difference-in-differences design. I am aware that the unmodified difference-in-differences approach forces the assumption that the impact of the program occurs immediately and is constant over time. Also, the pilot program under study was implemented for two consecutive academic years, after which the program was implemented nationwide until 2013. This progression of the project prevents me from using post-treatment periods in my analysis. Because the decentralization program was implemented nationwide at the same time following the pilot program, I was unable to include more post-treatment periods in my analysis as all states that constitute the control group were effectively treated at the same time. However, it should be noted that education policies take time to assume its full effect. Therefore, it is highly likely that the effect of the funding decentralization program is cumulative in nature which is not captured in this study.

2.2 Empirical Strategy

The objective of this study is to estimate the effect of the 2008-2010 financial decentralization pilots in Egypt on grade retention to possibly capture one aspect of education quality. I'm also interested in identifying any gender differences and distributional effects of the program. Since the highest retention rates usually occur at the end of an education stage, I estimate the effect of the program on the 1996-1997 cohorts, who completed their primary education stage during the implementation of the pilot program. The program decentralized the funding in primary schools in 3 regions (Fayoum, Ismailia, Luxor). For this purpose, I use the 1994-1995 cohorts, who completed their lower secondary stage, to serve as a placebo group and also to trace spill-over effects of the program, if any, on other educational stages.

The fundamental issue with program evaluation is to successfully arrive at an unbiased estimate of the effects of a specific intervention, also known as the causality problem in the field economics. Interventions occur in a dynamic environment. If a time-trend exists in the outcome of interest, then

CEU eTD Collection

a simple differencing of the mean outcome before and after the intervention would result in a biased estimate. Potentially, the intervention not only causes changes in various factors but also in the structural ways those factors are related to each other, which in turn, makes isolating the direct effect of an intervention problematic. For example, if the average grade retention in a specific state was collected before the intervention and then observed again after the intervention, a simple differencing of both measurements is merely a gross effect of the intervention. The counterfactual condition is what would have happened to our outcome of interest (average grade retention) regardless of the intervention. Our inability to observe the counterfactual condition biases our estimate through unobserved effects that vary on an aggregate level such as region-specific effects or time-fixed effects.

One way to approach this problem of causality is the difference-in-differences (DD) approach. DD studies have become a standard technique used by researchers in economic policy, healthcare policy and other disciplines for program evaluation and impact assessment (economics: (Blundell, Brewer, and Shephard 2005), health-care: (Rao et al. 2014)). The popularity of this technique most probably comes from its flexibility and ability to reach an unbiased estimate of the program under study.

In principle, naturally occurring policy changes can be used as a quasi-natural experiment that splits the population into a well-defined treatment group and a control group that are observed in two time periods (or more). To isolate the effect of the policy, the method finds the difference between the differences in outcome of each group. The method relies on the control group to construct the treatment group's counterfactual situation. The unobserved time-invariant group-specific and group-invariant time-fixed effects are isolated from the estimate of interest through assuming that the average outcome in the treatment group would have demonstrated a similar trend to the control group.[7]

[7] I note here that in an RCT, group-specific effects are not a problem because both groups are identical to each other and only differ because one group receives the treatment while the other does not. However, in this study I am using a cross-sectional data where the composition of the groups varies across space and time.

CEU eTD Collection

In this paper, I would like to compare the grade retention outcome of a student attending schools in the decentralized regions to the counterfactual situation i.e. the same student's grade retention outcome if they attended a school in a non-decentralized region. Since we can only observe outcomes of an observation in one situation, I could use student's education outcomes in non-decentralized regions to reflect this counterfactual situation if the identification assumptions hold. Formally, following (Wooldridge 2007) I estimate the following regression:

$$Y_{ist} = \beta_0 + \beta_1 D_{is} + \beta_2 after_{it} + \beta_3 (D_{is}.after_{it}) + \lambda_t + \delta_s + \beta_4 X_{ist} + \beta_5 Z_{st} + \varepsilon_{ist} \qquad (1)$$

Where; Y_{ist} is the outcome of interest (grade retention, in our case) for individual i = 1, …, N; in state s = 1, …, N; in each time period t = 1, …, N. D_{is} is a treatment dummy which takes the value of 1 if the observed student *i* is from a state *s* that has been included in the pilot program. $after_{it}$ is a post-treatment dummy that takes the value 1 if the student is observed after the intervention took place. λ_t and δ_s are a vector of time period effects and region effects respectively. X_{ist} and Z_{st} include individual covariates (such as rural/urban household, education of head of household, age of head of household…) and regional covariates (such as sub-national hdi, income index, health index…) respectively. Finally, $(D_{is}.after_{it})$ is an interaction between the treatment and the post-treatment dummies. It takes the value 1 if the student is from a treated region after the pilot program.

β_0 captures common trends (baseline average) in the control group i.e. the control group specific effect. β_1 captures the difference in average grade retention between observations in the treatment and control groups prior to the intervention, i.e. the treatment group specific effect. β_2 captures the time trend in the control group. λ_t is a vector of year dummies which captures time-fixed effects common to both groups. δ_s is a vector of region dummies which captures time-invariant region-specific effects. ε_{ist} is the unobserved "error" term which contains all determinants of Y_{ist}omitted from our model. Also, by construction, the "error" term is assumed to be randomly distributed and independent of the right-hand variables. Using the above regression formulation, I can estimate β_3 which is the estimate

CEU eTD Collection

of interest that represent the average effect of the pilot program on grade retention in the treatment regions post-decentralization which did not exist prior to the decentralization pilot program.

However, for the coefficient of interest to have a casual interpretation, the key identifying assumption is that important unobserved/unmeasured variables are either time-invariant group-specific effects or group-varying time-fixed effects (Wing, Simon, and Bello-Gomez 2018). This implies that outcome trends in each group should demonstrate a common trend before the intervention. Therefore, a (relative) common trend prior to treatment between both groups lends some credibility to the difference in difference design. A violation to the common-trends assumption would occur when some variables other than treatment change in one group but not in the other. Therefore, a researcher should have sufficient pre-treatment data in order to assess if the common trends assumption is satisfied.

In my case, due to data limitations, I can only observe the target cohorts for 1 period prior to the intervention. The (conditional) common trends assumption, following (Callaway and Sant'Anna 2018), states that, conditional on covariates, the average outcome for the treated group and the control group would have followed a common trend in the absence of treatment. This is a simple modification to the standard common trends assumption as it is assumed that the two groups follow a common trend only after conditioning on covariates, hence the conditional common trends assumption.

In order to formally test this assumption, we can run the following regression (Pischke 2005):

$$Y_{ist} = \beta_0 + \beta_1 D_s + \lambda_t + \sum_{j=a}^{c} \beta_j \left(D_{is}.time\_period_{ij}\right) + \beta_4 X_{isj} + \beta_5 Z_{sj} + \varepsilon_{ist} \qquad (2)$$

where now the variable $time\_period_{ij}$ is entered into the regression as a series of dummy variable for each time period ($j = a, b, c\ for\ each\ time\ period\ respectively$). The variable $\left(D_{is}.time\_period_{ij}\right)$ is an interaction term between the treatment variable D_{is} and $time\_period_{ij}$. This form of the regression produces a vector of treatment effects β_j for each lead i.e. we allow the treatment effect to vary for

CEU eTD Collection

each time period. As before, I also include a group and time period fixed effect. $\beta_3$and β_4 are a vector of coefficients for each individual and state-level covariate respectively.

I perform a joint coefficient test and test for the following null hypothesis;

$$H_o: \beta_a = \beta_b = 0,$$

Failing to reject the null hypotheses allows me to conclude that indeed the pre-treatment trends in the control group are similar to the treatment group.[8] In other words, I conclude that the trends are common if the F-statistic and corresponding p-value generated from the joint coefficient test indicate that the result of the test are not significant according to the common significance levels.

Another threat to identification is the violation of the strict exogeneity assumption. The outcome trend may change in anticipation of the intervention. For example, the education quality may worsen prior to the intervention because the education system in treated states expects the intervention to improve the outcome variable of interest. If the treatment group alters their behavior prior to treatment (also known as the Ashenfelter's dip, see (Heckman and Smith 1995), then my estimates will be biased as it is a violation of the common trends assumption. I argue that since the grade retention variable is observed 3 years before the intervention begins, it is plausible to assume exogeneity. Additionally, based on (F. H. Healey, Crouch, and Hanna 2014), the financial education decentralization pilot programs were perceived as "unexpected new money" by the subnational units. As discussed above, subnational allocations were transferred to the central government and subsequently delivered to the subnational units as goods and services. Therefore, it is unlikely that any behavior changes happened.

Another issue that should be noted is violation of the exogeneity assumption through time-varying covariates. Since I am working with a cross-sectional data, it is likely that the trend in covariates such

[8] Further elaboration and a discussion can be found in (Autor 2003) where he includes leads and lags of the treatment effect to study the dynamics of the impact of employment protection regulation on outsourcing by firms.

CEU eTD Collection

as total wages of the household vary. Additionally, since there are inherent differences between groups, non-randomness of treatment is usually a threat to identification.

In order to address these issues of endogeneity and baseline selection, researchers usually implement the propensity score matching method to handle confounding (Stuart et al. 2014). The propensity score is simply defined as the likelihood for an observation to be selected into treatment. This is a two-stage process where I first estimate the propensity score based on individual covariates using a logistic regression. I then match observations based on the nearest neighbor followed by a weighted regression in order to compare observations that have approximately similar values of the propensity score within a caliper.

The benefits of performing the propensity score matching is that it helps to make a more robust inference as it reduces dependence on the outcome model specification (Ho et al. 2007). Second, it effectively deals with the problem of the high dimensionality as the method reduces the many dimensions that explain selection into one propensity score. Therefore, it becomes easier to balance both groups based on a single scalar summary score (Rosenbaum 2010).

Finally, it should be noted that although some assumptions are verifiable it is impossible to fully check the assumptions in the model as they are essentially made about unobservable factors.

2.3 Results

In this section I present the estimates of the effects of the decentralization pilot program on student's grade retention. I begin by estimating the effect using two samples; the 1996/1997 cohort as the main effect of the pilot program, and on the 1994/1995 cohort to test for any spillover effects on other stages of education. As discussed in the data section, I observe the successful completed years of education of the 1996/1997 cohort in three time periods that coincide with the academic years 2005/2006, 2006/2007, and 2009/2010. The pilot program was implemented in all primary schools in three treatment states (Fayoum, Luxor, and Ismailia). After a year of preparation work, the

CEU eTD Collection

Ministry of Education successfully decentralized the funding for all primary schools in the treated states for the academic years 2008/2009 and 2009/2010.

The 1996/1997 cohorts should have completed the primary education stage during the pilot program. I assume that the experience of retention is more common at the end of an education stage primarily because the students should sit for a state-wide examination to get promoted to the next education stage. For the other grades (2nd or 3rd), the exams are conducted by the school administration.[9] It should be expected then that retention is experienced relatively more at the end of an education stage where schools have less control over the examination compared to other education stages.

According to decentralization theory, fiscal decentralization is expected to improve service delivery as decision making powers moves away from central authority towards local authority, which is assumed to be more in touch with local needs (Oates 2008). I expect that the sign of the interaction coefficient to be negative suggesting a positive effect of the pilot program on the probability of experiencing grade retention.

On the other hand, the 1994/1995 cohort completed the lower secondary education stage during the pilot program. Since the data source I'm using doesn't include any school identifiers, I can't observe the type of school (public vs. private) for each student or if the school also includes the lower secondary stages. I assume that public primary schools and public lower secondary schools are not mixed. Therefore, I use the 1994/1995 cohorts to test for spillover effects of the pilot program in lower secondary schools in the treatment states.

In a highly centralized education system, if reforms are implemented in one education stage, it is common for information to propagate through informal channels to schools of other stages of

[9] Article 19, Law No. (139) year 1981.

CEU eTD Collection

education (Besley and Coate 2003). If this was true, I expect that the untreated schools in treated states would alter their education processes in anticipation for treatment in a future period. Alternatively, a positive effect of the pilot on the untreated lower secondary schools could also be due to an increased supervision from the central authorities in treated regions. However, I do not test these additional hypotheses in this paper.

2.3.1 The effect of the decentralization pilot program on grade retention

Tables 2.3 and 2.4 present my estimates of the main effect and spillover effect of the pilot program on grade retention. The tables are structured as two big columns for each cohort grouping; 1994/1995 (I) and 1996/1997 (II). Each big column includes three additional columns for the three models used. On the left of the table, each panel indicates the sample used to estimate the effect, in addition to the total number observations and RMSE. I run the standard regression formulation presented in equation (1) progressively and include additional covariates for each of the three models. Firstly, model 1 includes a group fixed effect which captures fixed effects common to the treatment group relative to the control group. It also includes a post-treatment effect which captures effects that are common to all observations after the treatment period. Model 2 adds a second layer of detail as it includes several state specific fixed effects and time-period specific fixed effects common to each state and time period, respectively. Finally, I conclude with Model 3 by adding several individual-level and state-level characteristics to minimize confounding.[10]

I am interested in correctly estimating both the direction and magnitude of the effect of the pilot program on grade retention. I report robust standard errors clustered at the level of the sampling unit to account for serial correlation of the error terms resulting from the survey nature of my data (Abadie

[10] Individual-level covariates are age, total number of household members, total number of male household members, total number 15+ household members in the labor force, total number 15+ male household members in the labor force, age of the head and spouse, & gender, literacy status, education status, and main activity status of the head, type and tenure of dwelling, and several dummy variables for dwelling amenities. State-level covariates are logarithm of population(thousands), subnational human development index (HDI), income-index, and health index.

CEU eTD Collection

et al. 2017). Although I follow the difference-in-differences strategy to control for confounding factors and perform several robustness checks (see below), I would caution from causally interpreting the magnitudes of my estimates as they might still be biased by unmeasured or unobserved variables that have different trends between the treatment group and the control group.

In Table 2.3, Panel A estimates the effect on the total population. For the 1996/1997 cohort, Model 1 (column 4) shows that the pilot program had no effect on the probability of experiencing grade retention for primary school students in the treated states. However, once separate state and time period effects are included the effect gains statistical significance and increases. Additionally, once all additional covariates are included in model 3 (column 6), the effect increases and indicates that the pilot program has led to a decrease in the probability of experiencing of grade retention by 5.2 percentage points for primary school students in treated states. For the 1994/1995 cohort, model 3 (column 3) shows that the pilot program was also associated with a decrease in grade retention by a comparable 5.4 percentage points. Both estimates in model 3 are significant at all common levels of significance.

The change in magnitude and significance of the estimate once the model progressively controls for other variables is an indicator that the estimate suffers from the omitted variable bias problem. Once our observations become more comparable by the inclusion of the different covariates, the estimate should reflect a less biased estimate as the covariates decrease the confounding between the interaction coefficient and grade retention. In other words, when we factor in the number of household members, for example, and compare two students coming from households with the same number of household members, we arrive at a less biased estimate because the number of household members controls for some of the observed variation in grade retention when it was first, wrongly, assumed that this variation belongs to the pilot program effect. I only include covariates that are not affected by the

CEU eTD Collection

policy, such that the correlation between the included covariates and the interaction term is assumed to be zero.

One plausible explanation of the estimate for the 1996/1997 cohort (column 6) is that it is in line with education decentralization theory, as mentioned above. When more autonomy was given to schools over their use of funds, service delivery improves as it becomes more in touch with the needs at the local level compared to a central authority that is less capable of understanding those needs. The report on the Egyptian pilot program (F. H. Healey, Crouch, and Hanna 2014) mentions that a year was spent on capacity building workshops, prior to implementation, to train local authorities on the new funding formulae and ways to effectively manage the funds. In this light, the estimated effect can be attributed to the pilot program.

On the other hand, another explanation is that the observed effect is, in fact, not due to the pilot program per se but due to being actively supervised by central authority. The report (F. H. Healey, Crouch, and Hanna 2014) also mentions that a new administrative unit was established in the treated regions. Unfortunately, testing the latter explanation is out of scope for this paper. However, since lower secondary schools were not treated, they could potentially serve as a placebo group. The significant effect for the 1994/1995 cohort gives us a taste that those effects could have been due to increased scrutiny by central authority in the treated states and should not be directly attributed to decentralized funding.

2.3.2 Gender Differences

Panels B and C in Table 2.3 present estimates of the effect of the pilot program separately on a sample of males and females respectively. In the 1996/1997 cohort sample (II), I compare 842 male students in the treated states to 14,001 male students in the control states. Only 45 male students experienced a grade retention in the treated states relative to 1,108 male students in control states. For females, I compare 770 female students in treated states to 12,718 female students in control states

CEU eTD Collection

out of which 59 students experienced a grade retention compared to 980 students in the treated and control states respectively.

The estimate from model 1 (column 4) shows a biased negative estimate of the effect of the pilot program on grade retention. In model 3 (column 6), the estimate becomes significant at the 10% and 5% levels but not at the 1% significance level. The pilot program is associated with a 3.8 percentage point decrease in the probability of experiencing grade retention for male students in the treated states. On the other hand, model 3 (column 6 in Panel C) reports that the estimate for female students in the 1996/1997 cohort (I) is a 6.7 percentage point decrease in grade retention due to the pilot program.

For the 1994/1995 cohorts (I) in Panels B and C (Table 2.3), the estimates from model 3 (column 6) show that the spillover effect of the pilot program on grade retention doesn't differ much when the model was run on male and female samples. Male students who completed their lower secondary education stage during the pilot program experienced a 5.8 percentage point decrease on average compared to female students who experienced a comparable 4.9 percentage point decrease in grade retention on average.

For the main treated cohorts (1996/1997), I observed that the pilot program impacted male students and female students differently. The results indicate that female students who completed their primary education stage during the pilot program benefited by more than 70% compared to their male counterparts. However, a similar conclusion can't be reached for the 1994/1995 cohorts as the decrease in the probability for experiencing retention is higher for male students but by less than 20%.

2.3.3 Distributional Effects and a break down by gender

I then proceed to break down the main results by total household income level.[11] Table 2.4 presents the interaction term coefficients from running regression equation (1) on different sub-samples as

[11] Household income level is calculated as the summation of the total wage of the head and the total wage of the spouse (if available). Missing values were given the average household income by state and time period for convenience.

CEU eTD Collection

indicated in the sub-headings of the panel to the left. Panel D estimates the effect of the pilot program on the sub-sample of students coming from households with below average income. For the 1996/1997 cohort (II), the main interaction coefficient on model 3 (column 6) shows that the pilot program decreased the probability of experiencing grade retention by 5.1 percentage points with a standard error of 0.022. The program also has a comparable effect on the 1994/1995 cohort (I) as indicated by model 3 (column 3) as it decreased the experience of grade retention by. 6.3 percentage points. I should also point out that there is higher selection in the 1996/1997 cohort relative to the 1994/1995 cohort based on the progression of the coefficients through the models.

Panel E in Table 2.4 estimates the effect of the pilot on the above average income sub-sample. Model 3 (column 6) shows that the pilot program also significantly decreased the experience of grade retention by 8.4 percentage points for the 1996/1997 cohort while having no effect on the 1994/1995 cohort.

Panels F-I in Table 2.4 breaks down the effects for each income group by gender. Panel F estimates the effect on the male students coming from households below average income. For the 1996/1997 cohort (II), model 3 (column 6) shows that the pilot program decreased grade retention by 3.7 percentage points with a standard error of 0.022 (significant at the 10% level but not at the 5% or 1% level) whereas for the 1994/1995 cohort (I) the program is associated with a 7.2 percentage point decrease with a standard error of 0.024.[12] Panel G shows that the grade retention for male students in the 1996/1997 cohort (II) coming from above average income households decreased by 6.8 percentage points due to the pilot program but was not affected for male students in the 1994/1995 cohort (I). Model 3 (column 6) in Panel H reports that female students, in the 1996/1997 cohort (II)

CEU eTD Collection

Households were considered to be below average income if their total household income is equal to or below the average household income by state and time period.

12 The estimate for male students coming from households below average income in the 1996/1997 cohort becomes significant at the 5% level when the matched sample is used (see Appendix II).

coming from below average income household experienced a 6.4 percentage point decrease in grade retention (significant at 10% and 5% significance levels but not 1%) and their 1994/1995 (I) counterpart experienced a comparable reduction of 5.6 percentage points (also significant at 10% and 5% significance levels but not 1%). Panel I report the estimates using the sub-sample for female students coming from above average income households. Model 3 (column 6) shows that the 1996/1997 cohort of female students experienced a 10.7 percentage point decrease in grade retention due to the pilots while the female students in the 1994/1995 cohort coming from above average income households were not affected by the pilot program.

The results demonstrate two things. First, comparing the magnitude of the results for the 1996/1997 cohort, for the full sample split by income group or for the gendered sample split by income group, indicates that my results are in line with (Galiani and Schargrodsky 2002) who show that decentralization policies can have distributional effects that theoretically depend on the extent communities can voice and defend their preferences. The pilot program impacted students coming from higher income households more than students coming from average to lower income households. Second, for the 1994/1995 cohort, the pilot program impacted the average to lower income students but not the above average income students. The same pattern also persists when the male and female sub-samples are each compared across income groups.

One explanation for this result relates to the preference of families and the structure of the school system in Egypt. Family income is one of the critical factors to be considered if a family wishes to send their child to a private school. Also, families in Egypt may prefer to pay the extra school fees for private schools for the later stages of education. Therefore, families who are relatively well-off and can afford private schools may choose to send their children to private schools only in the later education stages. As explained above, the pilot program may have spillover or anticipatory effects on other public schools in the untreated education stages through increased supervision for example. For

CEU eTD Collection

this reason, the pilot program had a positive impact on the grade retention for students in the lower secondary stage who are enrolled in public schools (proxied by below average household income). But the pilot program had no impact on students who completed the lower secondary stage who are enrolled in private schools (proxied by above average household income). Better data on the type of school can help uncover this effect further.

2.3.4 Regional Effects

For this section, I break down the sample by regional units and choose states that weren't included in the pilot program to create a control group for each treated state. The aim of this section is to investigate if the pilot program had heterogenous effects on the experience of grade retention for the 1996/1997 and 1994/1995 cohort in each state.

The pilot program was implemented in three states; Fayoum – North Upper Egypt, Luxor – South Supper Egypt, and Ismailia – Suez Canal region.[13] I exclude Ismailia from this part of the analysis because the Suez Canal regional unit is comprised mostly of frontier states that may have specific trends. The analysis also relies on the idea that students from neighboring states are relatively more alike than students from areas that are farther away (Hanchate et al. 2015; Somers et al. 2013).

Table 2.5 presents the interaction term coefficients from running the regression formulation in equation (1) on sub-samples of regional units. Column (I) and (II) presents the results for the 1994/1995 and 1996/1997 cohort respectively. Each column presents the results for each regional unit as indicated in the heading of the sub-column. Panel A uses the complete sample to obtain an average effect over the full sample. Panels B & C breaks down each regional unit's sample by gender. Finally, Panels D & E further explores the distributional effects in each regional unit.

CEU eTD Collection

[13] North Upper Egypt regional unit includes Menya, Beni Suef, and Fayoum. South Upper Egypt regional unit includes Sohag, Qena, Luxor, Aswan and the Red Sea. Suez Canal regional unit includes Ismailia, Sharkia, Port Said, Suez Canal, North Sinai, and South Sinai.

Column 3 in Panel A shows that the pilot program significantly decreased average retention by 9.9 percentage points for students who completed their primary school education (1996/1997 cohort) in Fayoum in North Upper Egypt. However, the results show that the pilot program had no any significant effects of the pilot program on the students completing their primary school education in Luxor in South Upper Egypt (Column 4). Additionally, for the 1994/1995 cohorts, no significant spillover effects were found in Fayoum in North Upper Egypt, but the pilots were associated with a 10.6 percentage point reduction in average grade retention for students in Luxor in South Upper Egypt.[14]

The same pattern repeats for each of the male and female samples in Panels B and C of Table 2.5. The pilots led to a relatively comparable decrease of 9.5 and 10.5 percentage points (significant at the 10% level but not 5% significance level) for males and females respectively in the 1996/1997 cohorts in Fayoum in North Upper Egypt (column 3). No significant effects were found for each of the male and female sample for students in Luxor in South Upper Egypt (column 4). For the 1994/1995 cohorts, the pilots had no effect on the placebo cohorts in both the male and female sample in Fayoum in North Upper Egypt (column 1). However, I found a significant 11.6% and 8.3% decrease in grade retention for the male and female sample in Luxor in South Upper Egypt (column 2).

Panels D and E of Table 2.5 show the distributional effects in both regions. For the 1996/1997 cohort, the pilot program was associated with a 13.9 percentage points decrease in grade retention for students from below average income but no significant effect on students from above average income households in Fayoum in North Upper Egypt (column 3). However, no significant associated effects were found on students from both income groups in Luxor in South Upper Egypt (column 4). On the other hand, the pilot program wasn't associated with any effects on grade retention on students

[14] I highlight here that the result for the full sample in Luxor in South Upper Egypt (Panel A of Table) must not be interpreted causally as the sample fails the common trends assumption test (see below).

CEU eTD Collection

from both income groups in Fayoum in North Upper Egypt (column 1) but was associated with decreasing significant effects in Luxor in South Upper Egypt (column 2).

The analysis of the distributive effects in each regional unit gives us a hint that the pilot program may have had heterogenous effects in each region. More importantly, since most of the sub-samples used to estimate the distributional effects failed the common trends assumption test for income groups in Luxor in South Upper Egypt (see below), the only conclusion that is reasonable to make is that the income structures in each region in Egypt is too distorted to be able to uncover real effects using this data. Also, the number of observations used, especially for the income group by gender break-downs are too small that it is difficult to reach an unbiased estimate.

2.3.5 Testing the Conditional Common Trends Assumption

As mentioned above, I test the common trends assumption using the approach found in (Pischke 2005). I run the regression formulation presented in equation (2) above on each sample and sub-samples to obtain the lead coefficients, $\beta_a \; and \; \beta_b$. I then perform a joint coefficient test in order to verify if the pre-treatment trends in the control group and the treatment group are not statistically significant from zero. Referring to equation (2) above, I formally test for $H_o: \beta_a = \beta_b = 0$. If I fail to reject the null hypothesis using the F-statistics generated from the joint coefficient test, I can then conclude that the trends between the control group and the treatment group are indeed common.

Tables 2.6 through 2.7 have the same format and only differ by the sample used. The aim of these tables is to present the lead coefficients of the pre-treatment periods and the F-statistic and corresponding p-value from the Wald test implemented to verify the common trends for each sub-sample. The estimates from the main effects should be read jointly using these tables. If the p-value of the Wald test for a particular sub-sample indicates that we don't fail to reject the null hypotheses with a 5% or a 1% significance level, then we can conclude that no common trends exist in this sub-

CEU eTD Collection

sample. Consequently, we should not interpret the effects from its corresponding previous panels as a causal effect but only consider the estimate as suggestive.

The tables are arranged as follows, column (I) refers to the 1994/1995 cohort while column (II) refers to the 1996/1997 cohort. Each column presents the lead coefficients from two models. Model 1 only includes a group specific effect and time period specific effects. Model 2, on the other hand, includes the full set of individual-level and state-level covariates. I refer mainly to the results from Model 2 to verify that the conditional common trends assumption holds in most sub-samples. I fail to reject the null hypotheses at the 5% and 1% significance level for all sub-samples in the main effects analysis (Table 2.6) and verify the validity of the conditional common trend assumption. Panel D in Table 2.7 shows that I was unable to verify the common trends for the 1996/1997 sub-sample of students from below average income households in Fayoum in North Upper Egypt. Table 2.8 shows that I'm unable to verify the common trends assumption in both Panels D and E for both of the cohort groups in Luxor in South Upper Egypt. This perhaps suggests that the income structures in south upper Egypt is too distorted to uncover any trends.

A second approach commonly used by researchers to validate the common trends is through presenting a visual representation of the trends in the outcome variable between both groups (Wing, Simon, and Bello-Gomez 2018). However, due to the low sample size in some of the sub-samples, the volatility of the estimated means is relatively high which makes it difficult to visually spot the common trend. Additionally, the plotted data is raw and only a few time periods are available. Therefore, the visual representation of the common trends is not very helpful. Therefore, besides relying mainly on the results from the statistical tests to confirm the presence of common trends, I also present the plots of grade retention for each sub-sample in figures A.1.1 to A.1.4 in Appendix 1.

CEU eTD Collection

2.3.6 Additional Robustness Checks

As mentioned earlier, my results indicate that my data suffers from selection bias. In order to analyze if the results presented above are robust to selection, I match observations based on the 5 nearest neighbors within the propensity score caliper. I use the logisitic regression to calculate the propensity score for observations in each cohort group using individual-level covariates. As expected, the method manages to effectively reduce the bias in individual level covariates. As a result, I end up with 6,976 observations from the original 26,881 observations in the 1994/1995 cohort group and 7,756 observations from the original 28,331 observations in the 1996/1997 cohort group. This indicates that my original data indeed suffered from selection. However, the estimated interaction coefficients obtained from the regressions using the matched sample didn't differ greatly from the ones presented above. More importantly, the signs of the significant coefficients were the same. On average, given the magnitude of the estimates obtained from the robustness check were higher, we can regard the estimates presented in the paper as lower bound estimates of the true effect of the pilot program.

In Appendix 2, Figure A.2.1 visualizes the balance of covariates after matching and the reduction in bias reached after matching on the 1994/1995 and 1996/1997 cohort respectively. Figure A.2.2 presents histograms for the propensity scores by treatment status for each cohort group. The two figures together indicate two things. First, the propensity score histograms show that ther is a wide area for common support. Second, the balancing exercise shows that the matching exercise was successful and given the observable characteristics at the individual level, selection into treatment is relatively random. Tables A.2.1 and A.2.2 report the results obtained from the regression formulation in equation (2) and the F-statistic and p-value from a joint coefficient test to verify the common trends assumption on the matched samples. Table A.2.3 and A.2.4 present estimates of the main effects of the pilot program obtained from running the main regression formulation in equation (1) on both

CEU eTD Collection

matched samples and sub-samples. I also present figure A.2.3 to show the trends in grade retention using the matched sample.

Additionally, I use a fake policy year to verify if selection into treatment is responsible for the observed effects. Using the fake pilot year, I find no significant effects on all of the results for both cohorts (not reported here). Finally, I find that my results are robust over varied calculations of the reference value used to determine retention. I obtain the same results when I use the average years of school for every state, time period and cohort instead of the floor value of the average completed years of school. These findings lend further credibility that the results presented above are true effects of the pilot program.

CEU eTD Collection

Section 3 Discussion & Policy Implications

3.1 Discussion

Overall, my results show a positive impact of the financial decentralization intervention on the experience of retention for primary school students. The results reported in this paper are in line with the literature on school-based management interventions (P. Gertler, Patrinos, and Rodríguez-Oreggia 2012; Jimenez and Sawada 2013; Pradhan et al. 2011; Skoufias and Shapiro 2006). A survey study by (Carr-Hill et al. 2015) finds that the reviewed impact assessments had an overall effect of -0.09 SMD (95%CI = -0.13, -0.04) for grade retention. In my analysis, I find that the intervention led to a 5.2 percentage point decrease on average in the probability of experiencing grade retention for students in schools participating in the pilot program relative to students in the control states.

The household surveys used also included data on students in other stages of education. This allowed me to test for spillover effects using the 1994/1995 cohort which completed the lower secondary education stage at the time. During the pilot phase, only primary schools were included in the intervention.

One limitation of the surveys is that it doesn't contain information on the type of school attended by each student. If I assume that the school structure in Egypt is separated and that each school only provides services for one education stage, then it should have been expected to find no significant effects for the 1994/1995 cohorts. However, my analysis found an impact of the intervention on the 1994/1995 cohorts that was significant in the three tested models demonstrating a positive spill-over effect on untreated educational stages. Two possible explanations are: (1) It is plausible that the grade retention outcomes from the lower secondary schools exhibit some form of reactivity, where their behavior is being altered due to being observed and not directly due to the intervention (Jiménez-Buedo 2015); (2) A significant impact would be expected if the public schools in the treated states were actually made up of the two basic education stages (primary and lower secondary). In that case

CEU eTD Collection

an impact on the 1994/1995 cohort should be expected. However, I do not analyze this observation further as it is out of scope for my paper.

I also observe that the intervention impacted females and students from households with above average income more than it impacted males and students from households with below average income respectively. This observation also resonates with the findings of (Galiani, Gertler, and Schargrodsky 2008) where they also find a positive impact of an education decentralization reform in Argentina and concluded that the benefits were accrued to students from advantaged communities more than students from disadvantaged communities. As noted earlier, males in poorer communities in Egypt experience the highest repetition rates, and are therefore disadvantaged (Roushdy et al. 2010). Hence, the findings of this paper supports the findings of (Galiani, Gertler, and Schargrodsky 2008).

A study by (Von Braun and Grote 2002) performs a cross-country analysis and provides a counterclaim. They find that decentralization interventions in fact helps the poor. They suggest that decentralization interventions be assessed as a whole paradigm including the political and administrative aspects surrounding the intervention. They show that an analysis of the sequencing and progress of each type of decentralization intervention is better than studying decentralization interventions as an "on/off switch". One possible area for future research is to find if this observation is common to decentralization interventions in different contexts and to understand the reasons for this disparity.

Finally, I also found that the pilot program had a heterogenous impact in both Fayoum and Luxor.[15]A positive effect was found in Fayoum whereas the pilot program had no effect in Luxor. A recent report by the Central Agency for Public Mobilization and Statistics (CAPMAS) found that Luxor, with a poverty rate of 55.3%, was ranked as the state with the 3rd highest poverty rate in Egypt.

[15] I was unable to perform a similar analysis on the third treated state because (1) The Suez Canal region included 3 frontier states which have completely different trends than Ismailia (treated state). (2) After excluding the frontier states from the analysis, the number of observations was too low to find any unbiased effects.

CEU eTD Collection

Fayoum, with a poverty rate of 26.4%, had a poverty rate lower than the national average.(32.5% of Egypt population live in poverty - Daily News Egypt 2019) Also, as discussed by (Carr-Hill et al. 2015), education decentralization interventions perform worse in disadvantaged communities especially when there are huge disparities between the teachers and the broader community they function in. However, I caution interpreting the regional results in a causal fashion as many sub-samples (especially the income sub-samples for the South Upper Egypt region) failed the common trends test as discussed above. Future research can further explore the administrative and political differences that existed in each state in order to arrive at a clearer understanding of the factors that can be attributed to positive outcomes of financial decentralization interventions in Egypt.

3.2 Policy Implications

In conclusion, I find that implementing a financial decentralization intervention could lead to a positive impact on student's grade retention as one aspect of education quality. However, in highly disadvantaged settings a financial decentralization reform is less likely to be successful. More specifically, it seems that male students who are already experiencing higher retention rates were impacted relatively less than female students. Additionally, the intervention had a greater impact on students coming from wealthier households relative to students from disadvantaged households.

SBM and decentralization interventions imply a greater role of the local community in the education process. If the local community is already suffering from disadvantages and barriers to participation, then it will be more difficult for those communities to improve their local school's education process. This suggests that policymakers are more likely to find positive results of decentralization reforms in more advantaged communities. However, this raises serious equity concerns in Egypt where average poverty rates have reached about 32.5% in 2017/2018.

Also, from a review of the global literature and the results presented in this paper it becomes obvious that a one-size-fits-all approach is more than likely to have negative effects. The exact design,

CEU eTD Collection

procedure, timing and institutional framework that the decentralization reform is embedded in, is crucial for the success of this intervention. In this case, the cooperation between state ministries to incorporate a needs-based funding formula and training of school staff appears to have had a positive impact on an aspect of education quality, grade retention.

As more reliable data becomes available it will be possible to study the subsequent years of this intervention in order to study the main objective of the program, which is achieving a more equitable distribution of school funds between states. Policymakers should also pay attention to the possibility of conducting empirical impact evaluations on future interventions, where the same students baseline characteristics should be observed before and after the intervention to better assess the effect of the intervention on different outcomes.

Given the widespread adoption of decentralization policies around the Arab world and the lack of impact studies, the topic of SBM and decentralization interventions appears to be a relevant area for future research. The uniqueness of each education decentralization initiative should be studied individually to assess the potential enablers and barriers of effects. The negative aspects of these programs should be studied in order to better understand the interactions these policies generate in different contexts.

CEU eTD Collection

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. *When Should You Adjust Standard Errors for Clustering?* National Bureau of Economic Research.

Autor, David H. 2003. "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing." *Journal of labor economics* 21(1): 1–42.

Besley, Timothy, and Stephen Coate. 2003. "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach." *Journal of public economics* 87(12): 2611–2637.

Blundell, Richard, Mike Brewer, and Andrew Shephard. 2005. "Evaluating the Labour Market Impact of Working Families' Tax Credit Using Difference-in-Differences."

Boex, Jamie. 2011. "Democratization in Egypt: The Potential Role of Decentralization." *Center on International Development and Governance, Urban Institute, Washington, DC, February*: 3.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2018. "Difference-in-Differences with Multiple Time Periods and an Application on the Minimum Wage and Employment." *arXiv:1803.09015 [econ, math, stat]*. http://arxiv.org/abs/1803.09015 (September 22, 2019).

Carr-Hill, Roy et al. 2015. "The Effects of School-Based Decision Making on Educational Outcomes in Low-and Middle-Income Contexts: A Systematic Review." *3ie Grantee Final Review. London: International Initiative for Impact Evaluation (3ie)*.

Cockx, Bart, Matteo Picchio, and Stijn Baert. 2019. "Modeling the Effects of Grade Retention in High School." *Journal of Applied Econometrics* 34(3): 403–24.

Daily News Egypt 2019. "32.5% of Egypt Population Live in Poverty - Daily News Egypt." 2019. *Daily News Egypt*. https://wwww.dailynewssegypt.com/2019/07/29/32-5-of-egypt-population-live-in-poverty/ (November 22, 2019).

Dixon, John, Shahjahan Bhuiyan, and Yılmaz Üstüner. 2018. *Public Administration in the Middle East and North Africa*. Taylor & Francis.

Eide, Eric R., and Mark H. Showalter. 2001. "The Effect of Grade Retention on Educational and Labor Market Outcomes." *Economics of Education review* 20(6): 563–576.

El Baradei, Laila. 2015. "The Case for Decentralization as a Tool for Improving Quality in Egyptian Basic Education." *The Egyptian Center for Economic Studies, ECES Working Paper* 180.

CEU eTD Collection

ElRashidy, Ola. 2018. “School Management and Educational Quality in the Context of Centralized Educational Governance in Egypt: Examining Stakeholder Perceptions in Private Language Schools.”

Faguet, Jean-Paul, and Fabio Sánchez. 2008. “Decentralization’s Effects on Educational Outcomes in Bolivia and Colombia.” *World Development* 36(7): 1294–1316.

Galiani, Sebastian, Paul Gertler, and Ernesto Schargrodsky. 2008. “School Decentralization: Helping the Good Get Better, but Leaving the Poor Behind.” *Journal of Public Economics* 92(10): 2106–20.

Galiani, Sebastian, and Ernesto Schargrodsky. 2002. “Evaluating the Impact of School Decentralization on Educational Quality.” *Economia* 2(2): 275–314.

Gertler, P. J., P. Premand, and L. B. Rawlings. 2011. “WHO-Impact Evaluation in Practice, Chapter 6.” *Geneva: World Health Organization*.

Gertler, Paul, Harry Anthony Patrinos, and Eduardo Rodríguez-Oreggia. 2012. “Parental Empowerment in Mexico: Randomized Experiment of the" Apoyos a La Gestion Escolar (Age)" Program in Rural Primary Schools in Mexico.” *Society for Research on Educational Effectiveness*.

Ginsburg, Mark, Nagwa Megahed, Mohammed Elmeski, and Nobuyuki Tanaka. 2010. “Reforming Educational Governance and Management in Egypt: National and International Actors and Dynamics.” *education policy analysis archives* 18: 5.

Hammad, Waheed. 2013. “The Rhetoric and Reality of Decentralisation Reforms: The Case of School-Based Management in Egypt.” *International Studies in Educational Administration (Commonwealth Council for Educational Administration & Management (CCEAM))* 41(2).

Hammad, Waheed, and Nigel Norris. 2009. “Centralised Control: A Barrier to Shared Decision-Making in Egyptian Secondary Schools.” *International Studies in Educational Administration (Commonwealth Council for Educational Administration & Management (CCEAM))* 37(2).

Hanchate, Amresh D. et al. 2015. “Massachusetts Health Reform and Disparities in Joint Replacement Use: Difference in Differences Study.” *bmj* 350: h440.

Hanson, E. Mark. 1990. “Administrative Reform and the Egyptian Ministry of Education.” *Journal of Educational Administration* 28(4).

Hanushek, Eric A., and Ludger Wößmann. 2007. *The Role of School Improvement in Economic Development*. National Bureau of Economic Research.

CEU eTD Collection

Healey, F. Henry, Luis Crouch, and Rafik Hanna. 2014. “Education Finance in Egypt: Problems and a Possible Solution. Occasional Paper. RTI Press Publication OP-0017-1401.” *RTI International*.

Healey, Henry, Rafik Hanna, and Hany Attalla. 2016. *Decentralization and Decentralized Education Finance in Egypt (2007–2013)*. RTI Press. http://www.rti.org/publication/decentralization-and-decentralized-education-finance-egypt-2007-2013 (November 21, 2019).

Heckman, James J., and Jeffrey A. Smith. 1995. “Ashenfelter’s Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluation Strategies.”

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference.” *Political analysis* 15(3): 199–236.

Ibrahim, Ali S. 2010. “The Politics of Educational Transfer and Policymaking in Egypt.” *Prospects* 40(4): 499–515.

Jimenez, Emmanuel, and Vicente Paqueo. 1996. “Do Local Contributions Affect the Efficiency of Public Primary Schools?” *Economics of Education Review* 15(4): 377–386.

Jimenez, Emmanuel, and Yasuyuki Sawada. 2013. “Does Community Management Help Keep Children in Schools? Evidence Using Panel Data from El Salvador’s EDUCO Program.” *Economic Development and Cultural Change* 62(2): 307–338.

Jiménez-Buedo, María. 2015. “The Last Dictator Game? Dominance, Reactivity, and the Methodological Artefact in Experimental Economics.” *International Studies in the Philosophy of Science* 29(3): 295–310.

Khaleel, N. 2003. “A Comparative Study of Educational Administration in France and the USA: What Lessons Can Egypt Learn from Them?” *Al-Tarbia*.

King, Elizabeth M., and Susana Cordeiro Guerra. 2005. “Education Reforms in East Asia: Policy, Process, and Impact.” *East Asia decentralizes: Making local government work* 292: 179.

Krafft, Caroline. 2012. “Challenges Facing the Egyptian Education System: Access, Quality, and Inequality.”

Lechner, Michael. 2010. “The Estimation of Causal Effects by Difference-in-Difference MethodsEstimation of Spatial Panels.” *Foundations and Trends® in Econometrics* 4(3): 165–224.

CEU eTD Collection

Malen, Betty, Rodney T. Ogawa, and Jennifer Kranz. 1990. "What Do We Know about School-Based Management? A Case Study of the Literature—A Call for Research." *Choice and control in American education* 2: 289–342.

Manacorda, Marco. 2012. "The Cost of Grade Retention." *Review of Economics and Statistics* 94(2): 596–606.

Meng, Xin, and Chikako Yamauchi. 2015. "Children of Migrants: The Impact of Parental Migration on Their Children's Education and Health Outcomes."

Ministry of Planning, Monitoring and Administrative Reform. 2016. *Sustainable Development Strategy (SDS): Vision 2030*. Cairo, Egypt: Ministry of Planning, Monitoring and Administrative Reform. http://sdsegypt2030.com/category/reports-en/?lang=en.

Nasser-Ghodsi, Nadia. 2006. "What Is the Effect of Educational Decentralization on Student Outcomes in Egypt?" PhD Thesis. Citeseer.

OAMDI, 2016. Harmonized Labor Force Surveys (HLFS), http://erf.org.eg/data-portal/. Version 1.0 of Licensed Data Files; LFS 2011- Central Agency for Public Mobilization and Statistics (CAPMAS). Egypt: Economic Research Forum (ERF).

OAMDI, 2018. Harmonized Labor Force Surveys (HLFS), http://erf.org.eg/data-portal/. Version 1.0 of Licensed Data Files; LFS 2010- Central Agency for Public Mobilization and Statistics (CAPMAS). Egypt: Economic Research Forum (ERF).

OAMDI, 2019. Harmonized Labor Force Surveys (HLFS), http://erf.org.eg/data-portal/. Version 1.0 of Licensed Data Files; LFS 2008- Central Agency for Public Mobilization and Statistics (CAPMAS). Egypt: Economic Research Forum (ERF).

OAMDI, 2019. Harmonized Labor Force Surveys (HLFS), http://erf.org.eg/data-portal/. Version 1.0 of Licensed Data Files; LFS 2007- Central Agency for Public Mobilization and Statistics (CAPMAS). Egypt: Economic Research Forum (ERF).

Oates, Wallace E. 2008. "On the Theory and Practice of Fiscal Decentralization." *Institutional foundations of public finance: Economic and legal perspectives* 165: 176.

Parke, Mary. 2003. "Are Married Parents Really Better for Children? What Research Says about the Effects of Family Structure on Child Well-Being."

Patrinos, Harry Anthony, Felipe Barrera-Osorio, and Tazeen Fasih. 2009. *Decentralized Decision-Making in Schools: The Theory and Evidence on School-Based Management*. The World Bank.

Pischke, Jorn-Steffen. 2005. "Empirical Methods in Applied Economics: Lecture Notes." *Downloaded July* 24: 2015.

CEU eTD Collection

Pradhan, Menno et al. 2011. *Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia*. The World Bank.

Psacharopoulos, George, and Eduardo Velez. 1993. "Educational Quality and Labor Market Outcomes: Evidence from Bogota, Colombia." *Sociology of Education*: 130–145.

Rao, Mala et al. 2014. "Changes in Addressing Inequalities in Access to Hospital Care in Andhra Pradesh and Maharashtra States of India: A Difference-in-Differences Study Using Repeated Cross-Sectional Surveys." *BMJ open* 4(6): e004471.

Roderick, Melissa. 1994. "Grade Retention and School Dropout: Investigating the Association." *American Educational Research Journal* 31(4): 729–759.

Rosenbaum, Paul R. 2010. 10 *Design of Observational Studies*. Springer.

Roushdy, R. et al. 2010. "Survey of Young People in Egypt."

Santibañez, Lucrecia. 2006. "School-Based Management Effects on Educational Outcomes: A Literature Review and Assessment of the Evidence Base."

Savedoff, William D. 1998. *Organization Matters: Agency Problems in Health and Education in Latin America*. Idb.

Shehata, A. 1997. "Administrative Problems Facing Secondary Education Administrators: How Can They Be Overcome? Unpublished MA Dissertation." Cairo University: Institute of Educational Studies.

Skoufias, Emmanuel, and Joseph Shapiro. 2006. *Evaluating the Impact of Mexico's Quality Schools Program: The Pitfalls of Using Nonexperimental Data*. The World Bank.

Smits, Jeroen, and Iñaki Permanyer. 2019. "The Subnational Human Development Database." *Scientific data* 6: 190038.

Smoke, Paul J. 2001. *Fiscal Decentralization in Developing Countries: A Review of Current Concepts and Practice*. Citeseer.

Somers, Marie-Andrée, Pei Zhu, Robin Jacob, and Howard Bloom. 2013. "The Validity and Precision of the Comparative Interrupted Time Series Design and the Difference-in-Difference Design in Educational Evaluation." *MDRC*.

Stuart, Elizabeth A. et al. 2014. "Using Propensity Scores in Difference-in-Differences Models to Estimate the Effects of a Policy Change." *Health Services and Outcomes Research Methodology* 14(4): 166–82.

CEU eTD Collection

Von Braun, Joachim, and Ulrike Grote. 2002. “Does Decentralization Serve the Poor?” In *Managing Fiscal Decentralization*, Routledge, 84–112.

WENR 2019.“Education in Egypt.” 2019. *WENR*. https://wenr.wes.org/2019/02/education-in-egypt-2 (November 20, 2019).

Westhorp, Gill et al. 2014. “Enhancing Community Accountability, Empowerment and Education Outcomes in Low and Middle-Income Countries: A Realist Review.” *EPPI-Centre, Social Science Research Unit, Institute of Education, University of London*.

Wing, Coady, Kosali Simon, and Ricardo A. Bello-Gomez. 2018. “Designing Difference in Difference Studies: Best Practices for Public Health Policy Research.” *Annual review of public health* 39.

Wooldridge, Jeffrey. 2007. “Difference-in-Differences Estimation.” *Lecture Notes* 10.

World Bank. 2017. *World Development Report 2018: Learning to Realize Education’s Promise*. The World Bank. http://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-1096-1 (November 19, 2019).

World Bank. 2018. *Egypt - Supporting Egypt Education Reform Project*. The World Bank. http://documents.worldbank.org/curated/en/346091522415590465/Egypt-Supporting-Egypt-Education-Reform-Project (November 22, 2019).

CEU eTD Collection

Tables

Table 1.1: Grade Repetition

School Level	Ever repeated	Primary-school repetition	Preparatory-school repetition	Secondary-school repetition	Higher Education repetition
Currently attending					
Primary school	6.4	6.3			
Preparatory school	10.3	3.8	7.8		
Vocational secondary school	25.4	5.2	15.6	7.9	
General Secondary school	7.8	0.7	4.4	4.4	
Post-sec institute	22.3	0.0	2.6	13.1	6.5
University and above	15.8	0.6	2.7	5.1	10.7
Previously attended					
Primary school	27.9	27.9			
Preparatory school	39.4	9.1	34.6		
Vocational secondary school	19.5	3.8	12.4	6.5	
General Secondary school	22.3	1.7	10.0	12.2	
Post-sec institute	13.8	0.0	4.3	5.3	6.9
University and above	12.8	0.6	2.3	3.4	8.3
Total	17.1	5.9	11.4	6.0	9.1
Number of respondents	2,249	754	1,137	390	213

The figure shows the percentage of young people who ever attended school (previously or currently) that ever repeated, and repetition by level according to current school level, ages 10-29, Egypt 2009
(Source: Roushdy, R., Krafft, C., Harbour, C., Barsoum, G., & El-Kogali, S. (2010). Survey of young people in Egypt.)

CEU eTD Collection

Table 2.1: Structure used to compile the paper's dataset using LFS yearly data

Time Period	LFS Year	Survey round	Survey Months	Academic year
Pre-Treatment Period 1	2007	1	Jan. – Mar.	2005/2006
		2	April – Jun.	
Pre-Treatment Period 2	2007	3	Jul. – Sept.	2006/2007
		4	Oct. – Dec.	
	2008	1	Jan. – Mar.	
		2	April – Jun.	
Post-Treatment	2010	3	Jul. – Sept.	2009/2010
		4	Oct. – Dec.	
	2011	1	Jan. – Mar.	
		2	April – Jun.	

CEU eTD Collection

Table 2.2: Proportions and standard errors of key variables for the 1996/1997 cohorts

	Before Decentralization		After Decentralization	
	Control	Treatment	Control	Treatment
Number of Observations	16,999	982	9,720	630
A. Individual Characteristics				
Male	0.53	0.52	0.52	0.52
	(0.00)	(0.02)	(0.01)	(0.02)
Female	0.47	0.48	0.48	0.48
	(0.00)	(0.02)	(0.01)	(0.02)
Age	10.84	10.86	13.98	13.98
	(0.01)	(0.06)	(0.01)	(0.04)
Years of Education	4.72	4.84	6.91	6.88
	(0.02)	(0.04)	(0.02)	(0.04)
Primary Enrollment	0.96	0.96	0.07	0.07
	(0.00)	(0.01)	(0.00)	(0.01)
Lower Secondary	0.04	0.04	0.87	0.91
Enrollment	(0.00)	(0.01)	(0.00)	(0.01)
B. Parent's Characteristics				
Age of the head	44.0	43.7	46.8	48.0
	(0.11)	(0.30)	(0.12)	(0.34)
Age of the spouse	37.4	37.0	40.2	40.0
	(0.08)	(0.24)	(0.09)	(0.27)
Male head	0.53	0.52	0.52	0.52
	(0.00)	(0.02)	(0.01)	(0.02)
Female Head	0.47	0.48	0.48	0.48
	(0.00)	(0.02)	(0.01)	(0.02)
Head: No Education	0.47	0.67	0.45	0.55
	(0.02)	(0.05)	(0.01)	(0.04)
Head: Basic Education	0.10	0.05	0.11	0.09
	(0.00)	(0.02)	(0.00)	(0.01)
Head: Secondary	0.26	0.20	0.27	0.20
Education	(0.01)	(0.02)	(0.01)	(0.02)
Head: Post-Secondary	0.17	0.08	0.17	0.16
Education (university)	(0.01)	(0.02)	(0.01)	(0.03)
Head: Employed	0.93	0.96	0.90	0.92
	(0.00)	(0.01)	(0.00)	(0.01)
Head: Unemployed	0.002	0.001	0.011	0.014
	(0.00)	(0.00)	(0.00)	(0.01)
Head: Homemaker	0.03	0.02	0.03	0.02
	(0.00)	(0.01)	(0.00)	(0.01)
Head:	0.04	0.02	0.06	0.04
Pensioner/retired/disabled	(0.00)	(0.01)	(0.00)	(0.01)

CEU eTD Collection

Table 2.2: Proportions and standard errors of key variables for the 1996/1997 cohorts (contd.)

	Before Decentralization		After Decentralization	
	Control	Treatment	Control	Treatment
C. Household Characteristics				
All Household Members	5.79	6.24	5.80	6.54
	(0.06)	(0.11)	(0.04)	(0.12)
Male Household Members	2.95	3.20	2.98	3.52
	(0.04)	(0.09)	(0.03)	(0.08)
15+ members in labor force	1.50	1.67	1.57	1.95
	(0.02)	(0.07)	(0.02)	(0.08)
15+ male members in labor force	1.15	1.29	1.19	1.52
	(0.01)	(0.03)	(0.01)	(0.05)
Log (Household Income)	6.46	6.36	7.07	7.16
	(0.01)	(0.02)	(0.01)	(0.03)
Dwelling: House/Villa	0.35	0.64	0.28	0.62
	(0.03)	(0.07)	(0.02)	(0.07)
Dwelling: Apartment/Other	0.65	0.36	0.72	0.38
	(0.03)	(0.07)	(0.02)	(0.07)
Dwelling: Rented	0.16	0.06	0.12	0.07
	(0.02)	(0.03)	(0.01)	(0.02)
Dwelling: Owned	0.84	0.94	0.88	0.93
	(0.02)	(0.03)	(0.01)	(0.02)
Rural	0.64	0.78	0.64	0.72
	(0.05)	(0.09)	(0.03)	(0.09)
Urban	0.36	0.22	0.36	0.28
	(0.05)	(0.09)	(0.03)	(0.09)
D. Regional Characteristics				
Sub-national HDI	0.65	0.60	0.66	0.63
	(0.00)	(0.01)	(0.00)	(0.01)
Health index	0.77	0.77	0.77	0.77
	(0.00)	(0.00)	(0.00)	(0.00)
Income index	0.68	0.65	0.69	0.67
	(0.00)	(0.00)	(0.00)	(0.00)
Expected Years of Schooling: Children aged 6	11.9	10.7	12.0	11.2
	(0.07)	(0.17)	(0.04)	(0.15)
Mean Years of Schooling (25+)	5.86	4.45	6.39	5.54
	(0.10)	(0.20)	(0.06)	(0.23)
Log (Population Density)	7.92	7.43	7.91	7.46
Log (person over inhabited square km)	(0.08)	(0.11)	(0.05)	(0.18)

CEU eTD Collection

Table 2.3: The effect of the pilot program on grade retention by gender

	I. 1994/1995 Cohort			II. 1996/1997 Cohort		
	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 1	(5) Model 2	(6) Model 3
Panel A. Full Sample						
$(D_{is}.after_{it})$ (β_3)	-0.042**	-0.042***	-0.054***	-0.031	-0.037**	-0.052***
	(0.017)	(0.016)	(0.018)	(0.020)	(0.018)	(0.020)
# of Observations	26,881	26,881	25,823	28,331	28,331	27,150
Root MSE	0.299	0.297	0.295	0.268	0.266	0.265
Panel B. Male Sample						
$(D_{is}.after_{it})(\beta_3)$	-0.050**	-0.050**	-0.058**	-0.022	-0.029	-0.038**
	(0.022)	(0.021)	(0.024)	(0.021)	(0.018)	(0.020)
# of Observations	14,018	14,018	13,436	14,843	14,843	14,244
Root MSE	0.302	0.302	0.299	0.269	0.267	0.266
Panel C. Female Sample						
$(D_{is}.after_{it})(\beta_3)$	-0.033	-0.034*	-0.049**	-0.040	-0.047*	-0.067**
	(0.021)	(0.020)	(0.023)	(0.028)	(0.026)	(0.029)
# of Observations	12,863	12,863	12,387	13,488	13,488	12,906
Root MSE	0.294	0.293	0.290	0.267	0.265	0.264

Reported are the difference-in-differences interaction coefficients from the regression equation (1). Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars show the significance level of the coefficient where * 10%, ** 5%, ***1%. I also report the total number of observations used in the estimation and the Root MSE. Observations are weighted using survey individual weights.

Note: Each panel estimates the regression formulation on a different sample as indicated in the sub-heading. Model 1 includes a group effect and a post treatment effect. Model 2 additionally includes state specific fixed effects and time period specific fixed effect. Finally, Model 3 additionally includes various individual-level and state-level covariates.

CEU eTD Collection

Table 2.4: The effect of the pilot program on grade retention by household income and gender differences

	I. 1994/1995 Cohort			II. 1996/1997 Cohort		
	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 1	(5) Model 2	(6) Model 3
Panel D. Below Avg. Income						
$(D_{is}.after_{it})(\beta_3)$	-0.043**	-0.047***	-0.063***	-0.028	-0.035*	-0.051***
	(0.019)	(0.017)	(0.020)	(0.022)	(0.019)	(0.022)
# of Observations	21,934	21,934	20,893	23,002	23,002	21,831
Root MSE	0.304	0.302	0.300	0.271	0.269	0.269
Panel E. Above Avg. Income						
$(D_{is}.after_{it})(\beta_3)$	-0.033	-0.029	-0.018	-0.061**	-0.064***	-0.084***
	(0.028)	(0.031)	(0.034)	(0.024)	(0.022)	(0.024)
# of Observations	4,947	4,947	4,930	5,329	5,329	5,319
Root MSE	0.275	0.274	0.272	0.254	0.252	0.251
Panel F. Below Avg. Income - Male						
$(D_{is}.after_{it})(\beta_3)$	-0.057**	-0.059***	-0.072***	-0.020	-0.028	-0.037*
	(0.022)	(0.020)	(0.024)	(0.022)	(0.020)	(0.022)
# of Observations	11,445	11,445	10,874	12,080	12,080	11,488
Root MSE	0.309	0.307	0.305	0.273	0.271	0.270
Panel G. Above Avg. Income - Male						
$(D_{is}.after_{it})(\beta_3)$	-0.017	-0.009	-0.005	-0.048**	-0.054**	-0.068**
	(0.051)	(0.053)	(0.058)	(0.023)	(0.023)	(0.031)
# of Observations	2,573	2,573	2,562	2,763	2,763	2,757
Root MSE	0.273	0.272	0.269	0.252	0.249	0.247
Panel H. Below Avg. Income - Female						
$(D_{is}.after_{it})(\beta_3)$	-0.030	-0.035	-0.056**	-0.036	-0.044	-0.064**
	(0.024)	(0.023)	(0.026)	(0.032)	(0.029)	(0.033)
# of Observations	10,489	10,489	10,019	10,922	10,922	10,343
Root MSE	0.298	0.297	0.293	0.269	0.267	0.266
Panel I. Above Avg. Income - Female						
$(D_{is}.after_{it})(\beta_3)$	-0.047**	-0.044**	-0.026	-0.073**	-0.082**	-0.107***
	(0.022)	(0.022)	(0.028)	(0.033)	(0.032)	(0.034)
# of Observations	2,374	2,374	2,368	2,566	2,566	2,563
Root MSE	0.276	0.276	0.275	0.256	0.255	0.254

Reported are the difference-in-differences interaction coefficients from the regression equation (1). Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. I also report the total number of observations used in the estimation and the Root MSE. Observations are weighted using survey individual weights.

Note: Each panel estimates the regression formulation on a different sample as indicated in the sub-heading. Model 1 includes a group effect and a post treatment effect. Model 2 additionally includes state specific fixed effects and time period specific fixed effect. Finally, Model 3 additionally includes various individual and state covariates.

CEU eTD Collection

Table 2.5: The Effect of the pilot program on grade retention by region and gender

	I. 1994/1995 cohort		II. 1996/1997 cohort	
	(1) North Upper Egypt	(2) South Upper Egypt	(3) North Upper Egypt	(4) South Upper Egypt
Panel A. Full Sample				
$(D_{is}.after_{it})(\beta_3)$	0.005	-0.106***	-0.099**	0.022
	(0.052)	(0.034)	(0.038)	(0.046)
# of Observations	4,117	3,702	4,468	3,717
Root MSE	0.273	0.254	0.249	0.207
Panel B. Male Sample				
$(D_{is}.after_{it})(\beta_3)$	0.042	-0.116**	-0.095*	0.061
	(0.069)	(0.048)	(0.052)	(0.052)
# of Observations	2,260	1,991	2,371	2,041
Root MSE	0.277	0.258	0.246	0.201
Panel C. Female Sample				
$(D_{is}.after_{it})(\beta_3)$	0.082	-0.083**	-0.105*	-0.001
	(0.065)	(0.042)	(0.057)	(0.084)
# of Observations	1,857	1,711	2,097	1,676
Root MSE	0.269	0.249	0.252	0.213
Panel D. Below Avg. Income Sample				
$(D_{is}.after_{it})(\beta_3)$	0.036	-0.093**	-0.139***	0.044
	(0.058)	(0.042)	(0.043)	(0.052)
# of Observations	3,500	3,029	3,769	3,036
Root MSE	0.276	0.255	0.247	0.210
Panel E. Above Avg. Income Sample				
$(D_{is}.after_{it})(\beta_3)$	0.144	-0.143**	0.007	-0.067
	(0.105)	(0.080)	(0.069)	(0.041)
# of Observations	617	673	699	3,036
Root MSE	0.256	0.247	0.257	0.190

Reported are the difference-in-differences interaction coefficients from the equation 1 using group and post-treatment fixed effects, state and time period specific fixed effects and a full set of individual and household covariates. Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. I also report the total number of observations used in the estimation and the Root MSE. Observations are weighted using survey individual weights.

Note: Each panel estimates the regression formulation on a different sample as indicated in the sub-heading.

Note: The North Upper Egypt regional unit includes Menya, Beni Suef, and Fayoum where Fayoum was treated and the remaining states used as the control group for the period under study. The South Upper Egypt regional unit includes Sohag, Qena, Luxor, Aswan and the Red Sea where Luxor was treated by the pilot program and the remaining states used as the control group for the period under study.

CEU eTD Collection

Table 2.6: Lead coefficients and Common Trends Assumption tests

	I. 1994/1995 Cohort		II. 1996/1997 Cohort	
	(1) Model 1	(2) Model 2	(3) Model 1	(4) Model 2
Panel A. Full Sample				
Period (a) (β_a)	0.046**	0.041*	0.015	0.002
	(0.023)	(0.024)	(0.017)	(0.016)
Period (b) (β_b)	0.039**	0.041**	0.040	0.032
	(0.018)	(0.020)	(0.027)	(0.024)
# of Observations	26,881	25,823	28,331	27,150
F-statistic of Joint test	3.04**	2.58*	1.24	0.86
(p-values)	(0.049)	(0.077)	(0.291)	(0.423)
Panel B. Male Sample				
Period (a) (β_a)	0.061*	0.055*	0.022	0.002
	(0.032)	(0.033)	(0.022)	(0.019)
Period (b) (β_b)	0.044*	0.041	0.023	0.014
	(0.023)	(0.026)	(0.026)	(0.024)
# of Observations	14,018	13,436	14,843	14,244
F-statistic of Joint test	2.59*	1.94	0.69	0.18
(p-values)	(0.08)	(0.145)	(0.501)	(0.832)
Panel C. Female Sample				
Period (a) (β_a)	0.030	0.024	0.007	0.001
	(0.029)	(0.029)	(0.023)	(0.024)
Period (b) (β_b)	0.034	0.042*	0.058	0.051
	(0.023)	(0.024)	(0.036)	(0.034)
# of Observations	12,863	12,387	13,488	12,906
F-statistic of Joint test	1.24	1.51	1.32	1.24
(p-values)	(0.290)	(0.223)	(0.268)	(0.291)
Panel D. Below Avg. Income				
Period (a) (β_a)	0.038	0.035	0.011	-0.002
	(0.025)	(0.026)	(0.020)	(0.018)
Period (b) (β_b)	0.046**	0.050**	0.036	0.032
	(0.020)	(0.021)	(0.028)	(0.026)
# of Observations	21,934	20,893	23,002	21,831
F-statistic of Joint test	2.73*	2.77*	0.82	0.80
(p-values)	(0.066)	(0.063)	(0.440)	(0.450)
Panel E. Above Avg. Income				
Period (a) (β_a)	0.078*	0.043	0.049	0.034
	(0.042)	(0.044)	(0.030)	(0.038)
Period (b) (β_b)	0.007	-0.000	0.069**	0.057*
	(0.026)	(0.027)	(0.035)	(0.031)
# of Observations	4,947	4,930	5,329	5,319
F-statistic of Joint test	2.12	0.70	3.60**	2.21
(p-values)	(0.121)	(0.498)	(0.028)	(0.111)

Reported are the lead coefficients from the regression equation (2). Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. Also shown are the F-statistic generated from a Wald test and corresponding p-value. The Wald test verifies the common trend assumption if we fail to reject the null hypothesis: $\beta_a=\beta_b=0$.

Note: Model 1 controls for group fixed effects and post-treatment fixed effects whereas Model 2 additionally controls for individual, household, and state specific covariates. Each panel shows lead coefficients and runs the CTA test for a different sample as indicated in the sub-heading.

CEU eTD Collection

Table 2.6: Lead coefficients and Common Trends Assumption tests (contd.)

	I. 1994/1995 Cohort		II. 1996/1997 Cohort	
	(1) Model 1	(2) Model 2	(3) Model 1	(4) Model 2
Panel F. Below Avg. Income - Male				
Period (a) (β_a)	0.054	0.054	0.015	-0.005
	(0.035)	(0.036)	(0.023)	(0.021)
Period (b) (β_b)	0.058**	0.060**	0.023	0.016*
	(0.023)	(0.026)	(0.027)	(0.026)
# of Observations	11,445	10,874	12,080	11,488
F-statistic of Joint test	3.43**	2.94*	0.42	0.36
(p-values)	(0.033)	(0.054)	(0.656)	(0.695)
Panel G. Above Avg. Income - Male				
Period (a) (β_a)	0.086	0.065	0.066*	0.044
	(0.071)	(0.074)	(0.038)	(0.047)
Period (b) (β_b)	-0.023	-0.031	0.040	0.033
	(0.046)	(0.046)	(0.031)	(0.032)
# of Observations	2,573	2,562	2,763	2,756
F-statistic of Joint test	2.20	1.55	2.55*	0.88
(p-values)	(0.112)	(0.215)	(0.079)	(0.415)
Panel H. Below Avg. Income - Female				
Period (a) (β_a)	0.021	0.019	0.006	-0.002
	(0.030)	(0.030)	(0.027)	(0.026)
Period (b) (β_b)	0.034	0.044	0.051	0.047
	(0.028)	(0.029)	(0.040)	(0.038)
# of Observations	10,489	10,019	10,922	10,343
F-statistic of Joint test	0.80	1.17	0.87	0.96
(p-values)	(0.450)	(0.312)	(0.418)	(0.383)
Panel I. Above Avg. Income - Female				
Period (a) (β_a)	0.066	0.022	0.031	0.028
	(0.040)	(0.042)	(0.035)	(0.044)
Period (b) (β_b)	0.036	0.029	0.102**	0.085*
	(0.023)	(0.026)	(0.049)	(0.044)
# of Observations	2,374	2,368	2,566	2,563
F-statistic of Joint test	2.18	0.62	2.65*	2.00
(p-values)	(0.115)	(0.537)	(0.072)	(0.137)

Reported are the lead coefficients from the regression equation (2). Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. Also shown are the F-statistic generated from a Wald test and corresponding p-value. The Wald test verifies the common trend assumption if we fail to reject the null hypothesis: $\beta_a=\beta_b=0$.

Note: Model 1 controls for group fixed effects and post-treatment fixed effects whereas Model 2 additionally controls for individual, household, and state specific covariates. Each panel shows lead coefficients and runs the CTA test for a different sample as indicated in the sub-heading.

CEU eTD Collection

Table 2.7: Lead coefficients and Common Trends Assumption tests for North Upper Egypt

	I. 1994/1995 Cohort		II. 1996/1997 Cohort	
	(1) Model 1	(2) Model 2	(3) Model 1	(4) Model 2
Panel A. Full Sample				
Period (a) (β_a)	-0.029	-0.026	0.033	0.088
	(0.034)	(0.058)	(0.026)	(0.045)
Period (b) (β_b)	-0.031	-0.001	0.072**	0.088
	(0.026)	(0.051)	(0.031)	(0.042)
# of Observations	4,226	4,117	4,612	4,468
F-statistic of Joint test	1.02	0.16	2.98*	2.53*
(p-values)	(0.367)	(0.856)	(0.057)	(0.086)
Panel B. Male Sample				
Period (a) (β_a)	-0.009	-0.031	0.037	0.090
	(0.038)	(0.070)	(0.031)	(0.053)
Period (b) (β_b)	-0.023	-0.033	0.048	0.086
	(0.032)	(0.070)	(0.034)	(0.056)
# of Observations	2,318	2,260	2,452	2,371
F-statistic of Joint test	0.25	0.12	1.46	1.49
(p-values)	(0.783)	(0.886)	(0.238)	(0.232)
Panel C. Female Sample				
Period (a) (β_a)	-0.054	-0.110	0.030	0.086
	(0.043)	(0.071)	(0.032)	(0.066)
Period (b) (β_b)	-0.041	-0.057	0.099	0.094
	(0.032)	(0.064)	(0.043)	(0.059)
# of Observations	1,908	1,857	2,160	2,097
F-statistic of Joint test	1.38	1.23	2.64*	1.28
(p-values)	(0.257)	(0.300)	(0.079)	(0.284)
Panel D. Below Avg. Income Sample				
Period (a) (β_a)	-0.022	0.016	0.046*	0.120**
	(0.037)	(0.066)	(0.027)	(0.049)
Period (b) (β_b)	-0.025	0.044	0.069**	0.119**
	(0.026)	(0.055)	(0.030)	(0.046)
# of Observations	3,609	3,500	3,910	3,769
F-statistic of Joint test	0.61	0.47	3.34**	3.68**
(p-values)	(0.544)	(0.629)	(0.041)	(0.030)
Panel E. Above Avg. Income Sample				
Period (a) (β_a)	-0.082	0.169	-0.015	-0.013
	(0.064)	(0.109)	(0.058)	(0.079)
Period (b) (β_b)	-0.082	0.139	0.099*	-0.008
	(0.052)	(0.120)	(0.058)	(0.073)
# of Observations	617	617	702	699
F-statistic of Joint test	1.33	1.21	1.65	0.01
(p-values)	(0.273)	(0.304)	(0.201)	(0.987)

Reported are the lead coefficients from the regression equation (2). Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. Also shown are the F-statistic generated from a Wald test and corresponding p-value. The Wald test verifies the common trend assumption if we fail to reject the null hypothesis: $\beta_a=\beta_b=0$.

Note: In this table I use geographical proximity to find a control group for Fayoum. Fayoum is the treatment state whereas Menya and Beni-Sueif comprise the control group.

Note: Model 1 controls for group fixed effects and post-treatment fixed effects whereas Model 2 additionally controls for individual, household, and state specific covariates. Each panel shows lead coefficients and runs the CTA test for a different sample as indicated in the sub-heading.

CEU eTD Collection

Table 2.8: Lead coefficients and Common Trends Assumption tests for South Upper Egypt

	I. 1994/1995 Cohort		II. 1996/1997 Cohort	
	(1) Model 1	(2) Model 2	(3) Model 1	(4) Model 2
Panel A. Full Sample				
Period (a) (β_a)	0.086	0.081	-0.027	-0.036
	(0.032)	(0.033)	(0.034)	(0.037)
Period (b) (β_b)	0.081	0.082	0.024	0.033
	(0.036)	(0.040)	(0.036)	(0.044)
# of Observations	4,061	3,702	4,099	3,717
F-statistic of Joint test	4.91**	5.15***	1.23	0.98
(p-values)	(0.011)	(0.009)	(0.300)	(0.380)
Panel B. Male Sample				
Period (a) (β_a)	0.129**	0.098	-0.035	-0.062
	(0.064)	(0.062)	(0.044)	(0.048)
Period (b) (β_b)	0.099**	0.083*	-0.024	-0.041
	(0.039)	(0.042)	(0.044)	(0.058)
# of Observations	2,214	1,991	2,251	2,041
F-statistic of Joint test	3.98	2.92*	0.32	0.83
(p-values)	(0.024)	(0.062)	(0.728)	(0.440)
Panel C. Female Sample				
Period (a) (β_a)	0.034	0.028	-0.022	-0.026
	(0.041)	(0.043)	(0.058)	(0.062)
Period (b) (β_b)	0.064	0.071*	0.093	0.111*
	(0.039)	(0.041)	(0.068)	(0.064)
# of Observations	1,847	1,711	1,848	1,676
F-statistic of Joint test	1.35	1.52	3.20**	2.76*
(p-values)	(0.267)	(0.228)	(0.048)	(0.071)
Panel D. Below Average Income Sample				
Period (a) (β_a)	0.038	0.030	-0.075	-0.096**
	(0.034)	(0.029)	(0.025)	(0.036)
Period (b) (β_b)	0.104	0.096*	0.036	0.038
	(0.048)	(0.051)	(0.043)	(0.050)
# of Observations	3,388	3,029	3,418	3,036
F-statistic of Joint test	2.41*	1.98	7.33**	2.76***
(p-values)	(0.099)	(0.147)	(0.001)	(0.002)
Panel E. Above Average Income Sample				
Period (a) (β_a)	0.270***	0.270***	0.131***	0.137***
	(0.058)	(0.059)	(0.042)	(0.043)
Period (b) (β_b)	0.035	0.026	0.010	0.049*
	(0.029)	(0.041)	(0.039)	(0.072)
# of Observations	673	673	1,848	1,676
F-statistic of Joint test	10.49***	11.02***	5.01**	5.32****
(p-values)	(0.000)	(0.000)	(0.010)	(0.008)

Reported are the lead coefficients from the regression equation (2). Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. Also shown are the F-statistic generated from a Wald test and corresponding p-value. The Wald test verifies the common trend assumption if we fail to reject the null hypothesis: $\beta_a=\beta_b=0$.

Note: In this table I use geographical proximity to find a control group for Luxor. Luxor is the treatment state whereas Sohag, Qena, and Aswan comprise the control group.

Note: Model 1 controls for group fixed effects and post-treatment fixed effects whereas Model 2 additionally controls for individual, household, and state specific covariates. Each panel shows lead coefficients and runs the CTA test for a different sample as indicated in the sub-heading.

CEU eTD Collection

Appendix I Supplementary material

Table A.1.1: Proportions and standard errors of key variables of the 1994/1995 cohorts

	Before Decentralization		After Decentralization	
	Control	Treatment	Control	Treatment
Number of Observations	16,948	939	8,490	504
A. Individual Characteristics				
Male	0.52	0.54	0.52	0.50
	(0.00)	(0.03)	(0.01)	(0.03)
Female	0.48	0.46	0.48	0.50
	(0.00)	(0.03)	(0.01)	(0.03)
Age	12.79	12.82	15.98	15.99
	(0.01)	(0.05)	(0.01)	(0.03)
Years of Education	6.74	6.74	8.97	8.88
	(0.02)	(0.07)	(0.02)	(0.06)
B. Parent's Characteristics				
Age of the head	45.9	46.1	48.9	49.1
	(0.11)	(0.37)	(0.12)	(0.47)
Age of the spouse	39.3	39.3	42.2	41.6
	(0.09)	(0.41)	(0.10)	(0.37)
Male head	0.52	0.54	0.52	0.50
	(0.00)	(0.03)	(0.00)	(0.03)
Female Head	0.48	0.48	0.46	0.50
	(0.00)	(0.01)	(0.03)	(0.03)
Head: No Education	0.36	0.52	0.31	0.33
	(0.02)	(0.04)	(0.01)	(0.04)
Head: Basic Education	0.64	0.48	0.69	0.67
	(0.02)	(0.04)	(0.01)	(0.04)
Head: Secondary	0.50	0.68	0.45	0.51
Education	(0.02)	(0.04)	(0.01)	(0.04)
Head: Post-Secondary	0.09	0.05	0.10	0.09
Education (university)	(0.00)	(0.01)	(0.00)	(0.02)
Head: Employed	0.24	0.18	0.25	0.23
	(0.01)	(0.03)	(0.01)	(0.03)
Head: Unemployed	0.16	0.09	0.19	0.17
	(0.01)	(0.01)	(0.01)	(0.03)
Head: Homemaker	0.92	0.96	0.86	0.93
	(0.00)	(0.01)	(0.01)	(0.01)
Head:	0.002	0.001	0.010	0.011
Pensioner/retired/disabled	(0.00)	(0.00)	(0.00)	(0.01)

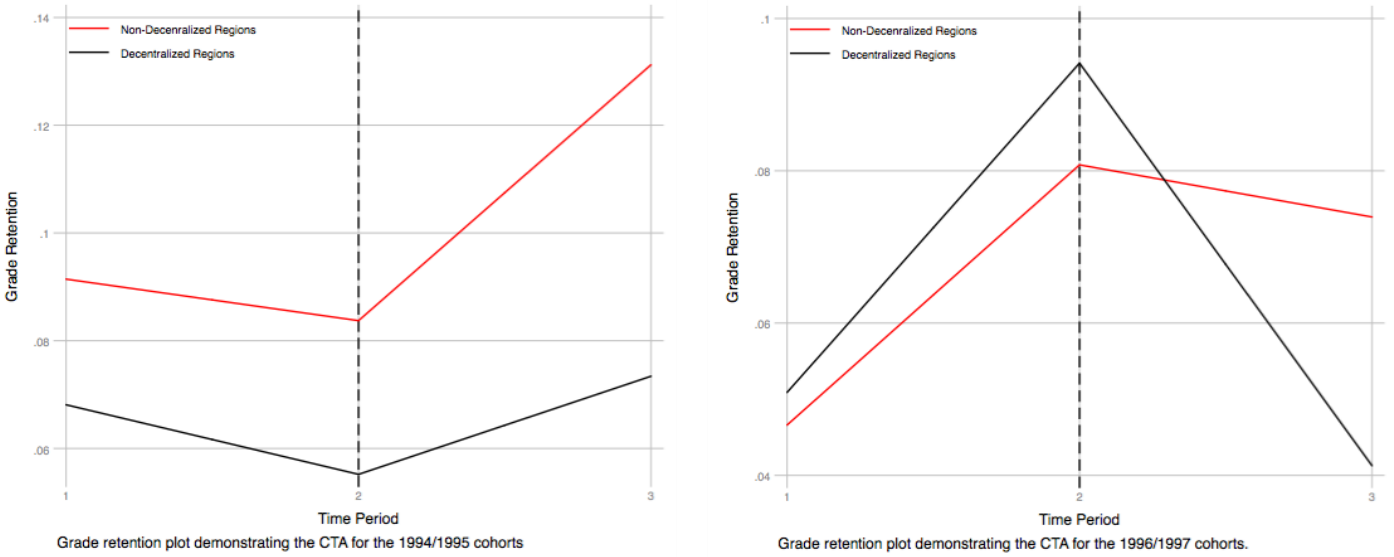
CEU eTD Collection

Table A.1.1: Proportions and standard errors of key variables of the 1994/1995 cohorts (contd.)

	Before Decentralization		After Decentralization	
	Control	Treatment	Control	Treatment
C. Household Characteristics				
All Household Members	5.89	6.49	5.75	6.40
	(0.06)	(0.13)	(0.04)	(0.12)
Male Household Members	3.03	3.39	2.99	3.35
	(0.04)	(0.11)	(0.03)	(0.10
15+ members in labor	1.58	1.88	1.68	1.95
force	(0.02)	(0.08)	(0.02)	(0.07)
15+ male members in	1.21	1.44	1.26	1.49
labor force	(0.01)	(0.04)	(0.01)	(0.06)
Log (Household Income)	6.48	6.41	7.12	7.01
	(0.01)	(0.02)	(0.01)	(0.03)
Dwelling: House/Villa	0.37	0.67	0.27	0.54
	(0.03)	(0.07)	(0.02)	(0.07)
Dwelling:	0.63	0.34	0.73	0.46
Apartment/Other	(0.03)	(0.07)	(0.02)	(0.07)
Dwelling: Rented	0.15	0.06	0.13	0.06
	(0.02)	(0.03)	(0.01)	(0.02)
Dwelling: Owned	0.85	0.94	0.87	0.94
	(0.02)	(0.03)	(0.01)	(0.02)
Rural	0.65	0.77	0.62	0.72
	(0.05)	(0.09)	(0.03)	(0.09)
Urban	0.35	0.23	0.38	0.28
	(0.05)	(0.10)	(0.03)	(0.10)
D. Regional Characteristics				
Sub-national HDI	0.65	0.60	0.66	0.64
	(0.00)	(0.01)	(0.00)	(0.01)
Health index	0.77	0.75	0.76	0.77
	(0.00)	(0.00)	(0.00)	(0.00)
Income index	0.68	0.65	0.69	0.67
	(0.00)	(0.00)	(0.00)	(0.00)
Expected Years of	11.9	10.7	12.0	11.3
Schooling: Children aged 6	(0.07)	(0.16)	(0.04)	(0.16)
Mean Years of Schooling	5.83	4.43	6.43	5.66
(25+)	(0.09)	(0.19)	(0.05)	(0.25)
Log (Population)	8.39	7.65	8.47	7.54
	(0.03)	(0.07)	(0.02)	(0.09)

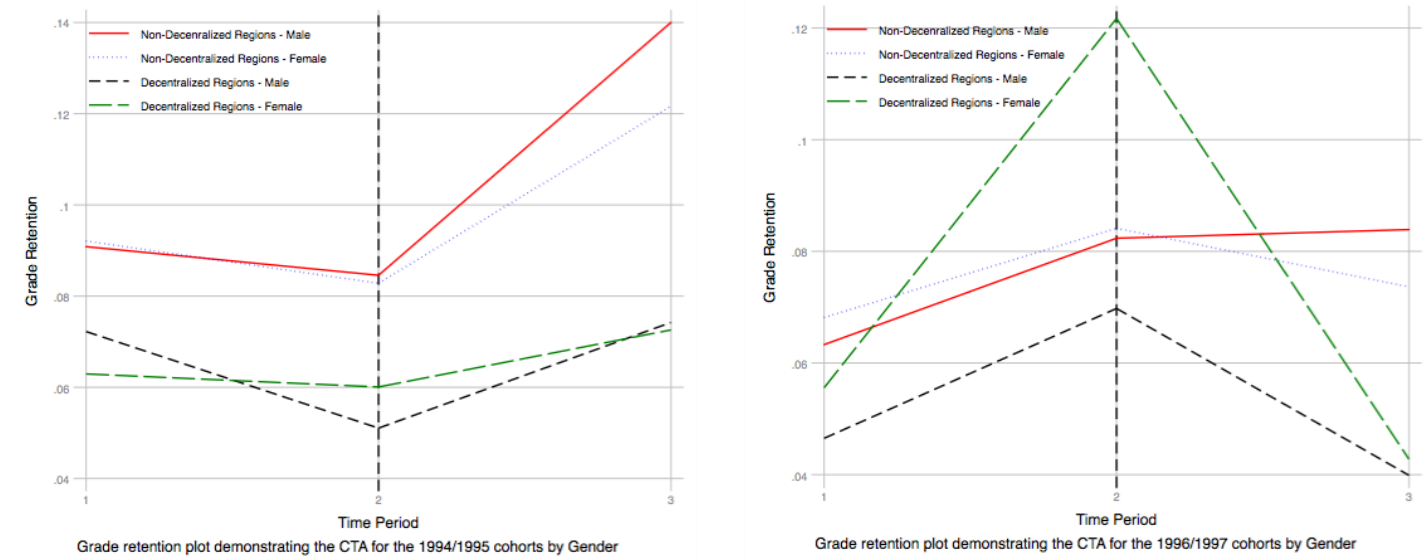
CEU eTD Collection

Figure A.1.1: Grade Retention plots for the 1994/1995 and 1996/1997 cohort by treatment



Note: The plots show average grade retention from a linear regression of grade retention for each treatment group. I used the full population for the 1994/1995 cohort (left) and the full population for the 1996/1997 cohort (right).

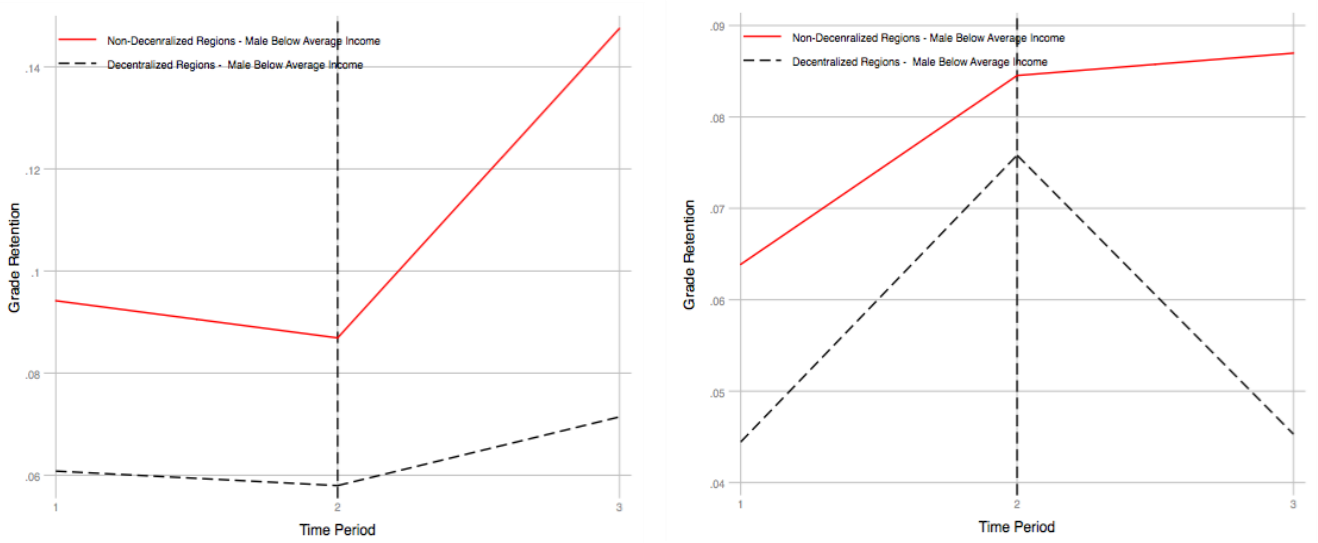
Figure A.1.2: Grade Retention plots for the 1994/1995 and 1996/1997 cohort by gender



Note: The plots show average grade retention from a linear regression of grade retention by gender and treatment status. I used the full population for the 1994/1995 cohort (left) and the full population for the 1996/1997 cohort (right).

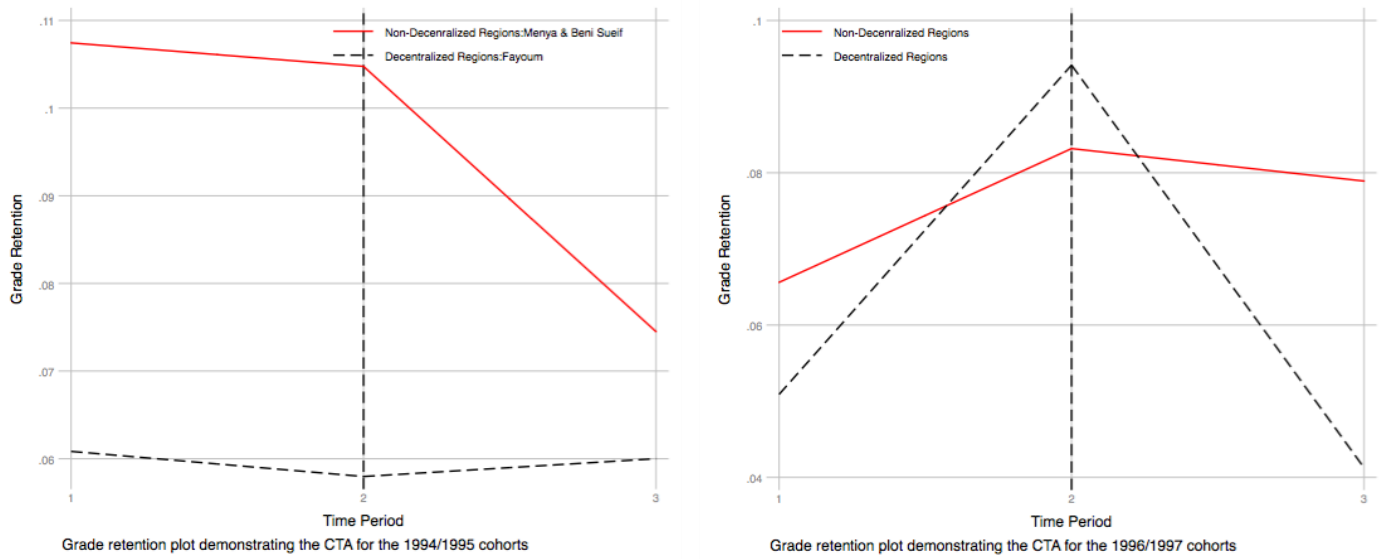
CEU eTD Collection

Figure A.1.3: Grade Retention plots for the 1994/1995 and 1996/1997 cohort – Male students – Below Average Income



Note: The plots show average grade retention from a linear regression of grade retention for male students coming from households at or below average income by treatment status. I restrict the full population to the male and below average income groups for both the 1994/1995 cohort (left) and the 1996/1997 cohort (right).

Figure A.1.4: Grade Retention plots for the 1994/1995 in North (left) and South (right) Upper Egypt

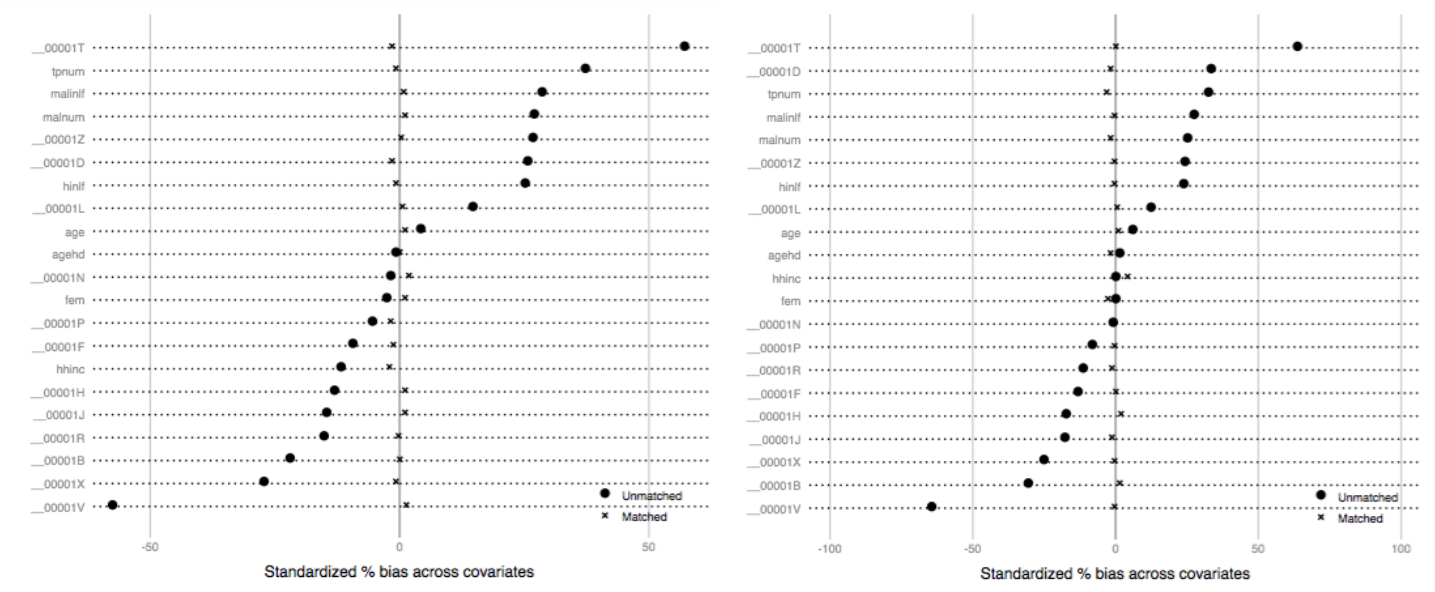


Note: The plots show average grade retention from a linear regression of grade retention by treatment status for the 1994/1995 cohort in each regional unit. The figure on the left show trends in North Upper Egypt and the figure on the right show trends in South Upper Egypt. The aim of the figure is to indicate the presence of spillover effects on the untreated education stages in the same state.

CEU eTD Collection

Appendix II 5-Nearest Neighbor Matching (within propensity score caliper)

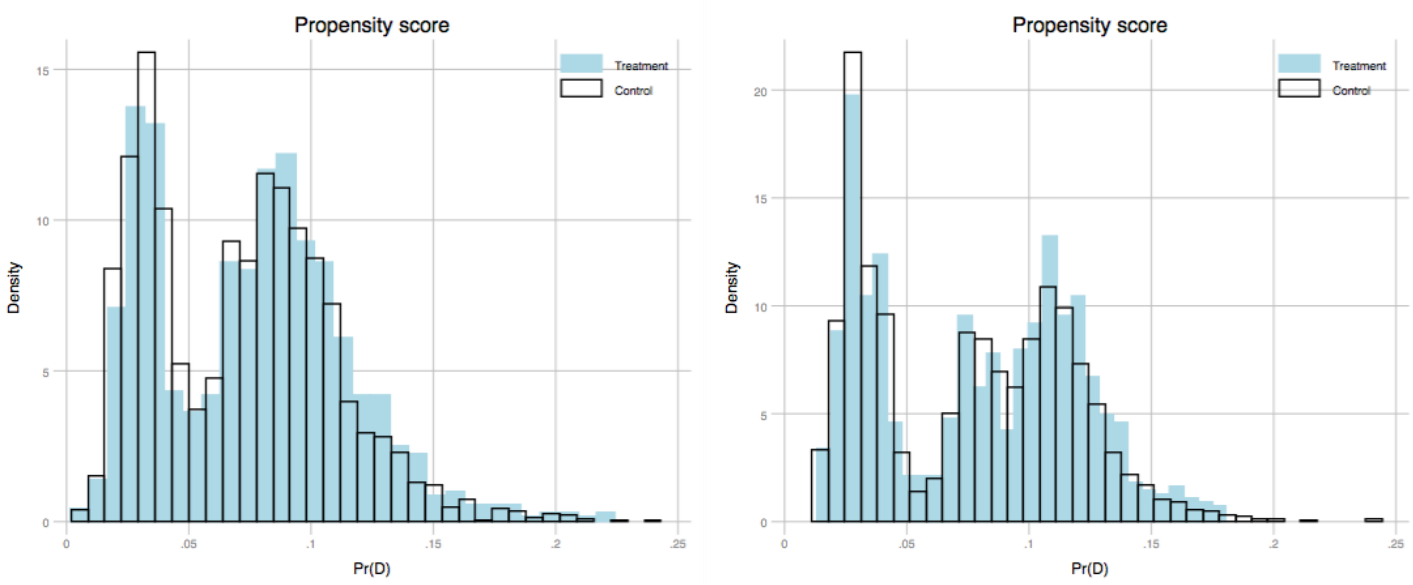
Figure A.2.1: Balance in individual level covariates after matching on 5 nearest neighbors



Note: A visual representation of the reduction in bias between the matched and unmatched sample achieved by matching using the 5-nearest neighbor within propensity score caliper. (Left) 1994/1995 cohort (Right) 1996/1997 cohort. The algorithm matches observations if their propensity scores falls within a pre-determined caliper. Propensity scores are obtained from a logistic regression that includes all individual level covariates and the main outcome variable is a dummy indicating treatment status.

Figure A.2.2: Histograms of the propensity scores for each cohort sample

(a) 1994/1995 cohort (b) 1996/1997 cohort



Note: The figure displays histogram plots of the propensity scores obtained from running a logisitic regression using all individual level covariates where the outcome variable is a dummy variable that indicates treatment. The figures show that for both cohort groups there is a wide range for common support.

CEU eTD Collection

Table A.2.1: Lead coefficients and Common Trends Assumption test by gender and household income groups – Matched sample

	I. 1994/1995 Cohort		II. 1996/1997 Cohort	
	(1) Model 1	(2) Model 2	(3) Model 1	(4) Model 2
Panel A. Full Sample				
$(D_{is}.time\_period_{it})$ (β_a)	0.036	0.032	0.031*	0.023
	(0.027)	(0.027)	(0.019)	(0.019)
$(D_{is}.time\_period_{it})$ (β_b)	0.032	0.034	0.044	0.037
	(0.021)	(0.021)	(0.027)	(0.026)
# of Observations	6,979	6,750	7,756	7,513
F-statistic of Joint test	1.37	1.44	2.30	1.89
(p-values)	(0.256)	(0.239)	(0.102)	(0.152)
Panel B. Male Sample				
$(D_{is}.time\_period_{it})$($\beta_a$)	0.044	0.046	0.032	0.023
	(0.036)	(0.034)	(0.024)	(0.021)
$(D_{is}.time\_period_{it})$($\beta_b$)	0.034	0.036	0.036	0.027
	(0.029)	(0.030)	(0.027)	(0.023)
# of Observations	3,740	3,606	4,011	3,879
F-statistic of Joint test	0.97	1.03	1.46	1.01
(p-values)	(0.379)	(0.356)	(0.233)	(0.364)
Panel C. Female Sample				
$(D_{is}.time\_period_{it})$ (β_a)	0.027	0.012	0.030	0.019
	(0.036)	(0.038)	(0.025)	(0.028)
$(D_{is}.time\_period_{it})$ (β_b)	0.028	0.055	0.054	0.049
	(0.028)	(0.027)	(0.038)	(0.033)
# of Observations	3,236	3,144	3,745	3,634
F-statistic of Joint test	0.56	0.44	1.39	1.14
(p-values)	(0.573)	(0.643)	(0.251)	(0.322)
Panel D. Below Avg. Income				
$(D_{is}.time\_period_{it})$ (β_a)	0.029	0.022	0.030	0.023
	(0.028)	(0.029)	(0.021)	(0.020)
$(D_{is}.time\_period_{it})$ (β_b)	0.038	0.040*	0.040	0.037
	(0.023)	(0.024)	(0.030)	(0.025)
# of Observations	5,944	5,720	6,536	6,294
F-statistic of Joint test	1.35	1.43	1.60	1.43
(p-values)	(0.260)	(0.241)	(0.203)	(0.242)
Panel E. Above Avg. Income				
$(D_{is}.time\_period_{it})$ (β_a)	0.046	0.047	0.042	0.020
	(0.058)	(0.060)	(0.048)	(0.049)
$(D_{is}.time\_period_{it})$ (β_b)	-0.006	0.008	0.081*	0.065*
	(0.044)	(0.041)	(0.041)	(0.035)
# of Observations	1,032	1,030	1,220	1,219
F-statistic of Joint test	0.64	0.37	2.15	1.79
(p-values)	(0.528)	(0.691)	(0.118)	(0.169)

Reported are the lead coefficients from a weighted regression where the main outcome variable is grade retention. Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%.

Shown are F-statistic and p-value from a joint test of the lead coefficients $\beta_a=\beta_b=0$.

Note: I implement a 5-neighbor matching within propensity score caliper on both populations using individual-level covariates

Note: Model 1 includes a group fixed effect and a time period specific fixed effect. Model 2 additionally controls for state level covariates. Each panel shows lead coefficients and runs the CTA test for a different sample as indicated in the sub-heading for each panel.

CEU eTD Collection

Table A.2.2: Lead coefficients and Common Trends Assumption test by interactions between gender and household income groups– Matched sample

	I. 1994/1995 Cohort		II. 1994/1995 Cohort	
	(1) Model 1	(2) Model 2	(3) Model 1	(4) Model 2
Panel F. Below Avg. Income - Male				
$(D_{is}.time\_period_{it})$ (β_a)	0.042	0.042	0.031	0.025
	(0.038)	(0.037)	(0.024)	(0.022)
$(D_{is}.time\_period_{it})$ (β_b)	0.047	0.044*	0.036	0.031
	(0.028)	(0.031)	(0.029)	(0.024)
# of Observations	3,207	3,074	3,415	3,284
F-statistic of Joint test	1.49	1.15	1.22	1.03
(p-values)	(0.226)	(0.319)	(0.298)	(0.359)
Panel G. Above Avg. Income - Male				
$(D_{is}.time\_period_{it})$ (β_a)	0.028	0.029	0.037	0.016
	(0.098)	(0.106)	(0.060)	(0.068)
$(D_{is}.time\_period_{it})$ (β_b)	-0.040	-0.064	0.057	0.047
	(0.076)	(0.065)	(0.042)	(0.041)
# of Observations	533	532	596	595
F-statistic of Joint test	0.58	1.14	1.04	0.65
(p-values)	(0.562)	(0.321)	(0.357)	(0.521)
Panel H. Below Avg. Income - Female				
$(D_{is}.time\_period_{it})$ (β_a)	0.016	-0.005	0.030	0.018
	(0.037)	(0.042)	(0.029)	(0.032)
$(D_{is}.time\_period_{it})$ (β_b)	0.028	0.028	0.046	0.042
	(0.033)	(0.031)	(0.041)	(0.036)
# of Observations	2,737	2,646	3,121	3,010
F-statistic of Joint test	0.35	0.52	0.88	0.70
(p-values)	(0.706)	(0.596)	(0.417)	(0.495)
Panel I. Above Avg. Income - Female				
$(D_{is}.time\_period_{it})$ (β_a)	0.064	0.037	0.048	0.016
	(0.062)	(0.062)	(0.064)	(0.062)
$(D_{is}.time\_period_{it})$ (β_b)	0.029	0.035	0.110*	0.085
	(0.041)	(0.045)	(0.060)	(0.054)
# of Observations	499	498	624	624
F-statistic of Joint test	0.59	0.33	1.77	1.24
(p-values)	(0.553)	(0.716)	(0.173)	(0.290)

Reported are the lead coefficients from a weighted regression where the main outcome variable is grade retention. Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%.

Shown are F-statistic and p-value from a joint test of the lead coefficients $\beta_a=\beta_b=0$.

Note: I implement a 5-neighbor matching within propensity score caliper on both populations using individual-level covariates.

Note: Model 1 includes a group fixed effect and a time period specific fixed effect. Model 2 additionally controls for state level covariates. Each panel shows lead coefficients and runs the CTA test for a different sample as indicated in the sub-heading for each panel.

CEU eTD Collection

Table A.2.3: Estimates for the effect of the pilot program on grade retention by gender and household income groups – Matched sample

	I. 1994/1995 Cohort			II. 1996/1997 Cohort		
	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 1	(5) Model 2	(6) Model 3
Panel A. Full Sample						
$(D_{is}.after_{it})$ (β_3)	-0.033*	-0.038**	-0.056***	-0.038*	-0.043**	-0.067***
	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.022)
# of Observations	6,976	6,976	6,750	7,756	7,756	7,513
Root MSE	0.279	0.278	0.276	0.258	0.257	0.256
Panel B. Male Sample						
$(D_{is}.after_{it})$ (β_3)	-0.038	-0.042	-0.066**	-0.034	-0.037*	-0.055**
	(0.027)	(0.026)	(0.028)	(0.021)	(0.021)	(0.024)
# of Observations	3,740	3,740	3,606	4,011	4,011	3,879
Root MSE	0.281	0.280	0.275	0.244	0.244	0.243
Panel C. Female Sample						
$(D_{is}.after_{it})$ (β_3)	-0.027	-0.032	-0.034	-0.044	-0.053*	-0.080***
	(0.026)	(0.025)	(0.027)	(0.029)	(0.028)	(0.030)
# of Observations	3,236	3,236	3,144	3,745	3,745	3,634
Root MSE	0.276	0.276	0.276	0.273	0.270	0.267
Panel D. Below Avg. Income Sample						
$(D_{is}.after_{it})$ (β_3)	-0.035	-0.042**	-0.058**	-0.037	-0.041*	-0.062***
	(0.022)	(0.020)	(0.023)	(0.023)	(0.021)	(0.023)
# of Observations	5,944	5,944	5,720	6,536	6,536	6,294
Root MSE	0.284	0.283	0.282	0.260	0.258	0.258
Panel E. Above Avg. Income Sample						
$(D_{is}.after_{it})$ (β_3)	-0.013	0.004	-0.029	-0.067**	-0.072**	-0.111***
	(0.044)	(0.043)	(0.041)	(0.032)	(0.033)	(0.035)
# of Observations	1,032	1,032	1,030	1,220	1,220	1,219
Root MSE	0.245	0.245	0.238	0.250	0.250	0.250

Reported are the difference-in-differences interaction coefficients obtained by running a weighted regression where the main outcome is grade retention. Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. I also report the total number of observations used in the estimation and the Root MSE.

Note: I implement a 5-neighbor matching within propensity score caliper on both populations using individual and household level covariates.

Note: Each panel estimates the regression formulation on a different sample as indicated in the sub-heading of each panel. Model 1 includes a group effect and a post treatment effect. Model 2 additionally includes a state specific fixed effects and a time period specific fixed effect. Finally, Model 3 additionally includes state-level covariates.

CEU eTD Collection

Table A.2.4: Estimates for the effect of the pilot program on grade retention by gender and household income groups – Matched sample

	I. 1994/1995 Cohort			II. 1996/1997 Cohort		
	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 1	(5) Model 2	(6) Model 3
Panel F. Below Avg. Income - Male						
$(D_{is}.after_{it})$ (β_3)	-0.045*	-0.051**	-0.070**	-0.034	-0.035	-0.051**
	(0.026)	(0.025)	(0.029)	(0.023)	(0.023)	(0.025)
# of Observations	3,207	3,207	3,074	3,415	3,415	3,284
Root MSE	0.283	0.282	0.277	0.247	0.246	0.245
Panel G. Above Avg. Income - Male						
$(D_{is}.after_{it})$ (β_3)	0.013	0.031	0.018	-0.053	-0.067**	-0.101**
	(0.078)	(0.074)	(0.066)	(0.035)	(0.035)	(0.040)
# of Observations	533	533	532	595	596	595
Root MSE	0.269	0.268	0.256	0.229	0.229	0.228
Panel H. Below Avg. Income - Female						
$(D_{is}.after_{it})$ (β_3)	-0.024	-0.032	-0.040	-0.041	-0.048	-0.069**
	(0.030)	(0.028)	(0.031)	(0.033)	(0.031)	(0.032)
# of Observations	2,737	2,737	2,646	3,121	3,121	3,010
Root MSE	0.286	0.286	0.286	0.273	0.271	0.268
Panel I. Above Avg. Income - Female						
$(D_{is}.after_{it})$ (β_3)	-0.041	-0.26	0.024	-0.083*	-0.098**	-0.130**
	(0.041)	(0.040)	(0.044)	(0.046)	(0.049)	(0.051)
# of Observations	499	499	498	624	624	624
Root MSE	0.216	0.217	0.207	0.269	0.270	0.270

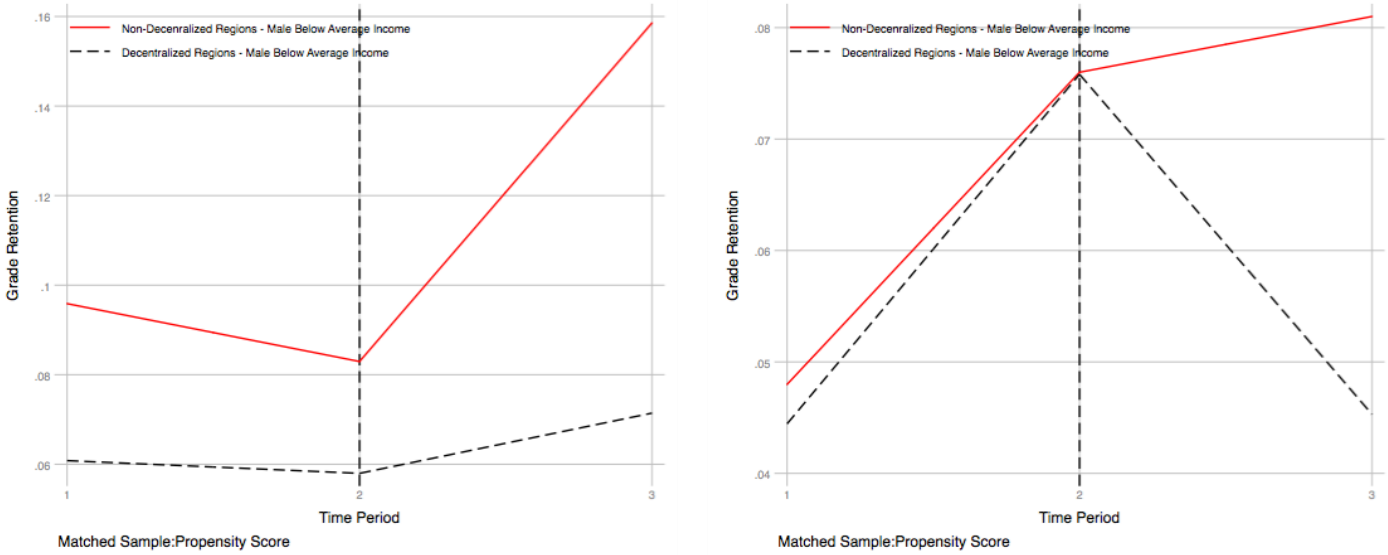
Reported are the difference-in-differences interaction coefficients obtained by running a weighted regression where the main outcome is grade retention. Standard errors are clustered at the level of the primary sampling unit and reported in brackets. The stars declare the significance level of the coefficient where * 10%, ** 5%, ***1%. I also report the total number of observations used in the estimation and the Root MSE.

Note: I implement a 5-neighbor matching within propensity score caliper on both populations using individual and household level covariates.

Note: Each panel estimates the regression formulation on a different sample as indicated in the sub-heading of each panel. Model 1 includes a group effect and a post treatment effect. Model 2 additionally includes a state specific fixed effects and a time period specific fixed effect. Finally, Model 3 additionally includes state-level covariates.

CEU eTD Collection

Figure A.2.3: Grade Retention plots for the 1994/1995 and 1996/1997 cohort – Male students – Below Average Income – using matched sample



Note: The plots show average grade retention from a linear regression of grade retention for male students coming from households at or below average income by treatment status. I matched observations using a 5-nearest neighbor matching within propensity score caliper to control for selection. (left) 1994/1995 cohort (right) the 1996/1997 cohort.

CEU eTD Collection