REPORT

FINAL REPORT

Design Report for the Impact and Performance Evaluation of the Education Quality Activity

December 19, 2018 (revised October, 24 2018)
(first version December 13, 2016)

Larissa Campuzano
Ivonne Padilla Espinosa
Camila Fernández

Submitted to:
Millennium Challenge Corporation
875 Fifteenth Street NW
Washington, DC 20005-2221
Project Officer: Anne Pizer
Contract Number: MCC-13-BPA-0040/MCC-14-CL-0003

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005
Project Director: Larissa Campuzano
Reference Number: 40349.HR4

This page has been left blank for double-sided copying.

CONTENTS

LIST OF ACRONYMS vii

I. INTRODUCTION 1

II. OVERVIEW OF THE EDUCATION QUALITY ACTIVITY 3

A. Background of the implementation 3

1. Strengthening the National Education System 3

2. The SI-EITP model 3

B. Updates in the implementation of the Education Quality Activity 5

C. Logic model 12

D. Evidence base for the SI-EITP model 15

1. System- or community-level interventions 15

2. School-level interventions 16

3. Teacher-level interventions 19

E. Gaps in the literature and contribution of the impact evaluation of the SI-EITP model 21

F. Economic rate of return 21

III. RESEARCH QUESTIONS AND EVALUATION DESIGN 23

A. Impact evaluation 23

1. Random assignment of potential Integrated Systems 23

2. Advantages of the experimental design 25

B. Performance evaluation 26

IV. DATA SOURCES AND OUTCOME DEFINITIONS 29

A. Key indicators for the impact evaluation of the SI-EITP 29

B. Data sources for the evaluation of the Education Quality Activity 33

1. Impact evaluation of the SI-EITP model 33

2. Performance evaluation of the Education Quality Activity 36

3. Sample sizes for the performance evaluation 41

4. Data collection schedule and implementation milestones 42

V. ANALYSIS PLAN 43

A. Estimating the impact of the SI-EITP model 43

1. Statistical power 44

2. Limitations and challenges of the impact evaluation of the SI-EITP model 48

B. Analytic approach for the performance evaluation 50

VI. ADMINISTRATIVE ASPECTS 51

A. Summary of data sources and IRB requirements 51

B. Preparing data files for access, privacy, and documentation 51

C. Dissemination plan 52

D. Evaluation team roles and responsibilities 52

E. Evaluation timeline and reporting schedule 52

REFERENCES 55

APPENDIX A SAMPLE SIZES AND TIMELINE PROPOSED 59

TABLES

Table II.1. Training offered to teachers in basic and secondary schools 6

Table II.2. Planned activities, targets, and objectives of the Education Quality Activity 10

Table II.3. MCC's assumed initial and projected values of the parameters used in the ERR calculation 22

Table III.1. Number of systems with and without a secondary school, by department and study group 25

Table IV.1. Descriptions and data sources of outcome indicators for the impact evaluation of the SI-EITP 31

Table IV.2. Research questions for the performance evaluation of the Education Quality Activity, and data sources by program component 38

Table IV.3. Sample sizes for the performance evaluation 41

Table V.1. Baseline values, targets, changes and minimum detectable impact (MDI) for selected outcomes (percentage points, unless otherwise indicated) 46

Table V.2. Minimum detectable impacts (MDIs) for full student sample or subgroups of half the students per Integrated System (percentage points, unless otherwise indicated) 48

Table A.1. Sample sizes and timeline proposed 61

FIGURES

Figure II.1. Geographic distribution of the Integrated Systems participating in implementation of the SI-EITP model of FOMILENIO II, by department 5

Figure II.2. Logic model for the Quality of Education Activity 14

Figure III.1. Random assignment of potential systems 24

Figure IV.1. SI-EITP evaluation data collection schedule, 2017–2025 42

This page has been left blank for double-sided copying.

LIST OF ACRONYMS

AENOR	Asociación Española de Normalización y Certificación (Spanish Association for Standardization and Centralization)
CDE	Consejo Directivo Escolar
CEI	Centro Educativo Integral
DDEs	Dirección Departamental Educativa (Department's Education Offices)
FOMILENIO II	Millennium Challenge Fund of El Salvador
GoES	Government of El Salvador
HML	Health Media Labs
ICC	Intra-class correlation
IRB	Institutional Review Board
LLECE	Latin American Laboratory for Assessment of the Quality of Education
MCC	Millennium Challenge Corporation
MDI	Minimum detectable impact
M&E	Monitoring and evaluation
MINED	Salvadoran Ministry of Education
NIE	Número de Identificación del Estudiante
OCCSI	Organismo Colegiado de Coordinación del Sistema Integrado
PAES	Prueba de Aprendizaje y Aptitudes para Egresados de Educación Media
PfG	Partnership for Growth
PNFM	Programa Nacional de Formación de Maestros
RCT	Randomized controlled trial
SAE	Sistema de Administración Escolar (School Information Management) System
SIGES	Sistema de Información para la Gestión Educativa Salvadoreña (Salvadoran Management Educational Information System)

SIRAI	Sistema Institutional de Registro Académico e Institucional (Academic Registry System for Secondary Education)
SBM	School-based management
TVET	Technical Vocational Education Training
TEAs	Technical education assistants
TIMSS	Trends in International Mathematics and Science Studies
USAID	United States Agency for International Development

I. INTRODUCTION

El Salvador's economic performance has been low in the past few years. From 2010 to 2014, the Salvadoran economy averaged just 2 percent annual economic growth—below the Latin American average. Under the Partnership for Growth (PfG) initiative, a joint U.S.-El Salvador technical team conducted a constraints analysis and identified two binding constraints to economic growth in El Salvador: (1) crime and insecurity, and (2) low productivity in the tradable sector (Joint U.S.-GOES technical team 2011).

To address these challenges, the Millennium Challenge Corporation (MCC) has partnered with the government of El Salvador (GoES) with the primary goal of promoting economic growth through large-scale investments. On September 30, 2014, MCC signed a US$267 million Compact with the GoES that is designed to increase productivity by improving the investment climate, strengthening human capital, and reducing the costs of transportation and logistics. Through intensive policy reforms and an integrated set of investments in the human, physical, and institutional capital of El Salvador, MCC and GoES expect the second Compact to help set the foundation for lasting economic growth and poverty reduction in El Salvador.

The Human Capital Project is one of three projects in the second Compact. The other two are the Investment Climate Project, which is focused on improving El Salvador's regulatory environment and institutional capacity and providing key public services in partnership with the private sector; and the Logistical Infrastructure Project, which is focused on reducing logistical and transportation costs and increasing investment and productivity in the trade of goods and services. The Millennium Challenge Fund of El Salvador (FOMILENIO II) is an entity established by the GoES that will use MCC funds to implement this Compact from September 2015 to September 2020.

The Human Capital Project, with total funding of $100.2 million, focuses on improving the quality of education and better matching the supply of skills to the demands of the labor market in El Salvador. The Human Capital Project consists of two activities:

1. The Education Quality Activity, with funding of nearly $85 million, included two sub-activities intended to improve the quality of the national education system:

The Strengthening of the National Education System sub-Activity, which is designed to improve the effectiveness and quality of the national education system in El Salvador

The Full-Time Inclusive Model sub-Activity ("SI-EITP" model, for its initials in Spanish), which is designed to strengthen and expand the implementation of a full-time inclusive school model in the Coastal Region of El Salvador

2. The Technical Vocational Education Training (TVET) System Reform Activity, with funding of nearly 15 million, is designed to strengthen the national TVET governance system and enhance its capacity to harmonize the skills supplied through education and training providers with the skills demanded by the labor market.

MCC has contracted with Mathematica Policy Research to conduct an evaluation of the activities under the Human Capital Project. Mathematica will conduct an impact evaluation of the SI-EITP model funded by MCC. We designed the impact evaluation in 2016, before

FOMILENIO II had completed the design of the intervention. As the implementation plans of the Education Quality Activity have been completed, we have updated the design to include a performance evaluation of the Activity. In this report, we present the updated evaluation design of the Education Quality Activity, including the design of both an impact and a performance evaluation.[1]

The report has six chapters. In Chapter II, we describe the interventions implemented under the Education Quality Activity, and report on a literature review of the evidence of effectiveness of similar interventions. We outline the research questions that the impact and the performance evaluations seek to answer and describe the design of both evaluation components in Chapter III. In Chapter IV, we outline the key indicators for the impact evaluation and discuss the data sources required for the impact and performance evaluations. In Chapter V, we describe the analysis plans for the impact and the performance evaluations of the Education Quality Activity. Finally, Chapter VI covers a number of administrative issues related to the evaluation, including the Institutional Review Board (IRB) requirements, data file preparation, dissemination plan, evaluation team, and timeline.

[1] We will also conduct a performance evaluation of the TVET System Reform Activity. The design of that activity will be discussed in a different report.

II. OVERVIEW OF THE EDUCATION QUALITY ACTIVITY

We begin this chapter by describing the original plans for the Education Quality Activity. We go on to summarize how implementation has evolved, then show the program logic and the available evidence on the effectiveness of some components of the Education Quality Activity. We finish by describing the ex-ante economic rate of return (ERR) that MCC calculated to compare the costs and expected benefits of the entire Human Capital Project, which includes the Educational Quality Activity and the TVET System Reform Activity.

A. Background of the implementation

Initially, the Education Quality Activity was to be implemented as two sub-activities:

1. Strengthening El Salvador's National Education System with the goal of improving its effectiveness and quality
2. Implementing the Full-Time Inclusive Model (SI-EITP model, for its initials in Spanish), which was designed to strengthen and expand the implementation of a full-time inclusive school model in the Coastal Region of El Salvador

1. Strengthening the National Education System

The Strengthening the National Education System sub-activity was planned to be implemented at the national level and to have the following key elements: (1) establishing the National Curriculum Evaluation Commission, (2) giving teachers across the nation professional development and training to enable them to provide quality academic education, (3) improving the measurement and use of national student learning and achievement data, and (4) developing a gender policy for the Ministry of Education (MINED).

2. The SI-EITP model

The implementation of the SI-EITP model focuses on the eight departments of the Coastal Region in El Salvador. The model organizes neighboring schools of all grade levels into a cohesive system (or cluster) in which representatives from the schools in the system work together to develop joint action plans, optimize and share resources, exchange expertise, and foster the involvement of families and the community. In the text box below we provide a brief description of the levels in El Salvador's education system. The overarching goal of the model is to expand and improve the learning opportunities offered to children in all primary and secondary grades in each system's geographic territory, and ultimately to improve students' labor market outcomes. Specifically, the SI-EITP model is designed to (1) increase enrollment, continuation, and completion of third cycle (grades 7 to 9) and secondary (grades 10 to 12) education; (2) extend the length of the school day from 25 hours per week to at least 36 hours (and up to 40) per week in grades 7 to 9; and (3) enhance students' development through the use of active learning methods in core subjects and the use of information technology; and (4) improve school management and sharing resources across schools in the Integrated Systems.

The SI-EITP model is based on a school-level model that MINED, with the support of the Italian government, began implementing in 2005. The original school-level model, Escuela Integrada de Tiempo Pleno (EITP), focused on individual schools with the goal of including

children with special needs and vulnerable children who have been excluded from the public schools. The model also promoted increased time in school. Later, MINED decided to implement this model at the system or cluster level (that is, in groups of schools) to better address some problems—for example, constraints in enrollment capacity in third cycle and secondary grades, and too many small schools offering primary education. By organizing neighboring schools into Integrated Systems that work together and share resources, including infrastructure, the SI-EITP model is used to balance the demand and supply for enrollment and to increase the feasibility of implementing longer school days, which were a key component of the original EITP model.

Levels in El Salvador's Education System:

Basic education: Offered to students aged 7 to 15. It comprises nine grades divided into three cycles.

- First cycle: comprises grades 1 to 3
- Second cycle: comprises grades 4 to 6
- Third cycle: comprises grades 7 to 9

Secondary education: Offers two programs. The general program has a duration of two years with 40 hours of classes per week. The technical program has a duration of three years with 44 hours of classes per week during the first two years and 30 hours per week on the third year that focuses on technical areas such as mechanics, nursing, accounting.

- General baccalaureate programs: comprises grades 10 to 11
- Technical baccalaureate program: comprises grades 10 to 12

In 2011, with support from the United States Agency for International Development (USAID), MINED piloted the SI-EITP model. Neighboring schools were reorganized into clusters or systems and worked jointly to implement the EITP model across the schools in the Integrated System (USAID 2012). The SI-EITP model was implemented in three municipalities of El Salvador under the pilot (Sonsonate, Zaragoza, and Nueva Granada). In 2012, MINED created the Adjoint Directorate for SI-EITP, which is responsible for the systematization, implementation, and expansion of the SI-EITP model across the country. This expansion started in 2013 and was financed with a loan from the World Bank. In 2015, changes in MINED leadership led to some modifications in the implementation of the model (Crespin 2015).

With MCC funding, FOMILENIO II will implement the SI-EITP model in eight departments covering 33 municipalities of the Coastal Region. FOMILENIO II has estimated that funds are available to implement the SI-EITP model in 45 of 147 potential Integrated Systems identified by MINED in the Coastal Region (see Figure II.1). FOMILENIO II's resources will mainly focus on education at third cycle (grades 7–9) and on general and technical secondary schools (grades 10–12). Each Integrated System will have a school designated as *"centro educativo integral"* (CEI)— a school with enhanced infrastructure and resources that will be accessible to all the other schools in the Integrated System.

Figure II.1. Geographic distribution of the Integrated Systems participating in implementation of the SI-EITP model of FOMILENIO II, by department

B. Updates in the implementation of the Education Quality Activity

Implementation plans for the Education Quality Activity began in 2015. FOMILENIO II's implementation plans for the SI-EITP model were based on the pilot and early experience during the model's expansion. In 2016 and 2017, FOMILENIO II conducted characterization studies in the 45 Integrated Systems that were selected to implement the SI-EITP model. The objective of the characterization studies was to describe current conditions of the schools, assess the needs of each system, and suggest relevant interventions for the Integrated Systems. FOMILENIO II used the findings from these studies[2] to finalize the implementation plans of the Education Quality Activity. During 2017, FOMILENIO II published the terms of reference to hire implementers for all the activities planned under the Education Quality Activity. At the beginning of 2018, FOMILENIO II contracted the entities that were selected to implement the components of the Education Quality Activity.

Some of the original implementation plans were adjusted in response to the characterization studies and changing priorities in MINED. For example, professional development and training, originally thought of as a national-level activity, was modified to focus on training teachers and principals from the 45 Integrated Systems selected for implementation, with the understanding that MINED will use this experience as models for the integrated systems in the rest of the country. As a result, the distinction between the two sub-activities became less clear, and stakeholders started referring to them as one activity, the Education Quality Activity. In this report, we describe the updated implementation plans for the Education Quality Activity, which

[2] The characterization reports of the Integrated Systems are available on FOMILENIO II's site: https://www.fomilenioii.gob.sv/los-informes-de-caracterizacion-de-los-si-eitp.

has 11 components. The components have elements that will be implemented at several levels. Elements of some components will be implemented both nationwide and in the systems; other components will be implemented only at the national level; and still others will be implemented only in the Integrated Systems financed by FOMILENIO II. Below, we describe the main activities in each component (Table II.2 provides more information).

Component 1: Professional development for specialists and teachers. The main goal of this component is to strengthen the content and pedagogical knowledge as well as the technological and social competencies of teachers in schools within the 45 Integrated Systems selected for implementation. The intervention will first provide training programs for specialists (teacher trainers) on the SI-EITP model, integrative methodologies, and evaluation of learning. Training for teachers of children in grades 7 to 9 is focused on content knowledge and will follow the national teacher development program (*Programa Nacional de Formación de Maestros*, or PNFM, see table II.1). But it will be provided by the trained specialists, so it will also teach integrative methodologies and evaluation of learning. Furthermore, as part of this intervention, the training modules of the six content areas of the PNFM will be re-designed as needed. Teachers of children in grades 1 to 6 will be offered training in language, mathematics, and strategies to teach multi-grade classrooms. Teachers of children in grades 1 through 12 will be offered training on socioemotional development skills, technological literacy, and gender inclusion. Teachers' participation in trainings is voluntary, but the intervention will offer a schedule that will make it easier for them to attend. This component also includes the design and implementation of workshops and technical assistance focused on improving principals' management skills.

Table II.1. Training offered to teachers in basic and secondary schools

	First and second cycle teachers (grades 1 to 6)	Third cycle teachers (grades 7 to 9)
Training in content knowledge for six content areas: mathematics, language, biology, chemistry, physics, and social science.		X
Training in language, mathematics, and strategies to teach multi-grade classrooms	X	
Training on socioemotional development skills, technological literacy, and gender inclusion.	X	X

Component 2: Strengthening pedagogical technical assistance and school management. MINED has designated technical education assistants (TEAs) in all the schools nationwide. TEAs are responsible for giving pedagogical and management assistance to schools. However, the effective time allocated for giving this assistance is limited because of the workload in many schools. The main goal of this component is to strengthen and provide timely and effective technical assistance on pedagogy and management for the schools in the Integrated Systems. To this end, 30 TEAs will be hired to focus on the 45 Integrated Systems selected for implementation.

At the national level, this component is designed to strengthen the management skills of MINED's central office staff and the staff of the 14 Department's Education Offices (DDEs for its acronym in Spanish). A training program will be offered to staff from MINED and the DDEs. The content of the training will focus on strategies to strengthen communication and integration between the central and departmental offices. In addition, team-building workshops designed to improve interpersonal relations and strengthen teamwork and communication will be offered to the TEAs from the 14 DDEs. These workshops will include strategic planning focused on results and collaborative work. The workshops methods include classroom hours and practice in the workplace.

Component 3: Governance and participation of SI-EITP. All schools in El Salvador have an educational council (Consejo Directivo Escolar, CDE), comprised of the principal, two teachers, three parents, and two students. The council's role is to plan, organize, administer, and supervise the resources allocated to the school in order to give students a quality education. However, the CDEs have not always been involved in decision making as envisioned. In addition to the CDE, the SI-EITP model will establish a school council that includes the principal and some teachers; the council will be in charge of developing a pedagogical proposal for the school, taking into account the plans of the other schools in the Integrated System. The pedagogical proposal describes of the adaptations the school will make to the national curricula along with their objectives. It describes the decisions taken by the school on what, how and when to teach and evaluate.[3]

The SI-EITP model also calls for resource sharing across the schools in the Integrated System and for greater involvement of teachers, principals, parents, and the local community in the education of students. The goal of this component is to contribute to the governance of the Integrated Systems by empowering existing and new councils and strengthening the capacities of the members. School management, leadership, and community participation workshops will be offered to (1) school management councils, (2) teacher councils, (3) technical education assistants, (4) student councils, and (5) parent councils.

Component 4: Strengthening MINED's technical capabilities in terms of learning and curricular evaluation. Besides strengthening the technical capabilities of MINED on curricular and learning evaluation, this component is also focused on facilitating student participation in international standardized tests at the national level and on reviewing and redesigning the language and literature curriculum and materials for grades 1–11.[4] To this end, training will be offered to specialists and MINED staff in areas associated with the application of standardized tests, such as evaluation models, preparation of items (questions for the test), development of tests, and communication of results. In addition, bibliographic resources and specialized software will be provided to MINED to facilitate processing data from standardized tests. Activities of this component also are designed to support the application of the Latin American Laboratory for

[3] The proposal also specifies the competencies that need to be attained for each grade and subject area. The school can modify order of the curriculum. Furthermore, the proposal describes the resources needed to achieve the objectives and corresponding budget. For more information, see the information available on FOMILENIO's site https://www.fomilenioii.gob.sv/modelo-de-escuela-inclusiva-de-tiempo-pleno-en-el-salvador--2

[4] Grade 12 is not included because grade 12 focuses on technical secondary education.

Assessment of the Quality of Education (LLECE) test in 2019. Finally, the curricula and materials for language and literature will be redesigned to focus on competencies. In the Integrated Systems, technical assistance will be offered to the TEAs to strengthen their technical capabilities in the use of standardized test results. This component is implemented at the national level.

Component 5: Development of curricula and implementation of the new technical offerings in technical programs and certificate programs (*diplomados*) in general programs. The goal of this component is to improve the alignment between the productive sector's needs and the technical programs offered in secondary schools in the Integrated Systems. The main element of this component is the implementation of new offerings in selected general and technical programs within the Integrated Systems. The characterization studies identified three technical areas on which to focus (services, agriculture, and innovation). As part of this component, new options for technical secondary education and certificates for the general programs within these three areas will be developed. The curricula of the programs will be designed and implemented, and teachers will be hired. Teachers in secondary school will receive technical assistance and professional development in pedagogical and management areas to help them implement the new offerings. This component will be implemented in some of the 45 Integrated Systems.

Component 6: Extension of the school day and life and work skills training for students in third cycle schools. The goal of this component is to extend the school day for at least 11 hours per week in grades 7 to 9 in the Integrated Systems. The school day will be extended through workshops, as has been done in past implementations of the SI-EITP model. At least three hours of the extended day will focus on English, two hours on sports, and six hours on work and life skills workshops. The implementation of this component also includes hiring and training staff, organizing the school to implement the workshops, and developing materials for the workshops. We should note that although the SI-EITP seeks to extend the school day in grades 7–9, there is no requirement or MINED directive for such an extension. The extension of the school day is therefore not mandatory and is planned to happen gradually, so not all schools will extend the school day at the same calendar time, for the same number of hours, or for all the students.

Component 7: Strengthening English-language teaching for third cycle and secondary schools. The goal of this component is to improve oral and written comprehension of the English language among students in the Integrated Systems. To achieve this goal, the English curriculum for third cycle (grades 7–9) and secondary students (grades 10–11) will be re-designed, and English teachers from the 45 Integrated Systems will be trained. During the first year of implementation, teachers will take English classes (tailored to their level). In the second year teachers will be trained in practices for teaching English. In addition, tablets designed to be used for learning English will be provided in grades 7–9, and English materials will be designed and provided.

Component 8: Development of reading communities in SI-EITP. The goal of this component is the promotion of reading and the development of communication skills among students in the Integrated Systems. The main activity of this component is to establish reading communities with students of all grade levels (grades 1–11), focusing on the primary grades and

progressively expanding to the secondary grades. Two libraries will be established in each of the 45 Integrated Systems. Implementers will also form 90 library committees and conduct training on library use. Groups of student leaders who promote reading will be formed. Teachers, student leaders, committees, and families will be trained on promoting reading. In addition, third cycle schools will receive tablets to promote reading.

Component 9: Implementation of gender equity and equality policy. At the national level, this component is designed to strengthen the implementation of inclusive, non-sexist teaching practices. It also seeks to improve MINED's response in cases of gender violence and to promote measures that eliminate inequalities and discrimination in school based on gender. To this end, 300 teachers will be trained to be specialists in non-sexist teaching practices. These specialists will be responsible for training 1,400 teachers and principals. In addition, MINED staff from its gender unit as well as key personnel will receive training in non-sexist education.

Component 10: Information systems governance and data quality assurance. This is a national level component whose goal is to have an information system that will automate processes, improve the quality of the data collected by MINED, and facilitate the monitoring of educational indicators. To this end, a functional web application, the Salvadoran Educational Management Information System (SIGES for its initials in Spanish), will integrate the information from all offices within MINED, will be designed and implemented. The core modules of the system will be the information collected from schools, teachers, and students. Complementary modules will include financial information from schools and other educational statistics. As part of the implementation, technical capacity will be offered to MINED to maintain the system in order to facilitate effective use of the information.

Component 11: School infrastructure. The goal of this component is to offer more options for secondary education through appropriate equipment and infrastructure in the 45 Integrated Systems. To this end, FOMILENIO II plans to construct or rehabilitate school facilities and provide equipment to offer an appropriate physical learning environment to the 45 Integrated Systems.

Table II.2. Planned activities, targets, and objectives of the Education Quality Activity

Activities	Implementation targets
Component 1: Professional development for specialists, teachers, and school principals	
Activities for SI-EITP schools • Design and provide training to specialists – 1 introductory module to the SI-EITP program—40 hours – 1 module on the methodology to train teachers—120 hours – 1 module on assessment of students' learning—120 hours	• 120–150 specialists
• Design and provide training to teachers in grades 1–6 – 2 training modules on language and mathematics—120 hours each – 2 training modules on strategies to teach multi-grade classrooms—120 hours each	• 1,948 first and second cycle teachers
• Redesign training modules for content knowledge in 6 content areas: mathematics, language, biology, chemistry, physics, and social science	• 6 training processes (48 modules)
• Provide training to teachers in grades 7 to 9 – 8 training modules on one content area: possible areas are Mathematics, Language, Biology, Chemistry, Physics, and Social Science—120 hours each	• 708 third cycle teachers
• Design and provide training to all teachers—basic (grades 1–9) and secondary (grades 10–12) schools – 7 workshops focused on socioemotional development skills—120 hours total – 3 technology literacy workshops focused on spelling and writing—80 hours each – 1 introductory module on education and gender—40 hours	• 2,656 teachers (all grades)
• Design and provide training and technical assistance to principals and vice principals – 2 training modules on management—120 hours each – Technical assistance on management skills—one visit per month	• 145 principals and 45 vice principals
• Design and provide the following materials – Learning guides for students in multi-grade classrooms – Material for students in multi-grade classrooms	
Component 2: Strengthening pedagogical technical assistance and school management	
National level activities • Design and provide team-building training for technical team and management staff from DDEs—16 hours	• 349 technicians and management staff
• Design and provide training for TEAs to provide management and pedagogical assistance—120 hours in 2018 and 120 hours in 2019	• 210 technicians in pedagogical assistance and 100 in management assistance
• Design and provide training in communication to staff from MINED's central offices and the 14 Department's Education Offices (DDEs)—64 hours	• 60 staff from MINED's central offices and 100 from DDEs
SI-EITP activities • Hire 30 technical education assistants who will provide technical assistance to the Integrated Systems	• 30 TEAs hired

Table II.2 (*continued*)

Activities	Implementation targets
Component 3: Governance of and participation in SI-EITP	
SI-EITP activities • **Design and provide w**orkshops to the following councils on school management, leadership and community participation: – Technical education assistants and principal councils—6 training sessions – School Management Council—12 training sessions – Teacher councils—4 training sessions in 2018 and 6 in 2019 – Student councils—4 training sessions in 2018 and 6 in 2019 – Parent councils—4 training sessions in 2018 and 6 in 2019 • Design and provide material for the participants of the councils: guides for the construction of the pedagogical proposal and the organizational manual for the Integrated Systems	• 210 TEAs • 2,700 participants • 2,846 teachers • 1,038 students • 692 parents and community members • 400 guides of the pedagogical proposal and 400 organizational manuals
Component 4: Strengthening MINED's technical capabilities learning and curricular assessment	
National level activities • Provide training to MINED staff in the following areas: psychometrics, multilevel analysis, and sampling—7 modules • Hire specialists in MINED's evaluation unit • Redesign language curriculum and materials for grades 1–11 • Provide logistical support for the application of LLECE test **SI-EITP activities** • Provide training and technical assistance to TEAs on the use of standardized evaluations • Design and provide materials to teach language	• 20 trainees • 2 specialists • 11 curricula, 11 methodological guides, and 3 workbooks • 210 TEAs • 98,000 books, workbooks, and guides
Component 5: Development of curricular and implementation materials for the new technical offerings in technical baccalaureate and certificate programs (diplomados) in general baccalaureate	
SI-EITP activities (secondary schools) • Update and implement certificate programs (diplomados) for **general** baccalaureate • Design and implement new technical programs for **technical** baccalaureate • Develop the profile of teachers needed for the design and redesign of programs • Hire teachers responsible to teach in the new/updated programs	• 6 certificate programs (diplomados) • 6 new technical programs
Component 6: Life and work skills training for third cycle students	
SI-EITP activities • Develop the strategy and implement the extended school day through workshops • Design the workshops to be offered • Design and develop material for the workshops • Hire staff responsible to conduct the workshops • Design and develop a volunteer program to teach in the workshops	• At least 11 hours added to school day per week • 32 workshops • TBD

Table II.2 (*continued*)

Activities	Implementation targets
Component 7: Strengthening English-language teaching and learning for third cycle and secondary school students	
SI-EITP activities • Provide English training for teachers and workshop providers • Provide tablets to third cycle schools and trained teachers • Redesign of English language curriculum for grades 7–12	• 200–400 teachers or workshop providers • 7,200 tablets
Component 8: Development of reading communities in SI-EITP	
SI-EITP activities • Establish reading communities in schools: reading club, library committee, young leaders • Establish libraries in all integrated systems • Provide tablets to third cycle schools • Provide training in the use of libraries for teachers, library committee, and families	• 349 reading communities • 2 libraries per integrated system (90 libraries total) • 2,700 tablets • 2,900 schools staff (approximately)
Component 9: Implementation of gender equity and equality policy	
National level activities • Provide training in non-sexist teaching practices for specialists—576 hours • Provide training in non-sexist teaching practices for teachers, and MINED staff—144 hours • Improve skills of MINED staff to respond adequately in case of gender-based violence in schools—120 hours	• 300 specialists nationwide • 1,400 teachers and principals for schools nationwide • 150 MINED staff
Component 10: Information systems governance and data quality assurance	
National level activities • Develop web application with student, teacher, and school information • Develop complementary modules with financial information • Develop web application to monitor information offline • Provide technical assistance for the management of the system	
Component 11: School infrastructure	
SI-EITP activities • Build or remodel a school in the integrated systems CEI to offer general or technical baccalaureates • Provide teaching material and equipment for schools: computers, libraries, and laboratories	• 45 integrated systems with at least one secondary school • 346 schools

[1] LLECE is the Latin American Laboratory for Assessment of the Quality of Education standardized test

C. Logic model

Figure II.2 summarizes how the Quality of Education Activity's components are intended to improve the educational outcomes of third cycle and secondary students. The logic model presented below is a high level depiction of the Quality of Education Activity, FOMILENIO II is working on the final detailed logic model. The implementation of the Education Quality Activity in the 45 Integrated Systems funded by FOMILENIO II includes the professional development of specialists and teachers. The goal is to strengthen the content knowledge, pedagogical knowledge, and technological and social competences of teachers in schools. The intervention will also provide training and technical assistance to principals and vice principals on management and leadership. In addition, a management structure will be established to promote joint decision making within schools on issues related to implementing the pedagogical proposal, and to share resources and make joint decisions with other schools in the Integrated Systems.

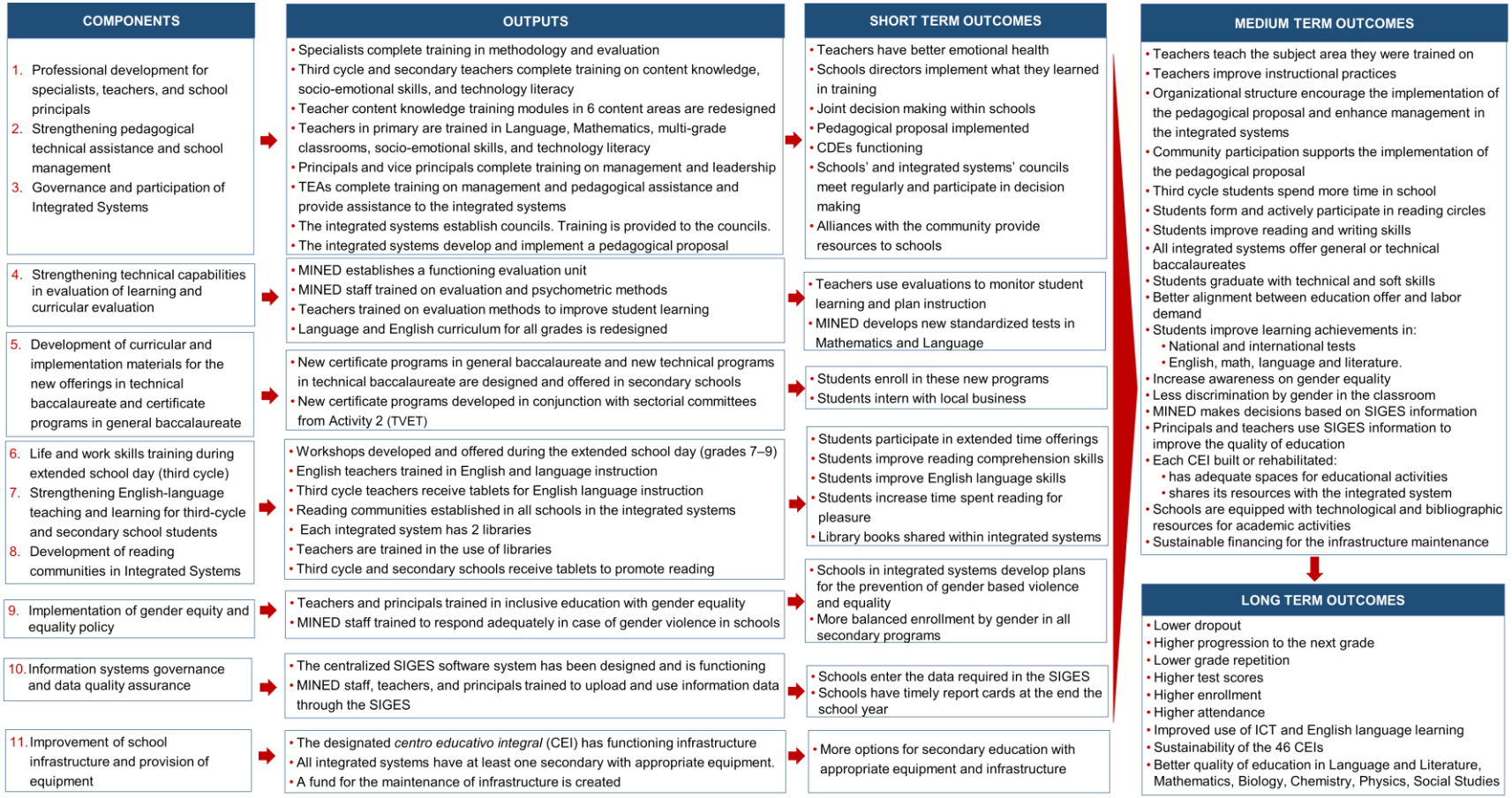
In third cycle schools in these Integrated Systems, extending the school day by at least 11 hours per week while offering workshops in areas such as English and socio-emotional competencies is an effort to increase the time in which students are engaged in academic tasks. In addition, reading communities and libraries established in all schools in the Integrated Systems are meant to improve students' reading and writing skills. These two initiatives are expected to translate into better student outcomes.

All schools in the Integrated Systems are expected to have a designated CEI with enhanced equipment and infrastructure to give youth better learning environments. Each system will also have new technical and general programs in secondary schools to give students more opportunities to learn. These programs are designed to strengthen the link between the skills gained in formal education and the needs of the productive sector. Furthermore, having secondary schools in clusters that did not have them before increases the educational opportunities offered at all grade levels and promotes the continuation and completion of secondary education.

At the national level, as part of the implementation of the gender equity and equality policy, training in non-sexist teaching practices will be provided to teachers and principals. MINED staff will also be trained to respond properly to any gender-based violence in schools. Schools are expected to develop plans to prevent gender-based violence and gender inequality. In addition, MINED is expected to have an evaluation unit, and teachers and MINED staff will receive training on evaluation methods. Teachers are expected to use evaluations to monitor student learning. Finally, a centralized information system will be available for MINED staff, schools, teachers, and parents in order to improve the availability of data on education. The goal is that with acquired capabilities and a high quality of information, MINED would make informed decisions about improving schools' educational quality.

The theory behind the overall implementation of the Quality of Education Activity is that all of its elements will improve the quality of education through more time in school, fewer dropouts, fewer repetition, and better achievement.

Figure II.2. Logic model for the Quality of Education Activity



Source: Authors' high level logic model depiction of the Quality of Education Activity.

D. Evidence base for the SI-EITP model

The SI-EITP model is designed to simultaneously implement several interventions for which evidence of effectiveness has been studied. In this section, we discuss the available evidence to support the potential effectiveness of key SI-EITP interventions such as decentralization of management and training for school principals, a lengthier school day, construction or rehabilitation of school infrastructure, and teacher training. The evidence presented comes from rigorous evaluations using both experimental and quasi-experimental methods. For some interventions, evidence from countries similar to El Salvador is limited. Therefore, interpretation of the evidence should consider this fact. In addition, the concurrent implementation of all these interventions as in the SI-EITP model has not been studied yet. Hence, while the results found in the literature provide guidance on the expected results for the individual interventions, we cannot assume the SI-EITP model will produce the same results.

1. System- or community-level interventions

Decentralization management interventions. The SI-EITP model aims to increase the participation of the community and other schools in the Integrated System to work together and develop joint action plans, optimize and share resources, share expertise, and foster family and community involvement. As part of the intervention, schools implementing the SI-EITP model will also receive trainings for the principals and members of the councils to improve their management capabilities. To our knowledge, there are no studies that have looked at shared management and resources across groups of schools as the SI-EITP model aims to do. However, there are a limited number of rigorous studies that have looked at the effects of school-based management (SBM) initiatives, in which there is a decentralization of authority from a centralized level to the school and community, on educational outcomes. The similarity of these SBM initiatives to the SI-EITP model relies in the fact that SI-EITP, as most SBM initiatives, relies on the decision making capacity of a school council comprising a principal, teachers, parents, and students. This aims to improve student performance and increase the involvement of the parents and their local communities. However, unlike some SBM models SI-EITP does not plan to give monetary grants to the school council. Barrera-Osorio et al. (2009) summarized the evidence on SBM initiatives across the world. Their study finds positive impact of SMB initiatives on the reduction of grade repetition and failure rates and increase in attendance rates. However, their findings on the impact of SMB initiatives on students' academic achievement is mixed. They report that changes in the schools' outcomes are observable about five to eight years after implementation of SBM initiatives.

For example, in Mexico, the Support to School Management Program or AGE (Apoyo a la Gestión Escolar) was implemented in primary schools in 1996. The program provided cash grants to parent associations to spend on any educational activities they considered appropriate (infrastructure improvement, provision of school equipment and materials for students, pedagogical training for teachers, and performance based monetary incentives for teachers). The monetary support varied from $500 to $700 per year depending on school size. Parents received training in the management of these funds and they also received information on the role of the school and the parent association on children's educational achievement, and recommendations on how to help children learn. Gertler et al. (2007) presented evidence of this program on student outcomes. The authors used a quasi-experimental design and found that the program reduced

repetition and grade failure rates by 4 percent in treatment group schools, but found no impact on dropout rates. These impacts were achieved in the first year of the program and impacts did not change with more years on the program.

Another recent systematic review of interventions that transfer decision making to schools (Carr-Hill et al, 2016) finds small to moderate positive effects in reducing repetition and the number of dropouts and in increasing test scores. Most important, this study concludes that impacts may take more time to develop than what most studies allow for. They, therefore, recommend longer follow-up periods. The study does not recommend a specific follow-up period, but most studies included in the review reported findings with a12 to 36 month follow-up period. This review concludes that parents' participation is crucial for success. School-based decision making reforms in highly disadvantaged communities, where parents have low education levels, are less likely to be successful. Another finding is that training is essential for success. However, training by government authorities seems to be less effective than non-governmental training. Furthermore, the specifics of the program design are important. Not all the elements of school-based management reforms are generally beneficial, but there is no clear answer as to which ones are and in which contexts.

Existing studies do not provide evidence on the potential effects of sharing resources across groups of schools as SI-EITP plans to do. In the context of the SI-EITP intervention, training will be provided both by the government and an independent implementer hired for this purpose. However, ex-ante, we cannot speculate whether the training will be more or less effective due to who provides the training. Therefore, we should use estimates from previous studies of SBM cautiously to forecast the potential impact of SI-EITP.

2. School-level interventions

Full-time programs and a lengthier school day. The SI-EITP model includes a component to increase the length of the school day in grades 7 to 9 from 25 hours per week to at least 36 hours and up to 40 hours. In the case of SI-EITP, the extended hours will be implemented through workshops offered to students targeting non-academic areas such as English, technology, socio-emotional competencies, and sports. In this section, we discuss the evidence from cases where the increase in the length of the school days has successfully improved education outcomes. A key difference with the SI-EITP model is that these interventions are typically designed on the premise that lengthier school days allow for increased instructional periods in which students are engaged in academic tasks. This additional instructional time could in turn translate into better student outcomes. In addition, less time out of school could decrease students' participation in high risk activities like gangs, drug use, or delinquency. Several countries in Latin America have already introduced longer school days. However, the implementation of the programs and the impacts on student outcomes vary greatly.

Holland et al. (2015) summarize the programs that have extended the school day in Latin America and their impact on student outcomes. The researchers classify programs extending the school day into three groups:

1. **Extended day model**: Extended the school day (throughout the week) without making major changes to pedagogical approaches or curricula.

2. **Partial extended day model**: Increased additional hours are in some days of the week but not all; and made no major changes to pedagogical approaches or curricula.
3. **Full-time schools model (FTS)**: Extended school time by switching from two school shifts—one in the morning and another in the afternoon—to an extended school session of longer duration (morning and half of the afternoon). Also, a new curriculum or pedagogy was implemented, and training for teachers and principals was provided.

Holland and collaborators classified the different country programs using these categories. For instance, Chile implemented an extended day program and Uruguay a FTS program. They classified El Salvador's program in the full-time category. However, studies of the implementation of the EITP and SI-EITP models in El Salvador, as well as our visits to schools implementing this model point to an implementation that has features of both the partial extend day and the FTS model (USAID 2012; Crespin 2015). The EITP and SI-EITP third cycle schools that serve grades 7 to 9 have been able to extend the school day, but many still have double shifts and are not consistent with the FTS model. The school day has been extended mostly through workshops but not all the students in the school are able to attend regularly. Furthermore, the day has not been extended for all days of the week. This is consistent with a classification of partial extension of the school day. Principals and teachers have attended trainings, which is consistent with the FTS classification. But these trainings have not consistently focused on specific pedagogies or curricula (Crespin 2015). The SI-EITP model, as implemented thus far, shares some features with the partial extended school day model but also incorporated some teacher training. In addition, the day extension is optional since there is no mandate from the MINED to implement it, and some schools already have extended the school day. FOMILENIO II's implementation of SI-EITP model plans to extend of the school day mainly for non-academic activities, but it does plan to provide teacher and principal training to implement an extended day as well as training in other areas. These issues need to be taken into account when using evaluation results from other models to inform expectations of what SI-EITP can accomplish. Below we describe evaluation results of school day extensions in Chile, Uruguay, and Colombia.

The implementation of the extended day program in Chile, started in 1997 and gradually expanded reaching 90 percent of schools by 2012. Schools replaced multiple shifts with one full-day shift that included the entire morning and half the afternoon. The implementation required extensive investment in school infrastructure and a permanent increase in the monthly per student expenditures. The schools decided how to use the additional time in terms of academic or other activities. Bellei (2009) presents the impact evaluation results of the extension of the school day on student academic outcomes at the high school level using a quasi-experimental design. The study uses a difference-in-differences model to estimate the effect of the program on students that attended 9th and 10th grade in schools that were starting to implement the extended the school day compared to students that attended schools that had not extend the school day. This study showed that the extension of school day in Chile had positive effects on students' academic achievement. Two years in an extended day school had positive effects ranging from 0.02 to 0.05 effect sizes in language and from 0 to 0.12 in mathematics. Stronger effects were found for rural and public schools than in urban and private schools. Further, Pirez and Urzua (2015) used propensity score matching to evaluate the effect of the FTS program in Chile on a sample of students interviewed about 10 years after attending full-day schools. They found that the full-day program reduced high school dropout and teenage

pregnancy by three percentage points. However, they found no effects on employment or wages at ages 25–26.

Uruguay started implementing the FTS model since 1990. Cerdán-Infantes and Vermeersch (2007) conducted an impact evaluation of the program in disadvantaged elementary schools. The study also presents a detailed description of the implementation of the program in Uruguay and the changes the program underwent up to 2005. In 1996, the FTS model included not only an extension of the school day but also additional inputs such as construction of new classrooms, reduction of class size, constitution of a teacher committee, increased participation of parents and community, provision of food and health care for students, teacher training on the FTS pedagogical model and on content areas, and provision of teaching materials. Findings from this study showed that the FTS program increased student test scores by 0.06 standard deviations per year in mathematics and 0.04 standard deviations per year in language. In addition, the program led to an additional 10 percent of students reaching the passing score on the third grade test in intervened schools. Despite these positive results, the study warns about the high cost of the program. The FTS model costs about 60 percent more than other schools due to additional teacher salaries, food provision, teacher training, and expenses related to extra-curricular activities.

In Colombia, schools switched from half-day schedules (four or five hours) to a full seven-hour school day between 2002 and 2009. In Colombia, there was no change in curricula or implementation of teacher training, as in the case of Uruguay, so this implementation shares more features with Chile's program. Hincapie (2016) used a school fixed effects model and estimated the impact of lengthier school days on student achievement for students in the 5th and 9th grades. The main model compared variation in average test scores across cohorts for schools that switched from a full-time to a half-day schedule and vice versa. Students in schools with a full-time schedule had test scores about 0.10 of a standard deviations higher than students attending schools with half-day schedules. The study also found larger effects for poor and rural schools. Bonilla-Mejia (2011) found positive effects on the scores of a high school exit exam. Students attending full school days had average test scores 2.5 percentage points higher than those attending half school days. Garcia, Fernandez, and Weiss (2013) conducted an impact evaluation using families with some children that attended a full-time elementary school and some that did not. They find that FTS reduces the probability of school dropout in elementary school by 1 to 2 percentage points and grade repetition by 2 to 5 percentage points.

Holland et al. 2015 review the evidence of full-time school models in Latin America and the Caribbean and concluded that these programs have positive impacts on student learning, repetition, dropout, and teenage pregnancy. However, there is large variation in impact sizes and program implementation. Furthermore, there is no evidence of effects on wages or employment. In their view there are other more cost-effective alternatives to obtain similar results on educational outcomes than FTS models. The authors suggest to extend the school day and use additional hours optimally, for example providing extra support to struggling students, peer learning among teachers, or project based learning activities for students. The research gaps they identify are better tracking of how additional hours are used, and isolating the impact of contributing factors such as time on core curriculum, time in extracurricular activities, teacher training, school lunches, and the role of the school principal.

While the evidence points towards positive effects of increasing the school day from half to full day on student educational outcomes, we should keep in mind that the SI-EITP model does not aim to increase the school day to that extent. As described above, the SI-EITP model is a partial extended-day not a full-day extension model and the increase in time is not focused on academic areas. Therefore, using the estimates provided by other studies to estimate what can potentially be accomplished by SI-EITP has to be done with caution.

School infrastructure. The SI-EITP model has an infrastructure component that aims to construct, improve, or repair existing school infrastructure and provide equipment such as computers and laboratories to the Integrated Systems. This component includes establishing a secondary school in the Integrated Systems that currently do not have one and improving existing facilities as needed. We found research on infrastructure interventions that consisted of the rehabilitation and repair of existing schools rather than construction of new schools. In Bolivia, the Social Investment Fund (SIF) repaired existing infrastructure and replaced classroom furnishings such as blackboards and chairs in rural primary schools. Using two different study designs—a randomized controlled trial (RCT) and quasi-experimental design—the study found that the infrastructure and furnishing improvements had little effect on enrollment, attendance, or academic achievement. Among student-level outcomes, impact was only found in dropout rates. The study estimated that SIF investments led to a reduction in dropout rates ranging from 3 to 3.8 percentage points in primary school (Newman et al. 2002).

In rural Georgia, infrastructure investments focused on improving school buildings: repairing roofs, windows, and floors; replacing indoor pipes; and installing sanitary and heating equipment. A controlled before-and-after study design suggested positive effects of the school rehabilitation intervention. Changes in school attendance—a change in the share of pupils missing more than 30 days in the school year—followed the school improvements. The share of students in project villages absent from classes dropped by 5.7 percent, whereas it increased by 2 percent in the matched control group (Lokshin and Yemtsov 2003).

The systematic review conducted by 3ie (Snilstveit et al. 2016) looked at the effects of construction of new schools in areas where there were no schools. They found two studies that studied the effects of construction of new schools in Niger (Dumitrescu et al. 2011) and in Afghanistan (Burde and Linden 2013). The studies found positive effects on enrollment and attendance. In Afghanistan, the program designed to place formal schools within villages has a significant effect on children's school participation. The program increased enrollment by 47 percentage points over villages not receiving a community-based school. In Nigeria, the effects on enrollment were small ranging from 3.7–4.3 percentage points, according to the household and school surveys and there were no effects on attendance. In both interventions, the effects were larger on girls than on boys. However, we should keep in mind that the context of Afghanistan and Niger are different than El Salvador. Furthermore, the construction of new schools in El Salvador will focus on high schools not elementary schools as the two mentioned studies.

3. Teacher-level interventions

Teacher training. The SI-EITP model includes several teacher training activities such as training for teachers in the upper grades (7 to 9) on content knowledge in their content area, training for teachers in the lower grades (1 to 6) on Language, Mathematics, and multi-grade

instruction. All teachers will also be trained on technology, gender inclusion, and socio-emotional development (Table II.2).

First, we discuss the available evidence on the impacts of pedagogical teacher training on student outcomes. Only a few rigorous evaluations of this type of intervention are available. The existing evidence suggests such training may be beneficial compared with business as usual. The evidence is limited, however, and the effects appear small. For example, a teacher training program was implemented in public primary schools in the Southeastern Katanga province of the Democratic Republic of Congo. The intervention consisted of two components: Teacher Learning Circles, which provided peer support and focused on effective teaching and instructional guides targeting social-emotional practices. This study used an RCT design that compared the intervention group to a comparison group of schools that received the intervention later. Twelve months after program implementation, Halpin et al. (2014) found positive but small impacts on learning outcomes, with a standardized mean difference of .04 in early grade math assessment, .07 in geometry, and .08 in Language Arts.

Another study that focused on teacher training on pedagogical methods found a positive impact of in-service training on children's reading and mathematics achievement in elementary schools in Jerusalem. The intervention came in the form of budget increases—money was spent mainly on teacher training but some resources were used for reorganization and for programs for children with special needs. The training was provided in the school on a weekly basis by outside instructors who focused on improving instruction techniques for language skills, mathematics, and English. Angrist and Lavy (2001) used a matched-comparison evaluation design, built on the premise that only a handful of public schools received additional resources for training. However, the results for different schools were not homogeneous. Only students in nonreligious schools improved their test scores: one-and-a-half years after the beginning of teacher training, reading scores increased by 0.62 standard deviations, and math scores increased by 0.46 standard deviations.

Second, we discuss the available evidence on content knowledge teacher training. Recently, rigorous evaluations of content knowledge teacher training have been conducted in the United States (Garet, Cronen et al. 2008; Garet, Wayne et al.2010; Garet, Heppen et al. 2016). These studies used random assignment designs in which some schools were assigned to the teacher trainings, and others were offered business-as-usual services. The content areas and grades covered by these studies were reading for 2nd grade and mathematics for 4th and 7th grades. The teacher training programs include training over the summer, meetings during the school year, and coaching during the school year. The programs were implemented as intended, and teachers participated in the trainings.

The studies found positive impacts on some but not all of the teacher knowledge outcomes examined. The size of the impacts on teachers' knowledge ranged from 0.38 to 0.55 standard deviations. The studies found one positive impact after examining changes in all the teachers' instructional practices. The size of the impacts on teachers' practices are in the range of 0.36 to 0.61 standard deviations. However, these improvements in teachers' knowledge and practice did not translate into positive impacts on students' academic achievement after one year of implementation.

E. Gaps in the literature and contribution of the impact evaluation of the SI-EITP model

Overall, there are several research gaps in research related to the effectiveness of interventions of the SI-EITP model. First, there is no evidence from interventions that implemented all the components of the SI-EITP model. The literature review presented above is from existing evidence supporting each of the components of the SI-EITP model independently. Second, for some components such as sharing resources across schools, existing studies do not provide evidence on the potential effects on education outcomes. Third, existing evidence focuses on the effects of increasing the school day from half to full day on student educational outcomes—to the best of our knowledge, there are no studies that assess the implementation of a partial extended-day similar to the SI-EITP model. Fourth, for some components, the existing evidence comes from interventions in primary schools, while SI-EITP will include secondary grades. Finally, evidence from countries resembling El Salvador is limited.

Our impact evaluation of the SI-EITP will contribute to addressing some of these gaps. The impact evaluation will provide rigorous estimates of the effects of the SI-EITP model on education outcomes of students in lower and upper secondary schools. However, is important to mention that we will measure the effects of the overall implementation of the Integrated Systems and we cannot estimate the impact of each of the components separately.

F. Economic rate of return

MCC uses the economic rate of return (ERR) models to assess whether its projects are sound investments. The ERR is a summary statistic that reflects the economic merits of an investment. Conceptually, it is the discount rate at which the benefits of an intervention are exactly equal to its costs; a higher ERR implies relatively higher benefits and lower costs. MCC developed an ERR model for the Human Capital project that included both the Educational Quality Activity and the TVET System Reform Activity. The ERR model focuses on the benefits and costs from both activities of the Human Capital project. MCC developed a model for the ERR in which the benefits of the implementation of the Human Capital Project come from an increase in the number of school years students complete as a consequence of fewer students dropping out and repeating grades and more students spending additional time in school. In turn, these will result in an increase the number of graduates of secondary education and the number of people employed.

The ERR model assumes that after about 10 years of implementation, the project will reduce dropout rates and repetition rates in grades 7 to 12. Table II.3 summarizes the assumptions related to reductions on dropout and repetition rates. MCC based these assumptions on MINED's goals. The model also assumes gains due to additional time in school and multiplies education years by 1.6, representing the increase from 25 hours to 40 (1.6 = 40/25). Using these goals and accounting for the infrastructure and construction costs, the model generates an ERR of 8.85 percent.

Table II.3. MCC's assumed initial and projected values of the parameters used in the ERR calculation

	Assumed initial value	Projected value after 10 years of implementation
Dropout rates		
7ht grade	12	5
8th grade	13	5
9th grade	27	15
10th grade	23	10
11th grade	20	10
12th grade	15	5
Repetition rates		
7ht grade	6	3
8th grade	4	2
9th grade	2	1
10th grade	6	3
11th grade	4	2
12th grade	2	1
Multiplier reflects gains due to more time in school (40 instead of 25)	1	1.6

The ERR model was based on assumptions that no longer hold after implementation plans have been developed. For example, the target for the extended school time in grades 7 to 9 is at least 11 hours a week, so weekly time in school will increase from 25 hours to 36 hours per week, not necessarily to 40 hours per week. Furthermore, the extension of the school day is not mandatory and some schools may already have extended the school day for some hours. The impact evaluation will provide estimates of dropout, repetition rates, effective use of time in academic activities, and effective number of hours the students engage in extended activities (instructional and recreational activities) that can be used to update the ERR model. In addition, if MCC decided to conduct a post-secondary survey as part of this evaluation, we could have data on wages and employment.

III. RESEARCH QUESTIONS AND EVALUATION DESIGN

In this chapter, we describe our proposed design for the impact evaluation of the SI-EITP and the performance evaluation of the Education Quality Activity. We begin by listing the research questions that the impact evaluation seeks to address and describe the methodological approach used to answer these questions. We also present the general approach of the performance evaluation, as well as our plans to answer each key research question.

A. Impact evaluation

The purpose of the impact evaluation is to determine whether students attending a school within a SI-EITP system are better off than they would have been without the intervention. Specifically, the evaluation will answer four questions:

1. What is the impact of the SI-EITP on dropout, continuation to the next grade, and graduation from secondary education?
2. What is the impact of the SI-EITP on the academic performance of the students as measured by the Prueba de Aprendizaje y Aptitudes para Egresados de Educación Media (PAES) in 11th grade [5]
3. What is the impact of the SI-EITP on the length of the school day? What is the impact of the SI-EITP on quality of education as measure by time-on-task?
4. What is the impact of the SI-EITP on postsecondary education enrollment, employment, and income[6]?

To answer key research questions about the SI-EITP model, we will use a rigorous impact evaluation based on a random assignment design. An impact evaluation estimates the impacts of a program by comparing outcomes of the individuals who were offered the program to what outcomes would have been in its absence. This last scenario, the counterfactual, cannot be observed. Thus, our objective is to simulate the counterfactual by identifying and tracking a group of schools (and students) that were not selected for the intervention but were similar to the selected schools (and students) before the intervention. The performance evaluation, discussed in greater detail later, will complement the impact evaluation by documenting to what extent the project was implemented as planned.

1. Random assignment of potential Integrated Systems

On October 24, 2014, we conducted a random assignment of systems for the SI-EITP impact evaluation. At that time, MINED provided a list of the 147 potential systems in the Coastal Region eligible for the intervention. Based on this list, we developed a software program to randomly assign each potential system to one of the following three groups:

1. Treatment group: Systems that will implement the SI-EITP model (45 Integrated Systems)

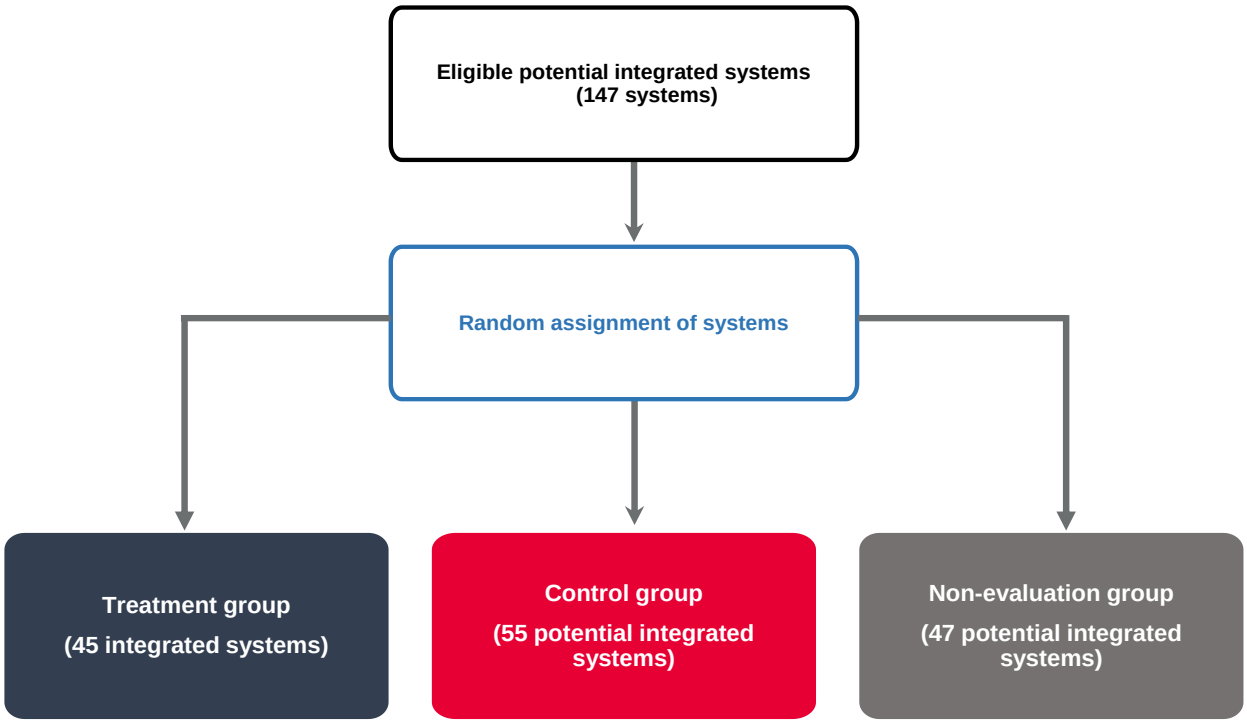
[5] If MINED implements national testing on another relevant grade level, we will consider assessing the impact on that test.

[6] This question is contingent on MCC's exercising the option of collecting postsecondary data.

2. Control group: A group of schools identified as potential systems that will not implement the SI-EITP model, but will participate in the evaluation (55 potential Integrated Systems)
3. Non-evaluation group: Systems that will not participate in the evaluation (47 potential Integrated Systems)

Data will be collected only for the treatment and control groups.

Figure III.1. Random assignment of potential systems



To allocate eligible systems to these groups, we considered two factors: (1) the administrative department in which the systems are located, and (2) the presence of at least one secondary school in each system. The use of the administrative department to randomly assign systems guaranteed that we assigned systems to either treatment or to control groups in all eight departments (Figure II.1). Accounting for the presence of a secondary school in each system was important in ensuring that treatment and control systems were balanced in terms of the presence of a secondary school in the system *after the intervention*. This factor is important for estimating the impacts of the model on secondary education outcomes. One of the goals of the SI-EITP model is to establish[7] a secondary school in those Integrated Systems that do not have one. Thus, all treatment systems will have a secondary school at the end of program implementation. In contrast, it is unlikely that all of the potential Integrated Systems in the control group will have a secondary school at the end of the program. The program assigned 55 potential Integrated Systems to the control group, 40 of which already had a secondary school. This allocation rule

[7] "Creation of secondary schools" refers to constructing and/or rehabilitating school facilities; however, it also conveys the idea that a secondary school could be established in an existing facility after a rearrangement within the system without the construction of a new school.

generated similar sample sizes across the groups: 45 Integrated Systems in the treatment group that either currently have a secondary school or will establish one during implementation, and 40 potential Integrated Systems in the control group that currently have a secondary school.[8]

An additional advantage of this evaluation design is that if more resources become available, FOMILENIO II could introduce the SI-EITP model to one or more of the 47 systems in the non-evaluation group. If additional funds do become available to serve more systems, MINED could serve them, following the order of the original assignment. We will assess the feasibility of including newly funded systems in the evaluation depending on their start date relative to our data collection schedule.

Table III.1 summarizes the sample sizes for the evaluation. In the treatment group, 34 systems have a secondary school, and 11 systems do not. In the control group, 40 potential Integrated Systems have a secondary school, and 15 systems do not. We expect that as part of the SI-EITP model, a secondary school will be established in the 11 treatment Integrated Systems that do not currently have one. FOMILENIO II estimates that during 2018, most new schools will be created, and by 2019 all of the secondary schools will be operating. We expect that by 2019 there will be 45 systems in the treatment group with a secondary school, and at least 40 potential Integrated Systems in the control group will have secondary schools.

Table III.1. Number of systems with and without a secondary school, by department and study group

	Treatment group		Control group	
Department	With secondary school	No secondary school	With secondary school	No secondary school
Ahuachapán	3	0	1	3
Sonsonate	5	3	8	2
La Libertad	4	1	3	1
San Salvador	1	1	3	0
La Paz	7	3	8	4
Usulután	7	1	8	2
San Miguel	2	1	3	1
La Unión	5	1	6	2
Total	**34**	**11**	**40**	**15**

Source: Censo Matricular Inicial 2014, MINED. Updated in 2017 with MINED administrative data.

2. Advantages of the experimental design

An experimental design has several advantages compared to a quasi-experimental design (for example, matching-of-students or regression discontinuity designs):

[8] See the memorandum submitted on August 5, 2015: "ESHC-24 Updated Statistical Power Calculations for SI-EITP Evaluation."

- Random assignment ensures there are no systematic differences between the systems slated to implement the SI-EITP model and those that will not implement it, either in observed or in unobserved characteristics.
- Because MCC can fund only 45 of the 147 eligible Integrated Systems, random assignment is a natural and fair way to allocate available resources.
- Random assignment ensures that the difference in means between the outcomes of the treatment and control groups is an unbiased measure of the impact of the model.
- Rich baseline data are essential for impact analysis using non-experimental designs, but are not crucial for random assignment designs.

As we will explain in the next section, we collected data on students who attended the study schools in 2017 and use these data as a baseline for educational outcomes. In addition, we will use existing data from available sources—enrollment data from SAE system for grades 7–9, and data from SIRAI for grades 10–12—to complement baseline information. The baseline data will be used to assess the balance between treatment and control groups on educational outcomes and to improve the statistical power of the analysis, as we will explain in Chapter V. Next, we discuss the performance evaluation and the data collection plan.

B. Performance evaluation

A performance evaluation study aims to describe the implementation of a program, identify the extent to which the program met its performance targets, and assess the contribution of the program to the intended outcomes, without using a counterfactual. MCC is interested in a small-scale performance evaluation to complement the impact evaluation of the SI-EITP model and enrich the interpretation of impact findings. Specifically, the goal of the performance evaluation of the Education Quality Activity is to describe how intervention activities for the 11 program components were rolled out and implemented, generate evidence on the factors that helped or hindered the implementation of the program components, assess the extent to which program targets were met, and identify what aspects of the program are working well and what could be improved in the future.

We propose a mixed-methods performance evaluation drawing on quantitative and qualitative data sources to answer these core research questions that were developed in coordination with MCC and FOMILENIO II:

1. How were the components of the Education Quality Activity implemented, were they implemented according to plan and which program targets were met and not met?
2. What were the key facilitators and challenges in the implementation of the Education Quality Activity's program components?
3. What plans or actions for the sustainability of the program components are in place or in progress?
4. In what ways the implementation of the SI-EITP model differs between high-performing from low-performing Integrated Systems?

To answer the first, second and third research questions, we will triangulate evidence from monitoring data, reports produced by implementers, and qualitative interviews with key stakeholders including implementation teams, MINED and FOMILENIO II staff, and teachers, principals, and students from a subset of Integrated Systems (case studies). Quantitative monitoring data will help us understand the extent to which targets for some program components were met. Qualitative interviews and a review of program documents will help us understand how program components were rolled out, which facilitators and challenges emerged during the implementation, and which actions have been taken toward promoting and ensuring the sustainability of the different program components.

To answer the fourth research question, we will conduct two comparative case studies on a few Integrated Systems. Case studies are a useful methodology to understand a complex program like the Education Quality Activity, as well as the role of the local context in the implementation of the program and its achievements (Yin 2014). Comparative case studies, in particular, cover two or more cases and emphasize comparison within and across contexts. The goal is to learn about factors associated with the success and challenges of an intervention, as well as variations in the implementation of an intervention that are likely to have affected key outcomes. Comparative case studies involve the analysis and synthesis of the similarities, differences, and patterns across two or more cases that share a common focus or goal. Dimensions for comparison between cases are central to this approach, and are defined by the key research questions (Goodrick 2014).

The case studies for this performance evaluation will focus on the implementation of the SI-EITP model, and will be designed to understand the implementation of SI-EITP activities, the local context, and other relevant factors that contributed to a set of results as defined by the program's logic model. Qualitative comparative analysis will help illustrate similarities and differences in outcomes across the contexts in which the interventions of the Education Quality Activity were implemented. To that end, we will select purposefully two pairs of Integrated Systems (high-performing and low-performing systems in two regions), and gather evidence from local stakeholders and beneficiaries through in-depth interviews and focus groups.

We will choose Integrated Systems strategically and in consultation with MCC, FOMILENIO II, and MINED. We will map the sample of Integrated Systems for the impact evaluation along the selection criteria listed below. Our plan is to first identify two pairs of systems that are comparable in size (for example, number of students enrolled) and located in the same geographic area, but markedly different in terms of the following school performance criteria: (1) gains in student academic achievement at the school level (for example, change in PAES scores over the past five years); (2) monitoring indicators related to school integration and functioning (for example, number of teachers participating in teacher training, establishment of shared resource centers, and number of students engaging in extension activities in grades 7 through 9); and (3) MINED's administrative data on school enrollment and dropout (2018–2019), if available. Using those performance indicators, we will classify systems into two categories: high performing and low performing. We will select two high-performing and two low-performing systems to use in building the comparative case studies.

The comparative case studies will enable us to describe the complex coordination of actions required to implement the 11 components across schools and to understand the conditions that

are likely to be associated with well-implemented and high-functioning systems. The two-pronged approach to the performance evaluation leverages the comprehensive view that implementing teams and stakeholders from MINED and FOMILENIO II offer about the entire Education Quality Activity, and will reveal the in-depth front-line perspectives of school directors, teachers, and students in a small set of purposefully selected systems. Although the findings of the comparative case studies will not be generalizable, their value lies in the in-depth contextual analysis that can be done with a limited number of purposefully selected cases to understand the conditions associated with the desired program outcomes. We propose including two pairs of integrated systems to meet the main goals of the performance evaluation while keeping within the level of effort that MCC has set forth for this performance evaluation. The comparative case studies will inform the four key research questions of the performance evaluation and supplement the quantitative analysis of monitoring data and the impact evaluation of Integrated Systems. The four key research questions presented earlier in this chapter are tailored to each program component, providing a framework for data collection and analysis. In Section IV we present the performance evaluation research questions tailored to each program component, and describe the data sources we will draw from to answer them.

IV. DATA SOURCES AND OUTCOME DEFINITIONS

The Education Quality Activity described in Chapter II is designed to improve the quality of education in El Salvador, in particular in the Coastal Region. In this chapter, we describe the key short- and medium-term indicators we propose to measure the impacts of the SI-EITP model, along with the corresponding data sources. We also present the performance evaluation research questions tailored to each program component, and describe the data sources we use to answer them.

A. Key indicators for the impact evaluation of the SI-EITP

The implementation of the SI-EITP model aims to expand and improve learning opportunities, and ultimately students' labor market outcomes. MCC and FOMILENIO II anticipate that activities such as improving infrastructure, extending school time, having better equipped schools and teachers with better instructional practices, and offering new technical options better aligned with employer's needs will increase access and enrollment in upper grades, grade progression, and grade completion, and provide improved competencies that can eventually generate increased employment and income.

The SI-EITP model as implemented by FOMILENIO II aims to improve outcomes in several domains such as Integrated Systems administration, engagement of parents and community, student time use, student educational outcomes, and finally postsecondary student outcomes. To capture outcome indicators across these diverse domains, we are collecting data from several sources. However, the number of outcomes studied raises concerns about a potential multiple comparisons problem. Conducting several tests increases the chance of erroneously finding a significant impact. The more tests conducted, the more likely we are to find a spurious statistically significant difference and make the wrong policy conclusions. While methods have been developed to correct for the for multiple comparisons problem, these methods reduce the statistical power. We do not plan use correction methods, but we will specify which domains aim to measure a common construct and limit the number of outcomes indicators to test within each domain. We will also categorize some outcomes as secondary, and those results will be interpreted as exploratory in order to reduce the number of tests that are expected to be confirmatory (Schochet, 2008).

As documented in the logic model (Figure II.2) the implementation of components of the Education Quality Activity is expected to affect outcomes in the short-, medium- and long-term. We will measure short-term outcomes to better understand how the SI-EITP model will affect schools' administration, teaching practices and use of time before we see changes in student learning outcomes (medium-term outcomes). To provide evidence on the impact of the intervention, we will examine short-term outcomes such as sharing of resources within the Integrated System, increased time spent by students in schools, and increased time spent on academic activities. We will also measure medium-term outcomes—dropout, grade progression, repetition, transition from 9th to 10th (this indicator identifies whether students enrolled in 9th grade in any school in the evaluation then enrolled in 10th grade in the subsequent year), and

academic achievement (PAES test scores).[9] We propose to report only two of the three possible outcomes within the school year (intra-year dropout in the same school and passed grade, failed grade will be excluded to avoid multiple comparison). We will report all three possible outcomes across two school years because all of these outcomes are part of the logic model (inter-year dropout and repetition, progression to the next year in the same school). Long-term outcomes will be postsecondary school attendance, employment, and income.[10] Table IV.1 presents a summary of the outcome indicators we will construct, the data sources that will be used, and the domain to which they belong. The next section provides more details on each indicator and the data sources.

MCC defined crime reduction as an outcome of interest because of the prevalence of violence in El Salvador and because crime could affect outcome indicators such as dropping out of school. However, during our visit to El Salvador in July 2017, MINED staff argued that the activities of the SI-EITP model are not focused on reducing crime, and therefore it should not be an outcome of this evaluation. We agree that is not clear how the SI-EITP will reduce crime, especially because only third cycle schools will increase time in school, they will probably not increase time for all students, and there will not be a specific amount of extra hours in school. We plan to measure perception of crime with survey data (from principals and students) and use this as contextual information. Furthermore, this information could be included in our analysis as moderators of final effects on key outcomes.

[9] For MCC, an outcome of interest to measure academic achievement is the Trends in International Mathematics and Science Study (TIMSS). However, TIMSS is not a measure that we have for all the schools in the sample, and the latest information available for El Salvador is from 2007. We could add other outcomes of quality of education once we have more information on how the implementers expect that quality education will be affected by the intervention.

[10] Depending on the schedule of the final survey, we could measure other outcomes such as dropout in postsecondary or graduation from postsecondary.

Table IV.1. Descriptions and data sources of outcome indicators for the impact evaluation of the SI-EITP

Outcome indicator	Description	Data source	Domain
Short-term outcomes			
Schools share resources	Indicator identifying whether schools in the systems shared resources during the school year	School principal survey	Integrated system administration
Joint decision making across schools in the Integrated System	Indicator identifying whether schools in the systems made joint decisions regarding the pedagogical proposal during the school year	School principal survey	Integrated system administration
Joint decision making within schools	Indicator identifying whether teachers within schools made joint decisions regarding the pedagogical proposal during the school year	School principal and teachers' surveys	School administration
Perceived parental and community involvement	Index that represents the perceived parental and community involvement	School principal and teachers' surveys	Community involvement
Time spent on academic activities	Index that represents the estimated effective use of time on academic activities through a systematic observation of classroom interactions	Classroom observation (Stalling Observation)	Time on task
Time in school for students	Average number of hours that students spend at school in instructional/academic activities, sports, and recreational or vocational activities organized by the school and with teacher/adult supervision (based on student self-report on a sub-sample of 7th and 9th grade students)	Student survey	Time in school
Instructional practices	Observation of teachers' use of instructional time, use of materials, core pedagogical practices, and ability to keep students engaged. This information will be used to create quantitative variables for general teachers' instructional practices.	Classroom observation (Stalling Observation)	Instructional practices
Alignment of teacher assignment with teacher training	Whether the teacher is teaching the subject he or she was trained for	Teacher survey	School administration
Medium-term outcomes			
Enrollment in grade	The grade the student was enrolled in each year. Secondary schools have data for grades 10–12; third cycle schools have data for grades 7–9.	Student records for third cycle (SAE) and secondary (SIRAI) SAE for grades 7–9 and SIRAI for grades 10–12	Go to school
Dropped out (inter-year in the same school)	Whether a student who was enrolled in one school year is enrolled in the next school year in the same school. This indicator can be constructed using two consecutive years of MINED student-level data for grades 7–9 and 10–12.		Leave school
Dropped out (intra-year in the same school)	Whether a student enrolled at the beginning of the school year is enrolled at the end of the school year in the same school. Using the MINED student-level records, we can only construct this indicator for secondary schools data for grades 10–12.		Leave school (secondary outcome)

Table IV.1 (*continued*)

Outcome indicator	Description	Data source	Domain
Repeated grade (in the same school)	Whether a student who was enrolled in a certain grade enrolled in the same grade the next year (in the same school). We propose to construct this outcome from student-level grade enrollment data in two consecutive school years for grades 9–12.		Repetition
Passed grade	Whether a student passed the grade in which he or she was enrolled that year. Using the MINED student-level records, we can only construct this indicator for secondary schools data for grades 10–12.		Progress in school (secondary outcome)
PAES test scores	Students' PAES grade 11 test scores for math, language, social studies, and science.	MINED records	Academic achievement
Transition from 9th to 10th grade (within or across schools inside or outside the integrated systems)	Binary outcome indicator from student-level enrollment data, identifying whether students who were enrolled in 9th grade in any school in the system enrolled in 10th grade in the next year, regardless of whether the secondary school is in the same integrated system or not. To construct the indicator, we will use the identification number of students enrolled in 9th grade in the system and track their enrollment in 10th grade in the following school year. In addition to the binary variable described above, we will construct another indicator for identifying students that enrolled in a secondary school belonging to the integrated system.	School records from third cycle schools; school records from secondary schools; visits to households SAE and SIRAI data from two consecutive school years	Progress in school
Secondary school graduation	Whether a student graduated from secondary school with a general degree, grade 11, or with a technical degree, grade 12.	SIRAI data	Completed school
Long-term outcomes[a]			
Secondary school graduation	Student self-reported graduation from secondary school with a general degree, grade 11, or with a technical degree, grade 12.	Post-Compact student survey	Completed school
Employment	Student self-reported employment status at the time of the survey, including part- and full-time employment.		Employment
Income	Student self-reported income in the 12 months preceding the survey, including formal and informal labor income, remittances, and other common sources of non-labor income.		Income
Postsecondary education	Student self-reported enrollment in any postsecondary institution after completing secondary education.		Postsecondary education

[a] Long-term outcomes will be measured only if MCC decides to exercise the option of the post-compact survey.

B. Data sources for the evaluation of the Education Quality Activity

1. Impact evaluation of the SI-EITP model

The main data sources for the impact evaluation of the SI-EITP model are in-school surveys of principals, teachers, and students and structured classroom observations collected by a third party, as well as student-level data from third cycle and secondary school records—that is, complete lists of students enrolled in each school and their status at the end of the year. After conducting a pilot data collection to confirm that MINED's data are reliable and consistent with records in schools, we agreed to construct the outcome indicators for the impact evaluation using three data collection methods: (1) in-school data collection of surveys and classroom observations conducted by a third party, (2) data directly provided by MINED such as school records at the student level and student test scores, and (3) if MCC decided to exercise this option, post-Compact student surveys collected by a third party.

Below, we summarize each data source and the outcome indicators for the evaluation. Table A.1 in the Appendix summarizes the sample sizes and timeline of data collection. We designed the in-school surveys in collaboration with FOMILENIO II and MCC, and pre-tested them in the Coastal Region before data collection.

a. In-school data collection.

In January of 2017, we started talking with FOMILENIO II about baseline data collection. The National Statistical Office in El Salvador (DIGESTYC, for its acronym in Spanish) made an agreement with FOMILENIO II to conduct the baseline data collection. In October 2017, DIGESTYC administered baseline principal surveys, teacher surveys, and student surveys, as well as baseline classroom and school observations made before the SI-EITP began implementation. In-school data collection will also include follow-up surveys to principals, teachers, and students along with school and classroom observations after the main components of the SI-EITP model are implemented (extension of the school day, teacher training, and secondary schools in each Integrated System).[11]

The initial plan was to collect data in 273 schools: two third cycle schools and one secondary school per Integrated System. However, the sample size was reduced because DIGESTYC did not have the capacity to collect data from the entire sample before the end of the school year. The sample size was reduced to 190 schools: 116 third cycle and 74 secondary schools. For the reduced sample, one third cycle school and one secondary school were randomly selected in each Integrated System. Follow-up data will be collected in the same schools that were in the baseline, plus the new secondary schools established in the Integrated Systems. We plan to collect follow-up data in the 2020 school year because this will be the last year of the compact implementation. See Table A.1 in the Appendix for details on the sample sizes for baseline and our suggestion for endline data collection.

[11] Principal, teacher, and student surveys focused on the respondents' perceptions of the incidence of crime in schools and their direct exposure to crime at the school or neighboring areas. These data are not outcomes of the intervention, but offer contextual information.

Principal survey. The primary goal of this instrument is to get information on how the SI-EITP model affects the dynamics in schools. For example, did the schools share resources between and within them and if so, how? What type of school management training was provided? Did the school staff make joint decisions about the pedagogical proposal? This survey will also reveal how many hours of school activities the school is offering (including instructional time and extracurricular activities) and how many students are attending those activities. We interviewed the principal in each school where data were collected at baseline, and plan to interview the principal in each school at follow-up.

Classroom observation. We also conducted a structured classroom observation using the Stallings Classroom Snapshot instrument. The Stallings instrument provides quantitative data on the interaction of teachers and students in the classroom. With this instrument, we measure teachers' use of instructional time, teachers' use of materials, core pedagogical practices, and teachers' ability to keep students engaged. We developed a protocol to select the classes for observation. DIGESTYC staff had to select three classrooms, one per area (language, math, and science) in grades 7–9 in third-cycle schools and grades 10–11 in secondary schools, following the suggested order when possible.[12] Then interviewers administered the teacher survey to the teacher whose classroom was observed.

Teacher survey. We developed a teacher survey to gather information on third cycle and secondary teachers' instructional practices and background characteristics, including education and teaching experience. This survey will also ask about the implementation of the pedagogical proposal and the extent to which resources are being shared across schools. Finally, the survey for third cycle teachers will also inform us about their involvement in the activities related to the extension of the school day and on the alignment of the course they are teaching and their area of study. At baseline, we surveyed the three teachers who were observed in the classroom observations at each of the schools.

Student survey. We developed a short student survey to understand students' time use in academic and non-academic activities and to gather information on their participation in extension activities offered throughout the Integrated Systems. Students reported on time spent on academic/instructional activities, as well as time spent on sports and recreation or vocational activities with or without teacher supervision, including recess time. Students also reported on their school engagement, perceived support from teachers, gender biases in the classroom, and school safety. At baseline, we interviewed 10 students per school enrolled in grades 9 and 11. These 10 students were randomly selected from lists we requested before visiting the schools where data was collected. All students were selected randomly within the grade, and, when possible, five male and five female students were randomly selected.

b. Data to be provided by MINED

Reporting on the educational outcomes we described in Table IV.1 requires student-level data. We met several times with MINED to better understand its management information systems and assess the possibility of using data provided directly by MINED to construct the

[12] We randomly selected the shift and order in which classes should be observed. For example, first language, then math, and then science. However, given the organization in some schools, it was not possible to follow the protocol.

indicators of interest to this impact evaluation. A third party, the Spanish Association for Standardization and Centralization (AENOR, for its acronym in Spanish), conducted a pilot data collection and quality checks at the beginning of 2017 to confirm that MINED's student-level data could be used to reliably construct the outcome indicators of interest. AENOR's pilot concluded that MINED's data were consistent with the data in school records. Therefore, we will rely on MINED's data to construct the student outcome indicators.

In El Salvador, the school year starts in January and ends in November. In June 2018, MINED delivered data from 464 third cycle and secondary schools for the school years 2016 and 2017. **Our original design relied on collecting administrative student-level data up to the 2021 school year for all the schools in the study sample**. However, we have agreed with MCC that we will collect administrative data up to the beginning of 2023 and include these educational outcomes in the final report in 2023.

Student-level data. We plan to use student-level data to construct measures of student enrollment, dropout, repetition, and grade progression for third cycle and secondary schools. MINED's student level data are available in two information management systems that use a unique student identification number (Número de Identificación del Estudiante, or NIE) provided by MINED. The first information management system is known as the School Management System (Sistema de Administración Escolar- SAE), and it has data for students in third cycle. The SAE is mainly an initial enrollment registry and does not collect information on the students' status at the end of year. However, if the NIE is used reliably to track student enrollment, we could use two consecutive years of student data for all the grades in the study schools to match students by NIE and identify the students who are repeating grades or have progressed to the next grade. Similarly, we can construct inter-year dropout (within school) by identifying the students who were enrolled in one school year but not enrolled in the next school year (neither repeating nor progressing). The outcome indicators we cannot construct with the SAE are on whether students are passing or failing a grade, because the system does not provide that type of information.

The second information management system is known as the Academic Registry System (SIRAI for it's acronym in Spanish), administered by the Office of Accreditation at MINED. SIRAI has enrollment data for students in secondary grades as well as information on the status of students at the end of the school year. With these data, we can construct student level indicators on dropout, passing a grade, repeating a grade, and graduating from secondary school (within school). Furthermore, we can construct the indicator for transition from grade 9 to 10 using two consecutive years of SAE (9th grade) and SIRAI (10th grade) data. We can merge both data sets by NIE and identify whether the 9th grade students made the transition to 10th grade or not. Because one of the components of Activity 1 aims to have an information system that will automate processes, and improve the quality of the data collected by MINED. The source of the follow-up data will be the new system, the Salvadoran Management Educational Information System (SIGES).

Table A.1 shows the years for which we will request MINED's SIRAI and SAE data. We have asked for information on all the students enrolled in the study schools in the school years 2016–2018. We plan to request these data yearly until the beginning of 2023 school year.

Student test scores. Each year, all students enrolled in 11th grade in El Salvador take the Prueba de Aprendizaje y Aptitudes para Egresados de Educación Media (PAES) test, a national test of language, mathematics, science, and social studies. We will request these student-level scores from MINED for all students in the sample schools to measure secondary school academic achievement (Table IV.I).. We will report students' scores for math, language, social studies, and science. MINED promptly shared student-level PAES scores for students enrolled in grade 11 in either of the school years 2016 and 2017. Our plan is to request student scores until early 2023 and include available education outcomes in the final report in 2023.

Transition to secondary. MCC and FOMILENIO II are particularly interested in the transition from third cycle (9th grade) to secondary school (10th grade). We propose to construct a binary outcome indicator from student-level enrollment data to identify students who enrolled in 9th grade one year and in 10th grade the next year. Our plan for constructing the transition indicator is to first identify the students in our sample who are enrolled in 9th grade, then identify which of them enrolled in 10th grade the next year in a secondary school either within or outside the Integrated System. The indicator will take the value of one if a 9th grade student was enrolled in a secondary school in the next school year and zero if the student was not enrolled in a secondary school. We will consider students enrolled in 10th grade in a secondary school regardless of whether the school belongs to the same Integrated System the student attended in 9th grade. One limitation is that we cannot identify whether students enrolled in 9th grade passed that school year, so the transition indicator will follow students enrolled in 9th grade at the beginning of the year and indicate whether they enrolled in 10th grade the following year (regardless of whether they passed 9th grade). In 2017, MINED provided nationwide data on students enrolled in 9th grade in 2016. Our plan is to ask MINED for nationwide enrollment data for grade 10 to construct the transition indicator until 2023.

c. Post-Compact student survey

MCC is also interested in long-term outcomes such as graduation from secondary school, continuation to postsecondary education, employment, and individual income. To measure these outcomes, a third party could conduct a student survey one or two years after the students complete secondary education, once the compact ends. MCC is interested in looking at outcomes two years after students complete secondary education in order to identify dropouts from postsecondary school. We have discussed with MCC the timeline to measure these outcomes. As of December 2018, we have agreed with MCC that 2021 is too soon to look for impacts in long term outcomes, such as employment and income. Impacts on postsecondary outcomes may be better assessed with a survey in 2024 (or 2025). The post-secondary survey has been excluded from the data collection schedule, because it is not a costed option at this time.

2. Performance evaluation of the Education Quality Activity

The performance evaluation will complement the impact evaluation of the Full-Time Inclusive Schools intervention, and will allow us to describe how various activities to strengthen the national education system were implemented, the extent to which program targets were met, and potential areas for program improvement. In this section, we present the performance evaluation questions and the main data sources we will draw on to answer them (Table IV.2). This set of research questions is not exhaustive, but it does illustrate the core topic areas that we will structure data collection and analyses on.

We plan to use a variety of methods to collect quantitative monitoring data from implementers and qualitative data from a range of stakeholders. These methods will include document review, semi-structured interviews, and focus group discussions. We next describe each type of data gathering in detail.

1. **Monitoring data.** We expect to have access to data on output indicators and targets from the teams in charge of implementing the 11 program components (FUSALMO and FHI). For example, we will request data on the number of specialists, teachers, and school principals who participated in professional development activities. We will also request enrollment rosters or attendance lists for curricular and extracurricular extension activities to measure changes in the length of the school day. Some of the monitoring data might also be provided by FOMILENIO II.
2. **Document review.** This review will include the diagnostic reports completed by FOMILENIO II (characterization of the Integrated Systems), progress reports produced by implementers, and work plans for the implementation of the components in the Integrated Systems and nationwide.
3. **Key informant interviews.** We will conduct semi-structured interviews with key respondents. We will interview staff from FOMILENIO II who have been in charge of leading design and implementation of the program; implementation team members; and MINED staff who have been involved in the decision making and implementation of the Education Quality Activity. For the comparative case studies, we will interview a wide range of regional and school-level key informants (Table IV.2).
4. **Focus group discussions.** For the case studies, we plan to conduct group discussions with teachers and students. These discussions will give us information about their overall perceptions of and level of satisfaction with the activities offered as part of the implementation of the Education Quality Activity. For example, we will explore teachers' perceptions on the activity's trainings and technical assistance, and students' perceptions about the reading communities and the workshops offered as part of the extension of the school day.

We will design data collection protocols for the performance evaluation using the four main research questions in Chapter III as a framework. To specify the type of information we will gather and the data sources we will rely on, each research question is tailored to each one of the 11 intervention components. As shown in Table IV.2, for each component we will gather evidence on (a) how the intervention component was implemented; (b) the facilitators and challenges encountered in implementing the component; (c) the plans or actions that were put in place or planned for the sustainability of the component; and (d) context and implementation factors that differentiate high-performing systems from underperforming ones.

Table IV.2. Research questions for the performance evaluation of the Education Quality Activity, and data sources by program component

Research questions	Data sources
Component 1. Professional development for specialists, teachers, and school principals	
How were the professional development activities for specialists, teachers, and school principals implemented, and which training targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the implementation of the professional development activities for specialists, teachers, and school principals?	Interviews with implementation teams and case studies
What plans or actions for the sustainability of the professional development activities for specialists, teachers, and school principals are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the implementation of the SI-EITP teacher professional development activities differentiate high-performing systems from low-performing ones?	Case studies: interviews with directors and focus groups with teachers in selected SI-EITP schools
Component 2. Strengthening pedagogical technical assistance and school management	
How were the activities to strengthen pedagogical technical assistance and school management at the national level and in SI-EITP schools implemented, and which targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the implementation of the activities to strengthen pedagogical technical assistance and school management at the national level and in SI-EITP schools?	Interviews with implementation teams; case studies
What plans or actions for the sustainability of the activities to strengthen pedagogical technical assistance and school management at the national level and in SI-EITP schools are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the implementation of the activities to strengthen pedagogical technical assistance and school management in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with directors and focus groups with teachers in selected SI-EITP schools
Component 3. Governance and participation of SI-EITP	
How were the governance and participation activities implemented in SI-EITP schools, and which targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the implementation of the governance and participation activities in SI-EITP schools?	Interviews with implementation teams; case studies
What plans or actions for the sustainability of the governance and participation activities in SI-EITP schools are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the implementation of the governance and participation activities in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with technical education assistants and members of the principal, school management, teacher, student, and parent councils
Component 4. Strengthening technical capabilities in learning and curricular assessment	
How were the activities to strengthen technical capabilities in evaluation of learning and curricular evaluation implemented at the national level and in SI-EITP schools, and which targets were met?	Interviews with implementation teams and MINED staff; document review; and monitoring data
What were the key facilitators and challenges in the implementation of activities to strengthen technical capabilities in evaluation of learning and curricular evaluation at the national level and in SI-EITP schools?	Interviews with implementation teams and MINED staff

Table IV.2 (*continued*)

Research questions	Data sources
What plans or actions for the sustainability of the activities to strengthen technical capabilities in evaluation of learning and curricular evaluation at the national level and in SI-EITP schools are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the implementation of activities to strengthen technical capabilities in evaluation of learning and curricular evaluation implemented in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies; interviews with SI-EITP school directors and TEAs
Component 5. Development of curricular and implementation materials for the new technical offerings in technical baccalaureate and certificate programs (*diplomados*) in general baccalaureate	
How were the materials for the new technical offerings in baccalaureate and certificate programs (diplomados) developed, and which targets were met?	Interviews with implementation teams; document review; and monitoring data
What were the key facilitators and challenges in the development of materials for the new offerings in baccalaureate and certificate programs (diplomados)?	Interviews with implementation teams
What plans or actions for the sustainability of the activities to develop materials for the new offerings baccalaureate and certificate programs (diplomados) are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the activities to develop materials for the new offerings in baccalaureate and certificate programs (diplomados) implemented in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: Interviews with SI-EITP school directors and teachers, focus groups with students
Component 6. Life and work skills training for third cycle students	
How were the life and work skills training activities for third cycle students implemented in SI-EITP schools, and which training targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the implementation of life and work skills training activities for third cycle students in SI-EITP schools?	Interviews with implementation teams, and case studies
What plans or actions for the sustainability of the life and work skills training activities for third cycle students in SI-EITP schools are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the life and work skills training activities for third cycle students implemented in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with SI-EITP school directors, and focus groups with students and teachers
Component 7. Strengthening English-language teaching and learning for third cycle and secondary school students	
How were the activities to strengthen English-language teaching and learning for third cycle and secondary school students implemented in SI-EITP schools, and which targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the implementation of activities to strengthen English-language teaching and learning for third cycle and secondary school students in SI-EITP schools?	Interviews with implementation teams; case studies
What plans or actions for the sustainability of activities to strengthen English-language teaching and learning for third cycle and secondary school students in SI-EITP schools are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the activities to strengthen English-language teaching and learning for third cycle and secondary school students in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with SI-EITP school directors; focus groups with students and teachers

Table IV.2 (*continued*)

Component 8. Development of reading communities in SI-EITP	
How were reading communities developed in SI-EITP schools, and which targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the development of reading communities in SI-EITP schools?	Interviews with implementation teams; case studies
What plans or actions for the sustainability of reading communities in SI-EITP schools are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the reading communities in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with SI-EITP school directors, and school librarians, and focus groups with students and teachers
Component 9. Implementation of gender equity and equality policy	
How was the gender equity and equality policy implemented at the national level, and which targets were met?	Interviews with implementation teams; document review; monitoring data; and case studies
What were the key facilitators and challenges in the implementation of the gender equity and equality policy?	Interviews with implementation teams; case studies
What plans or actions for the sustainability of the gender equity and equality policy are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the gender equity and equality policy implemented in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: Interviews with SI-EITP school directors, and specialists, and focus groups with students and teachers
Component 10. Information systems governance and data quality assurance	
How were the information systems governance and data quality assurance activities implemented, and which targets were met?	Interviews with implementation teams; document review; and monitoring data
What were the key facilitators and challenges in the implementation of the information systems governance and data quality assurance activities?	Interviews with implementation teams
What plans or actions for the sustainability of the information systems governance and data quality assurance are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the information systems governance and data quality assurance activities implemented in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with SI-EITP school directors; focus groups with students and teachers
Component 11. School infrastructure	
How were the school infrastructure activities implemented, and which targets were met?	Interviews with implementation teams; document review; and monitoring data
What were the key facilitators and challenges in the implementation of the school infrastructure activities?	Interviews with implementation teams
What plans or actions for the sustainability of the school infrastructure activities are in place or in progress?	Interviews with implementation teams and MINED and FOMILENIO II staff
What aspects of the school infrastructure activities implemented in SI-EITP schools differentiate high-performing systems from low-performing ones?	Case studies: interviews with SI-EITP school directors; focus groups with students and teachers

3. Sample sizes for the performance evaluation

We plan to conduct two rounds of qualitative data collection to answer the first three research questions. The first round will take place in 2019, and the second round in 2020. We propose to interview the coordinators leading the implementation of each component and their counterparts at MINED, as well as FOMILENIO II staff. For the case studies, we will conduct interviews and focus group discussions with personnel of four Integrated Systems in 2020. In Table IV.3, we summarize the data collection methods, respondents and sample sizes (number of interviews or focus groups), and the performance research questions those sources will be informing.

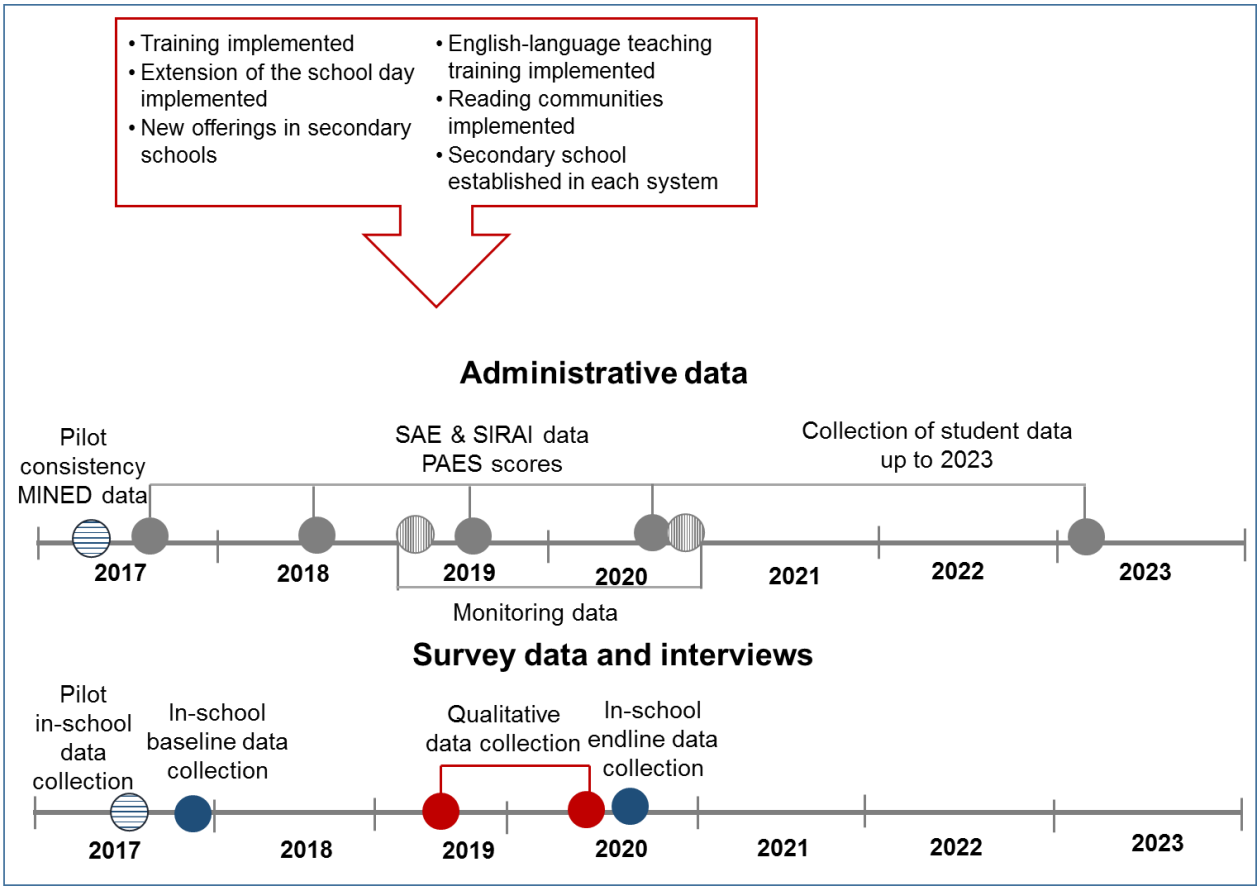
Table IV.3. Sample sizes for the performance evaluation

Data collection method	Respondent(s)	Research questions	Components	Sample size
Interviews	Implementation teams (1 interview per component)	1, 2, and 3	Components 1 to 11	First round: 11 Second round: 11
Interviews	FOMILENIO II staff (Infrastructure, gender, information systems, FHI, FUSALMO) (1 interview per topic)	1, 2, and 3	Components 1 to 11	First round: 5 Second round: 5
Interviews	MINED staff (1 interview per component)	1, 2, and 3	Components 1 to 11	First round: 11 Second round: 11
Case Studies	**Respondent(s)**	**Research questions**	**Components**	**Sample size**
Interviews	School director (1 interview for components 1,2, and 4; 1 for component 5; and 1 for components 6–11: 3 total)	4	Components 1, 2, and 4 Components 5 Components 6, 7, 8, 9, 10, and 11	Second round: 3 per school for a total of 12
Interviews	TEAs and members of the principal, school management, teacher, student, and parent councils (1 interview per council, 6 total)	4	Components 3 and 4	Second round: 6 per school for a total of 24
Interviews	Baccalaureate teachers	4	Component 5	Second round: 4
Interviews	School librarian	4	Component 8	Second round: 4
Interviews	Specialists	4	Components 9	Second round: 4
Focus groups	1 teacher focus group for components 1–2, and 1 teacher focus group for component 6–11, per school (Third cycle)	4	Components 1 and 2 Components 6, 7, 8, 9, 10, and 11	Second round: 2 per school for a total of 8
Focus groups	2 students focus groups for components 6–11, per school (Third cycle and baccalaureate)	4	Components 6, 7, 8, 9, 10, and 11	Second round: 2 per school for a total of 8

4. Data collection schedule and implementation milestones

The data collection schedule will be based on the current intervention's implementation timetable and other milestones. According to discussions with FOMILENIO II, in 2016, key implementation components such as the extension of the school day, teacher training, and establishment of secondary school will start in 2018, and most schools will be established by the 2019 school year. We planned the data collection schedule around these milestones. We plan to collect in-school data in 2020 on the data collection sample, regardless of different schools' implementation progress. Figure IV.1 presents these milestones, along with the data collection schedule. Please note that the figure below includes the timeline to collect MINED administrative data up to 2023 enrollment, this follow-up period of educational outcomes will allow us to estimate the impacts of the intervention on dropout until 2022 if the enrollment data for 2023 is available early in the year. A previous version of the data collection schedule included two options to collect data with a post-secondary survey. As of December 2018, MCC agreed that data collection in 2021 is too soon to look for effects in long term outcomes such as employment or income, but we have not agreed on when to collect data. We can include post-secondary data collection once MCC makes a decision on when to measure long-term outcomes.

Figure IV.1. SI-EITP evaluation data collection schedule, 2017–2023



Note: The original calendar included postsecondary survey data collection collect data in 2021. However, we have agreed with MCC that the impacts on postsecondary outcome may be better assessed with a survey in 2024 (or 2025). We excluded the post-secondary from the figure, because it is not a costed option at this time.

V. ANALYSIS PLAN

A. Estimating the impact of the SI-EITP model

Given the use of random assignment, the basic method to estimate impacts is to compare the mean outcomes of the treatment and control groups. We intend to use regression models to estimate impacts because the regression adjustment will enable us to account for design characteristics, such as the group-level randomized design, and improve statistical precision through the inclusion of covariates.

We thus will estimate impacts of the SI-EITP model using a regression specification that compares outcomes of students who attended a school that implemented the SI-EITP model (treatment group) with outcomes of students who attended schools that were not part of a potential integrated system (control group), controlling for idiosyncratic differences between the two groups. The basic model can be expressed as follows:

$$(1) \quad Y_{ipsd} = \alpha + \beta\chi_{ipsd} + \lambda T_s + \gamma z_d + \varepsilon_{ipsd}$$

In this equation, Y_{ipsd} is the outcome of interest for student i in educational program or school p in Integrated System s in stratum d used in the random assignment design; X_{ipsd} is a vector of baseline characteristics of student i in educational program or school p in Integrated System s in stratum d (baseline student-level variables are, for example, age and gender; school-level variables are enrollment, dropout, pass, and repetition rates before the program started); T_s is an indicator equal to one if the student is in an educational program or school that is part of an Integrated System s assigned to the treatment group (SI-EITP) and zero if the student is in a school that is part of an Integrated System assigned to the control group. The vector z_d includes indicator variables for each stratum d used in the random assignment design (department and the presence of a secondary school in the Integrated System). In addition, ε_{ipsd} is a random error term for student i in program or school p in Integrated System s in stratum d. The parameter estimate for λ is the estimated impact of the SI-EITP model on the outcome of interest. Finally, all of the impact estimates are weighted to account for differential assignment probability within stratum. Because we conducted random assignment at the Integrated System level, we will adjust the standard errors for clustering at that level.

The estimate is based on an intent-to-treat analysis, so the estimates described above will be based on the randomized sample. Students who drop out of school or transfer will still be treated as either intervention or control, based on their randomization outcome. Similarly, if a school assigned to an SI-EITP system does not participate in the intervention, it will be treated as part of the treatment group regardless of whether it implemented the model or not. However, we will report the number of schools in the treatment group that did not participate in the model and the number of schools in the control group that did participate. To the extent possible, we will also report this for students.

We will conduct an additional analysis of impacts by gender by adding an indicator variable for gender to equation (1) above and having it interact with the treatment variable. This approach

will allow us to determine whether impacts on graduation, enrollment, and income differ for males versus females.

1. Statistical power

We calculated the minimum detectable impacts (MDIs) for the design with randomization at the system level and analysis of outcomes. An MDI is the smallest impact that a study of a fixed sample size can detect statistically with high probability. It means that an impact larger than the MDIs can be detected statistically, but impacts smaller than the MDIs will not be statistically detected by the study. The assumptions used in the calculations are the following:

- There is 80 percent power and a statistical significance at the 5 percent level for a two-tailed test.
- For third cycle outcomes based on MINED's administrative data, we use the following sample sizes: 45 Integrated Systems in the treatment group and 54 in the control group (one system in the control group only has basic education), and 140 students per system. For secondary outcomes, we assume 45 Integrated Systems in the treatment group, 40 Integrated Systems in the control group, and 100 students per system on average. We should note that currently only 33 systems in the treatment group have a secondary school, but it is expected that all 45 systems in the treatment group will have a secondary school by 2019. We estimated the student sample sizes based on MINED's administrative records from 2016. For postsecondary outcomes based on student surveys, we assumed 50 students per system for a total student sample of 4250 students for the postsecondary survey.
- The percentage of variance explained by the use of system-level covariates in the regression model (R_1^2) is 0.1. In general, we expect to include as a regression covariate at the system level the baseline value of the outcome analyzed; for example, the baseline dropout for 10th grade for the Integrated Systems will be included as a covariate when we analyze the program's impact on 10th grade dropouts. We used data from a similar analysis done for the evaluation of secondary schools for the first Compact between MCC and El Salvador (Campuzano et al. 2016), and estimated R^2 values in the range of 0.03 to 0.20. Although the analysis we plan to conduct for this evaluation is similar to the one we used for the evaluation of Compact 1 in the sense that we have students clustered in schools, the sample of schools differs. Therefore, we also conducted regressions with the school-level data from Censo Escolar 2012–2014 on the sample of schools for the SI-EITP evaluation. In this analysis, we estimated values for R_1^2 that are between 0.05 and 0.15. Based on the two analysis conducted, we assume an $R_1^2 = 0.1$ in our statistical power calculations.
- The percentage of variance explained by the use of student-level covariates for student data in the regression model (R_2^2) is 0.1. Most likely, the only student-level characteristics we will include as covariates in the analysis are age and gender. Therefore, we do not expect to have student-level covariates that can help explain a large part of the outcome variance. We assume a value of 0.10 for the variance explained by student-level covariates. We also did calculations with a value of zero, and the MDIs stay constant up to the third decimal place.
- The intra-class correlation (ICC) is 0.025 for most student educational outcomes based on MINED's administrative records and 0.03 for outcomes based on postsecondary survey data. The ICC is a measure of the ratio of the variability of the student-level outcomes

between systems to the total between- and within-system variation for student-level data. We used an ICC = 0.025 for student educational outcomes because it is the value we calculated using the student-level educational outcome data clustered in schools from MINED's administrative records. We used and ICC= 0.03 for post-secondary outcomes because it is the value we estimated using survey data we collected for the evaluation of secondary schools for the first Compact between MCC and El Salvador (Campuzano et al. 2016). These ICCs are also in line with estimations of ICC's for binary outcomes in educational studies in the US (Schochet 2013).

Table V.1 presents our results. We calculated the MDI for educational outcomes such as dropout, repetition, and pass for 7th and 10th grades, on-time progression from 10th to 11th grade, PAES score, and on-time graduation of secondary; and for postsecondary outcomes such as enrollment in postsecondary, employment, and income.[13] The first column presents the outcomes of interest. The second column presents revised baseline values calculated by us using data from MINED's administrative records from SAE and SIRAI in 2016. The third column presents MCC's target value (from Annex III in the Compact).

Although MCC is interested in all of the outcomes presented in the table, it has only defined targets for four of those outcomes: repetition and pass rates for grades 7 and 10. The fourth column presents the expected changes if MCC's target is attained with respect to the revised baseline values (revised baseline value minus target value). The fifth column presents the MDIs we calculated with the assumptions presented above and the revised baseline values for each of the key outcomes presented. The sixth column presents an assessment of whether the study can detect effects of the size MCC is targeting with respect to the revised baseline value. In the last column, we present the range of values for statistically significant impacts found in previous literature and discussed in the literature review section. As we caution in that section, other studies have different contexts, the programs have different components, and many report findings for different grades than those we are interested in. However, we include them in the table in order to provide some information for the outcomes for which MCC has no specific targets.

[13] We did not include the calculations for time on task because we do not have reliable data to estimate the baseline mean, standard deviation, and ICC. We expect to obtain these data from the baseline data collection and conduct power calculations based on them.

Table V.1. Baseline values, targets, changes and minimum detectable impact (MDI) for selected outcomes (percentage points, unless otherwise indicated)

Outcome	Revised baseline value	Target value (MCC)	Expected difference with revised baseline value	MDI	Is the difference detectable?	Impacts on similar outcomes from literature review
	2	3	(2-3)			
Outcomes from MINED's administrative records	**Revised based on SAE and SIRAI 2016**					
School dropout between 7th and 8th	22.7	5	17.7	4	Yes	0 to 4
7th grade repetition	5.3	3	2.3	2.2	Yes, but close	2 to 5
Grade progression in the same school 7th to 8th	72	n.a.	n.a.	4.3	n.a.	n.a.
Transition from 9th to 10th in any school	72.5	n.a.	n.a.	6.9	n.a.	n.a.
School dropout between 10th and 11th	22.4	10	12.4	4.5	Yes	0 to 4
10th grade repetition	5.1	3	2.1	2.4	No, but close	2 to 5
10th grade pass rate	78.2	n.a.	n.a.	5.4	n.a.	4
Grade progression in the same school 10th to 11th	72.5	n.a.	n.a.	4.9	n.a.	n.a.
PAES score	5.0	n.a.	n.a.	0.31 (s.d.)	n.a.	0–0.12 (s.d.)
Graduation from secondary	92.7	n.a.	n.a.	3.4	n.a.	n.a.
Outcomes from student survey	**Revised based on compact 1 survey**					
Enrollment in postsecondary education	32.6	n.a.	n.a.	6.1	n.a.	0
Employment	35.7	n.a.	n.a.	6.2	n.a.	0
Income (dollars)	1,018	n.a.	n.a.	132	n.a.	0

Sources: Annex III Compact and MINED administrative records 2016, SAE and SIRAI.

n.a. = not available; s.d. = standard deviation units.

All our MDIs are within or very close to the targets set by MCC given the assumptions described above. The study will be able to detect impacts on educational outcomes of the size MCC is targeting when we use the data from MINED's administrative records from 2016 as baseline (in other words, we assume that the control group mean will be equal to the 2016 baseline value, and the treatment group mean will be the target value). The expected difference between the target and revised baseline value for 7th grade dropout is 17.5 percentage points which is greater than the MDI of 4. Thus, the study can detect it. The expected difference between the target and the revised baseline for 7th grade repetition is 2.3, and the MDI is 2.2; therefore, the study cannot detect it although it can come very close to detecting it. For 10th grade dropouts, the difference between the revised baseline and MCC's target is 12.4, the MDI is 4.5, and the corresponding values for 10th grade repetition are 2.1 and 2.4 percentage points. Therefore, the study will be able to detect impacts of the size targeted by MCC for 10th grade dropout but not for repetition, although we are close to detecting them.

The last column presents the range of impact values found in previous research (see Section II. D). We should keep in mind that these values come from evaluations done for programs that have been implemented differently than the way SI-EITP's implementation is planned to be implemented, and many apply to elementary school, whereas this implementation is designed for later grades. The studies can, however, provide preliminary information on the size of effects that other studies have found. In general, the MDIs for educational outcomes are within the range of impacts found in the literature, or slightly larger. For example, the MDIs for repetition for both grades 7th and 10th and dropout for 7th grade fall inside the range of previous impacts, but MDIs for 10th grade dropout, PAES score (which measures academic achievement), and pass rates are slightly higher than what the literature has found. However, previous literature has not found statistically significant impacts on postsecondary outcomes such as income and employment.

In sum, if the SI-EITP is as successful as MCC and MINED expect it to be, it is likely that we can detect impacts on most educational outcomes. Furthermore, the study is powered to detect effects that are slightly larger than or within the range of what other studies have found on educational outcomes. However, neither MCC nor MINED has defined goals for postsecondary outcomes, and there is no evidence that similar programs have had an effect on postsecondary outcomes. It is therefore possible that the sample sizes for this study could not detect impacts on postsecondary outcomes.

Subgroups. Table V.2 presents MDIs for the cases when we have half the number of students per system that are assumed in the above calculations (50 students for secondary outcomes instead of 100 (with the exception of the PAES score, where we assumed 43 students), 70 students instead of 140 for third cycle, and 25 instead of 50 for postsecondary survey outcomes). These calculations can inform us on the size of the impacts that the study can detect for subgroups of students; for example, females or males. As we can see in the table, in most cases, the MDI for subgroups differs by about 0.03 to 0.04 percentage points from the MDI for the full sample. For example, the MDI for the full sample of students is 3.6, and if we only have 50 students, the MDI is 4 percentage points. Therefore, if the program's effect on females (or males) is 0.04 percentage point higher than its effect on the full sample, we will be able to detect impacts for that subgroup. However, we should keep in mind that the SI-EITP model is, in

principle, not targeting any subgroup in particular. It is unclear whether any gender subgroup will benefit at a higher rate than the full sample does.

Table V.2. Minimum detectable impacts (MDIs) for full student sample or subgroups of half the students per Integrated System (percentage points, unless otherwise indicated)

Outcome	MDI for full sample	MDI for subgroup
7th grade dropout	4	4.5
7th grade repetition	2.2	2.4
Finished 7th grade and progressed to 8th	4.3	4.8
10th grade dropout	4.5	5.2
10th grade repetition	2.4	2.7
Finished 10th grade and progressed to 11th	4.9	5.5
PAES score (in standard deviations)	0.31	0.32
Graduation from secondary	3.4	3.9
Enrollment in postsecondary	6.1	7.2
Employment	6.2	7.4
Income (dollars)	132	157

Sources: Authors' calculations.

2. Limitations and challenges of the impact evaluation of the SI-EITP model

Although our impact evaluation design offers the best possible opportunity to inform the key research questions, it also faces some limitations and challenges that we will attempt to address to the extent possible.

No program has implemented all the components that SI-EITP will implement, so we should use existing evidence on individual components with caution. The SI-EITP model has various components that operate at different levels; for example, the component of sharing resources operates across schools within the Integrated System; improving school infrastructure and extending the school day operates at the school level; and training teachers operates at the teacher level. The literature review presents the evidence supporting each of these components. But in most cases, the studies implemented and evaluated only one component, and SI-EITP is a combination of those components. Furthermore, it is not clear how the implementation of each component in SI-EITP will compare to the implementation of the studies discussed. Therefore, caution should be exercised in using the results of these studies to estimate how big the effects of the SI-EITP will be. This study will, however, be novel in the sense that it will shed light on the impact of the particular combination of components implemented under the SI-EITP model, which can contribute to the literature of what works or does not in education.

We will not be able to separate the effects of each component implemented as part of SI-EITP, in particular the gender component. Because the schools in the treatment group will be able to implement all the components, we cannot assess the effect of each component separately. This is important to keep in mind, especially for the gender component, given that that all the schools in the treatment will be offered that component along with all others.

The period of exposure to the intervention is only two years. Implementation of a few activities began in 2018, but most activities will be starting implementation in 2019 and will continue through 2020. Therefore, we will be estimating the impact of about two years of the intervention. This period of exposure may not be long enough for the effects to materialize. As we discussed in the literature review, some of the interventions may need longer periods of exposure. For example, school-based management evidence says that effects could materialize in five years. Also, teacher training interventions have not had effects on student outcomes in short exposure periods, so we may need a longer follow-up period to see effects on student achievement.

There is a potential weak treatment contrast for the teacher training component of the SI-EITP. One of the challenges in the proposed evaluation design is that MINED is implementing a national teacher training focused on content knowledge rather than pedagogy, emphasizing teachers' knowledge of the main curricular areas. This teaching training is open to all teachers (including those in the comparison group). Although this training is enhanced by some additional trainings in methodology and evaluation that are only offered to treatment teachers, it is possible that teachers in the control group would also enroll in the teacher training program offered nationally (Programa Nacional de Formación de Maestros). If many control teachers also enroll in the national training program, the contrast between treatment and control teachers could be weak. In addition, as noted, teacher training is voluntary. Teachers in the SI-EITP will be encouraged to attend training, but they cannot be forced to attend, so there is no guarantee that all the teachers in the treatment group will participate in the training. In order to address this issue, we will carefully monitor the type of training received by teachers in both treatment and control groups. We will also be in continuous communication with the Ministry of Education, as we have been to this point, to assess the coverage of the national teaching training program

The implementation of the extension of the school day, in particular, and of other components of SI-EITP will be voluntary and will vary across schools. One of the goals of the SI-EITP model is to extend the length of the school day from 25 hours per week to at least 11 additional hours in grades 7–9. However, an important challenge in working to achieve this goal is that the extension will be gradual and voluntary because there is no requirement or MINED directive to extend the school day. If the engagement of students in activities during extended hours is low, or the school cannot offer many more additional hours, we might have small differences in hours that the model cannot detect. Similarly, in the previous implementations of SI-EITP, teacher training was in many cases voluntary. There will likely be variation in the extent to which this component is implemented in the study schools. The performance evaluation will document the extent to which schools implemented this component in order to better interpret the findings of the impact evaluation.

The statistical power of the impact evaluation is limited for some postsecondary outcomes and for subgroup analysis. If the SI-EITP model is as successful as MCC and MINED have assumed, it is likely that we can detect impacts on educational outcomes that will be as large as projected and be similar to what other studies have found. However, MCC has no targets for postsecondary outcomes, and no studies have found positive effects of similar interventions on employment and income. Therefore, it is likely that our study is not powered to detect long-term effects. In addition, MCC is interested in including an analysis of gender in the

study. We will only be able to detect impacts by gender subgroup if the model has larger effects on either gender group. However, it is unclear whether the effects on males or females will be larger than they are on the combined sample.

B. Analytic approach for the performance evaluation

Analyses of the quantitative monitoring and administrative data will help answer performance questions about implementation targets (for example, the number of teachers trained), and progress toward the outcome indicators that the program activities are collectively expected to affect (for example, school dropout rates). Qualitative analyses will help us achieve a nuanced understanding of program implementation and the progress toward the program's output and outcome goals.

For the semi-structured, in-person interviews and focus groups with key stakeholders (Table IV.2), we will develop and pre-test tailored qualitative data collection protocols before each round of qualitative data collection. We will obtain interview and focus group transcripts and conduct thematic analysis in two sequential steps. The first step involves *chunking* or data reduction, and mapping transcript segments to the core performance research questions. The second step involves *thematic framing*—the process of distilling themes and identifying patterns in qualitative data. We will develop a coding framework with a hierarchy of conceptual categories that are linked to the research questions and the program logic model, and complement those with interpretative thematic memos. We will also review several types of project documents including implementers' progress reports, operation plans, and program component designs. We will conduct a systematic process to extract relevant narrative and output indicator data from these and other similar sources.

Taking advantage of the complementarity that a mixed-method approach affords, we will integrate findings from descriptive quantitative analysis of monitoring and administrative data with findings from qualitative data sources, including the comparative case studies. We will analyze quantitative and qualitative data separately and then integrate the two to distill key lessons. We will use triangulation techniques to assess the consistency of the findings across methods and data sources, confirm patterns of findings, and identify potential discrepancies across the quantitative and qualitative data sources (Patton 2002). The findings of the performance evaluation will also provide contextual information that can help us interpret the impact findings.

VI. ADMINISTRATIVE ASPECTS

A. Summary of data sources and IRB requirements

For the baseline data collection, Mathematica requested IRB approval to ensure the protection of the rights and welfare of human subjects who were included in the study. For the SI-EITP evaluation, we requested IRB review for two data collection activities: (1) the pilot of instruments for the impact evaluation (that is, teacher, principal, and student surveys and classroom observations), and (2) actual data collection for the impact evaluation. We requested IRB approval through the Health Media Labs (HML) IRB, as we have done for similar studies.

There are two types of IRB reviews with different sets of requirements, depending on the type of data collection and the study population. The expedited review is permitted for studies whose activities pose no more than minimal risk to subjects—for example, survey research, focus group discussions, and observational studies. The full review is required when the research poses risks greater than the minimal risk typically encountered in daily life. Examples are clinical trials, strenuous exercise, stressful psychological tests, questions about illegal activities, blood or tissue removal, or research involving a vulnerable population such as children, the elderly, prisoners, pregnant women, students, and handicapped or mentally disabled persons.

Since the subjects from our study face no more than minimal risk, we requested an expedited review for the surveys and observations in both the pilot and the actual data collection for the impact evaluation. We received approval for the pilot study in April 2017. The pilot was conducted in July 2017. We also received approval for the revised instruments and observations in August 2017. Baseline data collection was conducted on October 2017. We plan to request IRB approvals for endline data collections in 2020.

B. Preparing data files for access, privacy, and documentation

After we produce the baseline, interim, and final evaluation reports, we plan to prepare corresponding de-identified data files, user manuals, and codebooks that may be made available to the public. These products will be de-identified according to the most recent guidelines set forth by MCC. The public-use data files will be free of personal or geographic identifiers that would permit unassisted identification of individual respondents or their households. In addition, we will remove or adjust variables that introduce reasonable risks of deductive disclosure of the identity of individual participants. Mathematica will remove all individual identifiers, including names, addresses, telephone numbers, government-issued identification numbers, and any other similar variables. We will also recode unique and rare data by using top and bottom coding or replacing affected observations with missing values. If necessary, we will also collapse into less identifiable categories any variables that make an individual highly visible as a consequence of geographic or other factors (such as ethnic classifications or languages spoken). These measures are designed to retain the usefulness of the data while preserving the privacy of survey respondents.

C. Dissemination plan

To ensure that the results and lessons from the evaluation reach a wide audience, we will work with MCC and FOMILENIO II to increase the visibility of the evaluation and its findings. Mathematica is committed to making the findings accessible through several venues:

- We will present key findings from the evaluation reports in Washington, DC, and El Salvador. The presentations will be targeted in particular to key stakeholders and policymakers, including El Salvador's Ministry of Education. We will inform stakeholders about the impacts, key findings from the implementation process, lessons learned, and results. The presentations will also give stakeholders an opportunity to engage directly with the research team and pose questions about findings.
- The evaluation reports will present findings in formats that are clear and accessible to nontechnical audiences.
- We will also write short summaries that can be translated into Spanish, depending on the audience.
- Mathematica will work with MCC to disseminate the findings through a variety of forums, including conferences focused on international education, in-country workshops and events, meetings at MCC offices, and other development conferences. Mathematica could also prepare journal articles to report impact findings.

D. Evaluation team roles and responsibilities

Dr. Larissa Campuzano, a senior researcher at Mathematica, will be the senior analyst for the impact evaluation of the SI-EITP model. Dr. Camila Fernández, a senior researcher at Mathematica, will lead data collection tasks. Since 20017, Dr. Campuzano has led all education evaluations under the first Compact. Dr. Fernández has substantial experience in measurement and has implemented surveys, observational measures, and standardized assessments for evaluations throughout Latin America. Dr. Evan Borkum will be the senior technical adviser for this contract. His responsibility will be to provide quality assurance for our deliverables. Ms. Ivonne Padilla will support data collection in El Salvador and organize data sets, prepare programming files, and produce all statistical output for quantitative analyses.

In addition, Mathematica will provide technical support for the in-country quantitative and qualitative data collection conducted by a third party. Throughout the evaluations, FOMILENIO II monitoring and evaluation (M&E) staff will supervise the data collection contract with the third party. MCC M&E staff will supervise Mathematica's evaluation work.

E. Evaluation timeline and reporting schedule

For the SI-EITP sub-Activity, FOMILENIO II envisions that the extension of the school day and teacher training will be completely implemented in 2018, and the secondary schools will be created in 2019-2020. Our findings will be summarized in three reports: baseline, interim, and final evaluation reports. We received the in-school surveys and observations collected in October 2017 in June 2018. Therefore, the baseline report will be ready by the first quarter of 2019. Our plan is to prepare a final report on educational outcomes by the end of 2023. This report will include follow up in-school data collected in 2020 and administrative data collected up to 2022

end of year status and if data is available on time 2023 enrollment. The timeline for a report focused on the outcomes of the post-secondary data collection will depend on MCC's decision on when to measure long-term outcomes.

This page has been left blank for double-sided copying.

REFERENCES

Angrist, J.D., and V. Lavy. "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools." *Journal of Labor Economics*, vol. 19, no. 2, 2001, pp. 343–369.

Barrera-Osorio, F., T. Fasih, and A. Patrinos. "Decentralized Decision-Making in Schools: The Theory and Evidence on School-Based Management." Washington, DC: World Bank, 2009.

Bellei, C. "Does Lengthening the School Day Increase Students' Academic Achievement? Results From a Natural Experiment in Chile." *Economics of Education Review*, vol. 28, no. 5, 2009, pp. 629–640.

Bonilla-Mejia, L. "Doble Jornada Escolar y Calidad de la Educación in Colombia." Documentos de Trabajo Sobre Economía Regional, no. 143, 2011.

Burde, D, and L.L. Linden. "The Effect of Proximity on School Enrollment: Evidence from A Randomized Controlled Trial in Afghanistan." Unpublished manuscript, 2009.

Campuzano L., R. Blair, V. Poggio, I. Padilla, S. Morgan, and M. Brannan. "Evaluation of the Formal Technical Education Sub-Activity of the Human Development Project, El Salvador" Washington, DC: Mathematica Policy Research, June 2016.

Carr-Hill, R., C. Rolleston, and R. Schendel. "The Effects of School-Based Decision Making on Educational Outcomes in Low and Middle Income Contexts: A Systematic Review Campbell Systematic Reviews 2016:9

Cerdán-Infantes, P., and C. Vermeersch. "More Time Is Better: An Evaluation of the Full-Time School Program in Uruguay." Washington, DC: World Bank, 2007.

Crespin, E. "Sistematización del Modelo de Sistemas Integrados de Escuela Inclusiva de Tiempo Pleno, SI-EITP. Informe Final." 2015.

Dumitrescu, A, D. Levy, C. Orfield, and M. Sloan. "Impact Evaluation of Niger's IMAGINE Program." Washington, DC: Mathematica Policy Research, 2011.

Garcia, S., C. Fernandez, and C. Weiss. "Does Lengthening the School Day Reduce the Likelihood of Early School Dropout and Grade Repetition? Evidence from Colombia." Paper presented at the Population Association of America 2012 Annual Meeting, San Francisco, CA, 2012.

Garet, M. S., S. Cronen, M. Eaton, A. Kurki, M. Ludwig, W. Jones, W., et al. "The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement (NCEE 2008-4030)." Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. 2008. Available at http://ies.ed.gov/ncee/pdf/20084030.pdf. Accessed August 8, 2018.

Garet, M. S., J. B. Heppen, K. Walters, J. Parkinson, T.M. Smith, M. Song, et al. "Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher Professional Development (NCEE 2016-4010)." Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. 2016, Available at http://ies.ed.gov/ncee/pubs/20164010/pdf/20164010.pdf. Accessed August 7, 2018.

Garet, M. S., A. Wayne, F. Stancavage, J. Taylor, M. Eaton, K. Walters, et al. "Middle School Mathematics Professional Development Impact Study: Findings After the Second Year of Implementation (NCEE 2011-4024)." Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. 2011. Available at http://ies.ed.gov/ncee/pubs/20114024/pdf/20114024.pdf. Accessed August 7, 2018.

Garet, M. S., A. Wayne, F. Stancavage, J. Taylor, K. Walters, M. Song, et al. "Middle School Mathematics Professional Development Impact Study: Findings After the First Year of Implementation (NCEE 2010-4009)." Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. 2010. Available at https://ies.ed.gov/ncee/pubs/20104009/pdf/20104010.pdf. Accessed August 7, 2018.

Gersten, R., J. Dimino, M. Jayanthi, J. Kim, and L. Santoro. "Teacher Study Group: Impact of the Professional Development Model on Reading Instruction and Student Outcomes in First-Grade Classrooms." *American Educational Research Journal*, vol. 47, no. 3, 2010, pp. 694–739.

Gersten, R., J. Dimino, M. Jayanthi, R. Newman-Gonchar, and M.J. Taylor. "Impact of the Teacher Study Group Professional Development Program on Student Vocabulary and Observed Teaching Practice: A Replication in First-grade Classrooms." Paper presented at the annual meeting of the Society for Research on Educational Effectiveness, Washington, DC, March 2013.

Gertler, P., H. Patrinos, and M. Rubio-Codina. "Empowering Parents to Improve Education: Evidence from Rural Mexico." Policy Research Working Paper 3935. Washington, DC: World Bank, 2006.

Gertler, P., H. Patrinos, and M. Rubio-Codina. "Methodological Issues in the Evaluation of School-Based Management Reforms." Unpublished manuscript. Washington, DC: World Bank, 2007.

Goodrick, D. Comparative Case Studies, *Methodological Briefs: Impact Evaluation 9*. Florence, Italy: UNICEF Office of Research, 2014.

Halpin, P. F., C. Torrente, and L. Aber. "The Early Grade Reading Assessment: Its Internal Structure and Use in Impact Evaluation." Working paper, 2014.

Hincapie, D. "Do Longer School Days Improve Student Achievement? Evidence from Colombia." IDB Working Paper Series no. IDB-WP-679, 2016.

Holland, Peter Anthony, Pablo Alfaro, and David Evans. "Extending the School Day in Latin America and the Caribbean." Policy Research working paper no. WPS 7309, Impact Evaluation series. Washington, DC: World Bank Group, 2015.

Joint USG-GOES technical team. "Partnership for Growth: El Salvador Constraint Analysis." 2011.

Lokshin, M., and R. Yemtsov. "Evaluating the Impact of Infrastructure Rehabilitation Projects on Household Welfare in Rural Georgia." Policy Research Working Paper Series 3155. Washington, DC: World Bank, 2003.

Modelo de Escuela Inclusiva de Tiempo Pleno en El Salvador. Documento de Sistematización. Ministerio de Educación. Gobierno de El Salvador. Available at https://www.fomilenioii.gob.sv/modelo-de-escuela-inclusiva-de-tiempo-pleno-en-el-salvador--2

Newman, J., M. Pradhan, L.B. Rawlings, G. Ridder, R. Coa, and J.L. Evia. "An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund." *The World Bank Economic Review*, vol. 16, no. 2, 2002, pp. 241–274.

Pires, T., and S. Urzua. "Longer School Days, Better Outcomes?" Working paper. 2015.

"Sistematización del Proceso de Elaboración del Sistema Integrado como Estrategia de Expansión del Modelo de Escuela Inclusiva de Tiempo Pleno (EITP), para que Sirva de Guía al Ministerio de Educación de El Salvador (MINED), en la Implementación de Dicha Estrategia. Informe Final." Programa de Fortalecimiento de la Educación Básica EQUIP2, El Salvador Sistema Integrado de la Escuela Inclusiva de Tiempo Pleno. USAID, 2012.

Schochet, P. Z. "Statistical Power for School-Based RCTs with Binary Outcomes," *Journal of Research on Educational Effectiveness*, vol. 6, 2013, pp. 263–294.

Snilstveit, B, J. Stevenson, R. Menon, D. Phillips, E. Gallagher, M. Geleen, H. Jobse, T. Schmidt, and E. Jimenez. "The Impact of Education Programmes on Learning and School Participation in Low- and Middle-Income Countries: A Systematic Review Summary Report." 3ie Systematic Review Summary 7. London: International Initiative for Impact Evaluation (3ie), 2016.

Yin, Robert K. *Case Study Research: Design and Methods*. Fifth edition. Los Angeles, CA: Sage, 2014.

This page has been left blank for double-sided copying.

APPENDIX A

SAMPLE SIZES AND TIMELINE PROPOSED

This page has been left blank for double-sided copying.

Table A.1. Sample sizes and timeline proposed

Instrument	Target sample respondents	Sample size	Data collection and timing of data request
In-school data collection			
Principal survey	• Principals from third cycle and secondary schools	• 116 principals in third cycle schools • 74 principals in secondary school at baseline and 85 principals at follow-up (assuming 11 new secondary schools will be established)	• Baseline data collection: second semester of 2017 • Follow-up data collection: second semester of 2020
Teacher survey	• Teachers from third cycle and secondary schools – 3 teachers of different grades or content areas will be randomly selected in each of the schools.	• 348 teachers in third cycle schools • 222 teachers in secondary schools at baseline and 255 at follow-up	• Baseline data collection: second semester of 2017 • Follow-up data collection: second semester of 2020
Student survey	• Students from third cycle and secondary schools. – 10 students will be randomly selected in each of the schools.	• 1160 students in third cycle schools • 740 students in secondary schools at baseline and 850 at follow-up	• Baseline data collection: second semester of 2017 • Follow-up data collection: second semester of 2020
Classroom observation	• Classrooms in third cycle and secondary schools – 3 classrooms will be randomly selected for observation in each of the schools.	• 348 classroom observations in third cycle schools • 222 classroom observations in secondary schools at baseline and at least 255 classroom observations at follow-up (assuming 11 new secondary schools will be established)	• Baseline data collection: second semester of 2017 • Follow-up data collection: second semester of 2020
Data provided by MINED			
School records (third cycle students)	• Enrollment records from students in third cycle schools at the beginning of the school year	• Approximately 39,000 student-level records for all students enrolled in all third cycle schools in study sample—about 383 schools	• Baseline/pilot: In May 2017, we requested students' records from the beginning of the school year from SAE 2016 and 2017. • Follow-up: Every May, we will request students' records from the beginning of the school year for 2018, 2019, 2020, 2021, 2022 and 2023

Instrument	Target sample respondents	Sample size	Data collection and timing of data request
School records (secondary school students)	• Records from students in secondary schools at the beginning of the school year (enrollment) and end of the school year (status such as pass, fail, dropout)	• Approximately 18,000 student-level records for all students enrolled in all secondary schools in study sample—about 90 schools at baseline and 110 at follow-up	• Baseline/pilot: In March 2017, we requested students' records from SIRAI for the status of students at the end of the 2016 school year and the beginning of the 2017 school year. We also requested beginning-of-the-year data from 2016. • Follow-up: In May of each year we will request students' records from the beginning of that school year and the records on students' status at the end of the previous school year for 2018, 2019, 2020, 2021, and 2022. We will also request records from the beginning of the 2023 school year.
School records for transition outcome	• Cohort of students enrolled in 9th grade who will make the transition to 10th grade in the next school year	• Approximately 12,000 student records for all students who enrolled in grade 9 in the study schools—about 383 schools • Approximately 101,500 student records for all students who enrolled in 10th grade in all the secondary schools in the country—931 schools	• Baseline/pilot: In May 2017, we requested SAE data for all students enrolled in grade 9 at the beginning of 2016. We also requested MINED the enrollment status of these students at the beginning of 2017 in all the schools in the country. • Follow-up: In early 2023, we will request data for all students enrolled in grade 9 at the beginning of 2022. We will also request data for all students enrolled in grade 10 at the beginning of 2023 in all the schools in the country.
Test scores	• Students in grade 11th	• Approximately 6,600 scores for all students enrolled in the secondary schools of the sample—90 schools at baseline and 101 at follow-up approximately	• Baseline/pilot: In May 2017, we requested PAES information for the 2017 school year • Follow-up: We will request annually PAES information for the 2018-2022 school years

Instrument	Target sample respondents	Sample size	Data collection and timing of data request
Data collection postsecondary by a third party			
Post compact student survey	• Randomly selected students that were enrolled in their last year of secondary school – 50 students in each of the systems	• 4,250 students (50 per system)	• TBD

Instrument	Target sample respondents	Sample size	Data collection and timing of data request
In-school records data collection by a third-party (pilot)			
Data collection of student records directly in a subsample of schools	• Randomly selected 32 third cycle schools • Randomly selected 8 secondary schools (sub-sample of 8% of the study)	• All student records from the 32 third cycle schools selected. • All student records from the 8 secondary schools selected.	• In the first semester of 2017, a third-party data collector visited 40 schools to collect student records for all students initially enrolled in the 2016 school year in each school and their status at the end of that school year.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ TUCSON, AZ ■ WASHINGTON, DC

MATHEMATICA Policy Research

Mathematica® is a registered trademark of Mathematica Policy Research, Inc.