Palacký University Olomouc
University of Clermont Auvergne
University of Pavia

MASTER THESIS

Evaluation of Public Policies to Support Education: the Case of Morocco

Elizaveta Rusakova

Supervisors:
Prof. Fouzi Mourji (University of Hassan II Casablanca)
Doc. Ing. Mgr. Jaromír Harmáček, Ph.D (Palacký University Olomouc)

GLODEP 2021

Evaluation of Public Policies to Support Education: the Case of Morocco

Submitted by
Elizaveta Rusakova

Erasmus Mundus Joint Master's Degree in International Development Studies
(GLODEP 2019-2021)

Supervised by
Fouzi Mourji
Professor of Econometrics
University of Hassan II Casablanca

&

Doc. Ing. Mgr. Jaromír Harmáček, Ph.D
Faculty of Science
Palacký University Olomouc

Date of Submission: 31 May 2021

Declaration

I, Elizaveta Rusakova, hereby declare in lieu of oath that this thesis has been composed by myself. I confirm that this work is presented for obtaining Erasmus Mundus Joint Master's degree in International Development Studies. Except where it is stated otherwise by reference or acknowledgment, the presented thesis is a result of my own work.

Signature:

Date: May 31st, 2021

UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	**Elizaveta RUSAKOVA**
Osobní číslo:	**R190774**
Studijní program:	**N1301 Geography**
Studijní obor:	**International Development Studies**
Téma práce:	**Evaluation of public policies to support education: the case of Morocco**
Zadávající katedra:	**Katedra rozvojových a environmentálních studií**

Zásady pro vypracování

According to UNESCO, in 2018 there are around 258 million children without access to school. Moreover, as UNICEF reports in 2018, approximately 58% of primary and lower secondary school-age children and adolescents do not reach minimum reading and mathematics proficiency levels. There are numerous projects implemented in order to improve education outcomes, and numerous studies that evaluate impact of these interventions. Programs' results and the effectiveness frequently depend on the period of implementation, modalities of execution, country-specific characteristics, and social context.

This study aims to analyze the impact of several public policies implemented in Morocco to support the educational system. Morocco is an interesting case-study as the expenditure on education has been increasing lately and is rather high for a developing country (6.5% of GDP in 2019), however the education outcomes (i.e. dropout rates, results of the TIMMS tests etc.) still show the need for considerable improvements.

Among the prospective programs for this case study are Tayssir (Cash Transfer program); Dar Taliba (construction of boarding houses for girls to support them in continuing education); „1 Million School Bags“ initiative (for kids from poor households) among others. The data sources will be national surveys (such as ONDH) and databases as well as from international databases such as UNESCO UIS Statistics. The study will mostly rely on quantitative methods for data analysis. Prior to utilization of various econometric techniques for modelling, we plan to use descriptive statistical tools such as cross-tabulation. The study will attempt to approximate the effects of each program, analyze the results for in-depth explanation and comparison of programs considering the specificity of Moroccan context. Our research is „policy-oriented“ as it will use the results to formulate recommendation regarding the impact of current programs and the possibilities for programs' performance optimization for students.

Rozsah pracovní zprávy:	**20-25 tisíc slov**
Rozsah grafických prací:	**dle potřeby**
Forma zpracování diplomové práce:	**tištěná**
Jazyk zpracování:	**Angličtina**

Seznam doporučené literatury:

Garcia, S., & Saavedra, J. E. (2017). Educational impacts and cost-effectiveness of conditional cash transfer programs in developing countries: A meta-analysis. Review of Educational Research, 87(5), 921-965

Liouaeddine, M., Elatrachi, M., & Karam, E. M. (2018). The analysis of the efficiency of primary schools in Morocco: modelling using TIMSS database (2011). The Journal of North African Studies, 23(4), 624-647.

Llorent Bedmar, V. (2015). Dysfunction and educational reform in Morocco. Asian Social Science, 11 (1), 91-96.

Saoudi, K., Chroqui, R., & Okar, C. (2019). Student Achievement in Moroccan Student Achievement in Moroccan Educational Reforms: A Significant Gap Between Aspired Outcomes and Current Practices. Interchange, 1-20.

UIS. UNESCO Institute for statistics. http://data.uis.unesco.org/#

Vedoucí diplomové práce:	**Ing. Mgr. Jaromír Harmáček, Ph.D.** Katedra rozvojových a environmentálních studií
Datum zadání diplomové práce:	**29. ledna 2021**
Termín odevzdání diplomové práce:	**31. května 2021**

L.S.

doc. RNDr. Martin Kubala, Ph.D. děkan	**doc. RNDr. Pavel Nováček, CSc.** vedoucí katedry

V Olomouci dne 29. ledna 2021

Acknowledgments

In the process of writing this thesis, I have received useful support and assistance.

I would like to express my gratitude to my supervisor, professor Fouzi Mourji, for his timely help and great encouragement during the process of writing this thesis. I am grateful to have an opportunity to learn more about the Economics of Education and impact assessment with your guidance. I would also like to thank my other supervisor, professor Jaromír Harmáček. You have always had insightful comments and useful questions that have helped me to improve my work to the level that I present here. I would also take an opportunity to thank Mr. Abdelfettah Hamadi from Observatoire National du Développement Humain (ONDH) for facilitating the acquisition of data.

I would like to thank all members of GLODEP Consortium for making this experience unique despite all the challenges. Your commitment to make things work, your hard work and your encouragement, especially during the COVID-19 pandemic, kept all our cohort motivated. Thank you to all the people from different parts of the world I have met during GLODEP for sharing this experience with me. Without the contribution of each one of you, these two years would not be the same.

Last but not the least, thank you to all members of my family and my friends. Despite the distance separating us, you are always here for me. I love you to the moon and back.

Abstract

This thesis is aimed at assessing the impact of public policies to support education in Morocco. Morocco is an interesting case study as educational investments are high but inequalities between urban and rural areas persist, repetition rate remains high while the transition to secondary education low. In this study, we assess how two programs (program of school supplies provision "One million schoolbags" and subsidized food provision "Canteens program") affect these outcomes for the main target group of the programs – rural area students of primary and lower secondary levels of education.

To assess the impact of programs, this study makes use of Quasi-experimental research design and Propensity score matching (PSM) to ensure that the found treatment effect is causality and not simply correlation. Certainty in the causality of the observed effect is maximized only if a counterfactual outcome is approximated well enough by the selected control group, which PSM is aimed at. Treatment effects are found by using logit models and calculating average marginal effects.

We have rejected the null hypothesis of no impact for the effect of One million schoolbags program and both programs on repetition: participation has considerably increased the probability to repeat the grade for program beneficiaries. The ability to benefit from free school supplies and subsidized meals (even in case of repetition) gives beneficiaries reassurance which might decrease fear of repetition and desire to exert maximum efforts for academic success. Effect of programs on transition to college was positive (participation increased the probability to transit), but estimations were statistically non-significant most likely due to the small sample size.

Key words: treatment effect estimation, quasi-experimental design, propensity score matching, education policies in Morocco, repetition rate, transition to secondary education

Contents

List of figures

List of tables

List of abbreviations

ATE	Average treatment effect on the population
ATT	Average treatment effect on the treated
CGIAR	Consultative Group for International Agricultural Research
DAMP	Dépense annual Moyenne par personne
DID	Differences-in-Differences
FE	Fixed-Effects
IA	Impact Assessment
INTRAC	International NGO Training and Research Centre
LASAARE	Laboratoire de Statistique Appliquée à l'Analyse et la Recherche en Économie
LEA	Logical framework analysis
LIC	Low-income countries
LMIC	Low-to-Middle-Income countries
MENA	Middle East and North Africa
MENFP	Ministère de l'Education Nationale et de la Formation Professionnelle
NGO	Non-governmental organization
ONDH	Observatoire National du Développement Humain
PSM	Propensity score matching
QE	Quasi-experimental
QED	Quasi-experimental design
RCT	Randomized control trial
RD	Regression discontinuity
SES	Socio-economic status
SMD	Standardized Mean Difference
TIMSS	Trends in Mathematics and Science Study
UAE	United Arab Emirates
UNESCO	United Nations Educational, Scientific and Cultural Organization

Introduction

Background: education and public policies evaluation

The importance of education for human development is beyond dispute. Providing equal and universal access to education is critical for country's development which is why it has constantly been a part of Millennium Development Goals (MDGs), Sustainable Development Goals (SDGs) for 2030, Global Education 2030 Agenda. Accessible and quality education is also prioritized at the national level as it is proved to be fundamental to address social disparities and accelerate economic growth. Mincer (1958) and Schultz (1961) are widely considered to be the first scholars to include education (human capital) as an explanatory part of gaps in economic performance between countries. Education is a human right according to the Universal Declaration of Human Rights (Article 26). It is a main component of equal and sustainable society and a transformation tool for long-term poverty reduction. Nevertheless, in 2018 around 263 million children (practically one out of five children) remain out of school, and the figure has not changed significantly between 2013 and 2018 (Otchet, 2018).

In the light of the crucial role of education for human and country development, evidence-based education policies have been attracting considerable attention of national and international policymakers. This led to an upsurge in public policy evaluations in education field at the end and beginning of the XXI century. According to Slavin (2002), evidence-based evaluations in education have intensified only since the turn of the XXI century: much later than in fields of medicine and agriculture where rigorous evaluations and their results led to unprecedented improvement (Shavelson & Towne, 2002). Slavin (2002) stated that before the XXI century the randomized experiments were not uncommon in education but tended to be short and theoretically focused lacking rigorous evaluations possible to serve as a solid base for policy and practice. Good evidence-based evaluations are not necessarily randomized controlled trials (RCT): quasi-experimental methods are also largely considered to be valid as they exhibit similar results to RCTs if constructed and designed well (Vivalt, 2015).

According to Glewwe and Muralidharan (2016), the quality and quantity of research increased even more after the formalization of guidance on research methodologies and design, and the publication of Handbook chapter on Economics of education in developing countries (Glewwe & Kremer, 2006). The boom in the number of impact studies led to meta-analysis of program's evaluations (Kremer & Holla 2009, McEwan 2015, Masino & Niño-Zarazúa 2016 among the most recent ones) and the creation of platforms with completed and ongoing impact evaluations such as World Bank platform (World Bank, 2016). Nevertheless, despite the existing pool of evidence-based evaluations, there is still disagreement (even in meta-analysis) on the effects of certain policies, so the need for education impact studies remains pressing. Notwithstanding academic research improvements and availability of different education policies and their proven efficiency, many developing countries nowadays still struggle to provide universal access to education, increase its quality and decrease drop-out and repetition rates. Despite tremendous improvement in the world visualized in Figure 1, around 58.4 million children were out of school in 2019 (World Bank, 2020). Even after increasing enrolment rates, there is still a problem of attendance, especially in developing countries (Figure 2).

Figure 1. Graph of number of out-of-school children in the world, 1998-2014

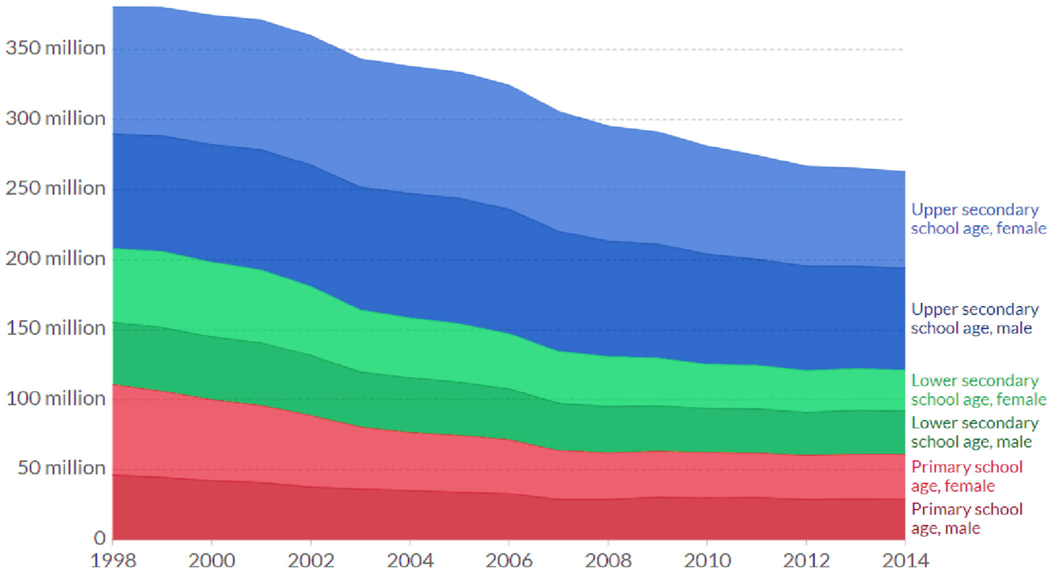
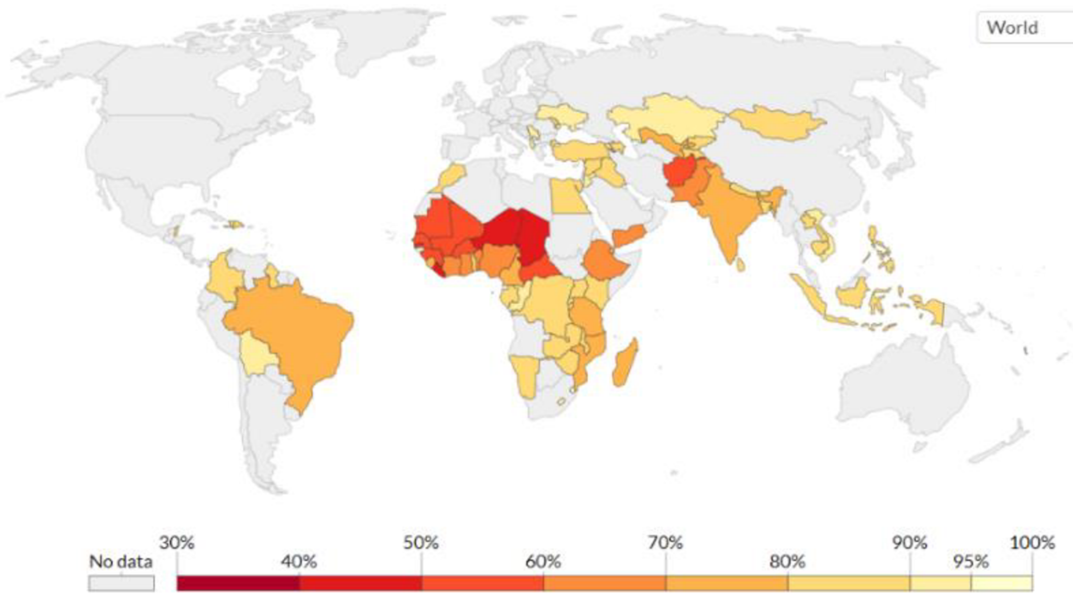


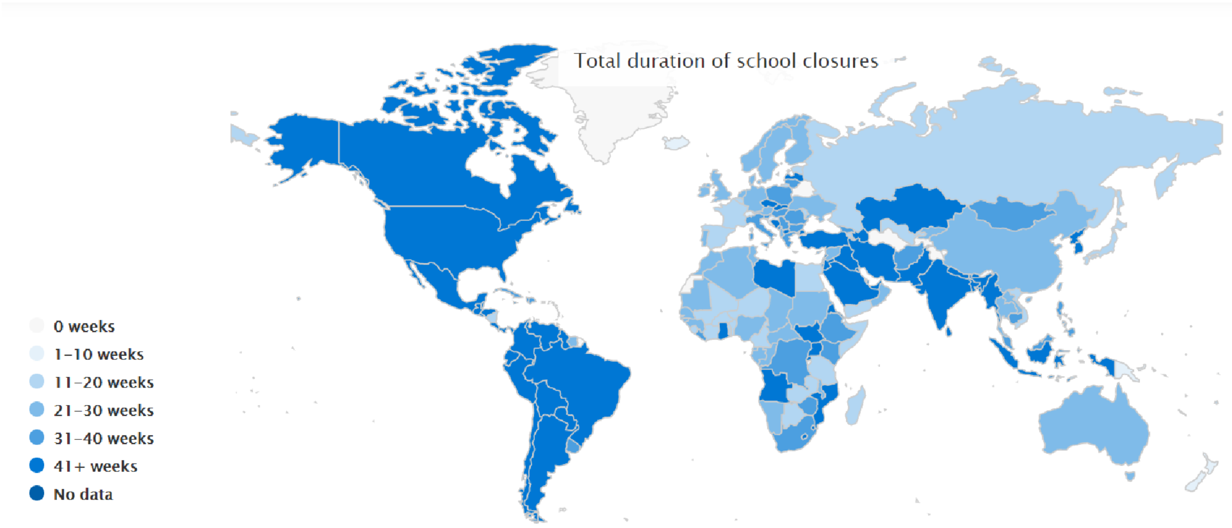
Figure 2. Map of net attendance rate in primary school in the world, 2015



Source: World Bank as cited in Roser M., & Ortiz-Ospina E. (2015)

Developing countries' national education systems are often disrupted by natural disasters, political conflicts, and the most recent outbreak of Covid-19 with forced confinement and school closure. UNESCO (2021a) estimated that at the peak of the pandemic, the number of out-of-school children was 1.6 billion from more than 190 countries. The crisis has also deepened pertaining inequalities affecting to a greater extent the most vulnerable students. As can be seen on the map (Figure 3), schools in many countries of South and North America, Africa and Middle East have remained closed for more than 41 weeks (UNESCO, 2021b). After a year of pandemic, it was predicted that over 100 million children are at risk of falling below the minimum proficiency requirements for reading while around 24 million students are at risk of dropping out (UNESCO 2021a). In the light of recent events, there is extra pressure to investigate existing educational programs and develop new ones to address existing educational inequalities and reduce educational system vulnerability.

Figure 3. Map of the total duration of school closures due to Covid-19 pandemic after a year of outbreak (as at March, 2021)



Source: UNESCO (2021b)

Country context: education in Morocco

The objective of this section is to provide a background of the Moroccan education system and its main challenges. The structure of school education system is presented in Table 1. Since 2000, basic compulsory education in Morocco is 9 years: primary and college (World Bank, 2020). A student can continue education in lower-secondary (equivalent of middle school; in Morocco – College[1]) and after in secondary (in many countries – equivalent of high school; in Morocco – Lyceum[2]).

Table 1. Education system in Morocco (school level)

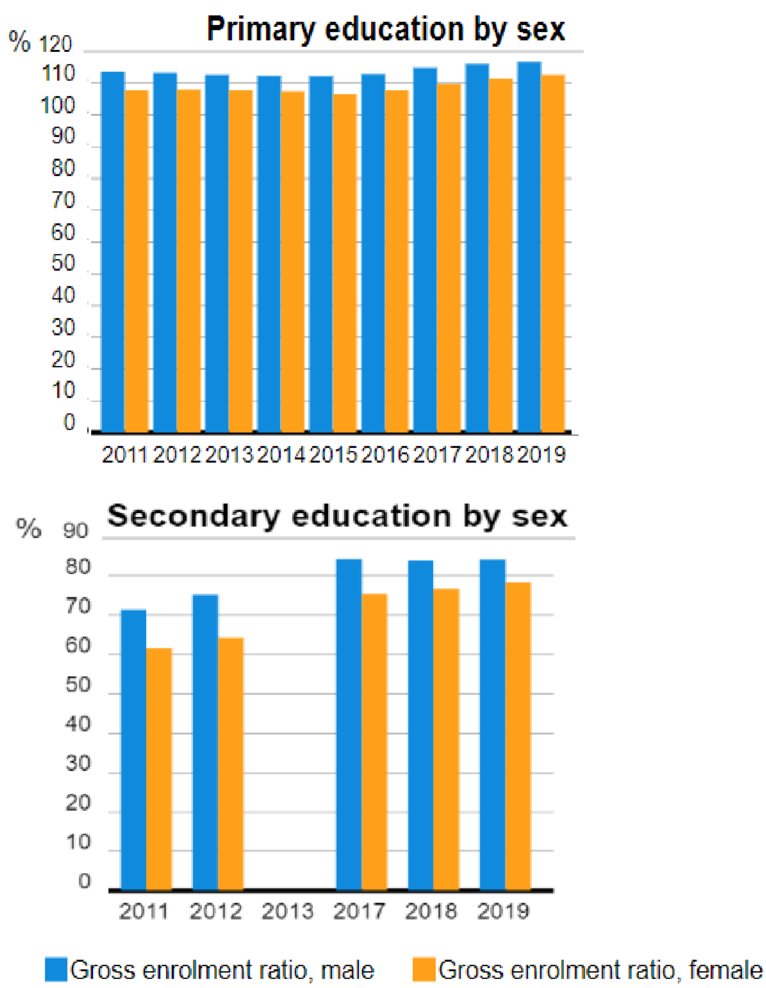
Education	Level	Grades	Age	Years
Primary	Primary school	1-6	6-12	6
Lower-secondary	Basic education (College)	7-9	12-15	3
Upper-secondary	General secondary (Lyceum)	10-12	15-18	3
	Technical secondary		15-18	3

Source: composed by author based on UNESCO UIS and https://www.scholaro.com/pro/Countries/Morocco/Education-System

[1] Further in the study this level of education is referred to as college
[2] Further in the study this level of education is referred to as lyceum

According to UNESCO Statistics, the number of out-of-school children has decreased significantly from 205.2 (2011) to 16.1 (2019) thousand (Annex I). According to the latest available data (2019), in primary school, the gross enrolment is 114.7% and net enrolment – 99.5% while in secondary it is 81.19% and 66.2%. In gross enrolment rates, there are apparent gender disparities in school access: in 2019 primary school there are 116.69% males and 112.73% females while in secondary school – 84.05% males and 78.18% females (Figure 4). Differences in gross and net enrolment ratios can often be explained by high repetition rates or late school enrolment. Despite good results regarding enrolment, some indicators still need to be worked on. For example, repetition rate in primary school in 2019 was 10% for boys and 8% for girls. Number is quite high for MENA region: in 2019, in neighboring Algeria repetition is 5.2% (UIS, 2020a), in Egypt – 1.4% (UIS, 2020b). The survival to last year of primary school was 94.26% while the transition rate from primary to college was 92.3% (93.9% boys and 90.6% girls) in 2018 (UNESCO, 2020). The transition rate has been decreasing in the last decade (from 88.6% in 2011) as well as gender disparities have been: in 2011 transition among girls was 84.5%, for boys – 92.3%.

Figure 4. Gender disparities in primary and secondary school gross and net enrolment



Source: UNESCO (2020). http://uis.unesco.org/en/country/ma

Regarding the quality of education in Morocco, the picture is not so bright. In TIMSS program in 2015, Morocco was ranked the last of all country participants with an average of 377 in math in the 4th year of primary school (Mullis, Martin, Foy, & Hooper, 2015). It is considerably lower than both the average international standard (500 points) and other MENA countries (Turkey – 483; UAE – 452; Iran – 431 points). According to Ikira (2021), these results might be partially explained by lack of parents' involvement in children performance: it is low among Moroccan parents (52%) while in countries with higher average performance, this indicator is higher (Turkey – 82.3%, UAE – 73.3%). Presented indicators can be a sign of positive but to a certain extent development of Moroccan educational system which however can be considered insufficient in the view of significant (for a developing country) national expenditure on education. For example, in 2008 educational expenditure composed around 25% of all public expenses (UNESCO UIS) while in the last decade, expenses level was constant at 5.5% (% of GDP) in 2010-2015 (Ifa, & Guetat, 2018) and then raised to 6.5% in 2020 (Abdessamad, 2020).

Educational reforms in Morocco have started slowly after country's independence in 1956. Some deep-rooted problems were born in the colonial period. First, teaching methods were honed to train civil servants, police, militants which made public sector employment preferable to the private sector. Secondly, in the authoritarian system, the evaluation of teachers and their work were highly uncommon. Thirdly, there was a separation of Morocco into rich agricultural plains privileged by colonists and remote mountain areas that were left less developed than the former. This separation had an impact on the underdevelopment of rural areas. Government historically has invested insufficiently in rural areas: for example, in the 80s only around 10% of educational expenses were directed to rural areas though around 52% population lived there (Khandker, Lavy, & Filmer, 1994).

Educational reform intensified in the 90s together with democratic transition: Special Commission for Education-Formation was created in 1999; the decade of 2000-2009 was called the national decade of Moroccan education reform (Chtatou, 2015). First-decade reforms have managed to decentralize education decision-making and adopt the practice of evaluation to a certain extent, but they cannot be called very successful. Many teachers were reluctant to training due to no financial incentives, parents were afraid of increasing costs, reforms lacked transparent evaluation schemes and implementation strategies. Consequent National Educational Emergency Programme (2008-2012) had a positive impact on gender equality and the increase of competition between schools which led to certain teaching quality improvements. Nevertheless, there are several persisting challenges such as unequal and incomplete access to basic education, exclusion of the most vulnerable, low level of knowledge and skills as well as low educational efficiency represented by high repetition rate, dropping-out and low transition rate to secondary education.

In the context of the abovementioned education issues, in the second decade of the XXI century, Moroccan administration still prioritized reforms aimed at the most disadvantaged, poorest and rural students. The main goal of the latest reforms was to decrease direct (supplies, transportation) and indirect costs of schooling (opportunity cost of a child being at school rather than contributing to family income). Government efforts are tangible and rewarding, however persisting as well as newly emerging inequalities and challenges reinforce the need for further assessment of implementing policies, their strengths and shortcomings. This study will focus on the programs with the broadest coverage: free lunch provision (Canteens program) and annual provision of school supplies "One million schoolbags" initiative.

Purpose of study

The aim of this study is to extend current evidence on the evaluation of education policies on education outcomes. This paper sheds the light on the impact of two different programs on studied outcomes in compulsory education which still require considerable improvements in Morocco: repetition rate and transition to secondary education[3]. The chosen programs, though different in context and approach, have similar goals to increase schooling and provide equal opportunities for the most disadvantaged children primarily from rural areas. It was decided to limit students to rural area only as these areas are more problematic in terms of providing educational access and quality historically and practically. Rural area students are also the main target of the chosen programs so the estimated impact should be clearer and more accurate. This study focuses on the most acute indicators (outcomes) for the most needed students at the level of compulsory education. This approach will allow seeing how programs address relevant educational challenges and the most needed population. It is important to note that the purpose is not to compare the programs and their effects, but to attempt to draw comprehensive conclusions about public policies to support education in Morocco and their impact on insufficiently studied outcomes.

Main research question: What is the impact of the chosen educational policies to support education on the chosen education outcomes in compulsory education in Morocco?

Sub-questions of the study are:

- Does participation in programs tend to reduce the repetition rate among students from rural areas?
- Do chosen educational policies help beneficiaries to transit from primary to college more often than non-beneficiaries?
- How can policies be changed, adjusted, or optimized to improve their effect?

Significance and relevance of the study

The public policies analyzed in this study have been widely implemented in developing countries though their impact is still not widely understood as in the case of school meals' provision and school supplies programs. In Morocco research studying their impact has not been conducted yet so this study is a pioneer to examine their effect. Previous assessment studies of other programs in Morocco are largely based on the data of 2015 and 2016 (Ikira, 2021; Gazeaud & Ricard, 2021). This study makes

[3] Enrolling into first level of secondary school (college) after obtaining primary school diploma

use of more recent data from 2017. In a broader context, in the area of educational policy evaluation, there have been numerous impact studies on such outcomes as enrolment, drop-out, achievement (Jomaa, McDonnell, & Probart, 2011; Sabarwal, Evans, & Marshak, 2014; García, & Saavedra, 2017). However, the following study attempts to fill in the literature gap and to contribute to the available literature pool on educational policies evaluation by continuing the research on rarely studied (in context of the chosen programs) outcomes – repetition and transition to college from primary school. Findings are expected to be useful and informative for national decision-makers such as the National Observatory of Human Development (ONDH) and Ministries as well as to academia in public policy evaluation.

Scope and limitations

The study is aimed at approximating the impact of two programs implemented in Morocco in education. The study uses individual data. The research focuses on three main groups of population: primary school students who are current beneficiaries of the chosen programs, previous beneficiaries limited by age and non-beneficiaries who are used as control group individuals. The study was constrained by the Covid-19 pandemic and inability to reach research institute LASAARE in Morocco: limitations with access to additional data, information, and resource, certain challenges with communication and supervision. As the research progressed, the topic had to be refined and adjusted due to the changes made in the used questionnaire and data of 2017[4].

The study uses propensity score matching (PSM) to assess the treatment effect of a program on the chosen outcomes. PSM is a quasi-experimental method aiming to reduce non-randomization bias by matching individuals with similar observed characteristics but different in treatment. PSM is based on the conditional independence[5] assumption which is challenging but crucial to test for. Though the tests are not well developed yet and not frequently used (Stuart, 2010), this study attempts to test the robustness of results and the validity of the made assumption for the estimated effects.

Organization of work

The study is subsequently divided into three chapters. Chapter I focuses on theoretical and empirical literature review outlining notion and evaluation of impact studies and existing evidence on evaluation of chosen public policies in the world and Morocco. Chapter II elaborates on chosen research design, methodology and study outcomes. Chapter III discusses the main findings, their robustness, limitations and potential explanations for the found effect. Lastly, the conclusion summarizes work and makes final remarks.

[4] The data used in the study is based on the questionnaire of 2017 where the questions from educational part were altered in comparison with three previous questionnaires of 2012, 2013 and 2015

[5] Treatment assignment (program participation) is independent of the potential outcome (in this study repetition and transition rate to college) conditional on observable characteristics (Rosenbaum & Rubin, 1983). It does not consider unobservable characteristics

Chapter 1. Literature review

This chapter is divided into theoretical and empirical literature reviews to comprehend better the complexity of the topic. In theoretical section 1.1, the first part studies the notion of impact evaluation, its theoretical grounding and evolution in the recent period. The second part focuses on evaluation studies in education, their methodology, research designs and most common studied outcomes. In empirical literature review, Section 1.2.1. first summarizes evidence-based studies describing the most effective interventions in developing countries. After, it looks into available evidence for the type of projects related to this study. Section 1.2.2. focuses on the existing evaluations of the projects further analyzed in this study in Morocco.

1.1. Theoretical literature review

1.1.1. Impact studies: notion and evolution

In the beginning, it is crucial to define what are evaluation and impact studies. There are numerous definitions and classification of public policy analysis what supports its multidimensional nature and complexity. According to Sabatier (1991), the literature pool on public policies research can be divided into four groups according to their focus:

(1) Substantive area research which is a comprehensive study aimed at understanding politics in specific policy area (for example, health and education);

(2) Evaluation and impact studies are evaluation analyses of policy from welfare economics that later started to include other criteria and went beyond traditional cost-benefit analysis. Among the criteria are citizen participation, distributional effects among others;

(3) Policy process studies focus on the particular factors affecting stages of policy formulation and implementation;

(4) Policy design studies investigate the efficacy of policy instruments.

It is essential to provide a definition though its number is abundant in this area. Notion and essence of impact assessments differ from one study or institution to another. As defined by Fitz-Gibbon (1996), “impact is …any effect of the service (or of an event or initiative) on an individual or group”. One of the widely accepted is the definition by Roche (1999, p.21): “impact assessment is the systematic analysis of the lasting or significant changes— positive or negative, intended or not— in people's lives brought about by a given action or series of actions”. For O’Flynn (2010), the impact assessment is about answering such questions as who has been affected by the change, how significant and long-lasting will the change be. CGIAR (2013) has added several more possible descriptions to Roger’s definition: impact is the overall and long-term effect that can be primary or secondary, direct or indirect.

It is debatable whether evaluation and assessment of the program/policy are interchangeable concepts. Lately some scholars and organizations have started to differentiate evaluation and impact assessment. INTRAC report (O’Flynn, 2010) suggests that impact assessment embraces a wider range of questions than evaluation. Generally, evaluation relates to outcomes and results of the performed intervention while the impact assessment concerns the overall intervention's effect on the life of the target population (more details in Table 2). In other cases, scholars do not draw a line between policy evaluation and impact assessment considering both parts of a general analysis of project/program impact.

Table 2: Difference between Evaluation and Impact Assessment

Evaluation	Impact Assessment
Measure performance against objectives	Assesses change in lives of people (positive/negative; intended/unintended)
Middle of end of project/program cycle	Can be included at all stages and/or specifically after the project/program
Focus on interventions and outcomes	Focus on affected population and impacts
"What has happened? Did we achieve what we planned?"	"What has changed? For whom? How significant is it for the population?"

Source: Reprinted from Impact Assessment: Understanding and assessing our contributions to change by O'Flynn, M. Copyright 2010 by INTRAC

Impact studies have changed over time. Roche (1999) states that since the 1950s several strategies were used to ex-ante evaluate the project. These were environmental impact assessment, social impact assessment, cost-benefit analysis, social cost-benefit analysis (Howes 1992). Later, by the end of the 1960s impact analysis for ex-post project assessment started to be used by the development sector. As Weiss (1998) assumes, Scriven for the first time, in his work of 1967, has introduced the notion of formative and summative evaluations. He has complemented concepts in his further works (Scriven, 1991; 1996). He has defended dichotomy as a reasonable way to classify evaluation and described formative evaluation as an analysis of a program in its implementation stage while summative evaluation as a measure of achievement of intended outcomes.

In the 1970s logical framework analysis (LFA) was developed: this approach is described as "a practice with relatively little accompanying theory" (Gasper, 2000, p.17) and is frequently used in the ex-post evaluation. LFA is often represented in a project matrix with goal, purpose and outputs activities. However, its main disadvantage is a focus on delivering outcomes and achieving the intended effect through planned routes and interventions missing out on identifying unintended effects. Approaches developed later focused on the inclusion of diverse points of view and participatory methods. Participatory methods in impact evaluation have gained momentum in the XXI century as development focused on localization, participation, and community empowerment. A good example of participatory approach application is Rogers (2009) Conditional Cash Transfer (CCT) evaluation. She suggests using participatory census mapping in representative communities, creating focus groups, taking interview, collecting stories, working with individual cases of outliers, and facilitating causal-linkage diagramming.

Other common evaluation methods are conformance-based (also called "goal-attainment" as it considers only the intended effects of the problem; Vedung, 1997) and performance-based evaluations. Laurian et al. (2004) explain the difference between them: conformance-based approach assumes that policy is implemented when it meets the set objectives while performance-based approach refers to the way of application of the policy, not its implementation. A theory to back up the evaluation process whether quantitative or qualitative is important (Weiss, 1998). White (2009) has highlighted the importance of developing theory-based impact evaluations, investigating not only what works or fails but why it does so. Evaluation should be ready to adapt to changing circumstances, apply competitive theories and analyze unintended consequences.

Impact assessment was described by Banhalmi-Zakar et al. (2018) as a tool causing much dissatisfaction lately. Authors debate whether evolution fixing some drawbacks of impact assessment (IA) is enough or revolution replacing IA as an evaluation approach is needed. Shahab, Clinch, and O'Neill (2019) have concluded that evaluation literature mainly focuses on conformance-based approach and evaluation of outputs instead of outcomes. Therefore, a new, more holistic approach called impact-based evaluation was suggested. It merges different properties of both policy evaluation and impact assessment mentioned before. Apart from conformance and performance as basic criteria, authors suggest including efficiency, equity, acceptability, and institutional arrangements (Ibid). To support the following arguments, they provide theoretical backgrounds (from welfare and institutional economics) according to which inclusion of these criteria into impact-based evaluation is crucial.

Assessing the impact of a program is important to demonstrate success to beneficiaries, donors, and audience; to justify spent funds and increase accountability practice; to increase awareness of advantages/disadvantages of a certain tool or practice; to use findings to advocate for further changes in behavior, attitudes, and legislation. Despite the complex nature of evaluation or impact assessment of programs/interventions, the only thing that remains indisputable is their relevance and importance.

1.1.2 Impact studies in education

As it was stated by Whitehead (1959): "Education should begin…and end in research… For its whole aim is the production of active wisdom". According to Vivalt (2015), the upsurge in number of impact evaluations in education was caused by increased interest of developing countries in evidence-based policies. According to Mertens (2014), the evaluation in education started in a postpositivist paradigm: researchers wished to study social world the way they study the natural by using experimentation and measurement which were often decontextualized. As mentioned by Madaus & Kellaghan (2000), one of the evaluation pioneers was Tyler who has developed objectives-based evaluation model in 1949. Its main focus was defining objectives and main activities to meet them, organizing activities and further assessing learning experience. Another scholar, Provus (1969) created discrepancy model for evaluation aimed at comparing actual program performance to the desired standards. Among other evaluations is the theory-based evaluation model of Campbell who proposes to create a theory to guide the program in the progress of solution of social problem (Donaldson, 2007).

One of the ideologists of the constructivist paradigm, Stake (2006) has developed the model of responsive evaluation that involved comparing the outcome of project with the certain standard criteria set by expectations and stakeholders involved in a program. The transformative paradigm from the beginning of the 2000s has proposed several methodologies: the inclusive evaluation paradigm by Mertens (2003) focused on including affected people in methodological decisions. Empowerment evaluation by Fetterman and Wandersman (2007) used evaluation to strengthen improvement and self-determination. Another approach to evaluation was the pragmatic paradigm which extended evaluation definition beyond simply reaching a set goal to providing useful information for decision-making (Stufflebeam, 1983). Some of the advocates were Patton (2008) with utilization-focused evaluation: he stated that the quality of evaluation is defined by the use of its findings. Real World evaluation was proposed by the international development field where evaluators are constrained by money and time. For example, Bamberger (2006) has developed an evaluation design usable in conditions of time constraints and considerate of cultural complexities.

As for the methodology for evaluation studies, Mertens (2014) has claimed that lately the evaluation researchers have been resorting to a pluralistic approach to methodology, merging quantitative and qualitative approaches. The methodologies are abundant so that even meta-analysis impact evaluations in education reach conflicting conclusions mainly due to different methodological approaches or inclusion criteria (Evans & Popova 2015**).** Nevertheless, as stated in the guide on educational practices (U.S. Department of Education, 2003, p.1): "Well-designed and implemented randomized controlled trials are considered "gold standard" for evaluating intervention's effectiveness". Unfortunately, evaluation studies in education do not resort to complex estimation methods since it requires sufficient resources: time, money, people, especially for long-term studies. The main difference between experimental and non-experimental studies is the lack of random assignment to treatment.

To address this issue, researchers use quasi-experimental (QE) studies which mimic randomized, true experiments in experimental structure but lack random assignment (Kirk, 2009). While experimental studies have treatment and control groups where the last one takes part in pre- and post-testing, quasi-experimental ones have simply a comparison group due to practical or ethical conditions (Plonsky, 2017). There are various quasi-experimental designs (QED) such as static-group comparison design assigns treatment to an experimental group and compares performance with control one on a post-test stage. Nonequivalent control group design also adds groups' comparison on a pretest stage. In QED key assumption is the ignorable treatment assignment meaning that treatment should be independent of potential outcomes given specific covariates.

As stated by Gopalan, Rosinger, & Ahn (2020), the most common quasi-experimental research designs for education evaluations are Regression discontinuity (RD), Differences-in-Differences (DID), Fixed-Effects (FE) and Propensity Score Matching (PSM). According to their estimations on the number of articles published in top education research journals in 1995-2018, there were 41 using natural experiments and 101 using QEDs. RD, DID and PSM have started to gain momentum in the education field in 2005-2009 with 9, 8 and 10 articles published correspondingly. After, according to authors' calculations, in 2015-2018 DID, FE, RD were more popular (in 59, 65 and 77 publications) while PSM was used in 35 research papers (Ibid, p. 223). Regression discontinuity design is used in education studies when "treatment eligibility is defined based on a cutoff on a continuous score or index" such as GPA point (Ibid, p.225). In research studies with multiple treatments where the order of treatment is important: there are ordering effects designs such as counterbalanced one in which one treatment is assigned first to one group and another treatment to other group first. Difference-in-Differences (DID) design is gaining popularity in education: it allows to trace the causal effects of policies affecting one group at a point in time while not affecting another group. DID compares pre- and post-treatment periods to find causal effects. According to the authors, in QED two-way fixed effects might be described as an extension of performing DID with larger flexibility for treatment adoption and time periods. QE studies usually measure the Average Treatment Effect on the Treated (ATT) or on the population (ATE).

Matching designs such as Propensity score matching assume the importance of matching two groups based on chosen variables to control for the impact of some extraneous factors. There might be 1:1 matching, weighting and subclassification mechanisms for matching. Though matching methods are used since the middle of the XX century, a theoretical basis for their use started to develop in the 1970s with the contributions of Cochran and Rubin (1973) and Rubin (1973) (Stuart, 2010). The main challenge is to find a perfect matching. In 1983 with the creation of propensity score defined as treatment assignment probability given specific covariates, matching on many covariates became easier (Rosenbaum & Rubin, 1983). It allowed for the matching of the groups even without an exact match on all individual variables.

Impact assessment in education focuses on effect estimation on various outcomes. As noted from numerous meta-analysis of impact studies in education, the most common study outcomes are participation and achievement outcomes (Glewwe, Kremer, Moulin, 2009; Damon, Glewwe, Wisniewski, & Sun, 2016; Masino & Niño-Zarazúa 2016). Enrolment is one of the most commonly researched outcomes: it attracted the interest of researchers due to enormous increase at the end of XX – beginning of XXI centuries: to over 100% net enrolment from 73% in East Asia and Pacific, 54% in the Middle East and North Africa, 56% in South Asia in the 1960s) (Glewwe et al., 2009). Other participation outcomes are competition, attendance, drop-out, and repetition. The achievement (performance) outcomes that are often used as indicators for assessing the quality of education and students' knowledge are test scores, acquisition or improvement of reading or mathematical skills.

Some outcomes might display the long-term effects of educational intervention but they are not often analyzed in literature such as transition rate (to college or higher education), student retention and persistence, educational attainment. The main difference is that retention is related to institutions and persistence to students: institutions retain students within the educational system and students persist (Hagedorn, 2005, p.6). As defined (Ibid, p.4), "retention is staying in school until completion of a degree". According to Spear (2020), retention is about students staying in the education system from one semester/year to another one. Student persistence in the education system has been studied before mostly in high-income countries such as Canada (Parkin, & Baldwin, 2009), Belgium (Vanthournout, Gijbels, Coertjens, Donche, & Van Petegem, 2012), Italy (Checchi, Fiorio, & Leonardi, 2013); one study South Africa (Sampson, 2011).

After observing the types of studies included in the biggest meta-analysis on impact studies in education, it can be noted that there are several types of studies. First, studies of the effect of some program in a particular country for a specific group of people such as the effect of studying in boarding school on achievements in China (Shu, & Tong, 2015) or the effect of providing textbooks in Sierra Leone on test scores (Sabarwal, Evans, & Marshak, 2014). Second, studies of the effect of one specific intervention type, comparison of its application and results in different countries. For example, there is abundant research on the effect of cash transfers on educational outcomes in different countries: García, & Saavedra (2017) have reviewed 94 studies from 47 countries. Third, meta-analyses on educational projects and programs to find the most efficient/cost-effective interventions. For instance, one of the latest available meta-analyses of 114 studies on the effects of various interventions (Damon, et al., 2016). There are rarely attempts to analyze several programs taking place within one country to analyze their impact and efficiency for a specific country.

1.2 Empirical literature review

1.2.1 Evidence from previous impact studies on education interventions

First, interventions that proved most effective will be discussed. Secondly, research papers and their results relevant to this study will be analyzed. Due to the abundance of impact studies, it is more convenient to refer to meta-research on education evaluations to analyze and summarize main outcomes.

Several education programs with proved efficiency and least conflicting evaluation conclusions in research literature will be mentioned. Starting with interventions aimed at increasing enrolment, the meta-study from 2009 has concluded that the "results on ways to increase schooling are remarkably consistent across settings" (Kremer & Holla, 2009, p.21). Providing subsidies or scholarships (Ibid), cash transfers (Glewwe & Muralidharan 2016), information on income differences proportionate to education levels to students (Ganimian & Murnane 2016); building more schools (Glewwe & Muralidhara, 2016) lead to an increase in attendance. The conclusions are supported by a rigorous meta-study of 223 impact evaluations by Ganimian & Murnane (2016): they conclude that reducing costs of direct schooling and complements, informing students and parents on long-term benefits of schooling generally increase enrollment and attendance.

As for effectiveness in improving student achievements, technology-assisted learning, remedial education, tracking or streaming, have proved effective in many studies (Kremer & Holla 2009, McEwan 2015, Masino & Niño-Zarazúa 2016). Also individualizing instructions to match learning needs, offering additional help to struggling students increased achievement (He, Linden & MacLeod

2008, Ganimian & Murnane 2016). Previous results are supported in the study of Evans and Popova (2016) who conducted an analysis of education meta-analysis and concluded that pedagogical interventions that match teaching to students' learning whether computer-assisted or teacher-led have the highest effect. From the supply side, strengthening teacher incentives, hiring teacher on short-term contracts (Kremer & Holla 2009; Muralidharan & Sundararaman, 2010; Bold et al. 2013; Dupas, & Kremer 2015; Ganimian & Murnane 2016), improving teachers' accountability (Glewwe & Muralidharan 2016) have a positive impact. As stated by Evans and Popova (2016), individualized and repeated teachers' training had the second-highest effect.

Moving to the interventions relevant to this study, two types of programs will be discussed: canteens or school feeding programs, supply of school supplements (backpacks, stationery). Most of the school-feeding impact studies investigate their effect on health outcomes (weight, height, nutrition habits). However, those studies that focus on education outcomes reach a consensus regarding the positive effect on participation outcomes (enrolment, attendance). Researches on achievement outcome conclude that there is effect on math scores but not very significant effect for other tests. To facilitate the perception of results, the conclusions are displayed in Table 3.

Table 3. Summary of the impact evaluations for school feeding, canteens construction and food provision programs

Outcome of study	Authors	Detailed results/ conclusions
School participation	Snilstveit et al. (2015); Review of 216 education programs in 52 LMIC	Positive impact: standardized mean difference +0.11 for school participation (enrolment, attendance, completion)
Enrolment	Jomaa, L., McDonnell, E., & Probart, C. (2011); 15 studies in developing countries in primary school level in 1990-2010	Almost all studies show positive effect
	Krishnaratne, S., & White, H. (2013); Overall systemic analysis of education interventions effectiveness in LMIC	Significant positive impact of programs on enrollment
	Gelli, A. (2015); Different feeding modalities in 32 countries of Sub-Saharan Africa	Statistically significant increase with effect size of about 10%: onsite meals more effective in 1st year; after – effective if combined with take-home rations. Higher effects for girls
Attendance	Kristjansson, B., Petticrew, M., MacDonald, B., et al (2007); 18 studies (9 from LA) on effect for disadvantaged students	Program participants' attendance improved (on average, 4-6 more days per year per participant).
	Jomaa, L., McDonnell, E., & Probart, C. (2011);	Almost all studies show positive effect
	Petrosino, A., Morgan, C., Fronius, T., Tanner-Smith, E., Boruch, R. (2012); Meta-analysis comparing 31 different interventions	Positive (mainly short-run) effects. Effect size for deworming was 0.29. Effect size of school meals provision 0.14. 5 largest effect sizes vary from 0.74 to 0.47 (asthma/ epilepsy treatment and new schools building)
	Krishnaratne, S., & White, H. (2013)	Overall, mostly insignificant impact

Achievement Learning	Kristjansson, B., Petticrew, M., MacDonald, B., et al (2007)	Math gains consistently higher for groups receiving meals (Standard mean difference 0.66).
	Jomaa, L., McDonnell, E., & Probart, C. (2011)	Consistent positive effect on arithmetic tests results, lower effects for reading, writing, spelling tests
	Krishnaratne, S., & White, H. (2013)	Overall, mostly insignificant impact
	Snilstveit et al. (2015)	Positive effect: SMD of +0.09 – for learning (cognitive, language, math tests)
	Bashir, S., Lockheed, M., Ninan, E., & Tan, J. (2018); World Bank study of feeding programs in Sub-Saharan Africa	Overall positive impact for improving students' learning though results vary from country to country. Better reading results in Burundi, Chad, Togo; better mathematics – in Burundi, Burkina Faso, Cameroon
Drop-out	Jomaa, L., McDonnell, E., & Probart, C. (2011)	School meals, take-out rations programs reduce the rate (greater benefits to girls)
	Krishnaratne, S., & White, H. (2013)	Significant decrease in dropout

Source: created by author

Moving to the analysis of the pool of impact studies on programs aimed at providing school supplies, it is important to mention that they are quite rare and focus primarily on the provision of textbooks. Main findings are summarized in Table 4. Evidence suggests that provision of textbooks has a positive effect on the test results of children. However, effects depend on the modality of the program: in Kenya, the effect was lower for poorer children since they struggled to use books in English as it was their third language (Glewwe, Kremer, & Moulin, 2009). In Sierra Leone, the program did not have an impact since books were stored at school and not handed out to children (Sabarwal, Evans, & Marshak, 2014).

Table 4. Effect of textbooks provision programs on different education outcomes

Outcome of study	Authors	Detailed results/ conclusions
Achievement Performance Test scores	Heyneman, S., Jamison, D., & Montenegro, X. (1984)	Positive impact in Philippines: higher for impoverished children and for science scores
	Glewwe, P., Kremer, M., Moulin, S., (2009)	Increased only for better-off students: weaker students had troubles understanding books in English (3rd language)
	Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., & Sundararaman, V. (2013)	Unanticipated provision of textbooks in Zambia, India leads to significant improvements in test scores
	Sabarwal, S., Evans, D., & Marshak, A. (2014)	Randomized trial of a public program providing textbooks to primary schools in Sierra Leone had no impact on test scores
Time in school	Glewwe et al. (2009)	Overall increase
Instructional quality	Conn, K. (2014)	Meta-analysis of different interventions in Sub-Saharan Africa: provision of school supplies has low average effect (0.02 standard deviations)

Source: created by author

1.2.2. Previous evaluations of projects in Morocco

The choice of the program was based on the available individual data from the questionnaire ONDH. All these programs are a part of the announced in 2014 National strategy for the support of schooled children and their families with a budget of over 2.1 billion Dirhams. The strategy includes "Tayssir" conditional transfers program, “1 million schoolbags”, program of providing subsidized school meals and improving canteens, providing school transportation as well as boarding schools (including Dar Taliba), (Royaume du Maroc, Ministère de la Culture de la Jeunesse et des Sports, 2014). These interventions are also considered essential according to Vision Strategique de la Reforme 2015-2030, Strategic Vision of Reforms 2015-2030 plan.

One of the strategies to ensure equal access to education and training in Morocco is believed to be State’s and society’s responsibility to “make available the resources likely to facilitate the education and training process; this should concern essentially the level of education and health infrastructure, adequate pedagogical and didactic tools, and integrated reception structures (boarding school, canteens school, school transport, etc.)” (Conseil Supérieur de l'Éducation, de la Formation et de la Recherche Scientifique, p.16). This study analyzes 2 programs in compulsory school (primary and college). It is crucial to understand the essence of programs and previous results of programs’ evaluations to be able to assess their impact. To our best knowledge, there have not been any academic evaluations of these programs, though the state agency for human development ONDH has evaluated them in a certain way.

School canteens program has been a project operated by the Department of National Education since 1997 when the government became responsible for managing and financing school meals program which was before done by WFP (Projet de développement Maroc, 2013). The program is aimed at promoting equal access to school for all Moroccan children, decreasing direct and indirect costs of schooling for disadvantaged families, and removing the obstacles resulting from the remoteness of schools. Main activities are extending the number of school canteens in rural areas, improving school feeding programs, and providing an allowance of 1.4 dirhams per day per child in primary school, 7 and 14 dirhams – in the college canteen and boarding school (Ibid). Main target population are students aged 6-15 from primary In 2007-2008 there were around 5870 canteens (89% of them in rural areas) which raised the proportion of schools with canteens to 28.4% (Ministère d’Education Nationale, de l’Enseignement Supérieur, de la Formation des Cadres et de la Recherche Scientifique, 2008). Around 12,855 schools served meals without having a canteen raising the percentage of schools serving meals to students to 62% (Ibid).

Estimation of the program is rare. From 2009 through 2012, out of planned 1.641 million beneficiaries in primary and 69 000 in secondary school, only 78% and 81% benefitted (Ministère de l’Éducation Nationale, 2018). Gueddari (2016) has mentioned that the absence of canteens at school correlates with the absence of students at school in the afternoon. According to ONDH (2017), benefiting from school canteens in rural areas had a positive though the insignificant effect on students’ performance. In 2020, the service of canteens was the least satisfying education social support program among all the discussed programs in this study (ONDH, 2020).

Another education support program, Initiative Royale “1 Million de Cartables” (One Million School Bags Royal Initiative) was launched in 2008 by King Mohammed VI and has been renewed each year since then (Ministère de l’Economie et des Finances, 2017). The goal is to reinforce mandatory schooling, guarantee equal opportunities, provide support to poor families through the provision of school material (school bags, books, notebooks, and school supplies) to children on yearly basis. It targets primary school children from rural and urban areas and secondary school students from rural areas. It is funded by the Ministry of the Interior through the National Development Initiative Human and the Ministry of Economy and Finance. This program has the broadest coverage in Morocco: in 2018-2019, it targeted 4.365 million students (Medias24, 2018). As estimated (Ministère de l’Éducation Nationale, 2018), in 2009-2012 the program was implemented almost to its full extent: 97% of the planned number of students were reached (3.934 million out of 4.051). As stated in ONDH (2020), thanks to participation in the program, the risk of dropping is 7 times smaller (in terms of odds ratios).

Chapter 2. Research framework and methodology

This chapter is dedicated to an explanation of the research framework. Section 2.1 describes data sources and variables available for analysis. Section 2.2 justifies the choice of research method, chosen policies, and outcomes. Section 2.3 gives an econometric outline for the implementation of the chosen method.

2.1. Data description and source

2.1.1. Data source

The data used in this study comes from National Observatory for Human Development (L'Observatoire National du Développement Humain, ONDH). It was selected as a source because since 2010, ONDH has been conducting regular surveys and collecting individual data for dynamic analysis of human development. ONDH's main goal is to assess the impact of development programs and propose further actions to enhance human development in Morocco. They collect data through surveys; develop indicators and analyze collected data to evaluate impact, adjust policies or propose other measures.

Data is based on the household survey conducted in 2017 (Enquête panel de ménages, vague-2017). The survey is conducted every 2-3 years since being set up in July 2010. There are several parts in the survey: socio-demographic data; individual and household information on education; employment; health status and access to medical service; total spending on alimentary and non-alimentary consumption. The survey and data were translated as they were originally in French: the variables, values and labels were recoded in English. Data was altered after being checked for consistency with the command "*assert*" (ex., presence of school level value in observation if a child is not in school now); for the unique identifier with the command "*isid*"; for duplicates and missing values. Some new variables were generated using the available dataset such as education of household head, household size, number of children and adults in the household, dependency ratio.

2.1.2. Choice of programs and population focus

Education variables include the information on current enrolment and reason for non-enrolment, the last participated school level and year, highest obtained degree, sector of education (public or private), number of grade repetition, language skills. The education section questions have been altered after 2015: this fact is relevant to this study, as in 2017 more data on the different education programs and their beneficiaries became available. In 2012, 2013 and 2015, the only program studied in the survey was the Cash Transfer program – Tayssir. In 2017, there was information on beneficiaries of 6 education programs. I have used the latest available data from the 2017 questionnaire and focused on 2 out of 6 active programs: School canteens program (provision of free meals and construction of canteens in rural areas) and One million schoolbags (distribution of school supplies) program. These programs were chosen as they have the broadest coverage (Annex III) and their impact has never been evaluated before.

Information acquired in the survey: "Are you or have you ever benefited from this program?", so all variables on participation in each of these programs are binary where

1 – yes, benefit currently or have benefitted before;

0 – have never benefited.

As can be noticed from Annex III, the number of observations for each project vary greatly. Being aware of possible spillovers (one individual can participate in several programs at the same time), I control for them [6]. As there might be individuals taking part in both programs at the same time (overlap of beneficiaries), this study has three treatment arms: (1) participation only in One million schoolbags program; (2) participation only in school canteens program; (3) participation in both programs. This study focuses on compulsory education which consists of primary school and college and has a duration of 9 years. Though education is compulsory and free, many students from rural areas fail to attend for different reasons. As it was stated before, students from rural areas face more challenges in education: they tend to drop out and repeat grade more often as well as go to secondary school less frequently than urban area students (Mansouri & Moumine, 2017). As mentioned in the literature review, rural students are the main target of two chosen programs and are the target of evaluation. Focusing only on the rural population of primary and college levels allows seeing a more distinctive impact.

2.2. Research design

2.2.1. Choice of research design

The literature review has investigated possible research designs and methods for the evaluation studies. Out of three research designs most used in education (RD, DID and PSM), the most suitable for this study is Propensity score matching due to peculiarities of the data and chosen programs. Due to treatment assignment specificity in chosen programs in Morocco: there are often no particular thresholds or indexes defining the assignment, regression discontinuity design is not suitable. For example, for the canteen program, the assignment of the treatment usually depends on the area or the location of the school (not the personal characteristics of the student or his/her family). Differences-in-Differences is not possible to perform due to the absence of data for both groups to conduct the pre-test analysis. As the data is not panel, there is no time series to assess groups in the period before the treatment.

The propensity score matching method estimates the probability (propensity score) of an individual to get treatment (program) based on observed characteristics (covariates). After scores are used for the matching of actual beneficiaries (treatment group) with non-beneficiaries (control group): individuals with similar scores (observed characteristics) are matched and results are compared. The matching addresses problem of observing counterfactual: as it is impossible to observe the outcome of the treated individual as if he/she had never received the treatment, it is matched with control group individual who has never received the treatment with the same/similar set of characteristics. Use of PSM

[6] When estimating for every separate program, individuals benefiting from all other programs are excluded from the analysis

originated in the work of Rosenbaum and Rubin (1983) whose aim was to balance nonequivalent groups and reduce bias. In true randomized experiments, individuals have an equal likelihood to be assigned to the group and systematic differences are controlled in the experimental design process (Lane, & Henson, 2010). However, in non-randomized experiments, the probability to be assigned to the group is unequal and unknown. According to Rosenbaum and Rubin (1984), this probability can be assessed using covariates and calculating a probability value – a propensity score. Creating a scalar (a variable with propensity score) that summarizes the information on a set of covariates helps to identify the probability to be assigned to the treatment while balancing groups based on propensity score reduces the bias caused by non-randomization (Fan & Nowell, 2011). Bias is minimized if relevant and appropriate covariates are chosen, treatment and control groups are balanced and located in a common support area (overlap in scores), and the violation of the assumptions is minimized. There are several steps in applying propensity score which will be discussed and specified for this study 2.3. Econometrics design section.

2.2.2. Choice of educational outcomes

The two educational outcomes studied in this program will be repetition (grade retention) and transition to college (second compulsory level of education). These outcomes bring some novelty into the existing pool of studies on impact studies in education as they are rarely studied in impact assessment. In the context of Morocco, repetition and transition to college are highly relevant problems in comparison with other outcomes (such as enrolment) as it was seen in the introduction: in 2019 the repetition rate was still 10% and transition was 92.3% in 2018 (UNESCO, 2020). Repetition rate has been a big issue in Morocco. In the 70s-80s the repetition rate increased the level of new enrolment 30.8% in 1975 32.3% in 1983, at the beginning of XX century one-third of students have repeated a year at least once in their school career cycle with the situation being aggravated in rural areas (Mansouri & Moumine, 2017). Repetition was called the "main symptom of school wastage" (Ibid) and one of its major defaults (Altinok, 2011). Repetition or grade retention rates are used interchangeably.

Repetition rate is the number of repeaters in a given grade/education level in a given school time in percentage of enrolment in that grade the previous school year (UNESCO, 2020). Students are often retained due to unpreparedness for the next level, which is often perceived as a failure and a lack of support from school and teachers (Ruff, 2016). Despite high relevance of the problem for Morocco, the phenomenon of repetition has been studied insufficiently (Latifi, Soulaymani, Ahami, Mokhtari, Aboussaleh, & Rusinek, 2009; Altinok, 2011; Benbiga, Hanchane, & Idir, 2013). The chosen programs are supposed to have an indirect impact on repetition rate as the provision of food in school and the provision of school supplies are aimed at facilitating the acquisition of knowledge which in its turn should reduce repetition (Kremer & Holla, 2009).

Few known studies in Morocco assess the impact on the transition rate (Khandker, Lavy, & Filmer, 1994; Angrist & Lavy, 1997; Khaoula, Taoufik, & Wahbi, 2020; Gazeaud & Ricard, 2021). Transition is the number of new entrants to the level of secondary education in a given year as a percentage of the number of students who graduated from primary education in the previous year (UNESCO, 2004). Transition to secondary education can be considered an indicator of the long-term success of a project as it allows to see how the exposure to a program during primary school has affected the decision of students to continue education further and enroll in secondary education. Program participation in primary school and the potential possibility to benefit from programs in college (as the programs are also available in college) can act as extra motivation to continue compulsory education.

For treatment effect on repetition, a dummy variable *"repetitionbin"* was created: it takes the value of (0) if an individual has never repeated a grade; the value of (1) if an individual has repeated grade at least once in process of studying. Dummy variable *"transit"* to find the effect of transition was created using two other dummy variables: "*college"* for enrollment into college (secondary compulsory school level) and "*primdip"* for the primary school level being the only currently finished education cycle and the only obtained diploma. Variable *"transit"* takes the value of (1) if an individual who has obtained a primary school diploma has enrolled into college (*primdip*=1; *college*=1); and (0) if an individual has obtained primary school diploma not enrolled in college (*primdip*=1; *college*=0). In PSM individuals will be further matched by age to minimize the age-effect difference. Our null hypothesis is that a program has some effect on an outcome (Table 5). The treatment (participation in programs) is expected to have a negative effect (decrease) the repetition (students who benefit from program repeat grade less often than non-beneficiaries) and a positive effect (increase) on transition rate (students who have benefited from program enroll in college more often than those who have never benefitted).

Table 5. Description of the outcomes of study, hypothesis and expected treatment effect

Outcome of the study	Variable in the data	Values	Hypothesis	Expected treatment effect
Repetition rate	*repetitionbin*	Binary: (1) – repeated; (0) – have not repeated	H0: Program has no effect on an outcome H1: Program has some effect on an outcome	Negative
Transition rate to college	*transit*	Binary: (1) – Student enrolled into college after graduation from primary school; (0) – Student did not enroll into college		Positive

Source: created by author

After defining research design, treatment arms (chosen programs), outcomes, amount of observations used for matching for treated and control groups is presented (Table 6). The population sample that this study focuses on are individuals (beneficiaries and non-beneficiaries of programs) who are studying currently and who are participating, have participated before or have never participated in the program. There is a significant number of individuals who are not studying now but have studied before and dropped out or graduated from the cycles of interest (primary and college). It was decided to include a group of these individuals into the estimation of the outcome repetition. This allows to get more precise results and see a better effect as there can be many individuals who have dropped out and are not current students. However, for outcome transition, only those individuals who have obtained primary diploma are considered. It significantly decreases the number of observations, but it is important to ensure the validity of the results. Control groups include only those individuals who have never benefitted from any program to ensure clear effect. For each outcome, there are two different control groups. For outcome repetition, all individuals from primary school and college from rural areas. For outcome transition, a control group is smaller: all individuals who have primary school diploma as the highest obtained diploma and who have or have not entered college at the moment of data collection.

Table 6. Number of available observations per each treatment (program) and number of observations in control group (untreated)

Names of programs	Repetition	Transition
1 Million Schoolbags program	3168	871
Canteens program	370	167
Overlap of 2 programs	758	187
Control group	4443	1914

Source: created by author

2.3 Econometric design

2.3.1. Implementation of PSM

It is important to distinguish research design (Propensity score matching) and the methods to acquire estimates of the treatment effect (found via logit models) so here the main purpose of both is briefly described. PSM is a quasi-experimental method that allows to better select the control group, which is required for evaluating the impact of different treatments in the absence of fully randomized treatment assignment and sampling. PSM creates a control group as similar as possible to the group created in RCT, "the golden standard of impact evaluation" by matching the samples on a vector of observable characteristics. However, PSM may still be biased as there is always a bias-variance trade-off and potential bias due to unobservable covariates. Thus, to add more rigidity, after achieving success in PSM, logit models are run as it allows to find more accurate results after creating a comparable population sample with a good control group. Moreover, it allows to use robust standard errors when evaluating the effects and calculate marginal effects at means to quantify the change in the probability of repetition/transition attributed to participation in the program.

This study has followed an outline of Thoemmes and Kim (2011) (Annex II). PSM is a complex multistep process. First, the main characteristics for matching are selected. Second, based on them, propensity scores for each observation are calculated and used to find the best match. Third, the quality of matching and sensitivity of results to some unobserved covariate is assessed. Fourth, for each treatment arm to observe treatment effects, logit regressions are run. Lastly, for better interpretation of the results, marginal effects are calculated. Further, the steps are described in more details:

1) Covariates selection

Covariates are the factors that influence the likelihood of being selected for treatment. They need to be responsive to initial differences between treated and control groups. Control group should be as similar as possible to the treated group as it allows to closely mimic randomized control trials. Covariates that might potentially influence an examined outcome must be identified based on theory, past studies, treatment characteristics (Tanner-Smith, & Lipsey, 2014). As stated by Lane and Henson (2010), there is no limit in the covariates number so any covariate improving predictability should be included. Stuart (2010) claims that including variables unassociated with treatment has little cost while excluding them is costly in terms of bias so a researcher should be liberal. The main objective for all programs is to improve access to education and fight causes of abandonment, mainly among disadvantaged families and rural areas. There is a good proxy for socio-economic status – average annual expenditure per person in a household (*DAMP*, Dépense annuelle moyenne par personne). Average expenses were also divided into deciles to facilitate matching *(DAMP deciles)*[7]. However, according to Cabrera, Karl, Rodriguez, & Chavez (2018), one proxy is not enough so this study will also use education of household head and dependency ratio as SES proxies. Table 7 provides a list of variables used in the matching. For the effect on the outcome repetition, 14 covariates are used while for transition – 12 covariates.

Table 7. Type and list of covariates used for matching

Type of variables	Covariates
Socio-demographic	Age, gender
Education related	Distance from housing to closest school
Socio-economic variables	Average expenses per capita[8]; average expenses per capita in deciles; household head highest education; dependency ratio
Household characteristics	Longitude and latitude of household location; household size; availability of WC in the house; availability of bathroom; availability of kitchen;
Education levels[9]	Primary school as a last cycle of education; primary diploma obtained

Source: created by author

2) *Propensity score estimation and matching process*

Propensity score (π_i), defined as the probability (*P*) of an individual to be in treatment or control group (*T*) given specific covariates combination (*X*), is calculated as:

$$\pi_i = P(T_i = 1|X_i)$$

In propensity score matching, the scores are defined automatically by Stata for each observation through the probit model and they are saved as a new variable (scalar). Matching and statistical control techniques are aimed at removing "selection bias from causal effect estimates by equating treatment and control units on a sufficient set of measured covariates" identified before and managed with the help of propensity score (Tanner-Smith & Lipsey, 2014). After defining range of scores, common support area (the shared area of propensity scores distribution between two groups) is determined (Lane, To, Shelley, & Henson, 2012). Individuals from treatment group are further matched to control group individuals with the nearest score through chosen matching technique. Matching method choice is abundant but little academic guidance is provided for it. To mention a few techniques: intuitive stratification

[7] Division of average expenses per capita into deciles according to variable distribution in the population
[8] in the process of matching average expenses per capita is preferred over the average expenses in deciles as it gives more precise matching
[9] Added for the outcome repetition

(subclassification) creates subclasses based on score value: 5 subclasses were claimed to be enough to remove around 90% bias (Rosenbaum & Rubin, 1984). Nearest neighbor matching is matching of one unit to another with the closest score; radius matching matched treated unit to control within defined area and kernel method matches units based on the weighted average of all controls (Baser, 2006).

King and Nielsen (2019) have criticized propensity score matching, mainly pair matching without replacement as it involved random pruning (a lot of data and information being cut down). According to Jahn (2017), despite their criticism, PSM performance remains good with matching algorithms such as kernel matching. It matches with the statement of Frolich (2004) that the kernel is more precise than other matching techniques. Kernel is also considered to effective if a group of untreated individuals is bigger as estimates gain more precision (Caliendo & Kopeinig, 2008) which is the case of this study. It is beyond the scope of this paper to analyze pros and cons of matching mechanisms, but so far Kernel matching method is considered to effective and the least biased. Thus, this study is inclined to use of Kernel matching technique, but it will attempt other mechanisms and choose the method which brings the best balance. This study aims to find ATT (average treatment effect on the treated) with Kernel matching and the formula is as follows (Heckmann, Ichimura, & Todd, 1998):

$$ATT = \frac{1}{n_1} \sum_{i \in (T=1)} \left(y_{i1} - \sum_{j \epsilon (T=0)} w(i,j)\, y_{o_j} \right)$$

n_1 – number of individuals participating in program (in treatment group used for matching);

y_i – outcome for a treatment group individual

$w(i,j)$ – weight on each individual from control group (*j*) for specific treated individual (*i*)

y_j – outcome for a control group individual

3) Assessment of matching balance and sensitivity analysis

As the matching is performed based on propensity scores (not on individual variables), it is crucial to check the distribution of variables used for matching in treated and control group. If the matching method creates highly unbalanced groups, a researcher should reject it. There are several techniques, but this study will use Rosenbaum and Rubin's balance assessment (Rosenbaum & Rubin, 1985; Rubin 2001). The indicators are (Rubin, 2001, p. 174):

a) Absolute standardized difference of means of propensity scores in treated and control group (Rubin's B). It should be less than 25% (highest acceptable).

b) Ratio of variances of scores (treated to untreated group) (Rubin's R). Standard values [0.5; 2];

c) For each covariate: (1) standardized difference of means (% bias reduction) and (2) ratio of variance of residuals post-matching (orthogonal to the linear index of propensity score in treatment over control group) (V(T)/V(C)). As Hagen (2016) states, desirable remaining bias should be 3-5% while 20% is considered the largest acceptable. As for variance, standard values are close to 1: [4/5;5/4].

While assessment of matching balance is a necessary step to further proceed to results interpretation, sensitivity analysis is a part of the robustness check procedure. One of the issues with quasi-experimental design is that it allows controlling only for observable characteristics. Sensitivity analysis is a method of testing the sensitivity of the estimations to the presence of "hidden bias" which might raise from excluding unobserved covariate. The goal of the analysis is to check results' robustness by testing if the conditional independence assumption[10] is violated. There is no direct way to test the existence of unobserved covariates (Becker & Caliendo, 2007), but the bounding approach of Rosenbaum (2002) permits to check how strong unobserved covariate has to influence selection process to question the robustness of the found treatment estimations. In other words, the probability of the estimated results to be undermined/altered by unobserved covariate(s).

4) *Results estimation and interpretation*

After creating a comparable sample (treated and control group) and achieving the balance, I proceed with defining treatment effects. Propensity score matching only allows to see the means of the outcome in the treated and control group. Logit models[11] are run for each treatment arm for the following reasons. First, it allows incorporating the results of the Kernel matching procedure by adding the respective weights to the observations (based on probability to receive the treatment). Second, the main purpose of logit after matching is to be sure that found estimation is the causality and not simply correlation. One can be certain in this only if a counterfactual outcome is approximated well enough by the selected control group. PSM allows to improve this control group selection by ensuring that observations match on the vector of observable characteristics and the averages in the groups are not statistically significantly different from each other. Thus, the found effect can be attributed to the program, and not just treated as a contribution program adds to the overall change (the latter would be the case of logit without matching). Third, it allows calculating marginal effect (not just the difference between means of treated and control groups) to quantify the found treatment effect accounting for the clustering and adding robustness to the whole procedure. Fourth, the distribution of the effect size between gender and education level (for outcome repetition) can also be seen by running logit with interaction terms (*treatment#gender; treatment#primary*).

2.3.2. Key assumptions

There are two key assumptions that PSM relies upon:

- Strong ignorability (conditional independence; unconfoundedness assumption): the treatment assignment is independent of the potential outcome conditional on observable characteristics (covariates) (Rosenbaum & Rubin, 1983). If treatment assignment is strongly ignorable given covariates, it is strongly ignorable given any balancing score (Ibid, p.43). Some unobserved covariates might be correlated with observed ones so matching for observable characteristics implies a certain degree of matching for unobserved ones depending on the level of correlation (Stuart, 2010). The only

[10] Explained in detail in Section 2.3.2
[11] Not linear regression as outcomes are binary

concern is unobserved covariates unrelated to observed ones. The problem might be partially addressed by careful covariates selection and assessment of results sensitivity to the existence of unobserved covariates performed;

- Stable Unit Treatment Value Assumption (Rubin, 1980). The treatment assignment of one individual does not affect another individual's outcome. It means that the response of an individual depends only on treatment assigned to him/her: other individuals and their treatment assignment have no effect. This assumption might not hold if there is an interaction between experimental and control groups possibly leading to "spillover" effects.

2.3.3. Stata implications

"Psmatch2" package was installed in Stata. *Psmatch2* calculates propensity scores for each observation by probit regressions and shows the number of observations in both groups. Option "*common*" runs PSM with common support region. Matching mechanisms are applied with "*kernel*" – for Kernel matching. *"Pstest", "psgraph", "psmatch density"* check matching quality and show it graphically. "*Mhbounds"* is used for sensitivity analysis. "*Logit"* with *[pweight=_weight]* (weights being propensity scores) is used run to obtain results. *"Margins"* are further used to quantify the effect.

Chapter 3. Results and policy implications

This chapter analyzes matching performance and discusses estimated treatment effects. Section 3.1 provides descriptive statistics of the studied population sample. In Section 3.2 the balance between treated and control groups is assessed. Section 3.3 estimates the overall quality of matching performance and conducts a robustness check and sensitivity analysis. I proceed with an interpretation of treatment effects obtained after running logit models for each treatment on the matched population sample in Section 3.4. It also makes use of calculated marginal effects to interpret results in a clear ordinary way. Lastly, limitations and results are analyzed in Section 3.5 and Section 3.6 correspondingly.

3.1. Descriptive statistics

This section observes trends of statistics, repetition and transition rate for a population sample used in this study. The sample is individuals (beneficiaries and non-beneficiaries of programs) who are studying currently, have studied before and dropped out or graduated from the cycles of interest (primary and college). There is a significant number of program beneficiaries who are not studying now: this group is limited to 24 years old as after this age the number of beneficiaries is small, so it is hard to match. In our sample, almost 70% are primary school students. Only 5 individuals are working while currently studying (the youngest working student is 12 years old). However, the number might be higher as students from rural areas usually work helping parents in agricultural activities households are in the first five deciles of distribution of average expenses per capita. Only around 15% of students have its household head with obtained education level. Though the number of adults or children in the household can reach as high as 22 or 13 individuals, the average is 4.2 and 2.4 correspondingly. As for basic needs, availability of the kitchen is almost ubiquitous among households of students from rural areas though 5.9% (around 500 individuals) have no access to WC and 44.7% (more than 3000 individuals) – no access to bathroom or shower in their accommodation.

Table 8. Descriptive statistics for primary and college students from rural areas[12]

Variable	Variable name	Values	Average (quant.); % (qual.)
Socio-demographic indicators			
Age	*age*	5-24	13.87
Gender	*gender*	Male (1)	52.27%
Marital status	*martstatbin*	Single (1)	80.89%[13]
Work status	*worknow*	Working now (1)	24.16%[14]
Education related indicators			
Last cycle enrolled	*primary*	Primary (1)	69.66%
Primary school diploma obtained	*primdip*	Primary school diploma obtained (1)	36.36%
School type	*schooltype*	Public (1)	98.91%

[12] Categorization of the variables is done according to our estimations though some variables might fit into different category or two categories simultaneously

[13] Most married individuals are not studying now, only 1 individual who is currently studying is married now

[14] Most of them work after dropping out or graduating, only 5 individuals are working while studying at the same time

Distance from housing to the closest school (m)	*schooldist*	0-50000 m	1342.87 m
Distance from house to the closest school (range)[15]	*schdistrange*	0-199 m (1)	18.16%
		200-500 m (2)	28.37%
		550-999 m (3)	26.20%
		1-1.99 km (4)	14.12%
		2-2.99 km (5)	5.59%
		3-4.5 km (6)	2.68%
		5-9.99 km (7)	3.31%
		10-29.99 km (8)	0.8%
		30-50 km (9)	0.36%
Socio-economic proxies			
Average annual expenditure per person in a household (DAMP)	*DAMP*	2196.297 – 98 666.96 Dirham	11 390.83[16]
Deciles of average annual expenditure per person in a household [17]	*DAMPdec*	1st quintile	22.88%
		2nd	18.65%
		3rd	15.97%
		4th	11.76%
		5th	10.02%
		6th	6.67%
		7th	5.47%
		8th	4.50%
		9th	3.04%
		10th	1.04%
Dependency ratio[18]	*dep_ratio*	0.083 – 5	0.76
Highest obtained diploma of household head [19]	*hh_educ*	No diploma (0)	84.72%
		Primary school (1)	9.94%
		Middle school (College) (2)	3.19%
		Upper-secondary (lyceum) (3)	1.02%
		Higher education (4)	1.02%
Household characteristics			
Geocode of household location (longitude and latitude)[20]	*longitude; latitude*		
Household size	*hhsize*	2-35	6.62
Number of adults in household [21]	*asize*	1-13, 15, 22	4.17
Number of children in household [22]	*csize*	1-8, 10, 13	2.41
Household head gender	*femhead*	Female (1)	7.27%
Availability of the kitchen in housing	*acckitch*	Have access (1)	98.07%
Availability of bathroom in housing	*accbath*	Have access (1)	55.33%
Availability of WC in housing	*accWC*	Have access (1)	94.13%
Number of observations			9637

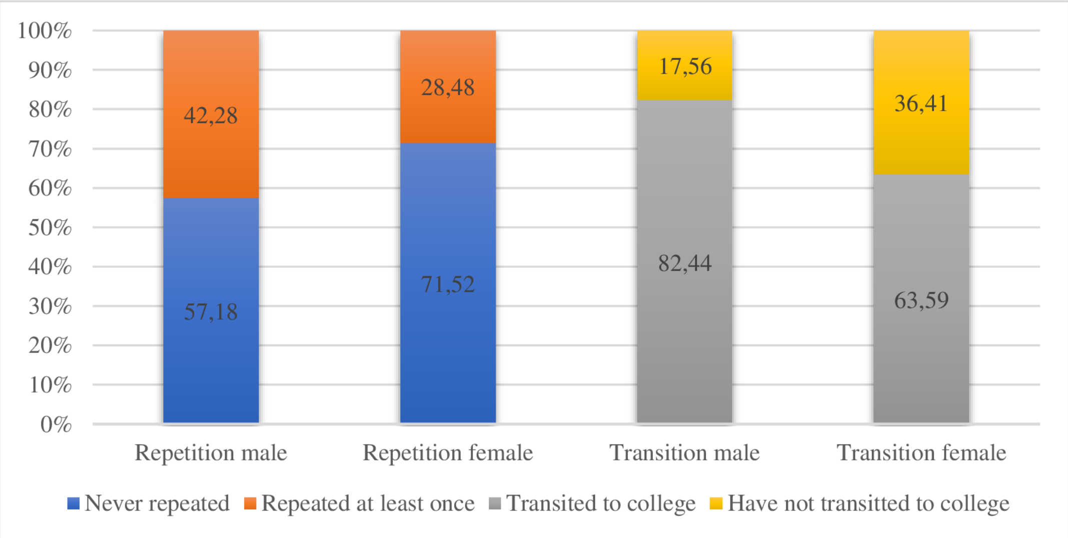
Source: created by author

When looking at the difference by gender (Figure 5 and Annex IV), male students tend to repeat a grade more often: 42.28% of boys have repeated a grade at least once while this figure is 28.48% among females. These results are similar to the official statistics of UNESCO (2020): in 2019, repetition for males was 2 percentage points higher though the difference is even more noticeable in the rural

[15] Created by separating variable distance to the closest school into several logical ranges
[16] Around 2580 euros as of May 15, 2021
[17] Division of variable average expenses per capita into ten equal parts and percentage of households in each decile
[18] Created by dividing number of children (less than 15 years old) over number of adults (of and over 15 years old)
[19] Created by extract "parent" and "child" using the family relationship "family head" and household id variables and matching each child and further assigning highest diploma obtained to each mother
[20] Not reported due to data confidentiality
[21] Created by using household ID and counting all individuals of and older than 15 years old according to UN definition of youth (age 15-24) for statistical reasons (Angel, 2015) and normally 15-64 is considered working age population (OECD)
[22] Created by using household ID and counting all individuals younger than 15 years old according to UN definition

population sample. Gender disparities are also evident in transition rate. Girls tend to transit to college after primary school less often than boys: 63.59% of girls and 82.44% of boys from rural areas have transited. This can be explained by the conclusion of Bouoiyour and Miftah (2015) that in rural Morocco parents are inclined to invest more in sons rather than daughters' education. Gender disparities are considered and controlled for when interpreting the results. Looking at outcomes distribution among deciles of average expenses (Annex IV), the repetition rate surprisingly increases with increase in deciles: in 1st, 2nd and 3rd deciles it is 33.25%, 32.71% and 32.74% while the highest rate is in 8th, 9th and 10th deciles (42.77%, 44.28% and 50.40%). As for the transition, it is the lowest in 1st and 10th quintiles (only 69.35% and 68.33% of individuals correspondingly transit to college) and the highest is 76.9% and 78.6% (2nd and 4th quintiles). There seems to be no distinct trend in repetition or transition rates that correlates with socioeconomic status.

Figure 5. Gender disparities for the chosen outcomes: repetition and transition



Source: created by author

3.2. Common support and matching balance diagnosis

Before proceeding to results interpretation, common support region and matching balance for each treatment arm need to be assessed. I have tried matching with different mechanisms but the best balance was achieved with Kernel matching (which coincides with our choice of matching method in Section 2.3.1.). Annex VII shows that for each treatment effect estimation, there is sufficient overlap of common support regions between treated and control groups given the chosen set of covariates. Only a few observations are disregarded from treated groups (maximum is 17 individuals), which, according to Bryson, Dorsett and Purdon (2002) does not appear to be a hurdle for further analysis.

Table 9 summarizes the numerical assessment of balance based on Rubin (2001) (Section 2.3.1), Annex V and Annex VI give statistical and graphical balance. Matching is always a trade-off between variance and bias. Overall, balance for each matching for each treatment arm was achieved. Standardized differences of means for overall matching is below highest possible 25%, only for treatment Canteens program it is quite high: 25% for repetition outcome and 18.4% - for transition (Annex V.B; Annex VI.B). Ratios of variance for all matching are within range. Bias was significantly reduced for all covariates in all programs after matching (on average 70-80% bias reduction). There were some issues with matching on individual covariates. For example, for canteens program and both programs school distance range (*schdistrange*) variable had a variance of 0.00 though the after-matching bias was only 0.2%. The remaining percentage bias for each covariate in all treatment arms was lower than the highest possible 20%, in most cases lower than 10% with several exceptions such as latitude (11%), access to WC (17.7%) for repetition (Annex V.B) and access to kitchen (10.6%) for transition for canteen program (Annex VI.B).

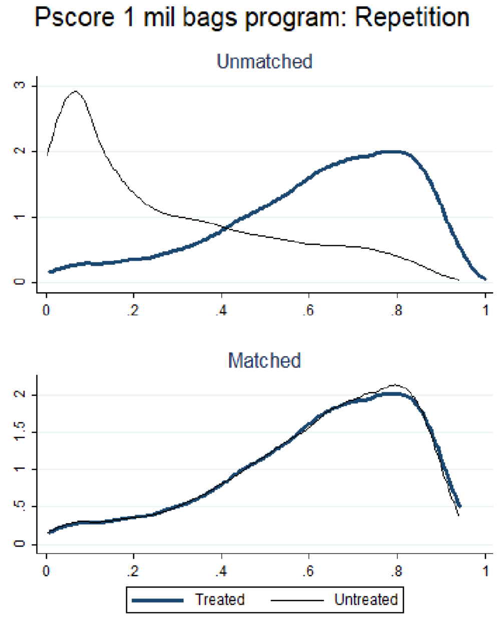
Table 9. The assessment of matching and balance between Treated and Control groups

	B	R	V(T)/V(C) and % bias for each covariate[23]	Annex
Standard value	<25%	[0.5; 2]	[4/5; 5/4]	
		Repetition		
One million schoolbags	12.6	0.83	All[24]	Annex V.A
Canteens	25.0	0.99	All	Annex V.B.
Both programs	11.7	1.27	All	Annex V.C
		Transition to college		
One million schoolbags	9.3	1.12	All	Annex VI.A
Canteens	18.4	0.98	All	Annex V.B.
Both programs	12.4	1.00	All	Annex VI.C

Source: created by author
B - Standardized difference of means of propensity score; R – ratio of variances of scores; V(T)/V(C) – ratio of variance of residuals

Graphically in histograms of the propensity score distribution, the scores for treated individuals are higher than for control ones so they have a higher probability to receive treatment. As seen in Figure 6 (example for One million bags program and variable repetition), propensity scores were matched well, in this line graph individuals off common support area can be observed (other graphs in Annex VIII). As the balanced was achieved and assessed as successful, further step is to proceed with results interpretation.

Figure 6. Distribution of propensity scores for treatment and control groups pre and post matching



[23] If some covariate is out of standard range of V(T)/V(C), we will look at percentage of remaining bias (%) after matching which should not be larger than the highest acceptable 20%
[24] Marked "All" if all covariates are within range of standard values V(T)/V(C) or/and percentage of remaining bias (%) is acceptable

3.3. Propensity score matching: Treatment effect and robustness check

Table 10; Annex IX show estimated means of outcomes in treated and control groups found in the performance of PSM matching. Though it is not their main results, it can be mentioned that only two estimations were statistically significant though the results are the opposite of the expected ones. The mean of repetition rate for beneficiaries of One million schoolbags program and both programs is higher than the mean of repetition rate for non-beneficiaries: the difference in percentage points is 4.68 and 9.46 correspondingly. The mean for repetition rate for beneficiaries of Canteens program is also bigger than for non-beneficiaries (by 6.97 ppts) but this result is statistically non-significant. The other three remaining estimations for transition rate are non-significant though mean of transition for beneficiaries of programs is higher than their peers in control groups. The non-significance might be due to the small sample size. The difference in means in pre- and post-matching stages is an illustrative example of the importance of matching: estimations have substantially changed after matching. It epitomizes the significance of matching prior to treatment effect estimation.

Table 10. Estimated average treatment effect for each program and outcome (Annex IX)

Program	Var name	Sam.[25]	Treated group mean	Control group mean	Diff. (ppts)	St.error	T-stat[26]
Outcome: repetition rate *(repetitionbin)*							
One million schoolbags	*bagsbin*	U	25%	41.87%	-16.87	.010911	-15.46***
		M	25.07%	20.39%	4.68	.015755	2.97***
Canteens programs	*cantbin*	U	42.43%	41.86%	0.56	.026703	0.21
		M	42.55 %	41.95%	0.59	.026926	0.22
Both programs	*cantbags*	U	27.18%	41.18%	-14.69	.019125	-7.68***
		M	27.26%	17.79%	9.46	.02306	4.10***
Outcome: transition rate *(transit)*							
One million schoolbags	*bagsbin*	U	76.80%	74.59%	2.21	.018448	1.20
		M	77.40%	77.14%	0.26	.022806	0.12
Canteens programs	*cantbin*	U	74.26%	75.10%	-0.84	.044810	-0.19
		M	74.26%	73.73%	0.52	.046163	0.11
Both programs	*cantbags*	U	79.68%	74.50%	5.17	.033188	1.56
		M	80.11%	77.05%	3.05	.034837	0.88

Source: created by the author

It is a standard procedure to assess matching and formed groups by bootstrapping standard errors (Annex XX) and sensitivity analysis (Annex XI). Bootstrapping is a resampling data procedure to re-estimate propensity scores and common support regions to approximate standard errors, confidence intervals and p-values. Though Abadie and Imbens (2008) mentioned that there is no formal justification for use of bootstrapping, it is widely applied in treatment effect estimations. In this study, after bootstrapping, standard errors decreased and coefficients have mainly increased (5 coefficients out of 6), intensifying the already found effect (Annex XX). This robustness check gives us some degree of confidence in matching validity.

[25] Sample: U for unmatched: coefficients before matching; M for matched: coefficients after matching
[26] ***; $p<0.01$

Sensitivity analysis is aimed at estimating matching and covariates set for the possible presence of unobserved covariate. It is done with "*mhbounds*" command using Mantel and Haenszel test-statistics based on gamma (γ) which is odds of differential assignment due to unobserved covariate, namely effect of the unobserved covariate ($\gamma = 1$, absence of unobserved selection bias). With $\gamma > 1$, bounds start to move apart and there are two scenarios: Q^{+}_{MH} (treatment effect was overestimated; positive unobserved selection – values going downwards) and Q^{-}_{MH} (underestimated effect; negative unobserved selection – values going downwards); and their significance levels p^{+}_{MH} and p^{-}_{MH}. As it is impossible to assess the existence of unobservables, the main idea is to measure the hypothetical value of gamma which will be required to undermine found matching and coefficients. If results become statistically not significant, even with small values of gamma (looking at p^{+}_{MH} or p^{-}_{MH}), the estimated treatment effect might be sensitive to the presence of unobserved covariate, thus, results should be treated with caution.

Sensitivity analysis results are presented in Annex XI. Looking at results of statistically significant treatment estimations, for One million schoolbags program and outcome repetition, there is a negative unobserved selection, so the assumption that results were underestimated is observed. Estimated results become statistically non-significant at 1.15 level of gamma so results might be slightly sensitive to the presence of unobserved covariate and need to be treated with caution. For results for the beneficiaries of both programs on outcome repetition, the observed estimates are stronger to the presence of unobserved covariate. There is a positive selection bias and results become statistically non-significant at a larger value of gamma (1.3); thus, found estimates are not very sensitive to the possible unobserved covariate. As for results that are not statistically significant, their sensitivity analyses show that they are sensitive to the presence of unobserved covariate.

3.4. Results assessment and interpretation

As it was seen, PSM is a bias-variance trade-off and some estimations might be sensitive to unobserved covariates so to add more rigidity, I run logit models to obtain results on a matched population sample. As logit model just shows the sign of the effect, average marginal effects are calculated to quantify the effect. As mentioned in descriptive statistics, there is a need to control for the effect of gender so this study also runs logit models with interaction terms of treatment and gender, treatment and school level (only for outcome repetition) to observe the distribution of effect. Running logit models after performing PSM allows attributing the effect to the program and not to other observable factors. Table 11 summarize the logit model results for each treatment and outcome. Logit models with interactions are presented in Annex XII and XIII. All logit models are statistically significant. All treatments had a positive effect on outcomes: participation in programs increased repetition and transition rates. There are only two statistically significant (at 1% level) treatments: participation in the One million bags program and in both programs simultaneously increased the repetition rate. As mentioned above, logit models facilitate results interpretation as they allow for the calculation of margins to quantify the estimated effects.

Table 11. Logistic regression results: impact of treatment on outcome repetition rate and transition to college						
	(1)	(2)	(3)	(4)	(5)	(6)
	Onemilbags	Canteenprog	Bothprog	Onemilbags	Canteenprog	Bothprog
VARIABLES	repetitionbin	repetitionbin	repetitionbin	transit	transit	transit
bagsbin	0.312***			0.0509		
	(0.107)			(0.230)		
cantbin		0.0458			0.0232	
		(0.140)			(0.274)	
cantbags			0.736***			0.217
			(0.145)			(0.344)
age	0.193***	0.0827***	0.255***	0.0107	-0.128***	-0.0527
	(0.0222)	(0.00856)	(0.0217)	(0.0374)	(0.0239)	(0.0372)
gender	0.400***	0.451***	0.501***	1.075***	1.098***	1.488***
	(0.0828)	(0.153)	(0.112)	(0.135)	(0.252)	(0.235)
DAMPdec	-0.0213	0.0357**	-0.0119	0.0458	-0.0449	0.0108
	(0.0212)	(0.0176)	(0.0268)	(0.0469)	(0.0382)	(0.0461)
dep_ratio	-0.194*	-0.415***	-0.206	-0.226	-0.369	-0.240
	(0.104)	(0.0964)	(0.140)	(0.178)	(0.305)	(0.196)
hhsize	-0.0335	-0.0269	-0.0619***	-0.00517	-0.0337	-0.0549
	(0.0209)	(0.0190)	(0.0218)	(0.0252)	(0.0326)	(0.0341)
hh_educ	-0.238***	0.00947	-0.693***	0.275**	0.0928	0.387
	(0.0624)	(0.0876)	(0.107)	(0.123)	(0.113)	(0.283)
schooldist	2.17e-05*			-3.34e-05**		3.49e-05
	(1.24e-05)			(1.46e-05)		(3.97e-05)
accbath	-0.0960	-0.445**	-0.315	-0.207	-0.0600	-0.415**
	(0.131)	(0.194)	(0.214)	(0.215)	(0.381)	(0.200)
acckitch	0.696**	-0.0403	0.0665	0.667	-0.0420	0.232
	(0.293)	(0.552)	(0.299)	(0.599)	(0.623)	(0.747)
accWC	-0.250	-0.0291	-0.0287	0.345	-0.0982	0.679*
	(0.175)	(0.216)	(0.140)	(0.512)	(0.417)	(0.355)
longitude	1.10e-05***	4.13e-06	3.34e-06	-1.87e-05**	-4.37e-05***	-1.67e-05
	(3.12e-06)	(4.25e-06)	(4.60e-06)	(9.11e-06)	(1.40e-05)	(1.04e-05)
latitude	-2.06e-05***	-8.05e-06	2.04e-06	-3.52e-05***	2.02e-05	-2.28e-05
	(5.65e-06)	(6.87e-06)	(7.68e-06)	(1.17e-05)	(1.41e-05)	(1.52e-05)
primary	-0.432***	-1.281***	-1.154***			
	(0.137)	(0.121)	(0.186)			
primdip	-0.329**	-0.516***	-1.364***			
	(0.153)	(0.133)	(0.252)			
region	-0.172***	0.00278	0.0468	-0.0181	0.506***	0.0519
	(0.0509)	(0.0473)	(0.0691)	(0.0805)	(0.121)	(0.104)
worknow	0.0172	-0.00341	-0.290*	-0.294**	-0.118	-0.378**
	(0.0774)	(0.0753)	(0.158)	(0.117)	(0.104)	(0.173)
martstatbin	-1.265***	-0.566***	-1.832***	-0.899***	-0.251	-0.936**
	(0.233)	(0.127)	(0.346)	(0.310)	(0.257)	(0.366)
schdistrange		0.00375	4.67e-05**		0.0421	
		(0.0482)	(2.30e-05)		(0.0825)	
Constant	3.387	1.844	-3.847	12.62***	-2.647	9.551*
	(2.065)	(2.467)	(2.909)	(4.261)	(4.747)	(5.322)
Observations	7,207	4,688	5,010	2,489	1,456	2,093

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Average marginal effects for two treatments that were found statistically significant are presented in Figure 7 and Figure 8, while other results are in Annex XIV though all main findings are reported here. Figure 7 shows marginal effect calculation to assess the effect of participation in One million schoolbags program on repetition rate. Participation in the program in primary school and college increases the probability to repeat the grade by 4.84 percentage points (significant at 1%). When variable participation in the program changes from 0 (baseline, not participating in the program) to 1 (participation in the program), it increases the probability of males to repeat a grade by 6.76 ppts (1% significance level) and of females – by 2.77 ppts (non-significant). Being a participant of the program and studying in primary school increases the probability to repeat a grade by 5.9 ppts (1% significance level) while being in college – by 2.57 ppts (non-significant). To conclude, participation in One million schoolbags program significantly increases the probability to repeat the grade: the effect is higher and statistically significant for male students and primary school level students.

Figure 7. Average marginal effects of participation in One million schoolbags program on outcome repetition, distribution of effect between gender and cycle of education

```
Average marginal effects                    Number of obs     =      7,207
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin	.0484278	.0167511	2.89	0.004	.0155962	.0812594

```
Average marginal effects                    Number of obs     =      7,207
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.bagsbin						
gender						
Female	.0276922	.0232797	1.19	0.234	-.0179351	.0733195
Male	.0675574	.0182576	3.70	0.000	.0317732	.1033416

```
Average marginal effects                    Number of obs     =      7,207
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.bagsbin						
primary						
0	.0257459	.0194787	1.32	0.186	-.0124317	.0639235
1	.0591532	.0191204	3.09	0.002	.0216779	.0966284

As for the impact of participation in both programs (One million school bags and canteen programs) on outcome repetition, effects are similar to One million schoolbags program but stronger (Figure 8). Participation in both programs increases the probability to repeat grade overall by 11.19 ppts; for male beneficiaries – by 13.66 ppts while for female – by 8.66 ppts; only for primary school students – by 11.13 ppts while for college students – by 10.87 ppts. All coefficients are significant at 1% significance level. Participation in both programs increases the probability of repetition with the effect being higher for males and for primary school students but significant for all estimations.

Figure 8. Average marginal effects of participation in both programs on outcome repetition, distribution of effect between gender and cycle of education

```
Average marginal effects                        Number of obs     =      5,010
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : cantbags

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    cantbags |    .111888   .0212857     5.26   0.000     .0701687    .1536072
------------------------------------------------------------------------------

Average marginal effects                        Number of obs     =      5,010
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.cantbags

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1.cantbags   |
      gender |
     Female  |   .0866314   .0204721     4.23   0.000     .0465068     .126756
       Male  |   .1366115   .0329789     4.14   0.000      .071974    .2012489
------------------------------------------------------------------------------

Average marginal effects                        Number of obs     =      5,010
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.cantbags

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1.cantbags   |
     primary |
          0  |   .1086774   .0297975     3.65   0.000     .0502753    .1670794
          1  |   .1113417   .0220819     5.04   0.000      .068062    .1546215
```

Almost none of the coefficients (except one) for the effect of the canteen program participation on repetition are statistically significant, but reporting the direction of effect is useful. Participation in program in 2017 tended to increase the probability of repeating a grade but slightly (0.9 ppts) (Annex XIV.B.a); while for males beneficiaries probability increases by 2.6 ppts, for females, it decreased probability of repeating (by 0.58 ppts, significant at 10%) (Annex XIV.B.b). Being in primary school decreases the probability of repetition by almost 3 ppts while being in college increases by 7.9 ppts (Annex XIV.B.c). As we have seen, for all programs the gender effect on the outcome repetition coincides with our findings from data analysis and official statistics of UNESCO (2020): the probability to repeat the grade for male students is higher than for female students. As for the difference by school level, the probability to repeat a grade for primary school students is higher than for college students.

As for the effect of treatments on transition rate, all estimated marginal effects are non-significant except one. The participation in programs increases the probability to transit to college: for beneficiaries of One million bags program – by 0.81 ppts (Annex XIV.D.a); for Canteens program – 0.4 ppts (Annex XIV.E.a); for both programs – by 3.16 percentage points (Annex XIV.F.a). The gender effect corresponds to our implication about the data and official statistics (UNESCO, 2020): girls tend to transit to college less frequently. The participation in programs increases the probability to transit for males and decreases for females: for One million bags beneficiaries – by 2.9 ppts and by 1.7 ppts (Annex XIV.D.b); for canteens program – increase by 1.19 ppts and decrease by 0.46 ppts (Annex XIV.E.b); for both programs beneficiaries – by 9.4 ppts (significant at 1%) and by 4.19 ppts (Annex XIV.F.b). The only statistically significant result is that the Canteen program increases the probability of male students to transit to college by 9 ppts.

In conclusion, the effect of treatment on repetition rate is more significant than on transition which might be due to smaller sample size when estimating effects on transition. Gender effect (despite some estimations being statistically non-significant) coincides with data inference and official statistics for both outcomes which makes results more valid and robust. In conclusion, participation in programs increases the probability to repeat the grade for beneficiaries of One million schoolbags and for beneficiaries of both programs, so for these outcomes, we can reject the null hypotheses of no effect. There is an effect but it is the opposite of the anticipated one: the participation in programs increases the probability of repetition, not decreases it. The effect is stronger for male students and primary school level students. For the other four treatment assessments, we cannot reject the null hypothesis of no effect as estimations are not statistically significant.

3.5. Limitations

There are several limitations regarding chosen research framework and its implication; results interpretation, data and outcomes. As for propensity score matching, firstly, it is criticized for sensibility to the choice of covariates: it was observed when balancing on some covariates such as longitude and latitude. They are extremely important for matching as they allow to match households and students from similar areas but achieving balance on them was challenging. Secondly, much criticism is related to its ability to account only for observed covariates. There is no way to control for unobserved covariates or detect their existence, but this study has tried to use the best way to address this issue – sensitivity analysis. The analysis is not always easy to interpret and in our case was useful only for two out of six treatments. One was found slightly sensitive, another – rather not sensitive to the presence of unobserved covariates.

Thirdly, there is little guidance on matching methods choice and how this choice influences the estimation of the results, which I tried to address in this study by choosing both the least criticized by academia and the most balanced (after controlling for balance between treated and control groups) method – Kernel matching. As it has been seen from robustness check, the limited number of observations (in case of participation in Canteens and both programs for the outcome transition), lack of consensus on the use of robustness check method (such as bootstrapping) does not always allow to collect compelling evidence for highly robust and statistically significant results, for example in case of Canteens program. Moreover, a small sample size for the outcome transition also poses certain limitations. Though the estimated effects were as expected, it cannot be concluded that programs had a positive effect on increasing transition as estimates were statistically non-significant.

There are some limitations regarding available data: more data on socio-economic status could be of great value. There is no data on school level: school characteristics, quality of teaching, student-teacher ratios, the efficiency of usage of funds provided to schools for program implementation. This data will ensure a better understanding of the situation and higher results robustness. Nevertheless, available data and set of covariates have allowed to generate balanced matching and draw compelling conclusions at least for two estimations. As for the outcomes, their limitations add strength and novelty to the study. There is still insufficient research on programs' impact on repetition and transition, especially in Morocco. Though there is no opportunity to compare with previous estimations made in Morocco, our results bring new evidence in the existing pool of studies on education intervention evaluation in this country.

3.6. Results discussion and policy implications

Propensity score matching and control for key characteristics have allowed reducing bias related to the lack of randomized selection in quasi-experimental study design. Found estimations were checked for robustness several times through bootstrapping of standard errors, sensitivity analysis. Two out of six treatment effects (benefitting from school supplies provision program and from meals and school supplies provision simultaneously on repetition rate) were constantly significant with the same direction of impact. Lack of significance for the other four treatments (canteen program effect on repetition; all effects estimations on transition) is likely to be related to small sample size. Among four non-significant treatments, there was the statistical significance of participation in both programs for male students: it increased the probability to transit to lower secondary education by 9%. Males generally transit to college more often than females but benefiting from two programs in primary school seems to give male students the necessary support and enough motivation to continue further education in college.

The correlation between participation in the program and repetition rate is much more noteworthy as, according to our estimations, program beneficiaries tended to repeat grades more often than non-beneficiaries with similar individual characteristics. On the one hand, results may seem counter-intuitive though robustness tests gives us a high degree of confidence in results validity. On the other hand, results are not novel: a similar treatment effect was found in Kenya where the provision of textbooks failed to decrease repetition rate but increased the probability of enrolment into secondary school (Glewwe, Kremer, Moulin, 2009). I can only assume why participation in programs led to an increase in repetition and further research is needed to explain this phenomenon. To begin with, students who receive school supplies at the beginning of school year and daily meals at school are expected to enroll and attend school as well as comprehend material better (Kremer & Holla, 2009). Nevertheless, awareness of the possibility to benefit from programs next year (even if a student repeats the grade) might undermine incentives to acquire material to the fullest and exert maximum effort to succeed at schooling and avoid repetition. In a certain way, reassurance of being able to benefit from the program(s) might decrease motivation to study at your maximum and might decrease the fear of repeating the grade.

Moreover, this reassurance might disrupt students' desire to attend school so repetition can be a result of poor school attendance (Mims, Stock, & Phinizy, 2001). If low attendance is an issue, adding conditionality on attendance as program component can be advised. Additionally, the possibility to benefit from program even if grade is repeated might contribute to parents' desire to leave children at home to help with family business or agricultural work. This explains higher effect (higher repetition rates) for male program beneficiaries as they are generally more involved in agricultural work than girls (International Labor Organization, 2013). In this context, increasing parents' literacy is a great option as this contributes to a better understanding of education importance and higher level of involvement in child's education. It is easy to explain higher treatment effect on repetition for primary school students: students tend to enroll in lower secondary education more consciously so studying more deliberately.

Besides, there can be some explanations on the supply-side: schools that offer programs and school administration. Quality of teaching instructions might be an issue. All programs from the 2017 questionnaire and currently in effect in Morocco are aimed at increasing schooling and decreasing existing educational inequalities. Assessed programs exist for a decade or more in the case of Canteens program so it might be the time to diversify efforts and resources and start investing in improving teaching practice. Firstly, one of the best evidence-based mechanisms is teacher training (Bowman, 2005). To our best knowledge, teachers' training was one of the educational reforms at the beginning of the XXI century but the participation rate was low due to the absence of any financial incentives and the need of a teacher to combine several workplaces at the same time due to low salaries as school (Chtatou, 2015). Secondly, the use of digital sources in the classroom or individual instruction for students experiencing troubles with a certain topic or subject can be advisable. As the digitalization of school might be challenging for rural areas, having some computers at school and using them for the needs of those who are falling behind can be a great advantage and step forward in improving performance (Masino & Niño-Zarazúa, 2016). Thirdly, considering the negative social-emotional impact of retention, social promotions (keeping students with their age group) through after-school or summer programs can be a great alternative to grade retention (Lincove & Painter, 2006). Finally, increase in repetition might be caused by inefficiency and irrationality of funds allocation provided for program implementation. To address this potential cause, a system of randomized school audits can be proposed to observe the efficiency of funds allocation and possible capture of funds by local elites (Reinikka & Svensson, 2004).

It is crucial to stress the need to adapt education in rural areas to economic, social and cultural realities. As children often help families with agricultural work, schedule and school breaks should be flexible. Knowledge and abilities taught in the rural areas should be altered to meet the demand of the area. For example, Mabrouk (2019) suggested 3 hours school day, curricula focused on practical skills, and recruitment of teachers from local youth.

Conclusion

In recent decades international policymakers have put a lot of effort in providing equal access to education and in improving its quality. As the demand for evidence-based policies increased, researchers got actively involved in the impact assessment of education projects. The goal of this research was to estimate the impact of public policies aimed at supporting education in Morocco. After conducting a literature review on previous evaluations (programs and outcomes of studies), potential methods and research design, propensity score matching was chosen as a research framework and logit models with interaction terms to estimate treatment effect. Two programs with the best coverage were selected for this study: program of school supplies provision "One million schoolbags" and program of subsidized food provision at school "Canteens program". The enrolment and attendance rates have been increasing constantly in Morocco, so this study has focused on more problematic outcomes: repetition rate and transition to college. The significance of this work lies in the fact that it is a pioneer not only in programs' evaluation on the chosen outcomes but in the overall assessment of these programs in Morocco. Study focused only on rural area students as they are the main target of the chosen programs as well as they tend to be more deprived of access to education due to historical and social reasons.

This research has highlighted the importance of matching and controlling for key observed characteristics when assessing the treatment effect. As we have seen from our results, some of the treatment effect estimations have varied significantly before and after matching so without using propensity score matching, the results could have been misleading. The application of quasi-experimental design and Propensity score matching has allowed us to find treatment effects that are attributed solely to the program and not only the effects that programs contribute to. It was observed that benefitting from One million schoolbags and two programs (One million schoolbags and canteen programs) has increased the probability to repeat the grade in compulsory school. The probability was higher for males and primary school students. Benefitting from the Canteens program had increased the probability of male students to transit to college. There was a noticeable effect of gender: for boys, the probability to repeat a grade as well as the probability to transfer to college after primary school was higher for all treatments (though significant only for two treatments). Gender effect is presented in pre-matching estimations of our data and in official UNESCO sources which increases the degree of confidence in the robustness of results found in this study.

Several potential shortcomings need to be considered. Firstly, there is still little guidance on choice of the matching mechanism while performing PSM. Secondly, in PSM there is an assumption of conditional independence (treatment is independent of the potential outcome conditional on observable characteristics). It is a strong assumption as there might be unobservable characteristics that were not included in process of matching. We have tried to address this limitation by assessing the sensitivity of our results to the presence of unobserved covariate. Two out of six estimations that were found significant, results of one estimation (impact of one million bags program on increase in repetition rate) were found moderately sensitive to the presence of unobserved covariate while another estimation was not very sensitive (impact of participation in both programs on increase in grade repetition). Nevertheless, sensitivity analysis does not prove the presence of unobserved covariate or reduce the validity of results, it just suggests treating some estimations with caution. There was a small sample size for estimations on the outcome transition to college which could be the reason for the statistical non-significance of found effect.

Our findings add to the growing literature pool on the impact evaluation of education public policies in education. We have provided evidence that even well-designed studies aimed at increasing schooling might cause some "side effects" such as an increase in repetition rate. There are only some grounds to assume the reasons for such an effect. For example, the ability to participate in a program in the future might decrease fear of repetition; distort motivation to exert maximum efforts in performance; reduce students' desire to attend school properly or parents' willingness to send children to school. Among potential policy implications are making programs conditional on attendance, increasing instruction efficiency by providing teachers with professional development courses, adding more flexibility to school organization, schedule, and curriculum in rural areas.

In the future, it might be of interest to further assess programs' effects using panel data. The questionnaire and individual data used in this study were of 2017 while the survey is conducted every two years. Panel data will allow to compare results, observe potential changes and make more comprehensive conclusions regarding the effectiveness of the programs. Adding more outcomes to programs' assessments will also allow for more holistic conclusions. Adding qualitative methods of analysis such as interviews of participants will enable both a deeper understanding of programs, their strengths and shortcomings and a more comprehensive analysis of potential reasons for the found unexpected effect of programs.

References

Abadie, A., & Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators. *Econometrica*, *76*(6), 1537-1557.

Abdessamad, H. (2020, October 20). *Maroc : Budget supplémentaire pour la santé et l'éducation et réduction du déficit en 2021*. Anadolu Agency. https://www.aa.com.tr/fr/afrique/maroc-budget-suppl%C3%A9mentaire-pour-la-sant%C3%A9-et-l-%C3%A9ducation-et-r%C3%A9duction-du-d%C3%A9ficit-en-2021-/2012456

Aikens, R. C., Greaves, D., & Baiocchi, M. (2020). A pilot design for observational studies: Using abundant data thoughtfully. *Statistics in Medicine*, *39*(30), 4821-4840

Akresh, R., De Walque, D., & Kazianga, H. (2013). *Cash transfers and child schooling: evidence from a randomized evaluation of the role of conditionality*. The World Bank.

Altinok, N. (2011). Analyse de la performance des acquis scolaires du Maroc à travers l'enquête TIMSS. *Cahiers de l'éducation et de la formation*, *4*, 35-44.

Andersson, C., & Johansson, P. (2013). Social stratification and out-of-school learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(3), 679-701.

Angel, W. A. (2015). Resolutions on Youth Rights Adopted by the General Assembly and Economic and Social Council of the United Nations: 1965–1993. In *The International Law of Youth Rights* (pp. 139-290). Brill Nijhoff.

Angrist, J. D., & Lavy, V. (1997). The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco. *Journal of Labor Economics*, *15*(1, Part 2), S48-S76.

Assembly, U. G. (1948). Universal declaration of human rights. *UN General Assembly*, *302*(2), 14-25.

Baird, S., Ferreira, F. H., Özler, B., & Woolcock, M. (2014). Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, *6*(1), 1-43.

Baird, S., McIntosh, C., & Özler, B. (2011). Cash or condition? Evidence from a cash transfer experiment. *The Quarterly journal of economics*, *126*(4), 1709-1753.

Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation*. Thousand Oaks, CA: *Sage*

Banhalmi-Zakar, Z., Gronow, C., Wilkinson, L., Jenkins, B., Pope, J., Squires, G., ... & Womersley, J. (2018). Evolution or revolution: where next for impact assessment? *Impact Assessment and Project Appraisal*, *36*(6), 506-515.

Barrera-Osorio, F., Bertrand, M., Linden, L. L., & Perez-Calle, F. (2008). *Conditional cash transfers in education: design features, peer and sibling effects evidence from a randomized experiment in Colombia*. The World Bank.

Baser, O. (2006). Too much ado about propensity score models? Comparing methods of propensity score matching. *Value in Health*, *9*(6), 377-385.

Bashir, S., Lockheed, M., Ninan, E., & Tan, J. P. (2018). *Facing forward: Schooling for learning in Africa*. The World Bank.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., & Pellerano, L. (2016). Cash transfers: what does the evidence say. *A rigorous review of programme impact and the role of design and implementation features. London: ODI*, *1*(7).

Becker, S. O., & Caliendo, M. (2007). Sensitivity analysis for average treatment effects. *The stata journal*, *7*(1), 71-83.

Behaghel, L., De Chaisemartin, C., & Gurgand, M. (2017). Ready for boarding? The effects of a boarding school for disadvantaged students. *American Economic Journal: Applied Economics*, *9*(1), 140-64.

Benadad, H. (2019, December 2). Cantines scolaires: L'etat met le paquet. *Le 360*. https://fr.le360.ma/societe/cantines-scolaires-letat-met-le-paquet-203740

Benbiga, A., Hanchane, S., & Idir, N. (2013). L'évaluation des acquis scolaires au Maroc: nouvelles approches. *Critique économique*, (30).

Benhassine, N., Devoto, F., Duflo, E., Dupas, P., & Pouliquen, V. (2015). Turning a shove into a nudge? A" labeled cash transfer" for education. *American Economic Journal: Economic Policy*, *7*(3), 86-125.

Blanden, J., & Gregg, P. (2004). Family income and educational attainment: a review of approaches and evidence for Britain. *Oxford review of economic policy*, *20*(2), 245-263.

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2013). Scaling up what works: Experimental evidence on external validity in Kenyan education. *Center for Global Development Working Paper*, (321).

Boulahya Z. (2018, September 20). Programme Tayssir: le detail des nouvelles mesures de la rentrée 2018-2019. *Medias 24.* https://www.medias24.com/MAROC/EDUCATION/185971-Programme-Tayssir-Le-detail-des-nouvelles-mesures-de-la-rentree-2018-2019.html

Bouoiyour, J., & Miftah, A. (2015). Migration, remittances and educational levels of household members left behind: Evidence from rural Morocco. *The European journal of comparative economics*, *12*(1), 21.

Bowman, L. J. (2005). Grade retention: Is it a help or hindrance to student academic success?. *Preventing School Failure: Alternative Education for Children and Youth*, *49*(3), 42-46.

Brisset C. (2018). Les enfants et les jeunes au Maroc: risques sociaux, réponses publique et pistes de réformes. Les premieres Assises Nationales de la Protection Sociale [Presentation]. https://www.unicef.org/morocco/media/1196/file/Les%20enfants%20et%20les%20jeunes%20au%20Maroc%20:%20risques%20sociaux,%20réponses%20publiques%20et%20pistes%20de%20réformes.pdf

Bryson, A., Dorsett, R., & Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies.

Bundy, D., Silva, N. D., Horton, S., Jamison, D. T., Patton, G. C., Schultz, L., ... & Filippi, V. (2018). Re-imagining school feeding: a high-return investment in human capital and local economies.

Cabrera, J. C., Karl, S. R., Rodriguez, M. C., & Chavez, C. (2018). Investigating Socioeconomic Status Proxies: Is One Proxy Enough?

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

CGIAR. (2013). *Monitoring, Evaluation, and Impact Assessment Strategy for CRP6 (2012 – 2016)* (Rep.). Bogor Barat, Indonesia: CIFOR.

Checchi, D., Fiorio, C. V., & Leonardi, M. (2013). Intergenerational persistence of educational attainment in Italy. *Economics letters*, *118*(1), 229-232.

Chtatou, M. (2015). *A Moroccan success story tainted with some shortcomings*. EFA Global Monitoring Report 2015, Education for All 2000–2015: achievements and challenges.

Chtatou, M. (2015). A Moroccan success story tainted with some shortcoming. *Paper Commissioned for The EFA Global Monitoring Report*, 2000-2015.

Cochran, W. G., & Rubin, D. B. (1973). Controlling bias in observational studies: A review. *Sankhyā: The Indian Journal of Statistics, Series A*, 417-446.

Conn, K. M. (2014). *Identifying Effective Education Interventions in Sub-Saharan Africa: A meta-analysis of rigorous impact evaluations* (Doctoral dissertation, Columbia University).

Conseil Supérieur de l'Éducation, de la Formation et de la Recherche Scientifique. (n.d.). *Pour une Ecole d'Équité, de la Qualité et de la Promotion: Vision Strategique de la Reforme 2015-2030* (National Education Plan). Royaume du Maroc - Ministère de l'Education Nationale, de la Formation Professionnelle, de l'Enseignement Supérieur & de la Recherche Scientifique. http://planipolis.iiep.unesco.org/en/2016/projets-de-la-vision-stratégique-2015-2030-6224

Curto, V. E., & Fryer Jr, R. G. (2014). The potential of urban boarding schools for the poor: Evidence from SEED. *Journal of Labor Economics*, *32*(1), 65-93.

Damon, A., Glewwe, P., Wisniewski, S., & Sun, B. (2016). *Education in Developing Countries-what Policies and Programmes Affect Learning and Time in School?*. Expertgruppen för biståndsanalys (EBA).

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., & Sundararaman, V. (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, *5*(2), 29-57.

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. Routledge.

Duflo, E., Dupas, P., & Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of public Economics*, *123*, 92-110.

Evans, D. K., & Popova, A. (2015). *What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews*. (Policy Research Working Paper No. 7203). The World Bank. Washington, DC.

Evans, D. K., & Popova, A. (2016). What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. *The World Bank Research Observer*, *31*(2), 242-270.

Evans, D., Kremer, M., & Ngatia, M. (2008). *The impact of distributing school uniforms on children's education in Kenya*. World Bank.

Faludi, A. (1989). Conformance vs. performance: Implications for evaluation. *Impact Assessment*, *7*(2-3), 135-151.

Fan, X., & Nowell, D. L. (2011). Using propensity score matching in educational research. *Gifted Child Quarterly*, *55*(1), 74-79.

Fetterman, D., & Wandersman, A. (2007). Empowerment evaluation: Yesterday, today, and tomorrow. *American Journal of Evaluation*, *28*(2), 179-198.

Fiszbein, A., & Schady, N. R. (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Publications.

Fitz-Gibbon, C.T. (1996) *Monitoring education: indicators, quality and effectiveness* London: Cassell

Foliano, F., Green, F., & Sartarelli, M. (2019). Away from home, better at school. The case of a British boarding school. *Economics of Education Review*, *73*, 101911.

Ganimian, A. J., & Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, *86*(3), 719-755.

García, S., & Saavedra, J. E. (2017). Educational impacts and cost-effectiveness of conditional cash transfer programs in developing countries: A meta-analysis. *Review of Educational Research*, *87*(5), 921-965.

Gasper, D. (2000). Evaluating the 'logical framework approach' towards learning-oriented development evaluation. *Public administration and development*, *20*(1), 17-28.

Gazeaud J., & Ricard C. (2020). "Present!" et maintenant? Tayssir et performance scolaires. *L'Education en questions*, 1, 82-92.

Gazeaud J., Ricard C. (2021). "Conditional Cash Transfers and the Learning Crisis: Evidence from Tayssir Scale-up in Morocco". *Études et Documents,* 8, CERDI.

Gelli, A. (2015). School feeding and girls' enrollment: the effects of alternative implementation modalities in low-income settings in sub-Saharan Africa. *Frontiers in public health*, *3*, 76.

Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education*, *2*, 945-1017.

Glewwe, P., & Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education* (Vol. 5, pp. 653-743). Elsevier.

Glewwe, P., Kremer, M., Moulin, S. (2009). Many children left behind? Textbooks and test scores in Kenya. Am. Econ. J. Appl. Econ. 1 (1), 112–135.

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: growth, promise, and challenges. *Review of Research in Education*, *44*(1), 218-243.

Gregg, M. T. (2018). The long-term effects of American Indian boarding schools. *Journal of Development Economics*, *130*, 17-32.

Gueddari, K. (2016). L'abandon scolaire en milieu rural marocain: une analyse interactionniste du point de vue des familles. Université de Montréal

Hagedorn, L. S. (2005). How to define retention. *College student retention formula for student success*, 90-105.

Hagen, T. (2016). Econometric evaluation of a placement coaching program for recipients of disability insurance benefits in Switzerland. *Frankfurt Research Institute for Business and Law Working Paper*, (10).

He, F., Linden, L. L., & MacLeod, M. (2008). How to teach English in India: Testing the relative productivity of instruction methods within the Pratham English language education program. *Working paper*.

Heyneman, S., Jamison, D., & Montenegro, X. (1984). Textbooks in the Philippines: Evaluation of the pedagogical impact of a nationwide investment. *Educational Evaluation and Policy Analysis*, *6*(2), 139-150.

Ifa, A., & Guetat, I. (2018). Does public expenditure on education promote Tunisian and Moroccan GDP per capita? ARDL approach. *The Journal of Finance and Data Science*, *4*(4), 234-246.

Ikira M. (2021). ÉVALUATION DE LA POLITIQUE EN MATIÈRE D'ÉDUCATION ET INTERFÉRENCE DES CARACTÉRISTIQUES SOCIOECONOMIQUES : Cas du programme Tayssir [Dissertation]. Laboratoire de Modélisation Appliquée à l'Économie et à la Gestion (MAEGE)

International Labor Organization. (2012, January 31). *Gender and child labour in agriculture*. https://www.ilo.org/ipec/areas/Agriculture/WCMS_172261/lang--en/index.htm#:%7E:text=Agriculture%20is%20still%20a%20significant,versus%2052.6%25%20for%20girls).

Internats: Amzazi dévoile les chiffres. (2020, January 15). *L'economiste*. https://www.leconomiste.com/flash-infos/internats-amzazi-devoile-les-chiffres

Jahn, B. (2017). *Kernel matching with automatic bandwidth selection* [Slides]. Stata.Com. https://www.stata.com/meeting/uk17/slides/uk17_Jann.pdf

Jomaa, L. H., McDonnell, E., & Probart, C. (2011). School feeding programs in developing countries: impacts on children's health and educational outcomes. *Nutrition reviews*, *69*(2), 83-98.

Khandker, S., Lavy, V., & Filmer, D. (1994). Schooling and cognitive achievements of children in Morocco. *World Bank Discussion Paper*, *264*.

Khandker, S., Lavy, V., & Filmer, D. (1994). Schooling and cognitive achievements of children in Morocco. *World Bank Discussion Paper*, *264*.

Khaoula, E., Taoufik, H., & Wahbi, B. E. (2020). Conceptual Evolution of the Sequences and Its Transition from High School to University in Morocco. *International Journal of Higher Education*, *9*(6), 26-33.

King, G., & Nielsen, R. A. (2019). Why propensity scores should not be used for matching. Political Analysis 27, 4 (May 2019): 435-454.

Kirk, R. E. (2012). Experimental design. *Handbook of Psychology, Second Edition*, *2*.

Kremer, M., & Holla, A. (2009). Improving education in the developing world: what have we learned from randomized evaluations?. *Annu. Rev. Econ.*, *1*(1), 513-542.

Krishnaratne, S., & White, H. (2013). *Quality education for all children? What works in education in developing countries* (No. 0000-0). International Initiative for Impact Evaluation (3ie).

Kristjansson, B., Petticrew, M., MacDonald, B., Krasevec, J., Janzen, L., Greenhalgh, T., ... & Welch, V. (2007). School feeding for improving the physical and psychosocial health of disadvantaged students. *Cochrane database of systematic reviews*, (1).

Lamlili N. (2016, April 07). Maroc: l'école des femmes. Jeune Afrique. https://www.jeuneafrique.com/mag/313526/societe/maroc-lecole-femmes/

Lane, F. C., & Henson, R. K. (2010). Using Propensity Scores in Quasi-Experimental Designs to Equate Groups. *Online Submission*.

Lane, F., To, Y., Shelley, K., & Henson, R. (2012). An illustrative example of propensity score matching with education research. *Career and Technical Education Research*, *37*(3), 187-212.

Latifi, M., Soulaymani, A., Ahami, A. O. T., Mokhtari, A., Aboussaleh, Y., & Rusinek, S. (2009). Comparaison des performances cognitives chez les adolescents consanguins et les non consanguins de la région nord Ouest marocain. *Antropo*, *19*, 57-65.

Laurian, L., Day, M., Berke, P., Ericksen, N., Backhurst, M., Crawford, J., & Dixon, J. (2004). Evaluating plan implementation: A conformance-based methodology. *Journal of the American Planning Association*, *70*(4), 471-480.

Le Siteinfo. (2021, March 09). Tayssir: le nombre de jeunes filles bénéficiaires a plus que triplé. https://www.lesiteinfo.com/maroc/tayssir-le-nombre-de-jeunes-filles-beneficiaires-a-plus-que-triple/

Lincove, J. A., & Painter, G. (2006). Does the age that children start kindergarten matter? Evidence of long-term educational and social outcomes. *Educational Evaluation and Policy Analysis*, *28*(2), 153-179.

Liouaeddine, M., Bijou, M., & Naji, F. (2017). The main determinants of Moroccan students' Outcomes. *American Journal of Educational Research*, *5*(4), 367-383.
Llorent Bedmar, V. (2015). Dysfunction and educational reform in Morocco. *Asian Social Science, 11 (1), 91-96.*
Mabrouk, E. (2019). Problems and strategy of educational development in Morocco. *New Trends and Issues Proceedings on Humanities and Social Sciences*, *6*(1), 430-439.
Macours, K., Millan, T. M., Barham, T., Maluccio, J., & Stampini, M. (2019). Long-term Impacts of Conditional Cash Transfers: review of the evidence. *World Bank Research Observer*.
Madaus, D. L. S. G. F., & Kellaghan, T. (2000). *Evaluation models: Viewpoints on educational and human services evaluation* (Vol. 49). Springer Science & Business Media.
Manley, J., Fernald, L., & Gertler, P. (2015). Wealthy, healthy and wise: does money compensate for being born into difficult conditions?. *Applied Economics Letters*, *22*(2), 121-126.
Mansouri, Z., & Moumine, M. E. A. (2017). Primary and Secondary Education in Morocco: From Access to School into Generalization to Dropout. *International Journal of Evaluation and Research in Education*, *6*(1), 9-16.
Marouane I. (2020). Pauvreté des enfants: Un rôle pour le programme Tayssir au Maroc? *L'Education en questions*, 1, 72-81.
Martin, A. J., Burns, E. C., Kennett, R., Pearson, J., & Munro-Smith, V. (2021). Boarding and Day School Students: A Large-Scale Multilevel Investigation of Academic Outcomes Among Students and Classrooms. *Frontiers in Psychology*, *11*, 3730.
Masino, S., & Niño-Zarazúa, M. (2016). What works to improve the quality of student learning in developing countries?. *International Journal of Educational Development*, *48*, 53-65.
McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, *85*(3), 353-394.
Medias24 (2018, September 14). Plus de 4,3 millions de beneficiares de l'initiative royale "Un million de cartables". https://www.medias24.com/MAROC/Quoi-de-neuf/185773-Plus-de-43-millions-de-beneficiaires-de-l-initiative-royale-un-million-de-cartables.html
Mertens, D. M. (2003). The inclusive view of evaluation: Visions for the new millennium. *Evaluating social programs and problems: Visions for the new millennium*, 91-107.
Mertens, D. M. (2014). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. Sage publications.
Mims, K., Stock, R., & Phinizy, C. (2001). Beyond grade retention. *eJournal of education policy*.
Ministère d'Education Nationale et de la Formation Professionnelle. (2014, June 5). Evaluation du programme Tayssir: Transferts monétaires conditionnels. Rabat: Ministère d'Education Nationale et de la Formation Professionnelle. http://www.ondh.ma/sites/default/files/1_tayssir_morocco_2014.pdf
Ministère d'Education Nationale, de l'Enseignement Supérieur, de la Formation des Cadres et de la Recherche Scientifique. (2008). Pour un nouveau souffle de la réforme de l'Education-Formation. [http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/morocco_programme_urgence_najah_rapport_detaille_version_projet.pdf
Ministère de l'Economie et des Finances.(2017). Projet de Loi de Finances pour l'année budgétaire 2017. Rapport Economique et Financier. Rabat: Ministère de l'Economie et des Finances [Report]. https://www.finances.gov.ma/Docs/DB/2017/ref_fr.pdf
Ministère de l'Éducation Nationale. (2018). *Rapport relatif à l'évaluation du programme d'urgence*. http://www.courdescomptes.ma/upload/MoDUle_3/File_3_613.pdf
Mullis, I., Martin, M., Foy, P., & Hooper, M. (2015). *TIMSS 2015 International Results in Mathematics*. http://timssandpirls.bc.edu/timss2015/international-results/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Mathematics.pdf
Muskin, J. A., Kamime, A., & Adlaoui, A. (2011). Empowered to Empower: A civil society–government partnership to increase girls' junior secondary school outcomes in Morocco. *Research in Comparative and International Education*, *6*(1), 129-146.
Newey W. (2007). Course materials for 14.386 New Econometric Methods. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. https://ocw.mit.edu/courses/economics/14-386-new-econometric-methods-spring-2007/readings/treatment_effect.pdf

O'Flynn, M. (2010). *Impact Assessment: Understanding and assessing our contributions to change* (Rep. No. M&E Paper 7). Oxford, UK: INTRAC.

Observatoire National du Développement Humain.(2017). *Enquete sur les Indicateurs de Prestation de Services en Éducation (IPSE) au Maroc* [Report]. Rabat: Complexe Administratif et Culturel de la Fondation Mohammed VI de Promotion des Oeuvre Sociales de l'Education-Formation. https://www.ondh.ma/sites/default/files/documents/rapport_ipse_vf.pdf

OECD. (2021). *Employment - Employment rate by age group - OECD Data*. (2021). OECD. https://data.oecd.org/emp/employment-rate-by-age-group.htm

ONDH. (2020). *Étude de l'impact de l'appui social sur la scolarisation en milieu rural 2018. Résultats fondamentaux.* http://www.ondh.ma/sites/default/files/documents/present_appui_social_20_fevrier_2020_ver3.pdf

Otchet, A. (2020, June 2). *One in every five children, adolescents and youth is out of school worldwide*. UNESCO. https://en.unesco.org/news/one-every-five-children-adolescents-and-youth-out-school-worldwide#:%7E:text=According%20to%20data%20from%20the,over%20the%20past%20five%20years

Parkin, A., & Baldwin, N. (2009). Persistence in post-secondary education. *The price of knowledge: Access and student finance in Canada*, 65-84.

Patton, M. Q. (2008). Utilization-Focused Evaluation. Thousand Oaks. *A: Sage*.

Petrosino, A., Morgan, C., Fronius, T. A., Tanner-Smith, E. E., & Boruch, R. F. (2012). Interventions in developing nations for improving primary and secondary school enrollment of children: A systematic review. *Campbell Systematic Reviews*, *8*(1), i-192.

Pfeifer, C., & Cornelißen, T. (2010). The impact of participation in sports on educational attainment—New evidence from Germany. *Economics of education review*, *29*(1), 94-103.

Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. *Studies in Second Language Acquisition*, *39*(3), 579-592.

Projet de développement maroc 200494. (2013). https://one.wfp.org/operations/current_operations/project_docs/200494.pdf

Reinikka, R., & Svensson, J. (2004). Local capture: evidence from a central government transfer program in Uganda. *The quarterly journal of economics*, *119*(2), 679-705.

Roche, C. J. (1999). *Impact assessment for development agencies: Learning to value change*. Oxfam.

Rogers, P. J. (2009). Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation. *Journal of development effectiveness*, *1*(3), 217-226.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association*, *79*(387), 516-524.

Roser M., & Ortiz-Ospina E. (2015). Primary and Secondary Education. https://ourworldindata.org/primary-and-secondary-education

Royaume du Maroc, Ministère de la Culture de la Jeunesse et des Sports, Département de la Communication. (2014, September 11). HM the King officially launches 2014-2015 school year and "One Million Schoolbags" operation. https://www.maroc.ma/en/royal-activities/hm-king-officially-launches-2014-2015-school-year-and-one-million-schoolbags

Royaume du Maroc, Ministère de la Culture de la Jeunesse et des Sports, Département de la Communication. (2020, December 10). Programme "Tayssir": Le paiement des bourses de digitalise. https://www.maroc.ma/fr/actualites/programme-tayssir-le-paiement-des-bourses-se-digitalise

Rubin, D. B. (1973). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics*, 185-203

Rubin, D. B. (1980). Bias reduction using Mahalanobis-metric matching. *Biometrics*, 293-298.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3), 169-188.

Ruff, R. R. (2016). The Impacts of Retention, Expenditures, and Class Size on Primary School Completion in Sub-Saharan Africa: A Cross-National Analysis. *International Journal of Education Policy and Leadership*, *11*(8), n8.

Saavedra, J. E., & García, S. (2012). Impacts of conditional cash transfer programs on educational outcomes in developing countries: a meta-analysis. *RAND Labor and Population Working Paper Series, WR-921-1*.

Sabarwal, S., Evans, D. K., & Marshak, A. (2014). *The permanent input hypothesis: the case of textbooks and (no) student learning in Sierra Leone*. The World Bank.

Sabatier, P. (1991). Political Science and Public Policy. *PS: Political Science and Politics, 24*(2), 144-147. doi:10.2307/419922

Saidi, A., Hamadi, A., & Elkartouti, S. (2020). Évaluation D'impact du Programme D'appui Social à la Scolarisation «TAYSSIR». *Revue de recherche en Droit, Économie et Gestion*.

Sampson, L. G. (2011). *Student persistence in higher education: A study of the challenges and achievements of a group of historically disadvantaged senior students studying at the University of the Western Cape* (Doctoral dissertation, Stellenbosch: University of Stellenbosch).

Saoudi, K., Chroqui, R., & Okar, C. (2019). Student Achievement in Moroccan Student Achievement in Moroccan Educational Reforms: A Significant Gap Between Aspired Outcomes and Current Practices. *Interchange*, 1-20.

Scriven, M. (1991). Beyond formative and summative evaluation. *Teachers College Record*, *92*(6), 18-64.

Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation practice*, *17*(2), 151-161.

Shahab, S., Clinch, J. P., & O'Neill, E. (2019). Impact-based planning evaluation: Advancing normative criteria for policy analysis. *Environment and Planning B: Urban Analytics and City Science*, *46*(3), 534-550.

Shu, B., & Tong, Y. (2015, April). Boarding at school and students' well-being: The case of rural China. In *Population Association of America 2015 Annual Meeting* (Vol. 30).

Snilstveit, B., Gallagher, E., Phillips, D., Vojtkova, M., Eyers, J., Skaldiou, D., ... & Davies, P. (2017). PROTOCOL: Interventions for improving learning outcomes and access to education in low-and middle-income countries: a systematic review. *Campbell Systematic Reviews*, *13*(1), 1-82.

Spear, E. (2020, March 25). *Persistence vs. Retention: Definitions & Improvement Tips*. Precision Campus. https://precisioncampus.com/blog/persistence-vs-retention/

Stake, R. E. (2013). *Multiple case study analysis*. Guilford press.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*(1), 1.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*(1), 1.

Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In *Evaluation models* (pp. 117-141). Springer, Dordrecht.

Tanner-Smith, E. E., & Lipsey, M. W. (2014). Identifying baseline covariates for use in propensity scores: A novel approach illustrated for a nonrandomized study of recovery high schools. *Peabody journal of education*, *89*(2), 183-196.

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate behavioral research*, *46*(1), 90-118.

Towne, L., & Shavelson, R. J. (2002). *Scientific research in education*. National Academy Press Publications Sales Office.

UIS. (2020a). *Algeria | UNESCO UIS*. UNESCO Institute for Statistics. http://uis.unesco.org/en/country/dz

UIS. (2020b). *Egypt | UNESCO UIS*. UNESCO Institute for Statistics. http://uis.unesco.org/en/country/eg

U.S. Department of Education. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. https://www2.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf

UNESCO. (2020). *Morocco | UNESCO UIS*. UNESCO UIS. http://uis.unesco.org/en/country/ma

UNESCO. (2021a). *Education: From disruption to recovery*. https://en.unesco.org/covid19/educationresponse#durationschoolclosures

UNESCO. (2021b, March 27). *One year into COVID-19 education disruption: Where do we stand?* https://en.unesco.org/news/one-year-covid-19-education-disruption-where-do-we-stand

UNESCO. 2004. Education for All: The Quality Imperative: EFA Global Monitoring Report, 2005

Vanthournout, G., Gijbels, D., Coertjens, L., Donche, V., & Van Petegem, P. (2012). Students' persistence and academic success in a first-year professional bachelor program: The influence of students' learning strategies and academic motivation. *Education Research International*, *2012*.

Vedung, E. (1997). Public Policy and Program Evaluation. New Jersey: Transaction Publishers.

Villa, J. M. (2014). The length of exposure to antipoverty transfer programmes: what is the relevance for children's human capital formation?. *Available at SSRN 2506109*.

Vivalt, E. (2015). *How much can we generalize from impact evaluations?* Unpublished manuscript. New York University. New York, NY.

Wang, S., & Zhang, D. (2020). The impact of time while boarding on students' academic achievement and social emotional competence: A propensity score matching analysis. *Studies in Educational Evaluation*, *65*, 100851.

Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies*. Pearson College Division.

White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of development effectiveness*, *1*(3), 271-284.

Whitehead, A. N. (1959). The aims of education. *Daedalus*, *88*(1), 192-205.

World Bank. (2010). *Morocco: Can cash transfers help a country reach universal primary school education?* https://www.worldbank.org/en/programs/sief-trust-fund/brief/morocco-can-cash-transfers-make-a-difference-in-childrens-schooling

World Bank. (2016, October 6). *Evaluations - Education*. https://www.worldbank.org/en/programs/sief-trust-fund/brief/evaluations-education

Annex

Annex I. Educational statistics in Morocco

Annex I.A. Number of out-of-school children and adolescents

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Out-of-school children										
Total	205,179	228,065	178,992	203,684	...	162,240	89,998	7,804	16,064	...
Female	109,009	118,850	93,749	104,023	...	81,149	46,615	...	...	...
Male	96,170	109,215	85,243	99,661	...	81,091	43,383	...	...	...
Out-of-school adolescents										
Total	351,863	258,347	...	...	...	...	199,288	194,138	170,904	...
Female	218,278	169,733	...	...	...	...	120,709	116,980	100,912	...
Male	133,585	88,614	...	...	...	...	78,579	77,158	69,992	...

Annex I.B. Gross and net enrolment rates in primary and secondary schools

PRIMARY EDUCATION	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Gross enrolment ratio (%)										
Total	110.74	110.72	110.3	109.96	109.52	110.4	112.4	113.88	114.76	...
Female	107.74	108.01	107.76	107.43	106.64	107.73	109.78	111.51	112.73	...
Male	113.6	113.29	112.73	112.36	112.26	112.92	114.88	116.11	116.69	...
Net enrolment rate (%)										
Total	93.2	92.5	93.2	93.3	...	94.6	96.9	99.1	99.5	...
Female	92.9	92.3	93	93.2	...	94.7	97	...	...	...
Male	93.5	92.7	93.4	93.3	...	94.6	96.8	...	...	...

SECONDARY EDUCATION	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Gross enrolment ratio (%)										
Total	66.52	69.75	...	...	...	...	79.86	80.23	81.19	...
Female	61.52	64.15	...	...	...	...	75.33	76.58	78.18	...
Male	71.31	75.11	...	...	...	...	84.18	83.7	84.05	...
Net enrolment rate (%)										
Total	53.8	56.7	...	...	...	...	63.5	64.5	66.2	...
Female	51	53.6	...	...	...	...	63.2	64.5	66.6	...
Male	56.5	59.6	...	...	...	...	63.7	64.5	65.8	...

Source: UNESCO (2020) http://uis.unesco.org/en/country/ma

Annex I.C. Gross intake ratio in first grade, survival to the last grade of primary school and transition rate to lower secondary education

	2011	2012	2013	2014	2015	2016	2017	2018	2019
Primary education									
Gross intake ratio into the first grade of primary (%)									
Total	104.3	101.9	101.9	103.3	103.6	106.8	109.4	109.8	110.2
Female	104.1	101.7	102	103.5	103.2	106.7	108.9	110.8	110.7
Male	104.4	102.1	101.8	103.2	103.9	106.9	110	109	109.7
Survival to the last grade of primary (%)									
Total	88.24	91.56	89.25	88.84	92.63	95.05	92.99	94.26	...
Female	87.79	91.32	88.79	86.95	93.2	95.4	93.66	95.05	...
Male	88.65	91.79	89.69	90.67	92.09	94.73	92.35	93.52	...
Gross intake ratio into the last grade of primary (%)									
Total	89.4	98	96	96.3	96.3	95	92.9	93.6	97.1
Female	87	96.4	95.5	95.4	95.3	94.6	93.5	94.4	97.7
Male	91.6	99.5	96.5	97.1	97.2	95.5	92.4	92.8	96.6
Number of pupils per teacher									
Pupil/teacher ratio	26.4	25.8	26	25.7	25.9	26.6	28	26.8	25.8
Secondary education									
Effective transition rate from primary to lower secondary general education									
Total	88.6	90.4	88.7	87.4	88.7	90.2	90.5	92.3	...
Female	84.5	86.8	85	83.9	86.2	87.9	88.3	90.6	...
Male	92.3	93.7	92.2	90.7	91.1	92.5	92.7	93.9	...

Source: UNESCO (2020) http://uis.unesco.org/en/country/ma

Annex II. Propensity score matching implementation guide

Annex II.A. Implementation steps of propensity score matching

Source: Caliendo, M., & Kopeinig, S. (2008).

Annex II.B. Elements to be reported in the research using propensity score matching

1. List of all covariates that were collected (with reliabilities)
2. List of all covariates that were used to estimate the propensity score
3. Method that was used to determine set of covariates used for estimation (e.g.,non-parsimonious model, predetermined significance threshold)
4. Inclusion of polynomial or interaction terms
5. Estimation method for propensity scores (e.g., logistic regression, regression trees)
6. Conditioning strategy (e.g., matching, stratification, weighting)
7. Region of common support (histograms, ranges)
8. Details on matching scheme, if applicable
   - 8.1 Type of matching algorithm (e.g., nearest neighbor, optimal, full, kernel)
   - 8.2 Number of treated and control units that were matched with each other (e.g.,1:many)
   - 8.3 Matching with or without replacement
   - 8.4 Caliper width, if applicable
9. Details on stratification, if applicable
   - 9.1 Number of strata
   - 9.2 Strategy to define strata (equal proportions, minimize variance)
10. Details on weighting, if applicable
    - 10.1 Type of weights used (inverse probability weights. odd weights)
    - 10.2 Distribution of weights, reporting of unusually large weights
11. Sample size before and after conditioning; report effective sample size if weights are used
12. Standardized difference before and after matching on the propensity score and all covariates, potentially also on interactions and quadratic terms
13. Point estimate of treatment effect and associated standard error
14. Inclusion of covariates in outcome model

Resource: Thoemmes, F. J., & Kim, E. S. (2011).

Annex III. Number of observations per each program, school level and type of beneficiary

(without excluding the overlapping beneficiaries taking part in two or more programs)

	Current beneficiaries		Beneficiaries before	
	Primary	College	Primary	College
1 Million Schoolbags	3923	419	1895	1886
School canteens	718	160	632	523
Tayssir (cash transfer)	223	42	627	365
Collective transport program	59	105	92	108
Boarding schools	31	51	95	78
Scholarship programs	11	9	68	33

Annex IV. Descriptive statistics

Annex IV.A. Repetition rate by gender

Key
frequency *row percentage* *column percentage*

1.2 Gender	Repetition of schoolyear: Never rep	Repeated	Total
Female	3,289 71.52 53.31	1,310 28.48 37.78	4,599 100.00 47.73
Male	2,880 57.18 46.69	2,157 42.82 62.22	5,037 100.00 52.27
Total	6,169 64.02 100.00	3,467 35.98 100.00	9,636 100.00 100.00

Annex IV.B. Transition rate by gender

1.2 Gender	transit: 0	1	Total
Female	552 36.41 61.27	964 63.59 37.03	1,516 100.00 43.26
Male	349 17.56 38.73	1,639 82.44 62.97	1,988 100.00 56.74
Total	901 25.71 100.00	2,603 74.29 100.00	3,504 100.00 100.00

Annex IV.C. Repetition rate by deciles of average expenses per capita

Deciles of average expenses per capita	Repetition of schoolyear: Never rep	Repeated	Total
1	1,401 66.75 22.71	698 33.25 20.14	2,099 100.00 21.79
2	1,158 67.29 18.77	563 32.71 16.25	1,721 100.00 17.86
3	1,015 67.26 16.45	494 32.74 14.26	1,509 100.00 15.66
4	699 62.86 11.33	413 37.14 11.92	1,112 100.00 11.54
5	619 62.91 10.03	365 37.09 10.53	984 100.00 10.21
6	404 57.96 6.55	293 42.04 8.46	697 100.00 7.23
7	352 61.11 5.71	224 38.89 6.46	576 100.00 5.98
8	269 57.23 4.36	201 42.77 5.80	470 100.00 4.88
9	190 55.72 3.08	151 44.28 4.36	341 100.00 3.54
10	62 49.60 1.01	63 50.40 1.82	125 100.00 1.30
Total	6,169 64.03 100.00	3,465 35.97 100.00	9,634 100.00 100.00

Annex IV.D. Transition rate by deciles of average expenses per capita

Deciles of average expenses per capita	transit: 0	1	Total
1	221 30.65 24.53	500 69.35 19.22	721 100.00 20.58
2	134 23.10 14.87	446 76.90 17.14	580 100.00 16.56
3	127 24.42 14.10	393 75.58 15.10	520 100.00 14.84
4	85 21.41 9.43	312 78.59 11.99	397 100.00 11.33
5	86 23.31 9.54	283 76.69 10.88	369 100.00 10.53
6	64 24.62 7.10	196 75.38 7.53	260 100.00 7.42
7	64 27.23 7.10	171 72.77 6.57	235 100.00 6.71
8	59 28.50 6.55	148 71.50 5.69	207 100.00 5.91
9	42 27.27 4.66	112 72.73 4.30	154 100.00 4.40
10	19 31.67 2.11	41 68.33 1.58	60 100.00 1.71
Total	901 25.72 100.00	2,602 74.28 100.00	3,503 100.00 100.00

Annex V. Matching balance assessment results for outcome repetition rate

Annex V.A. Matching balance assessment results for outcome repetition rate for beneficiaries of One million bags program

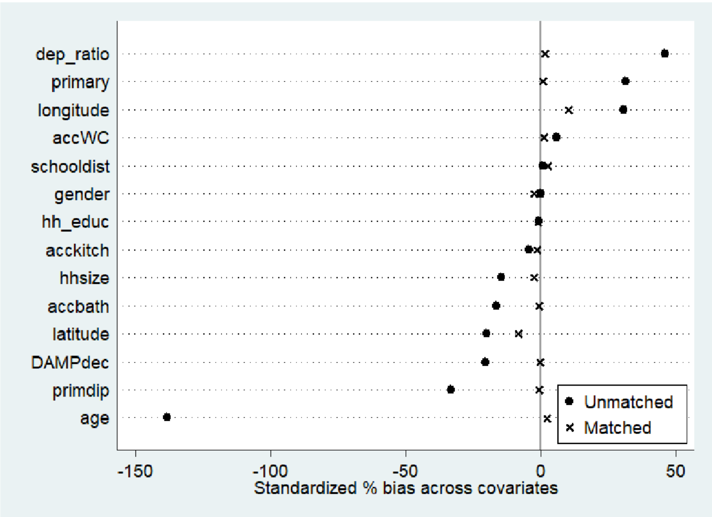
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
age	U	11.876	19.978	-138.0		-57.82	0.000	0.52*
	M	11.901	11.784	2.0	98.6	0.92	0.356	0.88*
gender	U	.5142	.51519	-0.2		-0.08	0.932	.
	M	.51444	.52767	-2.6	-1238.8	-1.05	0.293	.
DAMPdec	U	3.3516	3.8312	-20.6		-8.75	0.000	0.75*
	M	3.3374	3.3475	-0.4	97.9	-0.18	0.859	0.82*
dep_ratio	U	.88037	.62372	46.0		20.23	0.000	1.70*
	M	.87756	.87001	1.4	97.1	0.47	0.641	0.89*
hhsize	U	6.4018	6.7734	-14.5		-6.10	0.000	0.58*
	M	6.41	6.4792	-2.7	81.4	-1.10	0.272	0.63*
hh_educ	U	.21338	.21832	-0.8		-0.35	0.729	1.11*
	M	.21168	.21957	-1.3	-59.8	-0.50	0.619	1.00
primary	U	.77652	.63471	31.5		13.38	0.000	.
	M	.77563	.77292	0.6	98.1	0.26	0.797	.
schooldist	U	1328.9	1307.3	0.7		0.30	0.767	1.35*
	M	1334.9	1251.3	2.6	-287.2	1.12	0.265	1.93*
accbath	U	.49811	.58024	-16.5		-7.12	0.000	.
	M	.49889	.50324	-0.9	94.7	-0.35	0.730	.
accWC	U	.96275	.95116	5.7		2.43	0.015	.
	M	.96255	.96022	1.1	79.9	0.48	0.632	.
acckitch	U	.98169	.98695	-4.2		-1.84	0.065	.
	M	.98191	.98366	-1.4	66.6	-0.54	0.593	.
primdip	U	.27494	.43079	-33.1		-14.09	0.000	.
	M	.2761	.28003	-0.8	97.5	-0.35	0.728	.
longitude	U	68657	61365	30.8		13.29	0.000	1.12*
	M	68148	65772	10.0	67.4	4.26	0.000	0.78*
latitude	U	3.2e+05	3.3e+05	-20.1		-8.70	0.000	1.20*
	M	3.2e+05	3.2e+05	-8.5	57.4	-3.47	0.001	1.26*

* if variance ratio outside [0.93; 1.07] for U and [0.93; 1.07] for M

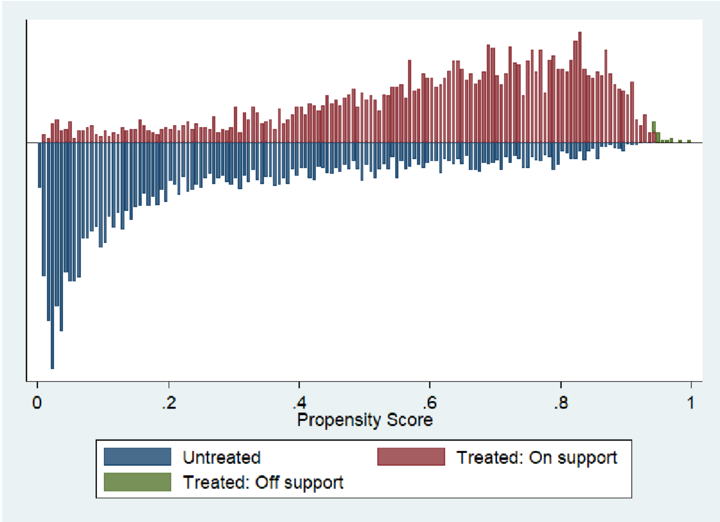
Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.289	2988.87	0.000	25.9	18.3	145.7*	0.60	100
Matched	0.003	25.10	0.034	2.6	1.4	12.6	0.83	88

* if B>25%, R outside [0.5; 2]

Annex V.A.b. Standardized bias of means across covariates before and after matching



Annex V.A.c. Histogram of propensity score distribution in treated and untreated groups



Annex V.B. Matching balance assessment results for outcome repetition rate for beneficiaries of Canteens program

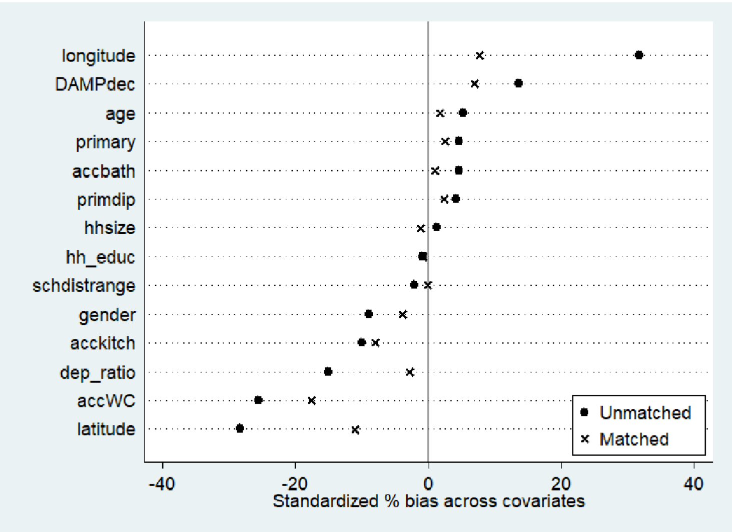
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
age	U	20.319	19.978	5.2		0.94	0.348	0.90
	M	20.352	20.239	1.7	66.9	0.24	0.812	0.95
gender	U	.47027	.51519	-9.0		-1.66	0.097	.
	M	.46883	.48819	-3.9	56.9	-0.53	0.599	.
DAMPdec	U	4.1676	3.8312	13.6		2.50	0.012	0.97
	M	4.168	3.9976	6.9	49.3	0.93	0.355	0.93
dep_ratio	U	.55807	.62372	-15.0		-2.56	0.011	0.65*
	M	.55823	.57079	-2.9	80.9	-0.43	0.666	0.94
hhsize	U	6.8081	6.7734	1.3		0.23	0.822	0.68*
	M	6.8103	6.8396	-1.1	15.7	-0.15	0.884	0.62*
hh_educ	U	.21351	.21832	-0.8		-0.15	0.882	0.91
	M	.21409	.21926	-0.9	-7.5	-0.12	0.905	0.89
primary	U	.65676	.63471	4.6		0.85	0.397	.
	M	.65583	.64407	2.5	46.7	0.33	0.738	.
schdistrange	U	3.2189	15.235	-2.1		-0.28	0.779	0.00*
	M	3.2249	4.1505	-0.2	92.3	-0.07	0.948	0.00*
accbath	U	.6027	.58024	4.6		0.84	0.400	.
	M	.60434	.59988	0.9	80.2	0.12	0.902	.
accWC	U	.88108	.95116	-25.5		-5.74	0.000	.
	M	.88076	.92935	-17.7	30.7	-2.26	0.024	.
acckitch	U	.97297	.98695	-10.0		-2.19	0.029	.
	M	.9729	.98424	-8.1	18.8	-1.06	0.288	.
primdip	U	.45135	.43079	4.1		0.77	0.443	.
	M	.45257	.44087	2.4	43.1	0.32	0.750	.
longitude	U	71399	61365	31.7		7.55	0.000	2.77*
	M	69561	67156	7.6	76.0	1.76	0.079	0.48*
latitude	U	3.2e+05	3.3e+05	-28.3		-4.60	0.000	0.49*
	M	3.2e+05	3.2e+05	-11.0	61.0	-1.53	0.126	0.52*

* if variance ratio outside [0.82; 1.23] for U and [0.81; 1.23] for M

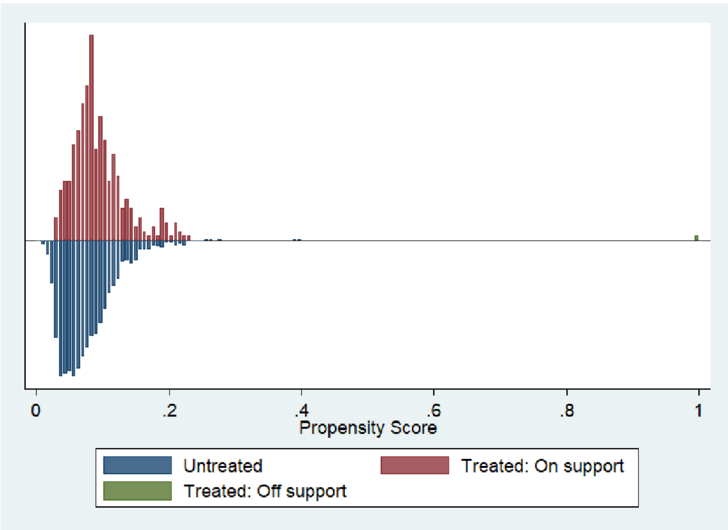
Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.038	99.75	0.000	11.1	7.1	48.2*	1.80	63
Matched	0.011	11.52	0.645	4.8	2.7	25.0*	0.99	50

* if B>25%, R outside [0.5; 2]

Annex V.B.b. Standardized bias of means across covariates before and after matching



Annex V.B.c. Histogram of propensity score distribution in treated and untreated groups



Annex V.C. Matching balance assessment results for outcome repetition rate for beneficiaries of both program

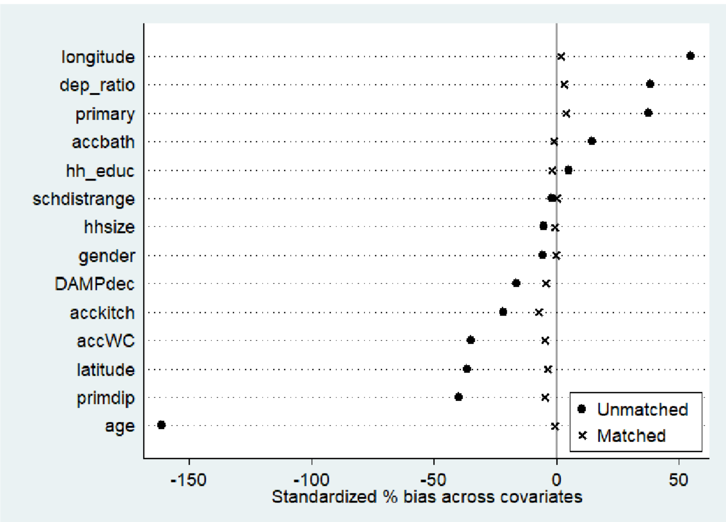
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
age	U	11.074	19.978	-161.3		-35.37	0.000	0.34*
	M	11.109	11.149	-0.7	99.6	-0.18	0.858	0.72*
gender	U	.48681	.51519	-5.7		-1.44	0.149	.
	M	.48537	.48803	-0.5	90.7	-0.10	0.918	.
DAMPdec	U	3.438	3.8312	-16.3		-4.06	0.000	0.87
	M	3.4428	3.5476	-4.4	73.4	-0.83	0.406	0.83*
dep_ratio	U	.83585	.62372	38.3		10.73	0.000	1.66*
	M	.82257	.80671	2.9	92.5	0.53	0.595	1.11
hhsize	U	6.6306	6.7734	-5.4		-1.29	0.197	0.66*
	M	6.641	6.6597	-0.7	86.9	-0.13	0.895	0.58*
hh_educ	U	.24934	.21832	5.0		1.30	0.194	1.18*
	M	.25133	.26507	-2.2	55.7	-0.39	0.697	0.84*
primary	U	.80079	.63471	37.5		8.98	0.000	.
	M	.7992	.78198	3.9	89.6	0.82	0.412	.
schdistrange	U	3.0211	15.235	-2.1		-0.41	0.684	0.00*
	M	3.0253	3.3515	-0.1	97.3	-0.06	0.956	0.00*
accbath	U	.6504	.58024	14.5		3.63	0.000	.
	M	.6516	.65683	-1.1	92.5	-0.21	0.831	.
acckitch	U	.94855	.98695	-21.9		-7.26	0.000	.
	M	.95612	.96909	-7.4	66.2	-1.33	0.185	.
accWC	U	.84697	.95116	-35.1		-10.95	0.000	.
	M	.85372	.86836	-4.9	86.0	-0.82	0.412	.
primdip	U	.2467	.43079	-39.6		-9.63	0.000	.
	M	.24867	.27138	-4.9	87.7	-1.00	0.316	.
longitude	U	72351	61365	54.9		12.59	0.000	0.50*
	M	72357	72054	1.5	97.2	0.30	0.765	0.54*
latitude	U	3.2e+05	3.3e+05	-36.5		-8.41	0.000	0.53*
	M	3.2e+05	3.2e+05	-3.8	89.5	-0.78	0.438	0.62*

* if variance ratio outside [0.87; 1.15] for U and [0.87; 1.15] for M

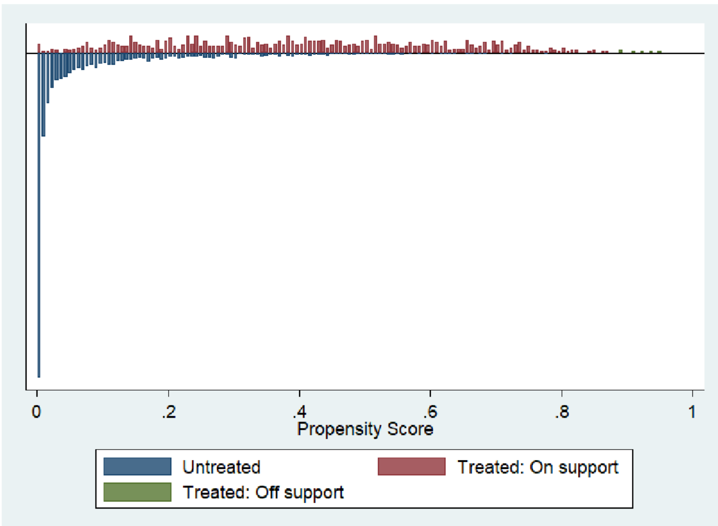
Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.345	1488.68	0.000	33.9	28.5	180.0*	0.40*	88
Matched	0.002	5.14	0.984	2.8	2.5	11.7	1.27	88

* if B>25%, R outside [0.5; 2]

Annex V.C.b. Standardized bias of means across covariates before and after matching



Annex V.B.c. Histogram of propensity score distribution in treated and untreated groups



Annex VI. Matching balance assessment results for the outcome transition to college

Annex VI.A. Matching balance assessment results for outcome transition rate for beneficiaries of One million bags program

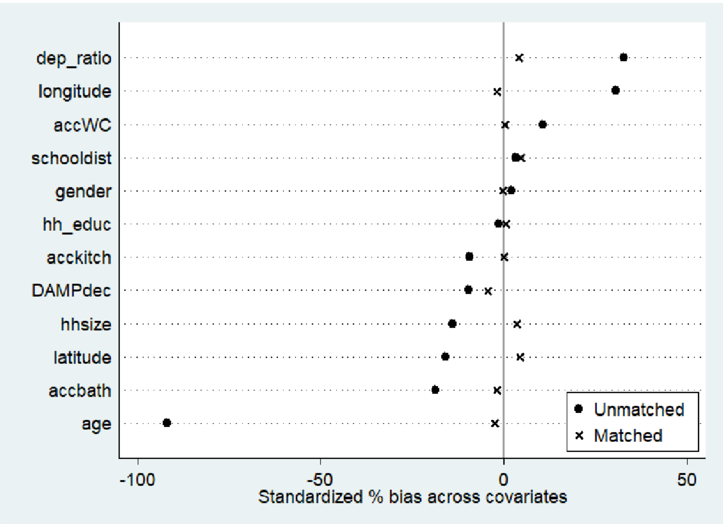
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
age	U	15.422	19.003	-91.9		-21.01	0.000	0.74*
	M	15.518	15.618	-2.6	97.2	-0.58	0.563	0.98
gender	U	.54438	.53411	2.1		0.48	0.631	.
	M	.54688	.54934	-0.5	76.0	-0.10	0.920	.
DAMPdec	U	3.5183	3.7479	-9.5		-2.19	0.028	0.82*
	M	3.5337	3.6436	-4.6	52.1	-0.94	0.345	0.87*
dep_ratio	U	.71857	.55983	32.7		7.97	0.000	1.77*
	M	.71129	.69242	3.9	88.1	0.74	0.457	1.17*
hhsize	U	6.5905	6.9552	-13.9		-3.10	0.002	0.50*
	M	6.5998	6.5087	3.5	75.0	0.79	0.432	0.68*
hh_educ	U	.2142	.22287	-1.4		-0.34	0.737	1.17*
	M	.21755	.21545	0.3	75.8	0.07	0.946	1.10
schooldist	U	1437.7	1309.5	3.3		0.80	0.425	1.82*
	M	1443.4	1267	4.5	-37.7	0.91	0.363	1.79*
accbath	U	.52071	.61274	-18.6		-4.37	0.000	.
	M	.52885	.53869	-2.0	89.3	-0.40	0.687	.
accWC	U	.97041	.94997	10.5		2.36	0.018	.
	M	.96995	.96959	0.2	98.2	0.04	0.966	.
acckitch	U	.97751	.9896	-9.5		-2.35	0.019	.
	M	.98077	.98077	0.0	100.0	0.00	1.000	.
longitude	U	69340	60839	30.6		7.48	0.000	1.86*
	M	68389	68970	-2.1	93.2	-0.52	0.606	0.75*
latitude	U	3.2e+05	3.3e+05	-15.9		-3.75	0.000	1.18*
	M	3.2e+05	3.2e+05	4.2	73.7	0.83	0.409	1.04

* if variance ratio outside [0.87; 1.14] for U and [0.87; 1.15] for M

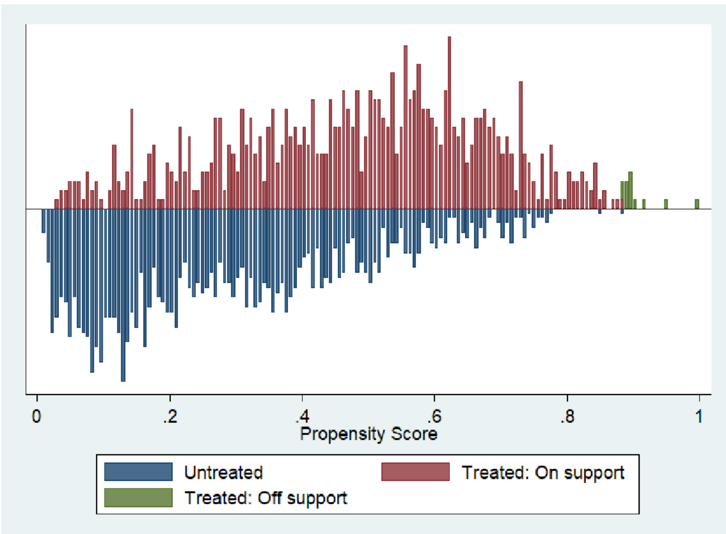
Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.170	528.01	0.000	20.0	12.2	101.4*	0.99	100
Matched	0.002	3.61	0.990	2.4	2.3	9.3	1.12	63

* if B>25%, R outside [0.5; 2]

Annex VI.A.b. Standardized bias of means across covariates before and after matching



Annex VI.A.c. Histogram of propensity score distribution in treated and untreated



Annex VI.B. Matching balance assessment results for outcome transition rate for beneficiaries of Canteens program

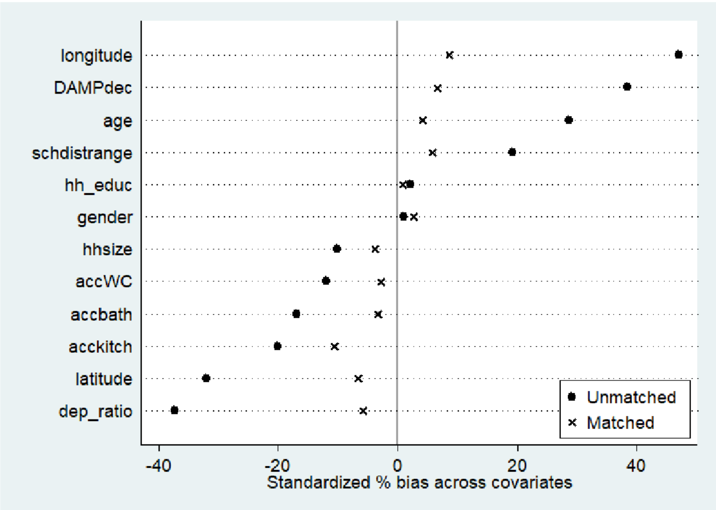
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
age	U	18.505	17.606	28.6		2.62	0.009	0.77
	M	18.505	18.376	4.1	85.6	0.30	0.766	0.82
gender	U	.52475	.51968	1.0		0.10	0.922	.
	M	.52475	.51181	2.6	-155.1	0.18	0.855	.
DAMPdec	U	4.4752	3.5237	38.5		3.73	0.000	1.02
	M	4.4752	4.3137	6.5	83.0	0.44	0.660	0.84
dep_ratio	U	.4378	.56331	-37.4		-3.09	0.002	0.40*
	M	.4378	.45766	-5.9	84.2	-0.51	0.611	0.73
hhsize	U	6.7624	7.0522	-10.1		-0.89	0.372	0.63*
	M	6.7624	6.8724	-3.8	62.0	-0.27	0.785	0.64*
hh_educ	U	.22772	.21365	2.1		0.22	0.823	1.64*
	M	.22772	.22271	0.7	64.4	0.05	0.958	1.65*
schdistrange	U	3.1485	2.8394	19.2		1.92	0.055	1.15
	M	3.1485	3.0546	5.8	69.6	0.40	0.687	1.03
accbath	U	.52475	.60884	-17.0		-1.66	0.097	.
	M	.52475	.5412	-3.3	80.4	-0.23	0.816	.
accWC	U	.92079	.9502	-12.0		-1.28	0.201	.
	M	.92079	.92767	-2.8	76.6	-0.18	0.854	.
acckitch	U	.9604	.99116	-20.0		-2.84	0.005	.
	M	.9604	.97669	-10.6	47.1	-0.66	0.510	.
longitude	U	70119	60735	47.1		3.98	0.000	0.46*
	M	70119	68406	8.6	81.7	0.62	0.537	0.48*
latitude	U	3.2e+05	3.3e+05	-32.1		-2.79	0.005	0.56*
	M	3.2e+05	3.2e+05	-6.6	79.3	-0.46	0.646	0.52*

* if variance ratio outside [0.67; 1.48] for U and [0.67; 1.48] for M

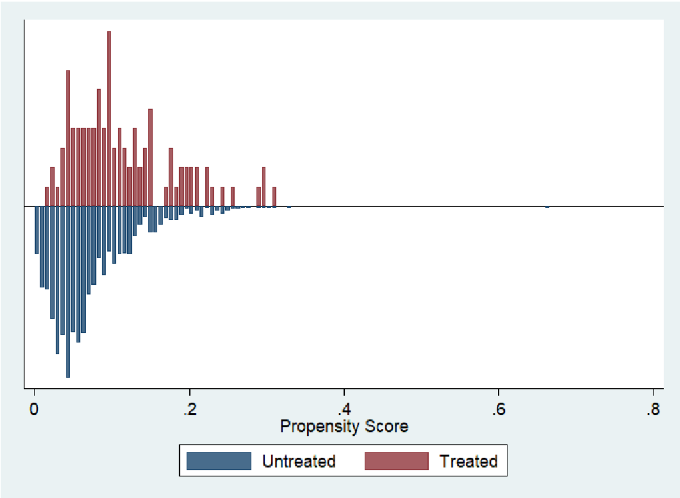
Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.079	56.63	0.000	22.1	19.6	80.9*	0.65	63
Matched	0.006	1.73	1.000	5.1	5.0	18.4	0.98	50

* if B>25%, R outside [0.5; 2]

Annex VI.B.b. Standardized bias of means across covariates before and after matching



Annex VI.B.c. Histogram of propensity score distribution in treated and untreated



Annex VI.C. Matching balance assessment results for outcome transition rate for beneficiaries of both program

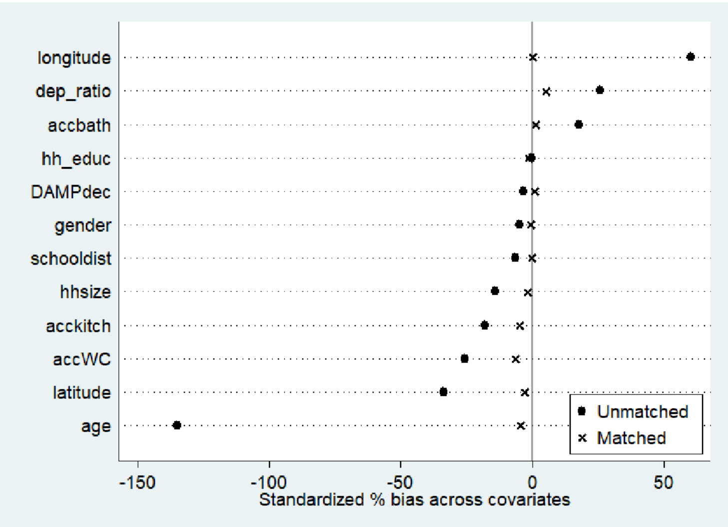
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
age	U	15.134	20.825	-135.0		-14.54	0.000	0.28*
	M	15.151	15.347	-4.7	96.6	-0.63	0.531	0.73*
gender	U	.53476	.55799	-4.7		-0.61	0.542	.
	M	.53763	.54013	-0.5	89.3	-0.05	0.962	.
DAMPdec	U	3.8235	3.908	-3.4		-0.44	0.664	0.85
	M	3.8333	3.8079	1.0	69.9	0.10	0.920	0.85
dep_ratio	U	.68172	.56267	25.9		3.65	0.000	1.42*
	M	.67822	.65566	4.9	81.0	0.45	0.653	1.13
hhsize	U	6.5241	6.8934	-14.2		-1.65	0.099	0.52*
	M	6.5215	6.5717	-1.9	86.4	-0.20	0.840	0.68*
hh_educ	U	.25134	.2534	-0.3		-0.04	0.964	1.11
	M	.25269	.26169	-1.5	-337.4	-0.13	0.896	0.80
schooldist	U	1151	1322.6	-6.5		-0.68	0.497	0.18*
	M	1146.4	1158.9	-0.5	92.7	-0.05	0.957	0.27*
accbath	U	.68984	.60554	17.7		2.26	0.024	.
	M	.69355	.68688	1.4	92.1	0.14	0.890	.
accWC	U	.88235	.95193	-25.4		-4.02	0.000	.
	M	.8871	.90458	-6.4	74.9	-0.55	0.582	.
acckitch	U	.96257	.98955	-17.7		-3.13	0.002	.
	M	.96774	.97549	-5.1	71.3	-0.45	0.654	.
longitude	U	72525	60732	60.5		6.79	0.000	0.40*
	M	72556	72508	0.2	99.6	0.02	0.981	0.41*
latitude	U	3.2e+05	3.3e+05	-33.6		-3.84	0.000	0.47*
	M	3.2e+05	3.2e+05	-2.9	91.4	-0.28	0.779	0.48*

* if variance ratio outside [0.75; 1.33] for U and [0.75; 1.34] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.251	316.56	0.000	28.7	17.7	156.3*	0.29*	75
Matched	0.003	1.44	1.000	2.6	1.7	12.4	1.00	63

* if B>25%, R outside [0.5; 2]

Annex VI.C.b. Standardized bias of means across covariates before and after matching



Annex VI.C.c. Histogram of propensity score distribution in treated and untreated

Annex VII. Propensity score matching: common support regions

Annex VII.A. Number of observations in the common support region. Repetition: 1 Million schoolbags program

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	4,443	4,443
Treated	17	3,151	3,168
Total	17	7,594	7,611

Annex VII.B. Number of observations in the common support region. Repetition: Canteens program

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	4,443	4,443
Treated	1	369	370
Total	1	4,812	4,813

Annex VII.C. Number of observations in the common support region. Repetition: both program

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	4,443	4,443
Treated	6	752	758
Total	6	5,195	5,201

Annex VII.D. Number of observations in the common support. Transition: 1 Million schoolbags program

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	1,914	1,914
Treated	11	860	871
Total	11	2,774	2,785

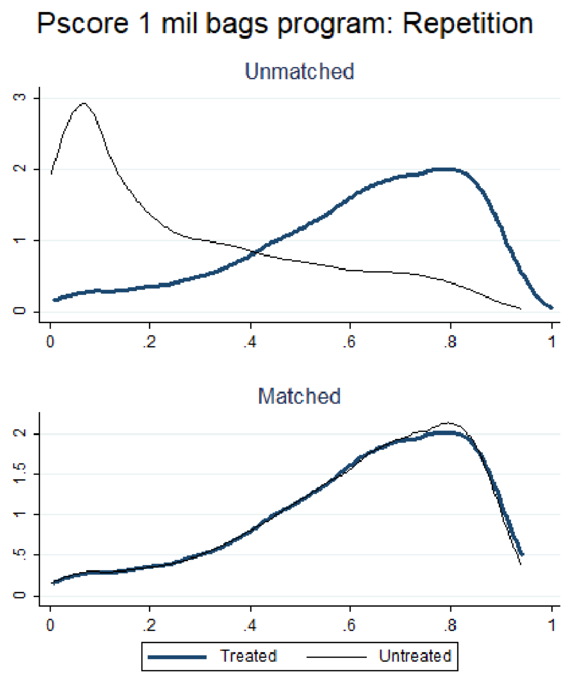
Annex VII.E. Number of observations in the common support region. Transition: Canteens program

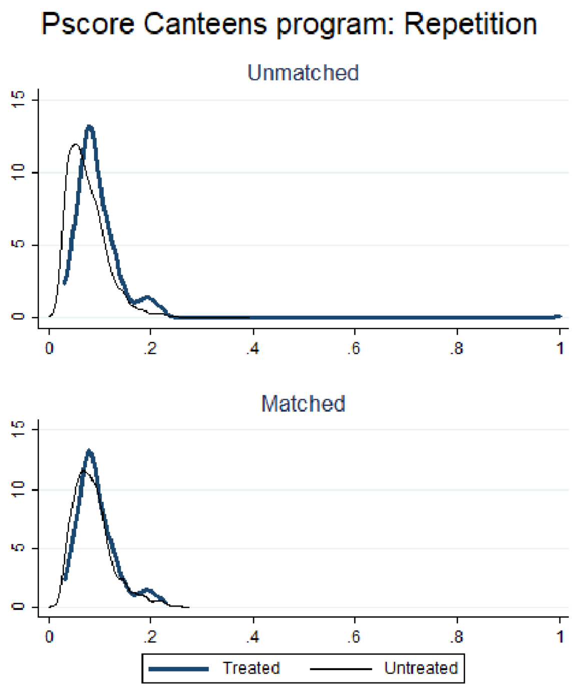
psmatch2: Treatment assignment	psmatch2: Common support On suppor	Total
Untreated	1,914	1,914
Treated	167	167
Total	2,081	2,081

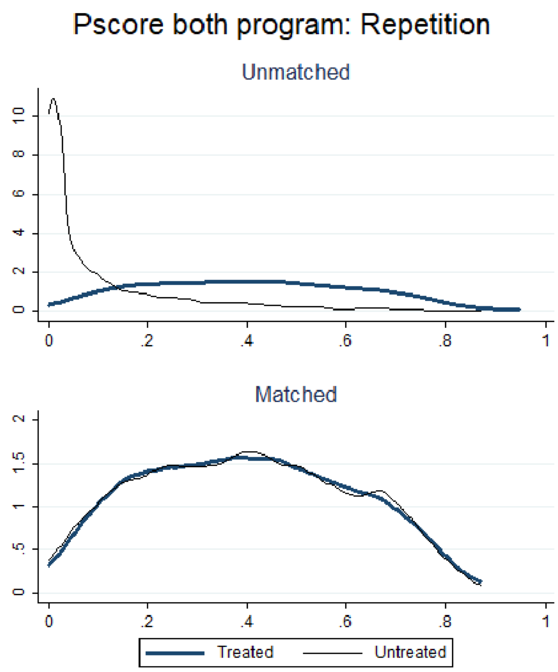
Annex VII.F. Number of observations in the common support region. Transition: both program

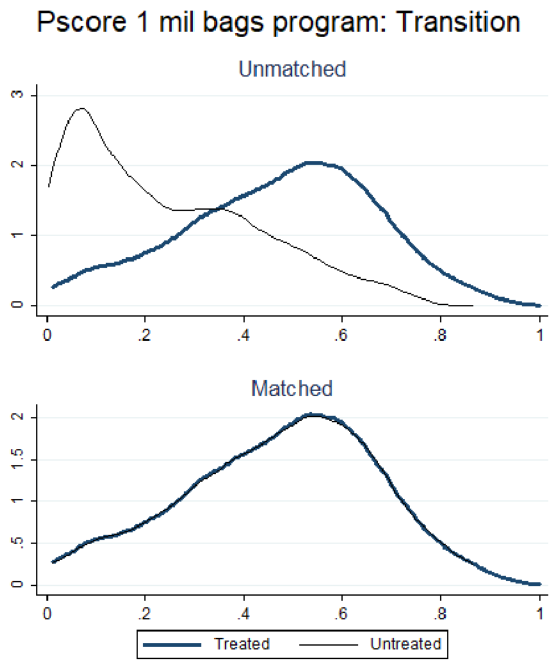
psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	1,914	1,914
Treated	1	186	187
Total	1	2,100	2,101

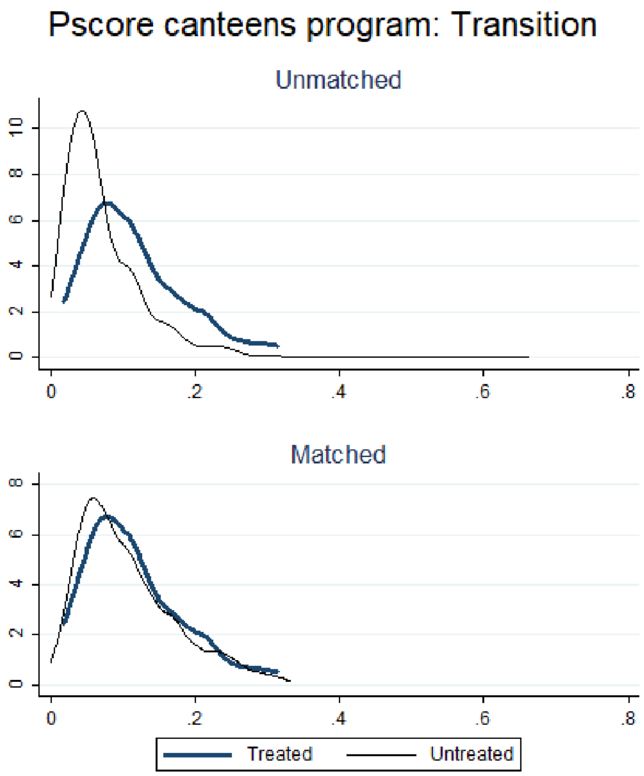
Annex VIII. Graphic results of propensity score matching

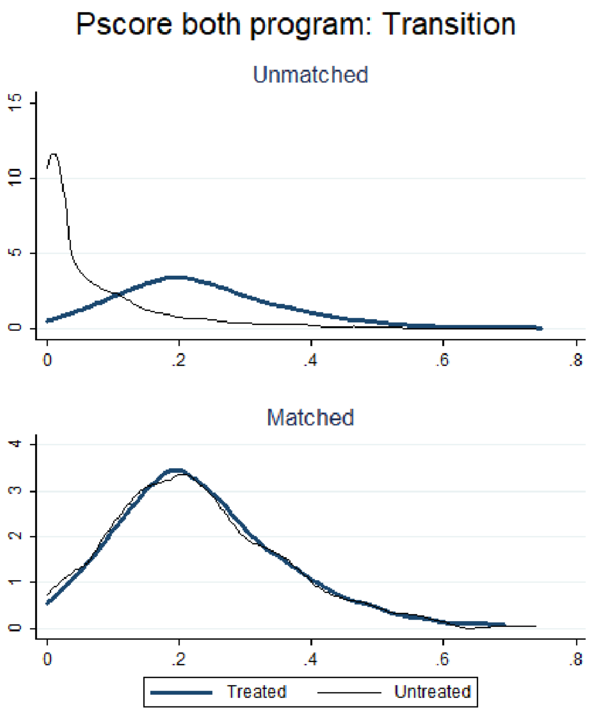









Pscore canteens program: Transition
Unmatched
0
2
4
6
8
10
0
.2
.4
.6
.8
Matched
0
2
4
6
8
0
.2
.4
.6
.8
Treated
Untreated


Pscore both program: Transition
Unmatched
0
5
10
15
0
.2
.4
.6
.8
Matched
0
1
2
3
4
0
.2
.4
.6
.8
Treated
Untreated

Annex IX. PSM estimation of average treatment effect on the treated

Repetition: 1 Million bags program

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
repetitionbin	Unmatched	.25	.418636057	-.168636057	.010911323	-15.46
	ATT	.250714059	.203948921	.046765138	.015755565	2.97

Repetition: Canteens program

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
repetitionbin	Unmatched	.424324324	.418636057	.005688268	.026703168	0.21
	ATT	.425474255	.419524226	.005950029	.026925912	0.22

Repetition: both programs

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
repetitionbin	Unmatched	.27176781	.418636057	-.146868247	.019125025	-7.68
	ATT	.272606383	.177943623	.09466276	.023064285	4.10

Transition to college: 1 Million bags program

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
transit	Unmatched	.768047337	.745938921	.022108416	.018447761	1.20
	ATT	.774038462	.771410628	.002627833	.022805561	0.12

Transition to college: Canteens program

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
transit	Unmatched	.742574257	.751004016	-.008429759	.044810482	-0.19
	ATT	.742574257	.737333692	.005240565	.04616261	0.11

Transition to college: both programs

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
transit	Unmatched	.982608696	.946341463	.036267232	.022024695	1.65
	ATT	.98245614	.946138636	.036317504	.018571976	1.96

Annex X. Robustness check results: bootstrapping of ATT and standard errors

Repetition: One million bags program

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
_bs_1	.0492424	.0159527	3.09	0.002	.0179758	.0805091

Repetition: Canteens program

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
_bs_1	.0108108	.0370644	0.29	0.771	-.0618342	.0834558

Repetition: both programs

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
_bs_1	.0963061	.0237073	4.06	0.000	.0498407	.1427714

Transition to college: One million bags program

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
_bs_1	.0130178	.0296342	0.44	0.660	-.0450643	.0710998

Transition to college: Canteens program

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
_bs_1	-.029703	.0658011	-0.45	0.652	-.1586708	.0992649

Transition to college: both programs

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
_bs_1	.0053476	.0595403	0.09	0.928	-.1113493	.1220445

Annex XI. Robustness check: sensitivity analysis results

Repetition: One million bags program

Mantel-Haenszel (1959) bounds for variable repetitionbin

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	2.49014	2.49014	.006385	.006385
1.05	3.12995	1.85171	.000874	.032034
1.1	3.74065	1.24325	.000092	.106887
1.15	4.32526	.662175	7.6e-06	.253929
1.2	4.88611	.105991	5.1e-07	.457795
1.25	5.42525	.351138	2.9e-08	.362743
1.3	5.94448	.863976	1.4e-09	.193801
1.35	6.44539	1.35761	5.8e-11	.087294
1.4	6.92938	1.83351	2.1e-12	.033364
1.45	7.39769	2.29297	6.9e-14	.010925
1.5	7.85146	2.73718	2.1e-15	.003098
1.55	8.29168	3.16716	1.1e-16	.00077
1.6	8.71926	3.58386	0	.000169
1.65	9.13501	3.98813	0	.000033
1.7	9.53968	4.38074	0	5.9e-06
1.75	9.93393	4.76237	0	9.6e-07
1.8	10.3184	5.13368	0	1.4e-07
1.85	10.6936	5.49524	0	2.0e-08
1.9	11.0601	5.84758	0	2.5e-09
1.95	11.4184	6.19119	0	3.0e-10
2	11.7688	6.52654	0	3.4e-11

Gamma : odds of differential assignment due to unobserved factors
Q_mh+ : Mantel-Haenszel statistic (assumption: overestimation of treatment effect)
Q_mh- : Mantel-Haenszel statistic (assumption: underestimation of treatment effect)
p_mh+ : significance level (assumption: overestimation of treatment effect)
p_mh- : significance level (assumption: underestimation of treatment effect)

Repetition: Canteens program

Mantel-Haenszel (1959) bounds for variable repetitionbin

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	.333757	.333757	.369281	.369281
1.05	.015424	.652619	.493847	.257001
1.1	.135101	.956516	.446266	.169406
1.15	.425298	1.24703	.33531	.106193
1.2	.703196	1.52534	.240967	.063588
1.25	.96983	1.79245	.166066	.03653
1.3	1.22611	2.04928	.110079	.020217
1.35	1.47284	2.29662	.070397	.01082
1.4	1.71073	2.53517	.043566	.00562
1.45	1.94042	2.76555	.026164	.002841
1.5	2.16248	2.98835	.015291	.001402
1.55	2.37742	3.20406	.008717	.000678
1.6	2.58571	3.41314	.004859	.000321
1.65	2.78776	3.616	.002654	.00015
1.7	2.98396	3.81303	.001423	.000069
1.75	3.17465	4.00456	.00075	.000031
1.8	3.36014	4.19092	.00039	.000014
1.85	3.54074	4.37238	.0002	6.1e-06
1.9	3.7167	4.54921	.000101	2.7e-06
1.95	3.88827	4.72166	.00005	1.2e-06
2	4.05567	4.88995	.000025	5.0e-07

Repetition: both programs

Mantel-Haenszel (1959) bounds for variable repetitionbin

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	3.01647	3.01647	.001279	.001279
1.05	2.67062	3.36565	.003786	.000382
1.1	2.3404	3.69829	.009632	.000109
1.15	2.02544	4.01701	.021411	.000029
1.2	1.72435	4.32301	.042322	7.7e-06
1.25	1.43592	4.61737	.075512	1.9e-06
1.3	1.1591	4.901	.123208	4.8e-07
1.35	.892938	5.17472	.185945	1.1e-07
1.4	.636615	5.43926	.262188	2.7e-08
1.45	.389394	5.69528	.348492	6.2e-09
1.5	.150616	5.94335	.440139	1.4e-09
1.55	-.061864	6.18399	.524664	3.1e-10
1.6	.161229	6.41768	.435957	6.9e-11
1.65	.377476	6.64485	.35291	1.5e-11
1.7	.587315	6.86589	.278496	3.3e-12
1.75	.79114	7.08115	.214431	7.1e-13
1.8	.989312	7.29096	.161255	1.5e-13
1.85	1.18216	7.4956	.118571	3.3e-14
1.9	1.36999	7.69537	.085345	7.1e-15
1.95	1.55308	7.89051	.060203	1.6e-15
2	1.73167	8.08124	.041666	3.3e-16

Gamma : odds of differential assignment due to unobserved factors
Q_mh+ : Mantel-Haenszel statistic (assumption: overestimation of treatment effect)
Q_mh- : Mantel-Haenszel statistic (assumption: underestimation of treatment effect)
p_mh+ : significance level (assumption: overestimation of treatment effect)
p_mh- : significance level (assumption: underestimation of treatment effect)

Transition to college: One million bags program

Mantel-Haenszel (1959) bounds for variable transit

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	.823895	.823895	.205	.205
1.05	.450012	1.19863	.326351	.115337
1.1	.093358	1.55596	.46281	.059859
1.15	.117196	1.89779	.453353	.028862
1.2	.443789	2.22552	.328598	.013023
1.25	.757134	2.54036	.224485	.005537
1.3	1.05832	2.84337	.144954	.002232
1.35	1.34832	3.13548	.088778	.000858
1.4	1.62797	3.41753	.051765	.000316
1.45	1.89804	3.69026	.028846	.000112
1.5	2.15919	3.95433	.015418	.000038
1.55	2.41204	4.21034	.007932	.000013
1.6	2.65713	4.45883	.00394	4.1e-06
1.65	2.89496	4.70027	.001896	1.3e-06
1.7	3.12597	4.93511	.000886	4.0e-07
1.75	3.35057	5.16374	.000403	1.2e-07
1.8	3.56913	5.38654	.000179	3.6e-08
1.85	3.78199	5.60383	.000078	1.0e-08
1.9	3.98947	5.81593	.000033	3.0e-09
1.95	4.19184	6.02311	.000014	8.6e-10
2	4.38937	6.22564	5.7e-06	2.4e-10

Transition to college: Canteens program

Mantel-Haenszel (1959) bounds for variable transit

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	.010597	.010597	.495772	.495772
1.05	-.149623	.17087	.559469	.432163
1.1	-.001857	.323698	.500741	.373083
1.15	.144225	.469798	.442661	.31925
1.2	.284116	.609766	.388161	.271009
1.25	.418342	.744123	.337849	.228401
1.3	.547368	.87333	.292063	.191242
1.35	.671602	.997788	.250919	.159191
1.4	.791406	1.11786	.214353	.131814
1.45	.907104	1.23386	.182176	.108628
1.5	1.01898	1.34608	.154105	.089139
1.55	1.12731	1.45477	.129807	.072866
1.6	1.2323	1.56017	.108918	.05936
1.65	1.33419	1.66248	.091071	.048208
1.7	1.43315	1.7619	.075908	.039043
1.75	1.52936	1.8586	.063087	.031542
1.8	1.62299	1.95273	.052296	.025426
1.85	1.71418	2.04445	.043248	.020455
1.9	1.80306	2.13388	.03569	.016426
1.95	1.88975	2.22115	.029396	.01317
2	1.97438	2.30637	.024169	.010545

Gamma : odds of differential assignment due to unobserved factors
Q_mh+ : Mantel-Haenszel statistic (assumption: overestimation of treatment effect)
Q_mh- : Mantel-Haenszel statistic (assumption: underestimation of treatment effect)
p_mh+ : significance level (assumption: overestimation of treatment effect)
p_mh- : significance level (assumption: underestimation of treatment effect)

Transition to college: both programs

Mantel-Haenszel (1959) bounds for variable transit

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	1.57554	1.57554	.057566	.057566
1.05	1.38728	1.76893	.082678	.038453
1.1	1.2059	1.95155	.113928	.025496
1.15	1.03287	2.12651	.150833	.01673
1.2	.86742	2.2945	.192856	.010881
1.25	.708898	2.45609	.239194	.007023
1.3	.556722	2.61182	.288859	.004503
1.35	.410383	2.76212	.340763	.002871
1.4	.26943	2.90742	.3938	.001822
1.45	.133462	3.04807	.446914	.001152
1.5	.002122	3.1844	.499153	.000725
1.55	-.12491	3.31669	.549703	.000455
1.6	-.010577	3.44521	.50422	.000285
1.65	.108455	3.5702	.456817	.000178
1.7	.223944	3.69188	.4114	.000111
1.75	.336107	3.81043	.368395	.000069
1.8	.445141	3.92605	.328109	.000043
1.85	.551227	4.03889	.290739	.000027
1.9	.654528	4.14911	.256386	.000017
1.95	.755197	4.25684	.225065	.00001
2	.853374	4.36222	.196726	6.4e-06

Annex XII. Logistic regression results for the interaction of treatment and gender: impact of treatment on outcome repetition and transition to college

VARIABLES	(1) Onemilbags repetitionbin	(2) Canteenprog repetitionbin	(3) Bothprog repetitionbin	(4) Onemilbags transit	(5) Canteenprog transit	(6) Bothprog transit
0b.bagsbin#1.gender	-0.405***			-0.242		
(not benefit) (male)	(0.112)			(0.236)		
1.bagsbin#0b.gender	0.194			-0.0881		
(benefit) (female)	(0.161)			(0.291)		
0b.cantbin#1.gender		-0.120			-0.0863	
		(0.167)			(0.302)	
1.cantbin#0b.gender		-0.0281			-0.0222	
		(0.189)			(0.381)	
0b.cantbags#1.gender			-0.827***			-1.067**
			(0.213)			(0.425)
1.cantbags#0b.gender			0.631***			-0.204
(female)			(0.158)			(0.433)
age	0.193***	0.0829***	0.255***	0.0108	-0.128***	-0.0519
	(0.0219)	(0.00853)	(0.0217)	(0.0375)	(0.0234)	(0.0360)
gender	0.689***	0.499**	1.211***	1.155***	1.127***	2.036***
	(0.137)	(0.214)	(0.194)	(0.277)	(0.287)	(0.434)
DAMPdec	-0.0219	0.0363**	-0.0117	0.0440	-0.0446	0.0150
	(0.0212)	(0.0177)	(0.0267)	(0.0473)	(0.0384)	(0.0463)
dep_ratio	-0.192*	-0.408***	-0.201	-0.226	-0.364	-0.188
	(0.103)	(0.101)	(0.142)	(0.177)	(0.306)	(0.212)
hhsize	-0.0329	-0.0264	-0.0613***	-0.00418	-0.0334	-0.0504
	(0.0208)	(0.0190)	(0.0219)	(0.0249)	(0.0322)	(0.0389)
hh_educ	-0.236***	0.00600	-0.692***	0.279**	0.0920	0.384
	(0.0629)	(0.0882)	(0.106)	(0.124)	(0.107)	(0.292)
schooldist	2.14e-05*			-3.42e-05**		2.63e-05
	(1.24e-05)			(1.43e-05)		(3.43e-05)
accbath	-0.0971	-0.447**	-0.317	-0.212	-0.0580	-0.410*
	(0.131)	(0.192)	(0.216)	(0.218)	(0.386)	(0.212)
acckitch	0.693**	-0.0307	0.0797	0.668	-0.0319	0.417
	(0.292)	(0.564)	(0.302)	(0.588)	(0.641)	(0.695)
accWC	-0.256	-0.0260	-0.0325	0.328	-0.0964	0.647*
	(0.173)	(0.212)	(0.140)	(0.506)	(0.420)	(0.350)

VARIABLES	(1) Onemilbags repetitionbin	(2) Canteenprog repetitionbin	(3) Bothprog repetitionbin	(4) Onemilbags transit	(5) Canteenprog transit	(6) Bothprog transit
longitude	1.10e-05***	4.25e-06	3.55e-06	-1.88e-05**	-4.36e-05***	-1.55e-05
	(3.11e-06)	(4.18e-06)	(4.63e-06)	(9.13e-06)	(1.37e-05)	(1.04e-05)
latitude	-2.04e-05***	-8.02e-06	1.98e-06	-3.51e-05***	2.01e-05	-1.96e-05
	(5.63e-06)	(6.85e-06)	(7.75e-06)	(1.17e-05)	(1.42e-05)	(1.56e-05)
primary	-0.430***	-1.280***	-1.144***			
	(0.136)	(0.122)	(0.177)			
primdip	-0.325**	-0.516***	-1.356***			
	(0.152)	(0.134)	(0.249)			
region	-0.170***	0.00205	0.0451	-0.0172	0.505***	0.0621
	(0.0509)	(0.0473)	(0.0705)	(0.0802)	(0.122)	(0.108)
worknow	0.0230	-0.00479	-0.283*	-0.291**	-0.115	-0.365**
	(0.0776)	(0.0750)	(0.163)	(0.115)	(0.109)	(0.152)
martstatbin	-1.267***	-0.566***	-1.831***	-0.902***	-0.251	-0.898**
	(0.232)	(0.126)	(0.342)	(0.309)	(0.253)	(0.361)
schdistrange		0.00371	4.53e-05**		0.0419	
		(0.0486)	(2.20e-05)		(0.0835)	
Constant	3.370	1.843	-3.796	12.68***	-2.625	8.333
	(2.066)	(2.460)	(2.942)	(4.283)	(4.812)	(5.504)
Observations	7,207	4,688	5,010	2,489	1,456	2,093

Robust standard errors in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

Annex XIII. Logistic regression results for the interaction of treatment and school level: impact of treatment on outcome repetition rate

VARIABLES	(1) Onemilbags repetitionbin	(2) Canteenprog repetitionbin	(3) Bothprog repetitionbin
0b.bagsbin#1.primary	-0.398***		
	(0.127)		
1.bagsbin#0b.primary	0.146		
	(0.111)		
0b.cantbin#1.primary		0.151	
		(0.146)	
1.cantbin#0b.primary		0.370*	
		(0.193)	
0b.cantbags#1.primary			-0.826***
			(0.174)
1.cantbags#0b.primary			0.546***
			(0.146)
age	0.193***	0.0824***	0.255***
	(0.0222)	(0.00849)	(0.0220)
gender	0.399***	0.450***	0.506***
	(0.0823)	(0.155)	(0.113)
DAMPdec	-0.0206	0.0314*	-0.0110
	(0.0212)	(0.0175)	(0.0265)
dep_ratio	-0.191*	-0.409***	-0.202
	(0.104)	(0.0960)	(0.138)
hhsize	-0.0332	-0.0270	-0.0627***
	(0.0208)	(0.0186)	(0.0219)
hh_educ	-0.237***	0.0106	-0.691***
	(0.0626)	(0.0871)	(0.106)
schooldist	2.20e-05*		
	(1.25e-05)		
accbath	-0.0949	-0.436**	-0.320
	(0.130)	(0.193)	(0.212)
acckitch	0.690**	-0.00795	0.0701
	(0.290)	(0.544)	(0.298)
accWC	-0.246	-0.0452	-0.0148
	(0.174)	(0.220)	(0.140)
longitude	1.09e-05***	4.11e-06	3.19e-06
	(3.08e-06)	(4.25e-06)	(4.55e-06)
latitude	-2.02e-05***	-8.85e-06	2.66e-06
	(5.61e-06)	(6.81e-06)	(7.68e-06)
primary	-0.165	-1.165***	-0.468***
	(0.151)	(0.172)	(0.180)
primdip	-0.329**	-0.495***	-1.351***
	(0.153)	(0.130)	(0.250)
region	-0.169***	-0.00327	0.0488
	(0.0508)	(0.0464)	(0.0692)
worknow	0.0172	-0.00682	-0.291*
	(0.0776)	(0.0750)	(0.155)
martstatbin	-1.269***	-0.554***	-1.831***
	(0.233)	(0.132)	(0.348)
schdistrange		0.000930	4.46e-05*
		(0.0469)	(2.31e-05)
Constant	3.312	1.972	-3.987
	(2.061)	(2.454)	(2.897)
Observations	7,207	4,688	5,010

Robust standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1

Annex XIV. Results of the logistic regressions and margins to estimate treatment effects

Annex XIV.A.a. Logistic model and margins to estimate treatment effect for outcome repetition for participation in One million bags program

```
Iteration 0:   log pseudolikelihood = -3184.6984
Iteration 1:   log pseudolikelihood = -2715.6355
Iteration 2:   log pseudolikelihood = -2697.0987
Iteration 3:   log pseudolikelihood = -2696.9888
Iteration 4:   log pseudolikelihood = -2696.9888

Logistic regression                             Number of obs     =      7,207
                                                Wald chi2(18)     =    1085.72
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -2696.9888               Pseudo R2         =     0.1531

                                    (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin	.3124138	.107207	2.91	0.004	.102292	.5225357
age	.1930219	.0222418	8.68	0.000	.1494287	.236615
gender	.3995778	.0828027	4.83	0.000	.2372875	.5618681
DAMPdec	-.0213187	.0211926	-1.01	0.314	-.0628553	.020218
dep_ratio	-.1938443	.1036459	-1.87	0.061	-.3969864	.0092978
hhsize	-.0334959	.0208932	-1.60	0.109	-.0744459	.0074541
hh_educ	-.2380039	.0624464	-3.81	0.000	-.3603966	-.1156112
schooldist	.0000217	.0000124	1.75	0.080	-2.56e-06	.000046
accbath	-.0959928	.1307965	-0.73	0.463	-.3523493	.1603637
acckitch	.6963438	.2929557	2.38	0.017	.1221612	1.270526
accWC	-.2504367	.174672	-1.43	0.152	-.5927875	.091914
longitude	.000011	3.12e-06	3.53	0.000	4.90e-06	.0000171
latitude	-.0000206	5.65e-06	-3.65	0.000	-.0000317	-9.56e-06
primary	-.4319603	.1371213	-3.15	0.002	-.7007132	-.1632074
primdip	-.3287779	.1534396	-2.14	0.032	-.629514	-.0280418
region	-.1718157	.0509463	-3.37	0.001	-.2716687	-.0719627
worknow	.0171922	.0773592	0.22	0.824	-.1344291	.1688135
martstatbin	-1.264875	.2325987	-5.44	0.000	-1.720761	-.8089903
_cons	3.386941	2.065274	1.64	0.101	-.6609226	7.434804

```
Average marginal effects                        Number of obs     =      7,207
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin	.0484278	.0167511	2.89	0.004	.0155962	.0812594

Annex XIV.A.b. Logistic model and margins to estimate treatment effect for outcome repetition for interaction of One million bags program participation and gender

```
note: 1.bagsbin#1.gender omitted because of collinearity
Iteration 0:   log pseudolikelihood = -3184.6984
Iteration 1:   log pseudolikelihood = -2714.3611
Iteration 2:   log pseudolikelihood = -2695.9119
Iteration 3:   log pseudolikelihood =  -2695.802
Iteration 4:   log pseudolikelihood =  -2695.802

Logistic regression                             Number of obs     =      7,207
                                                Wald chi2(19)     =    1100.16
                                                Prob > chi2       =     0.0000
Log pseudolikelihood =  -2695.802               Pseudo R2         =     0.1535

                                   (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin#gender						
0#Male	-.4054732	.1121285	-3.62	0.000	-.6252411	-.1857053
1#Female	.1937503	.1613747	1.20	0.230	-.1225383	.5100389
1#Male	0	(omitted)				
age	.1928876	.0219351	8.79	0.000	.1498956	.2358795
gender	.68917	.1368336	5.04	0.000	.4209811	.9573589
DAMPdec	-.021863	.0212038	-1.03	0.302	-.0634217	.0196956
dep_ratio	-.1917049	.1032218	-1.86	0.063	-.3940159	.0106061
hhsize	-.0329365	.0207533	-1.59	0.113	-.0736121	.0077391
hh_educ	-.2357973	.0628787	-3.75	0.000	-.3590374	-.1125572
schooldist	.0000214	.0000124	1.73	0.084	-2.90e-06	.0000457
accbath	-.0971162	.1312547	-0.74	0.459	-.3543707	.1601384
acckitch	.6928713	.2920096	2.37	0.018	.120543	1.2652
accWC	-.2557457	.1728539	-1.48	0.139	-.5945332	.0830418
longitude	.000011	3.11e-06	3.56	0.000	4.96e-06	.0000171
latitude	-.0000204	5.63e-06	-3.62	0.000	-.0000314	-9.37e-06
primary	-.4300884	.1363666	-3.15	0.002	-.6973621	-.1628147
primdip	-.3249901	.1520536	-2.14	0.033	-.6230098	-.0269704
region	-.1704163	.050923	-3.35	0.001	-.2702237	-.070609
worknow	.022962	.0776408	0.30	0.767	-.1292112	.1751352
martstatbin	-1.267001	.2316035	-5.47	0.000	-1.720936	-.8130666
_cons	3.370196	2.065551	1.63	0.103	-.6782103	7.418602

```
Average marginal effects                        Number of obs     =      7,207
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.bagsbin						
gender						
Female	.0276922	.0232797	1.19	0.234	-.0179351	.0733195
Male	.0675574	.0182576	3.70	0.000	.0317732	.1033416

Annex XIV.A.c. Logistic model and margins to estimate treatment effect for outcome repetition for interaction of One million bags program participation and cycle of study

```
note: 1.bagsbin#1.primary omitted because of collinearity
Iteration 0:   log pseudolikelihood = -3184.6984
Iteration 1:   log pseudolikelihood = -2714.7167
Iteration 2:   log pseudolikelihood = -2695.5626
Iteration 3:   log pseudolikelihood = -2695.4551
Iteration 4:   log pseudolikelihood = -2695.4551

Logistic regression                             Number of obs     =      7,207
                                                Wald chi2(19)     =    1088.30
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -2695.4551               Pseudo R2         =     0.1536

                                    (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin#primary						
0 1	-.3984876	.1273342	-3.13	0.002	-.6480581	-.1489171
1 0	.1462756	.111459	1.31	0.189	-.07218	.3647312
1 1	0	(omitted)				
age	.1934583	.0222413	8.70	0.000	.1498661	.2370505
gender	.3994869	.0823343	4.85	0.000	.2381146	.5608592
DAMPdec	-.0206442	.021175	-0.97	0.330	-.0621465	.0208581
dep_ratio	-.1913879	.1037266	-1.85	0.065	-.3946884	.0119125
hhsize	-.033158	.0208376	-1.59	0.112	-.073999	.007683
hh_educ	-.2374103	.0626006	-3.79	0.000	-.3601051	-.1147154
schooldist	.000022	.0000125	1.75	0.079	-2.57e-06	.0000465
accbath	-.0949067	.1300773	-0.73	0.466	-.3498535	.1600402
acckitch	.689882	.2898938	2.38	0.017	.1217007	1.258063
accWC	-.2463107	.1738784	-1.42	0.157	-.587106	.0944846
longitude	.0000109	3.08e-06	3.54	0.000	4.86e-06	.0000169
latitude	-.0000202	5.61e-06	-3.60	0.000	-.0000312	-9.20e-06
primary	-.1651229	.1505292	-1.10	0.273	-.4601547	.1299088
primdip	-.3291659	.1526316	-2.16	0.031	-.6283184	-.0300135
region	-.1691615	.0507675	-3.33	0.001	-.2686639	-.069659
worknow	.0171762	.0775777	0.22	0.825	-.1348733	.1692257
martstatbin	-1.268831	.2329625	-5.45	0.000	-1.72543	-.8122333
_cons	3.311676	2.060812	1.61	0.108	-.7274408	7.350793

```
Average marginal effects                        Number of obs     =      7,207
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.bagsbin						
primary						
0	.0257459	.0194787	1.32	0.186	-.0124317	.0639235
1	.0591532	.0191204	3.09	0.002	.0216779	.0966284

Annex XIV.B.a. Logistic model and margins to estimate treatment effect for outcome repetition for Canteens program participation

```
Iteration 0:   log pseudolikelihood =  -494.4583
Iteration 1:   log pseudolikelihood = -437.70072
Iteration 2:   log pseudolikelihood =  -437.4633
Iteration 3:   log pseudolikelihood = -437.46299
Iteration 4:   log pseudolikelihood = -437.46285
Iteration 5:   log pseudolikelihood = -437.46125
Iteration 6:   log pseudolikelihood =  -437.4604
Iteration 7:   log pseudolikelihood =  -437.4604

Logistic regression                               Number of obs     =      4,688
                                                  Wald chi2(18)     =    1125.66
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood =  -437.4604                 Pseudo R2         =     0.1153

                                  (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin	.0458075	.1401972	0.33	0.744	-.2289739	.320589
age	.0827129	.008561	9.66	0.000	.0659337	.0994921
gender	.4514821	.1534555	2.94	0.003	.1507148	.7522494
DAMPdec	.035657	.0176486	2.02	0.043	.0010664	.0702475
dep_ratio	-.4152688	.0964399	-4.31	0.000	-.6042874	-.2262502
hhsize	-.026892	.0189911	-1.42	0.157	-.0641138	.0103298
hh_educ	.0094698	.0876101	0.11	0.914	-.1622429	.1811825
schdistrange	.0037494	.0482498	0.08	0.938	-.0908184	.0983172
accbath	-.4453079	.1937268	-2.30	0.022	-.8250054	-.0656104
acckitch	-.0402849	.5519297	-0.07	0.942	-1.122047	1.041477
accWC	-.0291426	.2160674	-0.13	0.893	-.4526269	.3943417
longitude	4.13e-06	4.25e-06	0.97	0.331	-4.20e-06	.0000125
latitude	-8.05e-06	6.87e-06	-1.17	0.241	-.0000215	5.41e-06
primary	-1.281221	.1210734	-10.58	0.000	-1.51852	-1.043921
primdip	-.5158953	.132938	-3.88	0.000	-.776449	-.2553417
region	.0027819	.0472642	0.06	0.953	-.0898542	.095418
worknow	-.0034086	.0753471	-0.05	0.964	-.1510861	.1442689
martstatbin	-.5655516	.1272117	-4.45	0.000	-.814882	-.3162213
_cons	1.84429	2.467427	0.75	0.455	-2.991778	6.680357

```
Average marginal effects                          Number of obs     =      4,688
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : cantbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin	.0095556	.0292502	0.33	0.744	-.0477738	.066885

Annex XIV.B.b. Logistic model and margins to estimate treatment effect for outcome repetition for interaction of Canteens program participation and gender

```
note: 1.cantbin#1.gender omitted because of collinearity
Iteration 0:   log pseudolikelihood =  -494.4583
Iteration 1:   log pseudolikelihood = -437.59728
Iteration 2:   log pseudolikelihood = -437.36165
Iteration 3:   log pseudolikelihood = -437.36134
Iteration 4:   log pseudolikelihood =  -437.3612
Iteration 5:   log pseudolikelihood = -437.35966
Iteration 6:   log pseudolikelihood = -437.35879
Iteration 7:   log pseudolikelihood = -437.35879

Logistic regression                             Number of obs     =      4,688
                                                Wald chi2(19)     =    1280.55
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -437.35879               Pseudo R2         =     0.1155

                                   (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin#gender						
0#Male	-.1197134	.1673578	-0.72	0.474	-.4477287	.2083018
1#Female	-.0280527	.1892884	-0.15	0.882	-.3990512	.3429457
1#Male	0	(omitted)				
age	.0828765	.0085317	9.71	0.000	.0661547	.0995983
gender	.4990887	.2138141	2.33	0.020	.0800207	.9181566
DAMPdec	.0362784	.0176602	2.05	0.040	.0016651	.0708916
dep_ratio	-.4083802	.1005085	-4.06	0.000	-.6053733	-.2113871
hhsize	-.0263965	.0190488	-1.39	0.166	-.0637314	.0109384
hh_educ	.0059992	.0881811	0.07	0.946	-.1668325	.1788309
schdistrange	.003714	.0486052	0.08	0.939	-.0915504	.0989785
accbath	-.4467035	.1923831	-2.32	0.020	-.8237675	-.0696395
acckitch	-.0306929	.5639336	-0.05	0.957	-1.135982	1.074597
accWC	-.0259941	.2120074	-0.12	0.902	-.4415209	.3895326
longitude	4.25e-06	4.18e-06	1.02	0.309	-3.94e-06	.0000124
latitude	-8.02e-06	6.85e-06	-1.17	0.242	-.0000214	5.41e-06
primary	-1.279711	.1223527	-10.46	0.000	-1.519518	-1.039904
primdip	-.5159812	.1339076	-3.85	0.000	-.7784353	-.2535271
region	.0020453	.0472717	0.04	0.965	-.0906055	.0946961
worknow	-.0047947	.0750397	-0.06	0.949	-.1518699	.1422805
martstatbin	-.5656519	.1263642	-4.48	0.000	-.8133212	-.3179826
_cons	1.843245	2.460125	0.75	0.454	-2.978511	6.665

```
Average marginal effects                        Number of obs     =      4,688
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.cantbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.cantbin						
gender						
Female	-.0058062	.0391442	-0.15	0.882	-.0825274	.070915
Male	.0259989	.0365522	0.71	0.477	-.0456421	.09764

Annex XIV.B.c. Logistic model and margins to estimate treatment effect for outcome repetition for interaction of Canteens program participation and cycle of study

```
note: 1.cantbin#1.primary omitted because of collinearity
Iteration 0:   log pseudolikelihood =  -494.4583
Iteration 1:   log pseudolikelihood =  -436.5092
Iteration 2:   log pseudolikelihood = -436.28374
Iteration 3:   log pseudolikelihood = -436.28346
Iteration 4:   log pseudolikelihood = -436.28339
Iteration 5:   log pseudolikelihood = -436.28336
Iteration 6:   log pseudolikelihood = -436.28328
Iteration 7:   log pseudolikelihood = -436.28325

Logistic regression                             Number of obs     =      4,688
                                                Wald chi2(19)     =    1277.29
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -436.28325               Pseudo R2         =     0.1177

                                    (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin#primary						
0 1	.1509149	.1456213	1.04	0.300	-.1344976	.4363275
1 0	.3702642	.1929233	1.92	0.055	-.0078584	.7483868
1 1	0	(omitted)				
age	.0823783	.0084883	9.70	0.000	.0657414	.0990152
gender	.4498755	.1553217	2.90	0.004	.1454505	.7543005
DAMPdec	.0314412	.017512	1.80	0.073	-.0028817	.0657641
dep_ratio	-.4092152	.0960105	-4.26	0.000	-.5973923	-.2210381
hhsize	-.0270431	.0186132	-1.45	0.146	-.0635242	.0094381
hh_educ	.0106077	.0870747	0.12	0.903	-.1600555	.1812709
schdistrange	.0009302	.0469404	0.02	0.984	-.0910713	.0929316
accbath	-.4358311	.1934247	-2.25	0.024	-.8149367	-.0567256
acckitch	-.0079536	.5440514	-0.01	0.988	-1.074275	1.058368
accWC	-.045169	.2199364	-0.21	0.837	-.4762365	.3858984
longitude	4.11e-06	4.25e-06	0.97	0.334	-4.22e-06	.0000124
latitude	-8.85e-06	6.81e-06	-1.30	0.194	-.0000222	4.49e-06
primary	-1.165075	.171842	-6.78	0.000	-1.501879	-.8282711
primdip	-.4946753	.1298263	-3.81	0.000	-.7491301	-.2402204
region	-.0032702	.0464122	-0.07	0.944	-.0942365	.087696
worknow	-.0068178	.0750239	-0.09	0.928	-.1538619	.1402263
martstatbin	-.5536615	.1319367	-4.20	0.000	-.8122526	-.2950704
_cons	1.972294	2.453581	0.80	0.421	-2.836636	6.781223

```
Average marginal effects                        Number of obs     =      4,688
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.cantbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.cantbin						
primary						
0	.0798084	.0408273	1.95	0.051	-.0002117	.1598284
1	-.0304053	.029004	-1.05	0.294	-.0872521	.0264415

Annex XIV.C.a. Logistic model and margins to estimate treatment effect for outcome repetition for participation in both program

```
Iteration 0:   log pseudolikelihood = -755.81665
Iteration 1:   log pseudolikelihood = -633.84916
Iteration 2:   log pseudolikelihood =  -625.6792
Iteration 3:   log pseudolikelihood = -625.57729
Iteration 4:   log pseudolikelihood = -625.57711
Iteration 5:   log pseudolikelihood = -625.57711

Logistic regression                             Number of obs     =      5,010
                                                Wald chi2(18)     =    1544.83
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -625.57711               Pseudo R2         =     0.1723

                                   (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags	.7361252	.1453511	5.06	0.000	.4512423	1.021008
age	.2547127	.0217076	11.73	0.000	.2121667	.2972588
gender	.5006103	.1122956	4.46	0.000	.2805148	.7207057
DAMPdec	-.0119287	.0268142	-0.44	0.656	-.0644834	.0406261
dep_ratio	-.2058042	.1400522	-1.47	0.142	-.4803015	.0686931
hhsize	-.0619344	.0217788	-2.84	0.004	-.1046201	-.0192486
hh_educ	-.6931623	.106896	-6.48	0.000	-.9026747	-.48365
schdistrange	.0000467	.000023	2.04	0.042	1.73e-06	.0000917
accbath	-.3154924	.2141699	-1.47	0.141	-.7352577	.1042729
acckitch	.0664853	.2986318	0.22	0.824	-.5188222	.6517929
accWC	-.0286941	.1402105	-0.20	0.838	-.3035016	.2461135
longitude	3.34e-06	4.60e-06	0.72	0.469	-5.69e-06	.0000124
latitude	2.04e-06	7.68e-06	0.27	0.790	-.000013	.0000171
primary	-1.153589	.1856633	-6.21	0.000	-1.517482	-.7896954
primdip	-1.364025	.2520424	-5.41	0.000	-1.858019	-.8700312
region	.0468302	.0690587	0.68	0.498	-.0885225	.1821828
worknow	-.2901063	.1575488	-1.84	0.066	-.5988962	.0186835
martstatbin	-1.832249	.345967	-5.30	0.000	-2.510332	-1.154166
_cons	-3.8469	2.909439	-1.32	0.186	-9.549296	1.855496

```
Average marginal effects                        Number of obs     =      5,010
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : cantbags
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags	.111888	.0212857	5.26	0.000	.0701687	.1536072

Annex XIV.C.b. Logistic model and margins to estimate treatment effect for outcome repetition for interaction of participation in both program and gender

```
note: 1.cantbags#1.gender omitted because of collinearity
Iteration 0:   log pseudolikelihood = -755.81665
Iteration 1:   log pseudolikelihood = -633.56449
Iteration 2:   log pseudolikelihood = -625.45009
Iteration 3:   log pseudolikelihood = -625.34628
Iteration 4:   log pseudolikelihood =  -625.3461
Iteration 5:   log pseudolikelihood =  -625.3461

Logistic regression                             Number of obs     =      5,010
                                                Wald chi2(19)     =    2120.32
                                                Prob > chi2       =     0.0000
Log pseudolikelihood =  -625.3461               Pseudo R2         =     0.1726

                                    (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags#gender						
0#Male	-.826741	.2133026	-3.88	0.000	-1.244806	-.4086755
1#Female	.6308662	.1584138	3.98	0.000	.3203808	.9413516
1#Male	0	(omitted)				
age	.2546379	.0217074	11.73	0.000	.2120922	.2971836
gender	1.211078	.1943325	6.23	0.000	.8301937	1.591963
DAMPdec	-.0116528	.0267125	-0.44	0.663	-.0640084	.0407028
dep_ratio	-.2010246	.1423792	-1.41	0.158	-.4800828	.0780335
hhsize	-.0613166	.0218583	-2.81	0.005	-.1041581	-.0184752
hh_educ	-.6915212	.1058255	-6.53	0.000	-.8989354	-.4841071
schdistrange	.0000453	.000022	2.06	0.039	2.27e-06	.0000884
accbath	-.3173768	.2158976	-1.47	0.142	-.7405283	.1057747
acckitch	.0797341	.3019413	0.26	0.792	-.5120599	.6715282
accWC	-.0325467	.139623	-0.23	0.816	-.3062027	.2411093
longitude	3.55e-06	4.63e-06	0.77	0.443	-5.53e-06	.0000126
latitude	1.98e-06	7.75e-06	0.26	0.798	-.0000132	.0000172
primary	-1.144332	.1768742	-6.47	0.000	-1.490999	-.7976649
primdip	-1.355856	.248519	-5.46	0.000	-1.842944	-.8687675
region	.0450908	.0705148	0.64	0.523	-.0931157	.1832972
worknow	-.2828818	.1625133	-1.74	0.082	-.601402	.0356385
martstatbin	-1.830712	.3421696	-5.35	0.000	-2.501352	-1.160072
_cons	-3.795514	2.9421	-1.29	0.197	-9.561924	1.970897

```
Average marginal effects                        Number of obs     =      5,010
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.cantbags
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.cantbags						
gender						
Female	.0866314	.0204721	4.23	0.000	.0465068	.126756
Male	.1366115	.0329789	4.14	0.000	.071974	.2012489

Annex XIV.C.c. Logistic model and margins to estimate treatment effect for outcome repetition for interaction of participation in both program and cycle of study

```
note: 1.cantbags#1.primary omitted because of collinearity
Iteration 0:   log pseudolikelihood = -755.81665
Iteration 1:   log pseudolikelihood = -633.86472
Iteration 2:   log pseudolikelihood = -625.27649
Iteration 3:   log pseudolikelihood = -625.17121
Iteration 4:   log pseudolikelihood = -625.17104
Iteration 5:   log pseudolikelihood = -625.17104

Logistic regression                             Number of obs     =      5,010
                                                Wald chi2(19)     =    1643.32
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -625.17104               Pseudo R2         =     0.1729

                                    (Std. Err. adjusted for 50 clusters in town)
```

repetitionbin	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags#primary						
0 1	-.826498	.1737281	-4.76	0.000	-1.166999	-.4859972
1 0	.5463052	.1463397	3.73	0.000	.2594846	.8331258
1 1	0	(omitted)				
age	.2552396	.0219991	11.60	0.000	.2121222	.298357
gender	.5064056	.1129607	4.48	0.000	.2850067	.7278044
DAMPdec	-.0109806	.0265386	-0.41	0.679	-.0629953	.0410341
dep_ratio	-.2017915	.1379878	-1.46	0.144	-.4722425	.0686596
hhsize	-.062746	.0218893	-2.87	0.004	-.1056482	-.0198438
hh_educ	-.6910011	.1062842	-6.50	0.000	-.8993142	-.4826879
schdistrange	.0000446	.0000231	1.93	0.053	-6.59e-07	.0000899
accbath	-.3198931	.2124253	-1.51	0.132	-.7362391	.0964529
acckitch	.070088	.298183	0.24	0.814	-.5143399	.6545159
accWC	-.0148439	.1404485	-0.11	0.916	-.2901179	.2604302
longitude	3.19e-06	4.55e-06	0.70	0.483	-5.73e-06	.0000121
latitude	2.66e-06	7.68e-06	0.35	0.729	-.0000124	.0000177
primary	-.4681782	.1804689	-2.59	0.009	-.8218908	-.1144657
primdip	-1.351421	.2503618	-5.40	0.000	-1.842121	-.8607209
region	.0488155	.06916	0.71	0.480	-.0867356	.1843665
worknow	-.290648	.1546409	-1.88	0.060	-.5937386	.0124427
martstatbin	-1.830761	.3484602	-5.25	0.000	-2.513731	-1.147792
_cons	-3.986639	2.896695	-1.38	0.169	-9.664058	1.690779

```
Average marginal effects                          Number of obs     =      5,010
Model VCE    : Robust

Expression   : Pr(repetitionbin), predict()
dy/dx w.r.t. : 1.cantbags
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.cantbags						
primary						
0	.1086774	.0297975	3.65	0.000	.0502753	.1670794
1	.1113417	.0220819	5.04	0.000	.068062	.1546215

Annex XIV.D.a. Logistic model and margins to estimate treatment effect for outcome transition to college for One million bags program participation

```
Iteration 0:   log pseudolikelihood = -897.98324
Iteration 1:   log pseudolikelihood = -822.03839
Iteration 2:   log pseudolikelihood = -816.97489
Iteration 3:   log pseudolikelihood = -816.96444
Iteration 4:   log pseudolikelihood = -816.96444

Logistic regression                             Number of obs     =      2,489
                                                Wald chi2(16)     =     252.52
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -816.96444               Pseudo R2         =     0.0902

                                   (Std. Err. adjusted for 50 clusters in town)
```

transit	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin	.050926	.2303882	0.22	0.825	-.4006266	.5024787
age	.0107165	.0374276	0.29	0.775	-.0626403	.0840733
gender	1.07485	.1353288	7.94	0.000	.8096104	1.34009
DAMPdec	.045846	.0469437	0.98	0.329	-.046162	.1378539
dep_ratio	-.2261537	.1779195	-1.27	0.204	-.5748695	.1225622
hhsize	-.0051708	.0251529	-0.21	0.837	-.0544696	.044128
hh_educ	.2751808	.1228505	2.24	0.025	.0343983	.5159633
schooldist	-.0000334	.0000146	-2.29	0.022	-.000062	-4.79e-06
accbath	-.2071868	.2151951	-0.96	0.336	-.6289614	.2145879
accWC	.3445219	.5121002	0.67	0.501	-.659176	1.34822
acckitch	.6667443	.5988953	1.11	0.266	-.5070688	1.840557
longitude	-.0000187	9.11e-06	-2.06	0.040	-.0000366	-8.66e-07
latitude	-.0000352	.0000117	-3.02	0.003	-.0000581	-.0000124
region	-.0180697	.0805336	-0.22	0.822	-.1759126	.1397732
worknow	-.2942538	.1169262	-2.52	0.012	-.523425	-.0650826
martstatbin	-.8992997	.3098034	-2.90	0.004	-1.506503	-.2920963
_cons	12.62256	4.260793	2.96	0.003	4.271565	20.97356

```
Average marginal effects                        Number of obs     =      2,489
Model VCE    : Robust

Expression   : Pr(transit), predict()
dy/dx w.r.t. : bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin	.0080769	.036498	0.22	0.825	-.0634578	.0796117

Annex XIV.D.b. Logistic model and margins to estimate treatment effect for outcome transition to college for interaction of One million bags program participation and gender

```
note: 1.bagsbin#1.gender omitted because of collinearity
Iteration 0:   log pseudolikelihood = -897.98324
Iteration 1:   log pseudolikelihood = -821.38925
Iteration 2:   log pseudolikelihood = -816.10588
Iteration 3:   log pseudolikelihood = -816.09532
Iteration 4:   log pseudolikelihood = -816.09532

Logistic regression                             Number of obs     =      2,489
                                                Wald chi2(17)     =     249.10
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -816.09532               Pseudo R2         =     0.0912

                                    (Std. Err. adjusted for 50 clusters in town)
```

transit	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
bagsbin#gender						
0#Male	-.2421339	.2356199	-1.03	0.304	-.7039403	.2196726
1#Female	-.0880544	.290752	-0.30	0.762	-.657918	.4818091
1#Male	0	(omitted)				
age	.0108133	.037494	0.29	0.773	-.0626736	.0843002
gender	1.154989	.2771144	4.17	0.000	.611855	1.698123
DAMPdec	.0440233	.0472659	0.93	0.352	-.0486161	.1366628
dep_ratio	-.2259712	.1769688	-1.28	0.202	-.5728237	.1208812
hhsize	-.0041816	.0248886	-0.17	0.867	-.0529625	.0445992
hh_educ	.279342	.124008	2.25	0.024	.0362907	.5223933
schooldist	-.0000342	.0000143	-2.40	0.017	-.0000622	-6.25e-06
accbath	-.2120426	.2177529	-0.97	0.330	-.6388303	.2147452
accWC	.3278339	.5055445	0.65	0.517	-.663015	1.318683
acckitch	.6679501	.5878267	1.14	0.256	-.484169	1.820069
longitude	-.0000188	9.13e-06	-2.06	0.039	-.0000367	-9.09e-07
latitude	-.0000351	.0000117	-3.01	0.003	-.000058	-.0000123
region	-.0172055	.0802291	-0.21	0.830	-.1744517	.1400406
worknow	-.291285	.115179	-2.53	0.011	-.5170317	-.0655382
martstatbin	-.9023883	.3092274	-2.92	0.004	-1.508463	-.2963138
_cons	12.68188	4.283402	2.96	0.003	4.286563	21.07719

```
Average marginal effects                          Number of obs     =      2,489
Model VCE    : Robust

Expression   : Pr(transit), predict()
dy/dx w.r.t. : 1.bagsbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.bagsbin						
gender						
Female	-.0179875	.0594009	-0.30	0.762	-.1344111	.0984362
Male	.0292129	.0280569	1.04	0.298	-.0257776	.0842034

Annex XIV.E.a. Logistic model and margins to estimate treatment effect for outcome transition to college for Canteens program participation

```
Iteration 0:   log pseudolikelihood = -132.11628
Iteration 1:   log pseudolikelihood = -116.85013
Iteration 2:   log pseudolikelihood = -115.96813
Iteration 3:   log pseudolikelihood = -115.96266
Iteration 4:   log pseudolikelihood = -115.96266

Logistic regression                             Number of obs     =      1,456
                                                Wald chi2(16)     =     523.94
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -115.96266               Pseudo R2         =     0.1223

                                   (Std. Err. adjusted for 49 clusters in town)
```

transit	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin	.0232256	.2735477	0.08	0.932	-.5129181	.5593693
age	-.1284536	.0238912	-5.38	0.000	-.1752794	-.0816277
gender	1.097658	.2515662	4.36	0.000	.6045973	1.590719
DAMPdec	-.0449117	.0381963	-1.18	0.240	-.119775	.0299516
dep_ratio	-.3692745	.3046291	-1.21	0.225	-.9663365	.2277875
hhsize	-.0336922	.0325551	-1.03	0.301	-.097499	.0301146
hh_educ	.092766	.1129558	0.82	0.411	-.1286234	.3141554
schdistrange	.0421447	.0825176	0.51	0.610	-.1195868	.2038762
accbath	-.060041	.3810814	-0.16	0.875	-.8069468	.6868648
accWC	-.0982107	.417449	-0.24	0.814	-.9163957	.7199742
acckitch	-.04205	.6227583	-0.07	0.946	-1.262634	1.178534
longitude	-.0000437	.000014	-3.12	0.002	-.0000712	-.0000163
latitude	.0000202	.0000141	1.43	0.152	-7.41e-06	.0000478
region	.5057126	.1212683	4.17	0.000	.2680312	.743394
worknow	-.1175101	.1039825	-1.13	0.258	-.3213121	.0862919
martstatbin	-.2508472	.2566291	-0.98	0.328	-.7538309	.2521366
_cons	-2.647057	4.747079	-0.56	0.577	-11.95116	6.657046

```
Average marginal effects                        Number of obs     =      1,456
Model VCE    : Robust

Expression   : Pr(transit), predict()
dy/dx w.r.t. : cantbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin	.0039689	.0466524	0.09	0.932	-.087468	.0954058

Annex XIV.E.b. Logistic model and margins to estimate treatment effect for outcome transition to college for interaction of Canteens program participation and gender

```
note: 1.cantbin#1.gender omitted because of collinearity
Iteration 0:   log pseudolikelihood = -132.11628
Iteration 1:   log pseudolikelihood = -116.83603
Iteration 2:   log pseudolikelihood = -115.95447
Iteration 3:   log pseudolikelihood = -115.94913
Iteration 4:   log pseudolikelihood = -115.94913

Logistic regression                             Number of obs     =      1,456
                                                Wald chi2(17)     =     547.13
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -115.94913               Pseudo R2         =     0.1224

                                     (Std. Err. adjusted for 49 clusters in town)
```

transit	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbin#gender						
0#Male	-.0862816	.3023009	-0.29	0.775	-.6787805	.5062172
1#Female	-.0221873	.3810205	-0.06	0.954	-.7689737	.7245992
1#Male	0	(omitted)				
age	-.1282455	.0233891	-5.48	0.000	-.1740873	-.0824038
gender	1.127317	.2869061	3.93	0.000	.5649918	1.689643
DAMPdec	-.044578	.0384469	-1.16	0.246	-.1199325	.0307764
dep_ratio	-.3639552	.3056788	-1.19	0.234	-.9630747	.2351643
hhsize	-.0334486	.0322253	-1.04	0.299	-.096609	.0297117
hh_educ	.0919904	.1074215	0.86	0.392	-.1185518	.3025326
schdistrange	.0418627	.0834905	0.50	0.616	-.1217757	.2055012
accbath	-.0579893	.3862987	-0.15	0.881	-.8151209	.6991423
accWC	-.096417	.4197227	-0.23	0.818	-.9190583	.7262244
acckitch	-.0318771	.6413175	-0.05	0.960	-1.288836	1.225082
longitude	-.0000436	.0000137	-3.18	0.001	-.0000705	-.0000168
latitude	.0000201	.0000142	1.41	0.158	-7.80e-06	.000048
region	.5045445	.1219719	4.14	0.000	.2654839	.7436051
worknow	-.1146617	.1093558	-1.05	0.294	-.3289952	.0996718
martstatbin	-.2512327	.2530722	-0.99	0.321	-.747245	.2447796
_cons	-2.625405	4.811811	-0.55	0.585	-12.05638	6.805573

```
Average marginal effects                        Number of obs     =      1,456
Model VCE    : Robust

Expression   : Pr(transit), predict()
dy/dx w.r.t. : 1.cantbin
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.cantbin						
gender						
Female	-.0046236	.0795693	-0.06	0.954	-.1605766	.1513293
Male	.0118648	.0411447	0.29	0.773	-.0687773	.092507

Annex XIV.F.a. Logistic model and margins to estimate treatment effect for outcome transition to college for participation in both programs

```
Iteration 0:   log pseudolikelihood = -192.93023
Iteration 1:   log pseudolikelihood = -169.78654
Iteration 2:   log pseudolikelihood = -167.88968
Iteration 3:   log pseudolikelihood = -167.86115
Iteration 4:   log pseudolikelihood = -167.86113

Logistic regression                             Number of obs     =      2,093
                                                Wald chi2(16)     =     402.77
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -167.86113               Pseudo R2         =     0.1299

                                  (Std. Err. adjusted for 50 clusters in town)
```

transit	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags	.2173634	.3439223	0.63	0.527	-.4567119	.8914387
age	-.0527383	.0371972	-1.42	0.156	-.1256434	.0201668
gender	1.488251	.2347924	6.34	0.000	1.028066	1.948436
DAMPdec	.0108308	.0461371	0.23	0.814	-.0795963	.1012579
dep_ratio	-.2395802	.1957411	-1.22	0.221	-.6232258	.1440654
hhsize	-.0548973	.0341435	-1.61	0.108	-.1218173	.0120226
hh_educ	.3872262	.2828127	1.37	0.171	-.1670765	.9415289
schooldist	.0000349	.0000397	0.88	0.379	-.0000429	.0001126
accbath	-.4145171	.1995835	-2.08	0.038	-.8056935	-.0233407
accWC	.6785028	.3554758	1.91	0.056	-.0182169	1.375222
acckitch	.2323664	.7465863	0.31	0.756	-1.230916	1.695649
longitude	-.0000167	.0000104	-1.61	0.107	-.0000371	3.61e-06
latitude	-.0000228	.0000152	-1.50	0.134	-.0000526	7.03e-06
region	.0519001	.1043807	0.50	0.619	-.1526824	.2564825
worknow	-.3780442	.1732327	-2.18	0.029	-.717574	-.0385144
martstatbin	-.9364547	.3657233	-2.56	0.010	-1.653259	-.2196503
_cons	9.550581	5.321527	1.79	0.073	-.8794206	19.98058

```
Average marginal effects                        Number of obs     =      2,093
Model VCE    : Robust

Expression   : Pr(transit), predict()
dy/dx w.r.t. : cantbags
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags	.0316114	.0498528	0.63	0.526	-.0660983	.129321

Annex XIV.F.b. Logistic model and margins to estimate treatment effect for outcome transition to college for interaction of participation in both program and gender

```
note: 1.cantbags#1.gender omitted because of collinearity
Iteration 0:   log pseudolikelihood = -192.93023
Iteration 1:   log pseudolikelihood = -168.28898
Iteration 2:   log pseudolikelihood = -165.69117
Iteration 3:   log pseudolikelihood = -165.63082
Iteration 4:   log pseudolikelihood = -165.63077
Iteration 5:   log pseudolikelihood = -165.63077

Logistic regression                             Number of obs     =      2,093
                                                Wald chi2(17)     =     468.76
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -165.63077               Pseudo R2         =     0.1415

                                   (Std. Err. adjusted for 50 clusters in town)
```

transit	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
cantbags#gender						
0#Male	-1.067333	.425207	-2.51	0.012	-1.900723	-.2339429
1#Female	-.2041081	.4325133	-0.47	0.637	-1.051819	.6436023
1#Male	0	(omitted)				
age	-.0519074	.036002	-1.44	0.149	-.12247	.0186553
gender	2.035934	.4342145	4.69	0.000	1.184889	2.886979
DAMPdec	.0150231	.0462747	0.32	0.745	-.0756736	.1057199
dep_ratio	-.1878625	.2117879	-0.89	0.375	-.6029591	.2272341
hhsize	-.0504373	.0389165	-1.30	0.195	-.1267122	.0258377
hh_educ	.3840436	.291634	1.32	0.188	-.1875484	.9556357
schooldist	.0000263	.0000343	0.76	0.445	-.0000411	.0000936
accbath	-.4098978	.2115195	-1.94	0.053	-.8244684	.0046728
accWC	.646744	.3496505	1.85	0.064	-.0385583	1.332046
acckitch	.4167002	.6945958	0.60	0.549	-.9446827	1.778083
longitude	-.0000155	.0000104	-1.49	0.137	-.000036	4.92e-06
latitude	-.0000196	.0000156	-1.25	0.211	-.0000502	.0000111
region	.0621022	.1076381	0.58	0.564	-.1488647	.273069
worknow	-.3650811	.1521076	-2.40	0.016	-.6632064	-.0669557
martstatbin	-.897895	.3608055	-2.49	0.013	-1.605061	-.1907291
_cons	8.332912	5.503566	1.51	0.130	-2.45388	19.1197

```
Average marginal effects                          Number of obs     =      2,093
Model VCE    : Robust

Expression   : Pr(transit), predict()
dy/dx w.r.t. : 1.cantbags
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.cantbags						
gender						
Female	-.0418603	.0892187	-0.47	0.639	-.2167258	.1330052
Male	.0940971	.0344986	2.73	0.006	.0264811	.1617132