Kosovo Reliable Energy Landscape Project Evaluation

Contract No. 95332419C0104

Evaluation Design Report

Submitted To:

Cindy Sobieski, Project Officer
MCC—Headquarters
Millennium Challenge Corporation

Submitted By:

IMPAQ International, LLC
10420 Little Patuxent Parkway, Suite 300
Columbia, MD 21044

Date: February 5, 2021

TABLE OF CONTENTS

TABLE OF EXHIBITS

LIST OF ACRONYMS

AER	Apartment energy efficiency retrofits
BC&O	Behavior change and outreach
CBB	Consumption-based billing
CEA	Cost-effectiveness analysis
DHM	District heating and metering
EBRD	European Bank for Reconstruction and Development
EMC	Evaluation Management Committee
ERO	Energy Regulatory Office
FGD	Focus group discussion
GDP	Gross Domestic Product
HER	Household energy efficiency retrofits
HOA	Homeowners associations
IC	Implementing contractor
IPP	Independent power producer
IRB	Institutional Review Board
ITS	Interrupted time series
KCGF	Kosovo Credit Guarantee Fund
KEDS	Kosovo Energy Distribution Services
KII	Key informant interviews
kWh	Kilowatt-hour
MAB	Multi-apartment buildings
MCC	Millennium Challenge Corporation
MFK	Millennium Foundation of Kosovo
MW	Megawatts
ODC	Other direct costs
RELP	Reliable Energy Landscape Project
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SEEK	Subsidies for Energy Efficiency in Kosovo
USAID	U.S. Agency for International Development
WEE	Women in Energy Efficiency
WI-HER	Women Influencing Health, Education, and Rule of Law

1. INTRODUCTION

Kosovo gained independence from Serbia in 2008, and its economy has grown at a rate that is consistently higher than the Western Balkan average for the last two decades. Gross domestic product per capita grew from €1,000 in 2000 to €4,100 in 2019 (World Bank 2019). However, Kosovo still is one of the poorest countries in Europe. Using the domestic poverty line of €1.72 per day, as defined by the Kosovo Agency of Statistics, 29.7 percent of its population is considered poor (World Bank 2015), and its per capita income is one-tenth that of European Union levels (World Bank 2017a). An analysis of the binding constraints to sustainable economic growth and poverty alleviation in Kosovo by the Millennium Foundation of Kosovo (MFK) revealed three main barriers: an unreliable supply of electricity, weak rule of law and perceptions regarding the rule of law, and poor environmental services such as water, environment, and health *(*MFK 2017).[1]

To spur economic growth and reduce poverty in Kosovo, the Millennium Challenge Corporation (MCC) and the Government of Kosovo signed a USD$49 million, 4-year threshold program in September 2017 (MCC 2017).[2] The threshold program aims to address two key constraints to economic growth in Kosovo: unreliable supply of electricity, and real and perceived weakness in the rule of law, government accountability, and transparency. The program comprises two projects:

- **The Transparent and Accountable Governance Project** which supports the implementation of a case management information system to make judicial information publicly available, and improvements to the collection and reporting of environmental data to the public.
- **The Reliable Energy Landscape Project (RELP)** which aims to reduce the gap between energy supply and demand by encouraging greater household efficiency and bolstering private-sector participation in the power sector.

In 2019, MCC contracted with IMPAQ International (IMPAQ) to conduct an evaluation of RELP. Both electricity consumption levels and the share of electricity bills in household expenditures in Kosovo are higher than in other countries in the Balkans. Furthermore, the country lags in energy efficiency. Residential energy intensity in Kosovo is around 300 kilowatt-hours per square meter per year (kWh/m^2/year), almost double the European Union average of 174 kWh/m^2/year (Kokx 2017). Since higher energy intensities indicate a higher price or cost of converting energy, this implies that households in Kosovo are being heated inefficiently. Additionally, in 2016, Kosovo imported 555 gigawatt hours of electricity at a cost of €47 million—40 percent more expensive than the cost of domestic electricity consumption (MCC 2018a). Furthermore, based on outage indicators, a more reliable measure of the shadow cost of electricity, Kosovo ranks at the bottom among Western Balkan countries. Outages are estimated to reduce GDP by around 5 percent annually, and this unreliable electricity supply exacerbates the country's gap between energy supply and demand.

The objective of RELP is to reduce this gap, through three main activities: subsidies for energy efficiency in Kosovo (SEEK), district heating and metering (DHM), and independent power producers (IPPs). On the

[1] Binding constraints are those that, if relieved, would produce the largest gains in growth and entrepreneurship of any potential constraint areas (Hausmann et al. 2005).

[2] In December 2018, Kosovo was deemed eligible to also develop a compact. In general, MCC's threshold programs include small grants focused on policy and institutional reform in selected countries that come close to passing MCC's eligibility criteria and show a firm commitment to improving their policy performance. Compacts on the other hand are large, five-year grants for selected countries that meet MCC's eligibility criteria.

demand side, RELP will address two barriers to investments in energy efficiency: a general lack of consumer awareness about energy saving measures and their benefits, including the benefits of metering and other energy saving devices, and the inability of poor households to pay for energy efficiency measures. These objectives will be achieved by lowering energy use through piloting household investments in energy efficiency, metering existing district heating consumers, and switching new consumers to cost-effective district heating. On the supply side, RELP will aim to reduce barriers to renewable energy entrants to the market by stimulating the business environment.

The RELP evaluation will assess the extent to which project activities affected electricity and district heating consumption in targeted areas in Kosovo. The evaluation will also examine substitution between different energy sources. Furthermore, the evaluation will assess the implementation of each of the activities under RELP and verify the validity of the project logic across all three activities. Finally, the evaluation will generate evidence for lessons learned from implementation and provide recommendations for project sustainability. Adhering to MCC's monitoring and evaluation policy, this evaluation will help MCC, MFK, and its partners, including the Government of Kosovo, understand the effects of MCC's RELP on the energy sector in Kosovo. The study also will provide evidence to improve the performance of government and private stakeholders in designing and sustainably enforcing energy efficient behaviors.

This report describes IMPAQ's comprehensive mixed-methods design for the RELP evaluation and has three main objectives: (1) communicate the purpose and guiding evaluation questions behind the evaluation of the three activities under RELP; (2) describe the quantitative and qualitative methods chosen to respond to these evaluation questions, along with the challenges/limitations of these methods; and (3) outline IMPAQ's administrative approach to executing the evaluation, including the evaluation team, structure, and schedule. We will use a range of performance and impact evaluation methods to answer the evaluation questions. Our mixed-methods approach will seek to measure the impacts of and understand the changes related to RELP. Where relevant, we will disaggregate quantitative findings by key social groups, including gender, age, household economic well-being, and ethnicity, and we will use qualitative findings to provide further nuance about the heterogeneous impacts of RELP.

In the sections that follow, we provide context for the evaluation and describe its design in detail. **Section 2** provides an overview of the Kosovo threshold program and the proposed activities under RELP. We also describe the RELP theory of change and examine existing literature about energy efficiency at the household level and initiatives to increase renewable energy supply in developing countries. In **Section 3**, we present the evaluation questions and provide an overview of the quantitative and qualitative designs and data sources that will enable us to answer these questions. We also detail the evaluation timeline. Next, in **Sections 4**, **5**, and **6**, we detail the evaluation designs for each activity under RELP. In **Section 7**, we present the evaluation design for the women in energy efficiency (WEE) activity, and **Section 8** presents the evaluation design for broader sectoral-level outcomes. Finally in **Section 9**, we discuss overall implementation and evaluation challenges. We conclude in **Section 10** with a discussion of administrative details, including institutional review board requirements, data anonymization processes, dissemination plans, evaluation team roles and responsibilities, and the evaluation budget.

2. OVERVIEW OF KOSOVO THRESHOLD AND INTERVENTION

In this section, we describe each of the three activities included in the RELP and the mechanisms through which they are expected to affect outcomes as outlined in the theory of change. Specifically, **Section 2.1** provides an overview of the RELP and **Section 2.2** describes the RELP theory of change. Finally, **Section 2.3** summarizes key gaps in existing literature about energy efficiency and explains the contributions of this study to the energy sector in Kosovo and the energy field more broadly.

2.1 Overview of RELP

The overarching goal of the Kosovo MCC Threshold Program is to achieve poverty reduction through economic growth. According the "Monitoring and Evaluation Plan", the objective of RELP is to reduce the gap between energy supply and demand, by lowering energy use through piloting household investments in energy efficiency, switching to cost-effective non-electricity sources of heating, and reducing barriers to IPP entrants to the market. RELP aims to achieve these objectives through four main activities. We describe each of these below.

Activity 1.1 Subsidies for Energy Efficiency in Kosovo

The primary focus of the SEEK activity is the promotion of energy efficiency measures (including thermal insulation in walls and roofs, energy efficient windows, weather sealing, energy efficient water heaters, and energy efficient biomass stoves and furnaces) to reduce the consumption of electricity for heating.[3] SEEK aims to provide incentives to residential consumers to invest in retrofits to reduce their household electricity consumption, in addition to providing incentives aimed at increasing the involvement of women in the energy sector.[4] Behavior change and outreach (BC&O) is a crucial component to achieve the intended objectives of this activity. Finally, SEEK also encourages female employment through explicit requirements in the technical specifications for the IC to hire women. SEEK consists of two sub-activities:

1. **Household energy efficiency retrofits (HER).** This sub-activity will provide incentives to residential consumers to invest in energy efficiency retrofits. A key objective of HER is to test and evaluate the most cost-effective incentive delivery approaches for different beneficiary groups. Another key objective is to scale up the behavior change approaches found to be most effective, beyond the time horizon of the threshold program.

   Based on the study design presented in the "HER Operations Manual," the HER sub-activity will include at least two iterations, the first iteration will include only one model and the second iteration will include one to three models. The goal of the first iteration will be to test a streamlined model, with adjustments to grant levels received by households and other intervention package elements (e.g., eligibility criteria) to implement during the second iteration. After the end of the first iteration of HER, the implementing contractor (IC) will still evaluate the performance and lessons learned and propose adjustments (e.g., to the eligibility criteria) for the second iteration. Based on learnings from the first iteration, further measurements and models will be selected to monitor in the second iteration. The design of the second iteration of HER will also take into consideration identified behavior change elements, which may be related to performance (e.g., energy savings), communication (e.g., media modalities, message targeting), or knowledge (e.g., training programs).

[3] The SEEK activity was originally called PIEE – Pilot Incentives for Investment in Energy Efficiency. It was renamed "SEEK" in January 2020.

[4] Approximately 2,600 households and 25 multi-apartment buildings (MABs) will receive benefits under the SEEK activity.

The goal of the second iteration will be to evaluate the overall performance and lessons learned specific to each model and propose optimum one, to maximum three models for scaling up across various segments of the residential sector in Kosovo. Thus, both iterations of the HER sub-activity will test models, with learning occurring in between the two iterations with the goal of identifying a cost-effective iteration by the end of the threshold program.[5]

All 38 municipalities of Kosovo will receive the BC&O campaign of the HER sub-activity in both iterations, and households from across the country who are currently using electricity are eligible to participate. Detailed general, technical, and building criteria are presented in the "HER Operations Manual." The first iteration will include 500 households and the second iteration will target 2,100 households. Additional details of the HER iterative study, including incentive models and study sample, are outlined in **Appendix A**.

2. **Apartment energy retrofits (AER).** SEEK will provide grants to residential apartment buildings in selected municipalities to implement energy efficiency upgrades to common areas and the building envelope. To facilitate implementation, additional homeowners associations (HOAs) will be established, and municipalities will co-invest in retrofits in apartment buildings. Seven municipalities will be targeted under AER: Pristina, Mitrovica, Lipjan, Gjakova, Gracanica, Viti, and Novobrdo. The key target for the AER sub-activity is to retrofit 25 multi-apartment buildings (MABs) in the selected municipalities, with the goal of spreading out the 25 MABs between the seven municipalities.

Activity 1.2 District Heating and Metering

The DHM activity aims to support the transition of district heating residential customers on the Termokos network in Pristina from area-based billing to consumption-based billing (CBB). This intervention will target the 12,000 existing consumers living in residential buildings that are connected to district heating services in Pristina (based on the "DHM Inception Report"). The DHM measures in Pristina will significantly support scaling up district heating networks in Kosovo.[6] According to the "DHM Inception Report," the number of new consumers reached will depend on the energy savings received from installing meters in residences of existing district heating customers. Termokos and the Energy Regulatory Office (ERO) will work collaboratively on tariff design. Additionally, the project will test different modalities and strategies for packaging DHM and energy efficiency services to maximize customer value, reliability, and sector cost efficiency. This will entail feasibility work, including formative research to design a social and behavior change campaign. The DHM sub-activity is also expected to improve the performance of the Termokos network. Finally, similar to SEEK, DHM also includes explicit requirements for the IC to hire women and thus, encourages female employment.

Activity 1.3 Independent Power Producer Project Finance Facilitation

On the supply side, the IPP project finance facilitation activity directly addresses the problem of unreliable electricity supply, by facilitating entry into the renewable energy market. One of the key elements of the activity is that MFK will provide technical assistance and direct financial support to the Kosovo Credit

[5] This is in contrast to the original design wherein the second iteration was envisioned to implement the model being tested in the first iteration. The updated design will test models in both iterations, with learning in between the two iterations.

[6] Termokos supplies only 25 percent of Pristina's heating demand (covering the crowded suburbs of Dardania, Ulpiana, Sunny Hill, and City Centre). With savings from CBB under the DHM activity, Termokos can extend the network to other adjacent suburbs to connect new customers/demand, up to the available heat production capacity at the cogeneration unit ("DHM Inception Report").

Guarantee Fund (KCGF) to aid the expansion and build the internal capacity of KCGF to enable it to serve as a key domestic catalyst within the country to unlock commercial financing for small-scale renewable energy generation. Additionally, MFK will seek to develop a standardized renewable energy project-financing framework. The main goal of these complementary activities—referred to as the Kosovo Project Acceleration program—will be to generate a pipeline for the new KCGF renewable energy guarantees by providing targeted technical assistance to IPPs when they are submitting applications and to motivate them to bring renewable energy projects to a financial close against a clear timeline. Technical assistance will also be provided to banks for designing products on project finance.

Activity 1.4 Women in Energy Efficiency (WEE)

The overarching objective of the WEE activity is to promote female employment. The WEE activity seeks to increase women's awareness of opportunities in the energy sector, provide on-the-job and other relevant training to increase the pool of technically skilled women willing to work in energy-related jobs, and build linkages between energy employers and qualified women candidates in Kosovo. Improving the supply of Kosovar women with technical skills relevant to the energy sector will enable them to take advantage of economic opportunities that result from the implementation of sustainable energy initiatives. Approaches to achieve this include:

- *Scholarship Program.* This program will provide selected Kosovar women in high school, university, or with some work experience the opportunity to obtain a two-year technical degree in energy-related fields in international education institutions.
- *Internship Program.* This program provides current university students and recent university graduates with practical experience in relevant government agencies and in public and private companies in the energy sector.
- *Technical Assistance and Women Entrepreneurship Grant Program.* Under the program, technical assistance and incentives for women entrepreneurs will be provided to increase their activity in the energy sector and upgrade their businesses through energy solutions.
- *Women in Energy Summer Camp.* This initiative was created to encourage adolescent girls to pursue further education and potential careers in science, technology, engineering, and math (STEM) fields.

In the original design of RELP, the WEE activity was a sub-activity of SEEK. However, based on the MCC Evaluation Management Committee (EMC) meeting in July 2020, it was decided to treat WEE as a stand-alone activity given its objective of promoting female employment in the energy sector. While this is now a standalone activity, the IC for the SEEK activity will continue to implement the technical assistance and women entrepreneurship grant program, while MFK will implement the scholarship program, internship program, and the women in energy summer camp.

2.2 Overview of Theory of Change

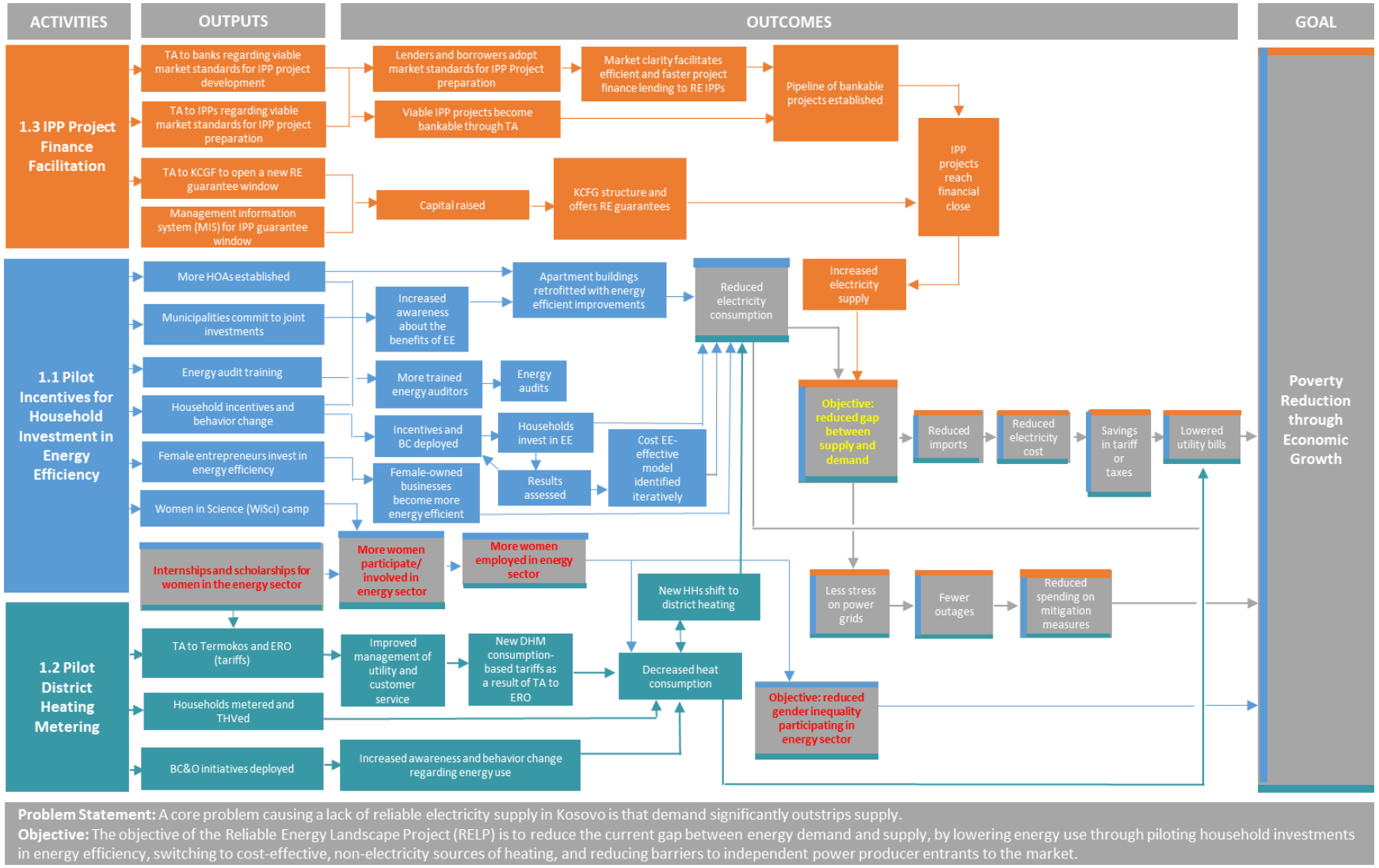
As per the threshold program agreement, the objective of RELP is to reduce the current gap between energy demand and supply by lowering energy use through piloting household investments in energy efficiency, switching to cost-effective non-electricity sources of heating, and reducing barriers to independent power producer entrants to the market. The RELP theory of change focuses on a combination of renewable energy investments on the supply side and energy efficiency incentives on the

demand side, to achieve the project's main objective. **Exhibit 1** illustrates RELP's original draft theory of change from December 2019, which includes the activities, outputs, and medium- and long-term outcomes related to RELP's three activities. It shows how planned activities under RELP are expected to lead to reduced electricity consumption, increased electricity supply, a smaller demand-supply gap, a reduced number and shorter duration of outages, fewer imports, and lowered utility bills. As a note, this theory of change includes WEE as a sub-activity of SEEK. As mentioned above, as of July 2020, WEE is no longer a sub-activity of SEEK and will be treated as a standalone activity. However, some aspects of the logic model lack clarity, and some important assumptions are missing. Furthermore, the latest version of the logic model is complex in terms of the number of nodes and interdependencies between them. Additionally, several outcomes must be achieved before reaching the primary objective of reducing the gap between energy demand and supply.

As discussed in IMPAQ's "Evaluability Assessment," IMPAQ modified the theory of change from MCC's original version and suggested revisions (see **Appendix B**). Based on extensive document review and the information gathered during a scoping trip, the "Evaluability Assessment" includes a discussion about whether the assumptions outlined in the RELP theory of change are realistic and whether there is evidence that the proposed activities can led to the intended outcomes. The evaluation design described in this report will allow us to assess whether some of these assumptions are accurate. For instance, a key assumption of the SEEK and DHM activities is that installation of energy efficiency retrofits (under SEEK) and installation of apartment-level meters (under DHM) will lead to reduced electricity consumption and district heating consumption, respectively. However, improvements in energy efficiency do not necessarily lead to one-to-one reductions in energy consumption (Galvin 2014), because energy efficiency gains alter the perceived cost of comfort and may generate shifts in consumption patterns—a "rebound effect" (Aydin et al. 2017). This discrepancy between the expected/realized energy savings and the optimal/actual investments in energy efficiency technologies is often referred to as the "energy efficiency gap" or "energy efficiency paradox," which has been illustrated and examined in multiple studies (Schleich and Gruber 2008; Chai and Yeo 2012; Allcott and Greenstone 2012; Ameli and Brandt 2015; Gerarden et al. 2017). That is, households and apartment buildings in the SEEK sub-activity might actually increase their energy consumption after the installation of retrofits. Similarly, households in the DHM sub-activity that receive new meters and thermal valves also may start heating larger areas of their households because of meter-induced energy savings.

Another important assumption is that retrofits and meters installed for the project will be maintained and function for its entire lifespan. Furthermore, there is an implicit assumption that activities such as internships, scholarships, and the energy summer camps will lead to not only a reduction in inequality of opportunities for women in the energy sector, but also to broad sectoral-level changes. Through quantitative and qualitative data collected from beneficiaries—including households, HOAs, ICs, MFK staff, Termokos staff, women beneficiaries, and ERO—we will assess the long-term sustainability and potential for sectoral-level changes as a result of RELP. As a note, no cost-benefit analysis was carried out under the Kosovo threshold program, and consequently none is discussed in this report.

Exhibit 1. RELP Logic Model Version 2[7]



[7] BC: behavior change, BC&O: behavior change and outreach, EE: energy efficiency, RE: renewable energy, TA: technical assistance and THV: thermal heat valves.

2.3 Literature Review

2.3.1 Status of Kosovo Energy Sector

Kosovo is one of the poorest countries in Europe, with almost one-third of the population living below the poverty line and 1 in 10 people living in extreme poverty (World Bank 2018). According to the United Nations Development Programme, Kosovo has the youngest population in Europe—half of its roughly 2 million population is younger than 25. The growing youth population, which has experienced high unemployment rates in the past decade, calls for a strong nationwide strategy to foster a faster growing economy.[8] A key contributor to poor economic growth in Kosovo is the large gap between electricity supply and demand.[9] The country's electricity system is outdated, inadequate, and undependable, which poses significant challenges to economic growth and development. The main obstacles faced by the energy sector are poor infrastructure, significant energy losses, and overdependence on lignite energy. Frequent power outages hinder investment and disrupt manufacturing, education, and health services. Without a reliable, affordable, and sustainable electricity supply, foreign and domestic firms are reluctant to invest in Kosovo and create jobs.

In this context, in 2017, the Government of Kosovo drafted its National Energy Strategy (2017–2026), which was based on a detailed energy sector analysis and defined five strategic objectives: (1) security of a sustainable, high-quality, safe, and reliable electricity supply with adequate capacities for stable power system operation; (2) integration into the regional energy market; (3) enhancement of existing thermal system capacities and construction of new capacities; (4) development of natural gas infrastructure; and (5) fulfillment of targets and obligations in energy efficiency, renewable energy sources, and environmental protection.

Electricity Supply and Demand in Kosovo

Demand for energy has been growing rapidly over the past decade. Today, more than 50 percent of electricity in Kosovo is used by the residential sector, 80 percent of which is used for space and water heating—much higher than the European Union average. In 2016, Kosovo imported 555 gigawatt hours of electricity at a cost of €47 million—more than 40 percent higher than the cost of domestic electricity consumption (MCC 2018). On the supply side, the basic primary energy source used in Kosovo to satisfy energy needs of different economic sectors is lignite (MED 2012). Kosovo lacks natural gas supply and its inability to import it (because of lack of infrastructure) limits the country's energy options and its ability to contain costs (MED 2017).[10] Recent repairs to Kosovo's two power plants have significantly lowered outages, but without further investment and a settling of Kosovo's participation in the European Network of Transmission System Operators for Electricity, reliability is expected to worsen again.[11]

[8] The overall unemployment rate is about 33 percent and unemployment for youth aged 15-24 years is about 60 percent. Kosovo has the highest unemployment rate in all of Europe. (World Bank 2017a).

[9] In 2017, MFK conducted a comprehensive analysis of constraints that hinder sustainable economic growth and poverty alleviation in Kosovo and identified unreliable electricity supply as a binding constraint for economic development in Kosovo.

[10] The Trans Adriatic Pipeline, connecting Greece, Albania, and Italy to supplies from Azerbaijan, is due to be completed in 2020. In theory, a spur could have been added to link up Kosovo, but for various reasons, it was not feasible.

[11] The European Network of Transmission System Operators represents 43 electricity transmission system operators from 36 countries across Europe and thus extends beyond European Union borders. Kosovo's participation in the group is held up by the country's disagreements with Serbia over Kosovo's official name.

Electricity Generation and Transmission in Kosovo

The electricity generation sector is dominated by the Kosovo Energy Corporation, which owns and operates the country's greatest share of generation capacities: Kosova A and Kosova B. Because of lack of maintenance and lack of necessary periodic rehabilitation, the technical availability of Kosova A and Kosova B is far below installed capacities and this has not been improved despite several overhauls performed in the last 10 years. The soundest element of Kosovo's energy system is transmission. Due to significant investments, it is considered one of the most efficient and stable systems in the region. The transmission system is managed by the Kosovo Transmission System and Market Operator.[12]

Renewable Energy in Kosovo

Residential space heating in Kosovo is largely based on firewood as fuel. Households and the services sectors are major consumers of energy for heating, whereas the agriculture sector accounts for a very small share of heating energy consumption (Kosovo Agency of Statistics 2015). After the Kosovo War ended, several district heating programs, the cogeneration program of thermal energy supply from Kosovo B, and the rehabilitation of thermal network and substations have enabled a higher quality of supply for existing consumers and possibilities for connection of new ones.

Renewable energy is witnessing a steady increase in Kosovo. Its share in final gross consumption in 2015 was 19.7 percent. The Government of Kosovo aims to increase the installed renewable energy to 401–470 megawatts (MW) by 2026, from 99.5 MW in 2017, depending on the development scenario (MED 2017). To encourage the use of renewable energy, Kosovo has set up a legal framework, as well as a support plan through feed-in tariffs for hydropower, wind energy, photovoltaic energy, and biomass.[13] One of the incentive measures for renewable energy makes it the legal obligation of the Kosovo Transmission System and Market Operator to purchase renewable energy generation with the regulated feed-in tariffs, as defined by the ERO. The ERO is also responsible for developing methodologies for regulated feed-in tariffs, issuing licenses for energy activities and authorizations for the construction of renewable energy capacities, and issuing certificates of origin for energy produced by renewable energy. Furthermore, to support and promote the use of renewable energy, the Ministry of Economic Development drafted a 10-year National Renewable Energy Action Plan with the support of the World Bank and the U.S. Agency for International Development (USAID), with targets and measures to be achieved by 2020. The third progress report on the National Renewable Energy Action Plan revealed that Kosovo registered a 22.9 percent share of renewable energy in 2017 and therefore is on track to achieve the target set for 2020 (MED 2018).

Gender-specific Energy Issues in Kosovo

A lack of a reliable, efficient, and secure energy supply can affect women and men differently (Rewald 2017). In Kosovo, this is a direct result of men's and women's different socioeconomic backgrounds, roles, responsibilities, and access to and control of resources. For instance, women and girls are inactive and unemployed at much higher rates than men (Kosovo Agency of Statistics 2019b). The limited energy sector data suggests that women in Kosovo lack access to information and decision-making power within the energy sector (IMPAQ 2018). Women are underrepresented in the field of energy studies, both as

[12] Because Kosovo does not participate in the European Network of Transmission System Operators for Electricity, the time between an unexpected outage and transmission system response is increased (Forbes 2013).

[13] The applied feed-in tariff is different for different technologies, as following: water energy (small hydropower plants): 67 EUR/ megawatt hours (MWh); wind energy: 85 EUR/MWh; and biomass energy 71.3 EUR/MWh and photovoltaic energy 136.4 EUR/MWh.

students and as teaching staff (MEST 2018). In the academic year 2017–2018, women comprised only 23 percent of all students graduating with bachelor's degrees in mechanical engineering from the University of Pristina. Other factors that are not exclusive to Kosovo but that contribute to this underrepresentation include occupational gender stereotypes and traditional perceptions about the suitability of energy-related occupations for women (IMPAQ 2018).

2.3.2 Evidence about Energy Efficiency

Even today, globally, nearly a billion people do not have access to electricity (International Energy Agency 2019).[14] Thus, access to energy has reemerged as a key priority for policymakers and donors in low-income countries. Extensive literature exists to advocate for programs focused on increasing energy access to generate meaningful long-term economic benefits (Kitchens and Fishback 2015). However, there is mounting concern that the poor will be worst hit due to climate change and thus there is a need to increase electrification in a *sustainable and inclusive* way. Governments around the world are pursuing a wide range of policies designed to narrow or close the energy efficiency gap (Loftus et al. 2015).[15]

Evidence about Impacts of Residential Energy Efficiency Investments

Energy efficiency investments are widely believed to offer a win-win opportunity, through several means: (a) the energy saved; (b) the reduction in energy consumption necessary to achieve a given level of services (e.g., indoor heating); and (c) the decrease in the greenhouse gas emissions that cause climate change and other pollutants that compromise human health (Granade et al. 2009). In a recent study, Fowlie, Greenstone, and Wolfram (2018) examine the impact on the energy efficiency gap of the United States' largest residential energy efficiency program, the federal Weatherization Assistance Program (WAP).[16] During the course of this 2011–2012 study, they failed to find significant benefits (economic or societal) for the program, as the upfront investment costs were about twice the actual energy savings. While participation in the Weatherization Assistance Program reduced energy consumption by 10 percent to 20 percent, efforts to persuade low-income households to take up the program were disappointing.

In another study, Allcott and Greenstone (2017) evaluate the returns on residential efficiency investments to measure energy savings in two large energy efficiency programs in Wisconsin, in the United States. The study finds that the returns on investments are negative socially.[17] Furthermore, improvements in energy efficiency do not regularly lead to one-to-one reductions in energy consumption (Galvin 2014), as energy efficiency gains alter the actual cost of comfort and may thereby generate shifts in consumption patterns—a "rebound effect" (Aydin et al. 2017; Sorell 2015). In contrast, studies from the United Kingdom and Germany (Adan and Fuerst 2016; Dowson et al. 2012) have found that energy efficient investments lead to reduced energy consumption.

[14] Furthermore, the World Energy Outlook estimates that in 2009, 25 percent of the population of the developing world lacked electricity, and the proportion was as high as to 37 percent in rural areas. In South Asia, the corresponding proportions are 32 percent and 40 percent, while in Sub-Saharan Africa, they were 69 percent and 86 percent.

[15] The efficiency gap is a large and persistent difference between the levels of investment in energy efficiency that are projected to save consumers money and the investments that individuals actually pursue.

[16] The Weatherization Assistance Program is the United States' largest residential energy efficiency program and has provided more than 7 million low-income households with weatherization assistance since its inception in 1976.

[17] In the two programs studied, the social welfare decreased by USD$0.18 per subsidy dollar spent, because the subsidies were not well-calibrated to estimated externality damages, and because of self-selection that attracted households whose participation generated low value (Allcott and Greenstone 2017).

Thus, the empirical evidence supporting the claims of energy efficiency investments being a win-win opportunity has been mixed, mostly due to varying contexts, initial conditions, and existing government regulations, which lead to vastly differing impacts of energy efficiency investments.

Behavioral Nudges

The mixed results from these studies highlight that there is still much to be learned about how well energy efficiency investments work in different contexts. While subsidies are a prevalent nudge tool used by governments, an increasingly popular public policy tool is choice defaults. This is based on consumer behavior theory and consists of a range of pricing options presented to consumers that states that all other things being equal, consumers are more likely to stick with their regular spending patterns, the default (Smith et al. 2013). While many studies point toward the positive effects of choice defaults, on the overall outcome (greenhouse emissions), there is little knowledge of the distributional consequences of such nudges for different groups in society (Dhami 2016).

One study in the residential electricity market in Switzerland analyzed the differences in impacts of these nudges on poor and rich. The study contrasted consumers' actual contract choices under an existing default option with the same consumer' active choices in a survey presenting a choice-set without any default (Ghesla et al. 2019). Poorer households were more prone to stick to the default option. Most importantly, these findings show that using defaults to trigger more environmentally friendly choices can effectively act as a hidden tax on the poor, while leaving untapped a significant willingness by richer households to pay for green electricity.

Evidence from Energy Efficiency Investments from Kosovo and Neighboring Countries

In light of the above evidence and in the context of RELP, it becomes important to examine the potential for energy efficiency in Kosovo. Nationwide, applying readily available insulation materials and double-glazed windows can amount to potential energy savings of 500-600 gigawatt hours/year, about 15 percent of total present demand. An additional 30-40 gigawatt hours/year could be saved in water heating applications through use of insulating jackets, low-flow heads, and timers (World Bank 2007). There is some preliminary evidence from neighboring countries about the effectiveness of energy efficiency retrofits in reducing the energy consumption in low-income households. The USAID-funded Residential Energy Efficiency for Low-income Households program in Armenia, Macedonia, and Bosnia and Herzegovina provided financial and behavioral incentives for energy retrofits and also delivered thermo-insulation for façades of buildings and repair or replacement of flooring, roofing, and windows.[18] It is estimated that retrofitting cut energy bills for low-income homeowners by as much as 50 percent, and reduced electricity consumption by 40 percent to 50 percent (Kakalejčíková 2017).

2.3.3 Gaps in Literature and Contribution of RELP Evaluation

There are several noticeable gaps in the literature about the impact of energy efficient investments. First, very few studies employ a counterfactual to assess the effect of household-level energy efficiency retrofits on measures of electricity consumption. The extensive review of literature in Bayer et al. (2020) notes that a total of 31 studies conducted statistical hypothesis tests to assess impacts of electricity access, among which only seven draw on a randomized experiment designed for causal inference. Second, evidence is particularly lacking in literature from low- and middle-income countries. One quasi-

[18] The Residential Energy Efficiency for Low-income Households project is one of the many assistance projects supported by USAID.

experimental study in Mexico analyzed a sample of new homes who were provided with insulation and other energy efficient upgrades (Davis et al. 2019). The upgrades had no detectable impact on electricity use or thermal comfort. Finally, many of the engineering models in literature about energy efficiency investments overestimate effects (Alcott and Greenstone 2017; Fowlie, Greenstone, Wolfram 2018). These results call for a reassessment of what we know about the impacts of energy efficiency investments.

The proposed study will provide a combination of rigorous and descriptive evidence about each of these areas of research. Specifically, we aim to estimate causal impacts of energy efficiency retrofits and metering on energy consumption and expenditure, using interrupted time series (ITS) approaches. Furthermore, since most evidence originates from programs in high-income countries, analysis of RELP activities will provide evidence for researchers, policymakers and governments that is more applicable for low and middle-income countries. Finally, there is a lack of literature and analysis of gender disparities in the Kosovo energy sector, and a lack of gender-disaggregated data. In particular, there is little research about factors that influence women's interest (or lack of interest) in energy sector opportunities, or detailed information about challenges faced by women in the sector. Our study will also add to the literature about women in the energy sector.

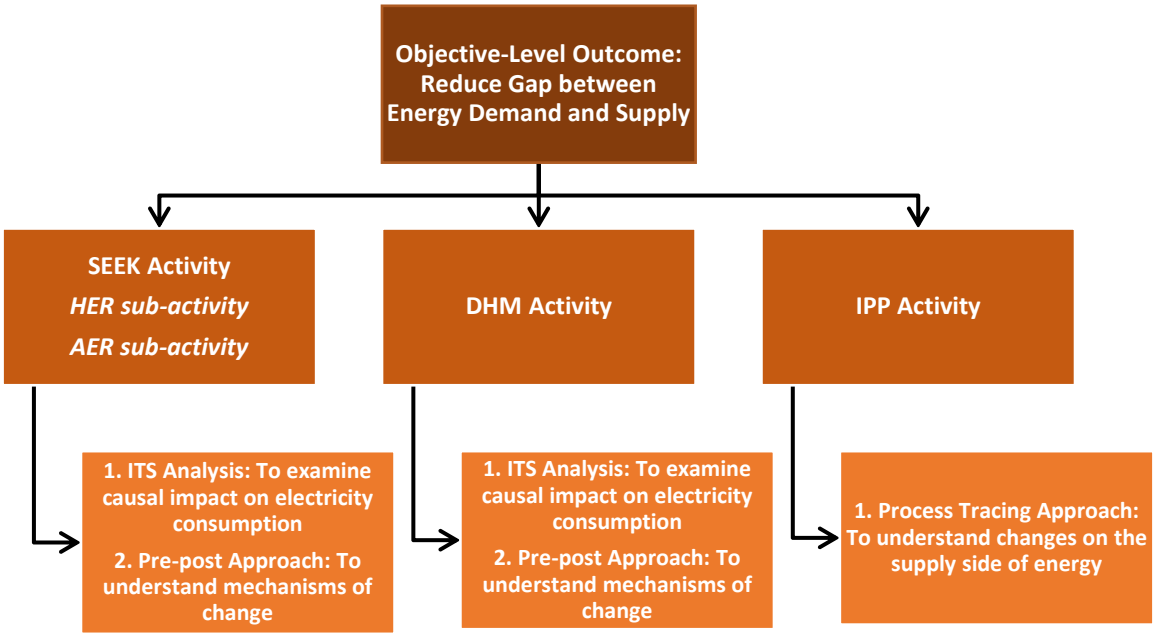
3. OVERALL EVALUATION DESIGN

In this section, we provide an overview of the RELP evaluation design. IMPAQ has proposed a comprehensive mixed-methods approach that balances methodological rigor with contextual and implementation realities. Section 3.1 provides an overview of the main evaluation questions, the proposed evaluation design, and the main data sources. Section 3.2 presents the overall evaluation timeline.

3.1 Overview of Evaluation

This evaluation is organized around the MCC RELP theory of change presented in Section 2.2 whose overarching objective is to reduce the gap between energy demand and supply. We will test the causal links in the theory of change using qualitative and quantitative data to determine whether the activities and inputs outlined in the RELP theory of change impact short- and medium-term beneficiary-level outcomes. Exhibit 2 depicts how the evaluation of different activities and sub-activities under RELP will help provide an understanding of the impacts on the main objective-level outcome of reducing the gap between energy demand and supply.

Exhibit 2. Connecting the RELP Evaluation to RELP's Objective-Level Outcome



In order to understand if the main objective-level outcome was achieved, the RELP evaluation will include descriptive analysis of broad trends and demographics of project beneficiaries and a combination of performance and impact evaluations. To understand if the objective-level outcome of reducing the gap between energy demand-supply was achieved, the evaluation of RELP will examine three activities: SEEK, DHM, and IPP. The evaluation of SEEK and DHM will focus on understanding the demand side of the energy demand-supply gap, while the evaluation of the IPP activity will focus on the supply side. For evaluating the extent to which SEEK reduced electricity demand (i.e. consumption) we will examine the impact of the household and apartment level retrofits using both impact and performance evaluations. The evaluation of DHM will follow a similar mixed-methods approach and examine the impact of household

level metering on district heating consumption in Pristina. On the supply side, the evaluation of the IPP activity will examine the changes in renewable energy production using a qualitative process tracing approach. Finally, our evaluation will also include the evaluation of the WEE activity, which is not linked to the main objective-level outcome of reducing electricity demand and supply, but instead focuses on increasing female employment in the energy sector.

Our proposed evaluation design incorporates findings from IMPAQ's scoping mission and subsequent conversations with all relevant stakeholders. Quantitative components of both performance and impact evaluations will use primary data, administrative data, and monitoring and evaluation data. Qualitative components will incorporate information from document reviews, key informant interviews (KIIs), and focus group discussions (FGDs) and will provide additional context to understand the quantitative impacts. Furthermore, as part of qualitative data collection, we will conduct women-only focus groups and ensure that other key marginalized groups are well represented to understand their experiences. Below, we outline in more detail our evaluation approach for each activity under RELP.

1. SEEK Activity

HER sub-activity

Quantitative Approach. We will estimate the causal impact of the HER sub-activity on electricity consumption using the quasi-experimental ITS method. This method will involve analysis of high-frequency (monthly) household-level administrative data on electricity consumption from the Kosovo Energy Distribution Services (KEDS). We will complement the ITS method with a pre-post approach using baseline and endline primary quantitative survey data to provide additional analyses on the consumption patterns related to other energy sources.

For the quantitative analysis, information on the main outcome variable of interest i.e. electricity consumption will be obtained from the KEDS administrative data. All other information will be obtained from primary quantitative data collection. Specifically, at baseline, the IC will be responsible, for procuring the data collection firm and IMPAQ will lead the enumerator training, survey instrument development, pre-testing, pilot testing, and quality checks. IMPAQ will design the baseline quantitative survey for HER and will include questions on demographic characteristics, physical structure of the house, energy consumption from different sources, and questions around behavioral change. The survey will also include a short module (designed by the IC) to capture the satisfaction levels of households at the end of the retrofit installation process. At endline, IMPAQ will be responsible for procuring the data collection firm and leading the data collection efforts and the survey will follow a similar format (excluding the satisfaction component) and will include additional questions to capture medium-term behavioral change around attitudes and practices around energy efficiency and questions related to the household's experiences with the SEEK activities.

Finally, the HER sub-activity will also include a cost-effectiveness analysis (CEA) for both iterations of the activity. The first iteration CEA will validate the findings of the IC on the cost-effectiveness of its model. The second iteration CEA will focus on the costs and benefits for beneficiaries under different models. This analysis will use data from both baseline and endline surveys.

Qualitative Approach. We will also utilize qualitative data collected through FGDs and KIIs to provide an in-depth understanding of the mechanisms behind changes in energy consumption. In addition to baseline

and endline, we will also collect qualitative data in the interim for continuous monitoring and to develop a deeper understanding of program effects.

AER sub-activity

Quantitative Approach. Similar to the HER sub-activity, we will use a combination of the ITS approach and a pre-post approach to understand the effect of this sub-activity on energy consumption in individual apartment units and the common areas of MABs. We will use an ITS approach to understand the impact on the main variable of interest: electricity consumption. The pre-post approach will complement the ITS analysis and provide evidence on the mechanisms of change.

Similar to the HER sub-activity, for quantitative analysis under the AER sub-activity, data will be collected through: (1) administrative data from KEDS; and (2) primary quantitative data collection. Data on one main variable i.e. electricity consumption will be obtained from the KEDS administrative database on a monthly basis throughout the project lifespan. All other variables of interest will be obtained from primary data collection at baseline and endline. Baseline and endline quantitative data will be collected at two levels for the AER sub-activity: (1) from HOA representatives, and (2) individual apartment units. At baseline, for data collection under the AER sub-activity, the IC will procure the data collection firm and IMPAQ will lead the enumerator training, survey instrument development, pilot testing, and quality checks. IMPAQ will design the baseline survey instrument which will include two separate modules: (1) the first module will include questions similar to the HER sub-activity, which will try to understand individual decision making for each apartment unit and will include questions around demographics, decision to contribute money to the MAB level retrofits, energy consumption from different sources, and behavior change questions among others; and (2) the second module will include questions at the HOA representative level including the physical structure of the MAB, buy-in of individual apartments, and behavior change questions, among others. Like the HER quantitative baseline survey, the AER survey will also include questions by the IC to capture the satisfaction of HOAs regarding the retrofit installation. At endline, IMPAQ will be responsible for procuring the data collection firm and leading the data collection efforts and the survey will follow a similar format (excluding the satisfaction component) to the baseline AER survey. Finally, we will also conduct a CEA for AER, measuring cost-efficiency at the MAB/apartment level.

Qualitative Approach. All quantitative findings in the AER sub-activity will be complemented with a qualitative case study approach, to provide more nuanced understanding of changes in energy consumption at the MAB level. Qualitative data will be collected at baseline and endline, in addition to three rounds of interim data for continuous monitoring.

2. DHM Activity

Quantitative Approach. We propose to estimate the causal impact of the DHM activity on the key outcome of interest—district heating consumption—through the ITS method, using high-frequency (monthly) administrative data for each beneficiary household from Termokos. Additionally, we will use a pre-post approach that will utilize primary survey data collected at baseline and endline to understand the impact of DHM on consumption of energy from other sources, including electricity. For the DHM activity, at baseline, the quantitative data collection will be undertaken by the IC. The IC will design the survey instrument and include question around physical structure of the apartment, demographics, consumption of energy from different sources, and behavior change questions. IMPAQ will add additional questions as

needed to answer the evaluation questions comprehensively. At endline, IMPAQ will be responsible for hiring the data collection firm and collecting endline quantitative data.

Qualitative Approach. Quantitative data collection will be accompanied with qualitative data collection at baseline, endline, and interim periods to understand how different program components affect district heating consumption after the installation of meters and thermal valves. Qualitative data will also help contextualize the quantitative findings.

3. IPP Activity

We will implement a process-tracing approach to assess the impact of the IPP activity. Our approach will be primarily qualitative, analyzing data collected through (a) document review and monitoring and evaluation data; and (b) semi-structured KIIs and small group interviews. The semi-structured KIIs will include quantitative questions, as appropriate to capture measurable outcomes like electricity production. However, he timeframe of our evaluation is too short to examine whether the IPP activity ultimately contributes to increasing the electricity supply in Kosovo. We will track the progress, successes, and challenges of the major implementation steps. We will assess whether the IPP activity was successful in reducing the perception of financial risk for investors and commercial lenders in the energy sector. We will also qualitatively assess whether investors and lenders are more willing to support IPPs.

4. WEE Activity

The implementation and effectiveness of this sub-activity will be analyzed primarily through qualitative collected through semi-structured KIIs (i.e. KIIs which include both qualitative and quantitative questions) and FGDs. We will also use document review and monitoring and evaluation data to complement the qualitative data collection. We will assess fidelity to program design by analyzing whether program targets and outputs have been achieved, and how identified challenges were overcome. We will also assess the effectiveness of program activities, through semi-structured KIIs and FGDs with beneficiaries. Additionally, we will use secondary data to assess the program's influence on women's academic and employment goals and access to job opportunities in the energy sector. However, trends in the secondary data will be largely exploratory and will not enable us assign attribution to the WEE activity. In the case of entrepreneur grants, we will determine whether grants led to more efficient use of energy, lowered energy costs, and improved profit margins through qualitative data as well as program monitoring data.

Each activity described above when feasible, will include analysis of key social groups, such as women-headed households, older and younger households, households of different ethnicities, and poor and vulnerable households, using both quantitative and qualitative data. We will integrate findings from the evaluation about the separate activities to provide a comprehensive assessment of RELP. We present the updated RELP evaluation questions in **Exhibit 3**, as discussed during the MCC EMC meeting in July 2020. For each evaluation question, we outline key themes explored, evaluation methods, and data sources.[19]

[19] As discussed with the MCC Evaluation Management Committee (EMC) in March 2020, one evaluation question cannot be answered through this evaluation. We cannot quantify and disentangle the impacts of different components of the SEEK activity as laid out in the original **EQ2.2** ***(i.e. what components of the program resulted in the largest change in energy consumption and why?)*** due to the change in the implementation design. Based on the information from the IC, during the implementation of the HER sub-activity, beneficiaries will not be randomly assigned to *certain* program components. All beneficiaries will receive *all* components of the program, including technical assistance and BC&O, which will preclude us from disentangling and quantifying impacts of each component. In July 2020, the wording of the evaluation questions was changed (as detailed in **Exhibit 3**) and this is no longer a concern.

Exhibit 3. Overview of RELP Evaluation Design

Evaluation Question	Key Outcomes	Evaluation Method	Data Source
▪ **EQ1** Were the activities implemented as designed and were the outputs outlined in the project logic achieved?	**Qualitative themes** ▪ Program design and changes over time ▪ Implementation progress, successes, and challenges **Quantitative themes** ▪ Trends in RELP outputs over time	▪ Comparison of planned vs. actual implementation design timeline ▪ Qualitative analysis ▪ Quantitative analysis tracking RELP outputs	▪ RELP monitoring and evaluation data (by MFK) ▪ FGDs and KIIs ▪ Site visits ▪ Document review (including the IC's Operations Manuals for each activity)
▪ **EQ2.1** Did the SEEK activity result in reduced electricity consumption in the intervention areas? ▪ **EQ2.2** What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and installation of energy efficient equipment)? ▪ **EQ2.3** Were energy savings previously envisioned during preliminary baseline studies and energy audits achieved, and why? ▪ **EQ3.1** Did the SEEK activity identify a cost-effective model for reducing household energy use that could be scaled up nationwide and what is the overall expected energy reduction?	**Qualitative themes** ▪ Quality and appropriateness of design ▪ Operational processes and systems; use of data management systems and communication procedures ▪ Progress, achievements, successes, and challenges ▪ Experiences of different groups, including women-headed households and different age groups, income levels, and ethnicities **Quantitative themes** ▪ Electricity consumption and expenditure *(ITS design)* ▪ Other energy sources consumption and expenditure *(Pre-post design)* ▪ Substitution between energy sources *(Pre-post design)* ▪ Willingness to pay *(Pre-post design)* ▪ Outages *(Pre-post design)* ▪ Types of energy efficiency retrofits installed and costs *(Pre-post design)* ▪ Disaggregating impacts by women-headed households, age groups, income levels, and ethnicities *(ITS and pre-post design)*	▪ Pre-post analysis ▪ ITS approach ▪ CEA ▪ Qualitative analysis ▪ Simulations of different policy counterfactual scenarios to understand impacts of scaling up the policy model	▪ *HER sub-activity*: o Household survey with direct beneficiaries *(for pre-post design)* o KEDS administrative data *(for ITS design)* o Monitoring and evaluation data (by MFK and IC) o FGDs and KIIs o Document review ▪ *AER sub-activity*: o HOA representative survey *(for pre-post design)* o Apartment owner survey *(ITS and pre-post design)* o Monitoring and evaluation data (by MFK and IC) o FGDs and KIIs o Site visits o Document review

Evaluation Question	Key Outcomes	Evaluation Method	Data Source
▪ **EQ4.1** Did the DHM activity result in a change in energy consumption in the intervention areas? ▪ **EQ4.2** What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and implementation of consumption-based billing)? ▪ **EQ4.3** How many new households shifted to district heating as a result of the DHM activity?	**Qualitative themes** ▪ Awareness of meters and their benefits ▪ Perceptions of affordability of district heating ▪ Experiences of different groups, including women-headed households, households of different age groups, income levels, and ethnicities **Quantitative themes** ▪ District heating consumption *(ITS design)* ▪ Other sources of energy used *(Pre-post design)* ▪ Monthly energy consumption and costs from different energy sources *(Pre-post design)* ▪ Decision to install meters and constraints to their installation *(Pre-post design)* ▪ New households shifting to district heating *(Pre-post design)*	▪ Pre-post approach ▪ ITS approach ▪ Qualitative analysis	▪ Household survey *(for pre-post design)* ▪ Termokos and KEDS administrative data *(for ITS and design)* ▪ Document review ▪ FGDs and KIIs
▪ **EQ5.1** Was the IPP activity successful in reducing the perception of financial risk for investors and commercial lenders in the energy sector, and why? ▪ **EQ5.2** Did the IPP activity contribute to increasing the electricity supply in Kosovo? If so, by how much and why?	**Qualitative themes** ▪ Engagement with technical assistance providers ▪ Readiness to apply for loans ▪ Experience with financing options ▪ Program milestones and benchmarks **Quantitative themes** ▪ Licensing and construction status ▪ Loans provided and capital raised ▪ Number and amount of guarantees offered	▪ Process-tracing approach ○ Qualitative analysis ○ Pre-post quantitative analysis	▪ Monitoring and evaluation data (by MFK and IC) ▪ KIIs and small group interviews ▪ Document review ▪ Site visits
▪ **EQ6.1** How effective were program activities at increasing women's employment in the energy	**Qualitative themes** ▪ Perception shift about women	▪ Pre-post quantitative analysis ▪ Qualitative analysis	▪ Semi-structured interviews (with quantitative and qualitative components)

Evaluation Question	Key Outcomes	Evaluation Method	Data Source
sector and at increasing investments in energy efficiency among women entrepreneurs?	▪ Satisfaction with internships, scholarships, summer camp, and grants ▪ Knowledge acquired ▪ Employer perceptions ▪ Expressed interest in or intention to shift career/academic pursuits to energy sector **Quantitative themes** ▪ Sources and consumption of energy ▪ Energy efficient measures invested ▪ Proportion of girls signing up for camps; proportion of women with internships; proportion of women who receive scholarships; proportion of entrepreneurs investing in energy efficiency ▪ Percentage of women finding full-time employment in the STEM field sector within two years after completing WEE internships/ apprenticeships ▪ Percentage of women finding employment with host company upon completion of WEE internship/ apprenticeship ▪ Percentage/number of participating companies that have added technical specifications in hiring guidelines to hire women ▪ Percentage of participating companies that employ intern as staff upon completion of WEE internship/apprenticeship ▪ Proportion of women/girls who enroll in energy-related fields of		▪ Monitoring and evaluation data (by MFK and IC) ▪ Secondary data about women's employment ▪ Document review ▪ Site visits

Evaluation Question	Key Outcomes	Evaluation Method	Data Source
	study/shift from non-energy academic tracks to energy-related academic tracks.		
▪ **EQ7.1** Did the program meet the stated objective of reducing the gap between supply and demand of electricity? ▪ **EQ7.2** How sustainable are critical outcomes of the program, and why?	**Qualitative themes** ▪ Likelihood of sustaining investments ▪ Sustainability of beneficiary outcomes and threshold outputs **Quantitative themes** ▪ Outage frequency and duration ▪ Electricity supply and demand ▪ Maintenance of retrofits and meters	▪ Pre-post analysis ▪ Qualitative analysis ▪ Simulations	▪ Household survey ▪ HOA survey ▪ Administrative data ▪ Monitoring and evaluation data (by MFK and IC) ▪ FGDs and KIIs ▪ Site visits

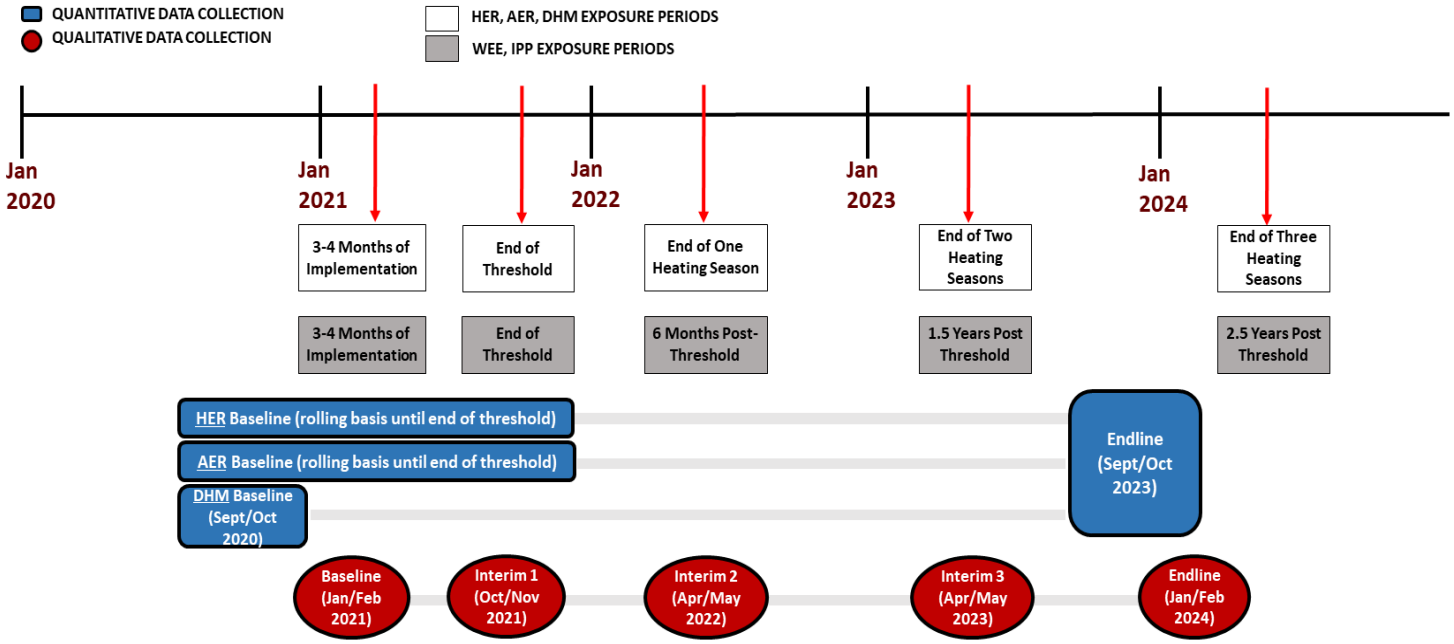
3.2 Evaluation Timeline

In this section, we present the detailed evaluation timeline. While administrative data and documentation will be collected on a regular basis, all primary data collection under RELP is intricately linked to implementation milestones and key exposure periods for different activities. In **Exhibit 4**, we describe the exposure periods for each activity. Primary quantitative data will only be collected for the HER, AER, and DHM activities. Primary qualitative data will be collected for all activities, including IPP and WEE activities. All dates presented are notional and will depend entirely on the implementation timeline. For instance, baseline qualitative data collection assumes at least 3-4 months of implementation activities. If that timeline shifts, baseline qualitative data collection will also shift.

Primary Quantitative Data (2 rounds). Primary quantitative survey data will be collected at two points in time: baseline and endline. For SEEK, the IC will collect the baseline quantitative data, with IMPAQ's support. That is, for SEEK, at baseline, the IC will procure the data collection firm and IMPAQ will lead the enumerator training, survey instrument development, pre-testing, pilot testing, and quality checks. For DHM, at baseline, the IC will be responsible for the data collection efforts including survey instrument development and IMPAQ will suggest additional questions, as needed. At endline for both SEEK and DHM, IMPAQ will be responsible for procuring the data collection firm and leading all aspects of the data collection efforts.

Primary Qualitative Data (5 rounds). With respect to qualitative data, primary data will also be collected at baseline and endline. Additionally, there will be three rounds of interim primary qualitative data collection for continuous monitoring and to examine changes that occur between baseline and endline. Thus, there will be two rounds of primary quantitative data collection and five rounds of primary qualitative data collection. IMPAQ will be responsible for all five rounds of primary qualitative data collection. For the WEE and IPP activities, primary qualitative data will include semi-structured KIIs which will include some quantitative questions, as needed.

Exhibit 4. RELP Exposure Periods



Given the rolling nature of baseline quantitative data collection for SEEK, the submission of the draft interim report is *tentatively* scheduled for December 2021, approximately 1-2 months after the baseline quantitative data collection is expected to finish. If implementation activities shift because of COVID-19 delays, then this will shift the submission of the draft interim report as well. The draft endline report is scheduled for May 2024. We summarize below the major data collection sources and timelines and provide a detailed evaluation work plan in **Appendix C**. In **Exhibit 5**, we describe all data collection activities over time, including RELP activity implementation milestones, exposure periods and the rationale for each data collection activity. As a note, "*Month X*" is a placeholder for the month in which implementation activities are expected to resume after the COVID-19 pandemic.

Exhibit 5. Linking RELP Exposure Periods to Data Collection Timing

Activity/Timeline	Ideal Exposure Period	Data Collection Source and Timing	Rationale
HER, AER: ***Retrofits installation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)*** ***(Sample Sizes:*** *Quantitative: 2,100 households (HER); 690 households and 25 MABs (AER)* *Qualitative: 4-6 FGDs and 4-10 KIIs depending on beneficiary type)* **DHM:** ***Meter and thermal valve installation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)*** ***(Sample Sizes:*** *Quantitative: 3,600 households* *Qualitative: 1-3 FGDs and 1-6 KIIs depending on beneficiary type)*	▪ **Energy Consumption Outcomes** ○ Short-run impacts can be observed in heating season right after retrofit/meter installation ○ Medium-run impacts can be observed after at least 1 full heating season ○ Long-run impacts can be observed after at least 2 full heating seasons ▪ **Behavior Change Outcomes** ○ Short-run impacts can be observed after at least 1 full heating season ○ Medium-run impacts can be observed after at least 2 full heating seasons	▪ **Baseline** ○ Quantitative data: Month X – September 2021 ○ Qualitative data: After the first iteration of HER and before the beginning of the second iteration. Qualitative for DHM and AER will happen at the same time ▪ **Interim** ○ Qualitative data: 1. After the end of the threshold period 2. At the end of 1 heating season 3. At the end of 2 heating seasons ▪ **Endline** ○ Quantitative data: After 2 heating seasons and before the start of the third heating season ○ Qualitative data: After the endline quantitative data collection	▪ **Baseline** ○ Quantitative data: 5 days after retrofit installation ○ Qualitative data: First iteration of HER households, some AER MABs, and some DHM apartments that have already completed the retrofits/meter installation and have experienced at least a few months of cold weather to provide formative feedback to MFK and the IC ▪ **Interim** ○ Qualitative data: Provide more information related to successes, challenges; explore effectiveness of BC&O; provide context to ITS analysis results; and track main outcomes over time since quantitative data collection will happen at only 2 points in time ▪ **Endline** ○ Quantitative data: Measure medium to long-run changes in energy consumption and behavior change outcomes ○ Qualitative data: After the endline quantitative data collection to put results in perspective

Activity/Timeline	Ideal Exposure Period	Data Collection Source and Timing	Rationale
IPP ***Implementation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)*** ***(Sample Sizes:*** *Qualitative:* *1-6 KIIs and/or small group interviews depending on beneficiary type)*	▪ **Readiness of IPPs** ○ Observed at 6 month intervals to ascertain progress ▪ **Progress of Renewable Energy Projects** ○ Observed at 6 month intervals to ascertain progress ▪ **Longer-term Outcomes** ○ Observed approximately 2-3 years after threshold implementation	▪ **Baseline** 3-4 months after start of implementation ▪ **Interim** 1. After the end of threshold period 2. 6 months after the end of threshold 3. 1.5 years after the end of threshold ▪ **Endline** ○ Approximately 2 years after end of threshold	▪ **Baseline** ○ *Qualitative data:* During start up activities to gain preliminary insights about whether the activity is on track to build a bankable renewable energy program pipeline ▪ **Interim** ○ *Qualitative data:* Crucial to assess the ongoing progress of each IPP against the theoretical assumptions of the program ▪ **Endline** ○ *Qualitative data:* To assess whether the program successfully completed the activities (established the Kosovo Project Acceleration committee, provided training, opened the IPP credit guarantee window); and whether these activities led to intermediate outcomes for the IPPs (licensing, reaching financial close, construction, connection to the grid)
WEE 1. **Internships** 2. **Scholarships** 3. **WiSci Camp** 4. **Grants** ***Implementation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)*** ***(Sample Sizes:*** *Qualitative:* *2-4 FGDs and 2-4 KIIs and/or semi-structured interviews depending on beneficiary type)*	**Internships, Scholarships, WiSci Camp, Grants** ▪ Perception shift about women, satisfaction with internships/scholarship/camp/grants, knowledge acquired, employer perceptions, sources and consumption of energy, energy efficient measures invested etc. will be observed every 6 months to determine gradual shift in behavior	▪ **Baseline** 3-4 months after start of implementation ▪ **Interim** 1. After the end of threshold period 2. 6 months after the end of threshold 3. 1.5 years after the end of threshold ▪ **Endline** Approximately 2 years after end of threshold	▪ **Baseline** ○ 3-4 months after start of implementation to gain preliminary insights about whether the activity is on track ▪ **Interim** ○ Crucial to assess the ongoing progress of each against the theoretical assumptions of the program ▪ **Endline** ○ To assess whether the program successfully completed the activities; and whether these activities led to intermediate outcomes for the beneficiaries including leading to long term behavior change and opportunities for women in the energy sector

In the sections that follow, we discuss the design for each activity of RELP separately. In **Section 4** and **Section 5**, we describe the quantitative and qualitative methods for the evaluation of the SEEK activity and the DHM activity, respectively. In **Section 6**, we describe the process-tracing approach of the IPP

activity. In **Section 7**, we detail the evaluation of the WEE activity. Finally, in **Section 8**, we detail the evaluation of broad sectoral-level changes. Each section describes the methodology, anticipated timing of outcomes, sampling, data collection and analysis plan.

4. SEEK ACTIVITY: EVALUATION DESIGN

In this section, we provide a detailed overview of the mixed-methods design for the evaluation of the SEEK activity. In Section 4.1, we outline the implementation status of the SEEK activity. In Sections 4.2 and 4.3, we outline the quantitative and qualitative approaches to the evaluation of HER and AER sub-activities. Finally, in Section 4.4, we detail the CEA for the SEEK activity.

4.1 Implementation Status of SEEK Activity

Exhibit 6 presents the implementation status for the HER and AER sub-activities, as of August 2020. The current implementation status has guided the proposed evaluation design.

Exhibit 6. Implementation Status of SEEK Activity

Activity	Implementation Status
HER sub-activity	▪ In the pre-implementation phase. ▪ 500 households (first iteration) were planned to be enrolled on a rolling basis from August 2020–December 2020. Only one model to be tested. However, due to COVID-19 implementation has been delayed and it is being determined when the implementation would start. ▪ 2,100 households (second iteration) expected to be enrolled on a rolling basis from March 2021–September 2021. One to three models to be tested. Implementation delays expected because of COVID-19. ▪ Both iterations will enroll nationwide in all 38 municipalities.
AER sub-activity	▪ In the pre-implementation phase. ▪ IC is working with MFK to identify a list of 25 MABs to be included in the sub-activity. This is proving to be challenging, as many MABs do not have a functional HOA. As of mid-April 2020, the IC together with the responsible municipality experts had prepared a long list of 40 MABs for the seven municipalities from which the final 25 MABs will be selected. Further, implementation delays are expected because of COVID-19.

4.2 Quantitative Approach: HER and AER Sub-activities

We propose to implement a quantitative performance evaluation and an impact evaluation to assess the impacts of the HER sub-activity and to use a quantitative performance evaluation to understand the impacts of the AER sub-activity. In Section 4.2.1, we describe the methods and in Sections 4.2.2, 4.2.3, and 4.2.4, we provide details about the statistical power, sampling strategy, and data sources. Section 4.2.5, outlines the analysis plan. Finally, Section 4.2.6 provides details about data collection.

4.2.1 Methods

The SEEK quantitative approach will estimate impacts of the installation of household and MAB-level energy efficiency retrofits by surveying/obtaining information from three different respondent groups:

1. Households enrolled under the second iteration of the HER sub-activity
2. Individual apartment owners under the AER sub-activity
3. Homeowner association (HOA) representative under the AER sub-activity

The main quantitative outcomes of interest for the SEEK activity are electricity consumption and expenditure. We will also examine additional outcomes, including: electricity reliability (measured through outages), energy consumption and expenditure from all other sources, awareness of and willingness-to-pay for energy efficiency, investment in electrical appliances, level of satisfaction with

retrofits, quality and maintenance of energy efficiency retrofits, shifts in energy efficiency usage in terms of additional investments, likelihood of sustaining investments, and impacts on peer networks. All outcomes will be disaggregated by key socioeconomic characteristics, including gender, age of head of household, ethnicity, and household's economic well-being, among others.

HER sub-activity. For the HER sub-activity, to ensure adequate gender representation, we will attempt to interview multiple people in the household with the most knowledgeable person answering the questions, as needed. This is because, since the major proportion of household heads in Kosovo are male, interviewing only household heads will exclude women's perspectives. The sample will include only the 2,100 households scheduled to be enrolled in the second iteration, and these will constitute our "main sample". Since at this stage it is unclear how many intervention models are proposed for the second iteration, the main goal will be to examine the combined effects of all the intervention models under the HER sub-activity, pooling households from all different models that might eventually be implemented. Data collected from the 500 households enrolled in the first iteration will be used primarily for the purposes of developing an understanding of the demographics, physical characteristics of the house, energy usage behavior; for the CEA described in **Section 4.4.1**.[20]

AER sub-activity. For the AER sub-activity, there will be two main respondent groups: (1) each individual apartment owner in a MAB (to study changes in apartment level energy consumption and behavior change); and (2) the HOA representative of an MAB (to study MAB-level decision making processes and to examine changes in electricity consumption of common areas). The unit of analysis in each of these groups will be different, with *individual apartment owners* being the main unit of analysis in the first group and the *MAB-level HOA representative* being the main unit of analysis in the second group. Similar to the HER sub-activity, while interviewing individual apartment units, we will interview multiple people in the household to ensure adequate gender representation.

1. ITS Approach

HER sub-activity. We will use the ITS approach to estimate the impacts of installing household level retrofits on household electricity consumption and expenditure for the HER sub-activity **(EQ2.1**[21]**)**. The ITS approach estimates the causal impact of a specific program activity—such as installing an energy efficiency retrofit—by analyzing administrative time-series data from KEDS before and after the program activity is completed, and by assessing to what extent household-level outcomes change immediately after the completion of the installation of energy efficiency retrofits, relative to a possible preexisting trend.[22] To predict the electricity consumption in the absence of the retrofit (the counterfactual), the ITS analysis relies on frequent observations before the retrofit installation is completed, to estimate a

[20] As per the implementation timeline proposed by the IC, households in the first iteration of the HER sub-activity will be studied over the period of 4-6 months only, before the second iteration is started. As such these households will only be administered a baseline survey and no endline survey and thus cannot form part of our "main sample". The primary aim of studying the behavior of the first 500 households will be to understand their electricity consumption over the 4-6 months and correlate with demographic characteristics, to draw lessons for program improvements before the second iteration.

[21] *EQ2.1: Did the SEEK activity result in reduced electricity consumption in the intervention areas?*

[22] An ITS approach is particularly useful when a randomized trial is infeasible or unethical. The approach usually involves constructing a time series of the outcome variable and testing statistically for a change in the outcome rate in the periods before and periods after implementation of a policy/program designed to change the outcome (Penfold and Zhang 2013).

possible preexisting trend. The ITS approach relies critically on the assumption that this preexisting trend would have continued over time in the absence of the retrofit.[23]

The ITS approach will estimate impacts of the HER sub-activity, by creating a longitudinal panel of households, and will aim to use monthly administrative data from January 2018 to May 2024.[24] In estimating the impacts using the ITS approach, we will account for other factors that could confound the impacts (including any seasonal sensitivities), including month-of-year effects and time-of-day effects (if available from the KEDS data). Finally, we will also control for demographic and socio-economic characteristics in the ITS analysis using the information collected on these variables from the primary quantitative baseline and endline surveys described in **Section 4.2.4**.

AER sub-activity. For the AER sub-activity, we will also use the ITS approach to estimate changes in the key outcome variables of interest i.e. electricity consumption and expenditure. We will measure changes in electricity consumption and expenditure for each individual apartment. Based on our discussion with KEDS in July 2020, each individual apartment unit in a MAB has an electricity meter. Thus, the ITS approach will measure the impact of MAB-level retrofits (as opposed to the household level retrofits that the HER sub-activity will measure) on electricity consumption of individual apartments. Similar to the HER sub-activity, we will use administrative data from the electricity meter in each individual apartment on a monthly basis from January 2018 to May 2024.

2. Pre-Post Approach

HER sub-activity. The ITS approach proposed above will only allow us to study the impact of the retrofits on electricity consumption and expenditure. In order to understand why electricity consumption would be impacted, we will need to explore additional outcomes and conduct further analyses. Hence, we will complement the ITS analysis with a pre-post longitudinal analysis of the same household and explore additional outcomes, such as substitution between energy sources, electricity reliability, and willingness to pay for energy efficiency retrofits, among other outcomes detailed in **Exhibit 6**, which will not be possible to examine through administrative data. We will compare outcomes during endline primary data collection in September–October 2023, with those measured during baseline primary data collection, to provide information about the evolution of household-level outcomes. To assess behavioral changes, both absolute and disaggregated by gender and socioeconomic status, we also will rely on the qualitative analysis described in **Section 4.3**.

AER sub-activity. The AER sub-activity includes 25 MABs and according to the IC's latest implementation schedule, the following seven municipalities will be selected for the AER implementation: Pristina, Lipjan, Novobrdo, Gracanica, Gjakova, Mitrovica, and Viti. While the ITS approach will only focus on individual

[23] In its simplest form, the ITS approach is modeled using a regression model (such as linear, logistic, or Poisson) that includes only three time-based covariates, whose regression coefficients estimate the pre-intervention slope, the change in level at the intervention point, and the change in slope from pre-intervention to post-intervention. The pre-intervention slope quantifies the trend for the outcome before the intervention. The level change is an estimate of the change in level that can be attributed to the intervention, between the time points immediately before and immediately after the intervention, and accounting for the pre-intervention trend. The change in slope quantifies the difference between the pre-intervention and post-intervention slopes (Kontopentalis et al. 2015).

[24] A sufficient number of time points before and after the intervention is needed to conduct segmented regression analysis. A general recommendation is for 12 data points before and 12 data points after the intervention (Wagner et al. 2002). An ITS approach can include observations at daily, weekly, monthly, or yearly levels. Similar to our study, which will use monthly observations, Penfold and Zhang (2013), Fretheim et al. (2013), and Albu et al. (2018) employ monthly data.

apartment units, we will conduct a longitudinal pre-post design at two levels for the AER sub-activity: (1) individual apartment units; and (2) HOA representatives. For individual apartment units the pre-post approach will be similar to that of the HER sub-activity with the main aim being to understand the mechanisms of change and the substitution between different energy sources. Specifically, we will capture similar outcomes including behavior change, substitution between energy sources, electricity reliability, and willingness to pay for energy efficiency retrofits, among other outcomes. For HOA representatives, the pre-post approach will provide information on the evolution of MAB level outcomes including behavior change at the MAB level, retrofit decisions at the MAB level, and quality and maintenance of retrofits, among others. However, the sample size for HOA representatives will be small. As a note, only one survey will be administered as part of the AER sub-activity. This survey will have two modules, one module for all apartment units and the second module which will only be asked if the household head is also an HOA representative. We will complement this with a qualitative case study approach described in **Section 4.3**, to provide further context into the MABs.

3. Simulations

We will use simulations to estimate the expected energy reduction if SEEK were to be implemented all over Kosovo **(EQ7.1**[25]**).** That is, we will forecast the anticipated energy reduction from widespread adoption of energy efficiency retrofits in the Kosovo population (Reyna and Chester 2017; Fleiter et al. 2011). Our models will account for changes in electricity use in the SEEK municipalities on direct beneficiaries, as well as key stakeholder opinions on the degree to which these changes might be due to the program. We outline the detailed approach in **Section 8**.

4.2.2 Statistical Power

HER sub-activity. The available sample for the analysis for the HER sub-activity is the total number of beneficiaries proposed to be enrolled in the second iteration, 2,100 households. We assume that we will be able to survey and obtain KEDS bill numbers for approximately 70 percent of these households (due to nonresponses) or a total of 1,470 households. *Given the projected number of completed surveys* of 1,470 households, we estimate and present the range of minimum detectable effects in **Exhibit 7**.[26]

We estimate minimum detectable effects for the key outcome variable, monthly electricity consumption, assuming a simple shift in the mean outcome after the intervention is implemented. In this case, the statistical power for an ITS model depends on the number of time points of data, the auto-correlation in outcomes, and the coefficient of variation (ratio of the standard deviation of the outcome to its mean). We use estimates of electricity consumption from Eurostat 2019 as the baseline value. Based on this study, the average electricity consumption of a Kosovar family is 312.5 kWh per month. However, we do not have estimates for the standard deviation of monthly electricity consumption. We present results using four ratios of standard deviation to mean encompassing a broad range—0.5, 1, 2, and 3. We also assume that through administrative data from KEDS, we will obtain 24 months of data before the retrofit installation and 24 months of data after the retrofit installation, and that the autocorrelation of outcomes is 0.2. Using these parameters, we provide minimum detectable effect estimates for the direct beneficiaries of the HER sub-activity in the second iteration.

[25] *EQ7.1: Did the program meet the stated objective of reducing the gap between supply and demand of electricity*?

[26] We conservatively estimate a response rate of 70 percent, given the low response rates for the baseline survey conducted by the IC for the SEEK activity in December 2019. Survey fatigue was a common reason for the low response rates.

Exhibit 7. Minimum Detectable Effects for HER Sub-activity

Outcome	Baseline Mean	Coefficient of Variation	Minimum Detectable Effect	% Change from Baseline	Number of Attempted Surveys	Projected Number of Completed Surveys	Time Periods (in terms of months)
HER sub-activity: Second Iteration							
Monthly Electricity Consumption (kWh)	312.5 kWh *(household level)*	0.5	3 kWh	1%	2,100 households	1,470 households	48 (pre: 24; post: 24)
		1	7 kWh	2.25%			
		2	13 kWh	4.16%			
		3	19 kWh	6.08%			

Note: We assumed a confidence level of 95 percent, two-tailed tests, 80 percent power, 30 percent sample attrition for the direct beneficiaries, and an autocorrelation of 0.20. Baseline mean value at the household level has been assumed from Eurostat 2019. We assumed that for the ITS analysis, the period will be split in the following way: pre-period: minimum 24 months; and post-period: minimum 24 months.

Assuming that of the 2,100 surveys attempted, we are able to obtain 1,470 completed responses (and their associated KEDS bill numbers), we would be able to detect a decline in *monthly consumption of electricity* by about 1 percent to 6 percent, depending on the coefficient of variation.[27] As shown, the minimum detectable effects can be quite large if the coefficient of variation (ratio of standard deviation to mean) is high. However, it drops as the standard deviation goes down relative to the mean. Our calculations suggest that if the coefficient of variation is 0.5, the autocorrelation is 0.2, and we can get data for 48 time periods, we could estimate impacts as small as 1 percent decline in electricity consumption. If the number of households for whom we obtain administrative data is higher than 1,470, then our statistical precision could improve.

We also calculate the minimum threshold of beneficiaries that need to be reached for a viable design effect. That is, in the event that the IC is unable to reach 2,100 beneficiaries, what would be the minimum number of beneficiaries (and corresponding minimum number of completed surveys) that would need to be reached and complete the survey to determine program impacts? We present this in **Exhibit 8.** Assuming a coefficient of variation of 0.5 and with all the other assumptions remaining similar to before, if only 400 beneficiaries sign up for HER and the corresponding number of completed surveys is 280 (assuming 70 percent response rate), then we would be able to detect a 2.08 percent decline in electricity consumption.[28] On the other hand, if 1,200 beneficiaries sign up (corresponding number of completed surveys is 840) then we can detect a 1.28 percent decline in electricity consumption.

Exhibit 8. Minimum Threshold Number of Surveys for HER Sub-activity

[27] Expected effect sizes are similar in other literature estimating impacts of energy efficiency retrofits. Adan, H., & Fuerst, F. (2016) find that cavity wall insulation, loft insulation, and new boilers, reduced total HH energy consumption by 10.5 percent. Similarly, Fowlie, M., Greenstone, M., & Wolfram, C. (2018), find that the weatherization program in the US, reduced monthly household energy consumption (electricity and gas) between 8-10 percent.

[28] Based on similar studies, the coefficient of variation is expected to close to 0.5; however, if changes in the coefficient of variation will change the minimum threshold number of beneficiaries needing to be reached will change as well.

Outcome	Baseline Mean	Coefficient of Variation	Minimum Detectable Effect	% Change from Baseline	Number of Attempted Surveys	Projected Number of Completed Surveys	Time Periods (in terms of months)
HER sub-activity: Second Iteration							
Monthly Electricity Consumption (kWh)	312.5 kWh *(household level)*	0.5	6.5 kWh	2.08%	400 households	280 households	48 (pre: 24; post: 24)
			5 kWh	1.60%	800 households	560 households	
			4 kWh	1.28%	1,200 households	840 households	
			3.5 kWh	1.12%	1,600 households	1,120 households	

Note: We assumed a coefficient of variation of 0.5, confidence level of 95 percent, two-tailed tests, 80 percent power, 30 percent sample attrition for the direct beneficiaries, and an autocorrelation of 0.20. Baseline mean value at the household level has been assumed from Eurostat 2019. We assumed that for the ITS analysis, the period will be split in the following way: pre-period: minimum 24 months; and post-period: minimum 24 months.

AER sub-activity. Under the AER sub-activity, all 25 MABs will be included in the beneficiary sample. We will sample all apartment units within each of the 25 MABs. Based on discussions with the IC, each MAB has at most 20-30 individual apartment units. Thus, the available sample for the AER sub-activity is at most 690 individual apartment units. Of these, similar to the HER sub-activity, we assume that we will be able to survey and obtain KEDS bill numbers for approximately 70 percent of these households (due to nonresponses) or a total of approximately 483 households. *Given the projected number of completed surveys* of 483 individual apartment units, we estimate and present the range of minimum detectable effects in **Exhibit 9**.[29]

We estimate minimum detectable effects for the key outcome variable, monthly electricity consumption, similar the HER sub-activity assuming a 70 percent response rate. Similar to before, we present results using four ratios of standard deviation to mean encompassing a broad range—0.5, 1, 2, and 3. We also assume that through administrative data from KEDS, we will obtain 24 months of data before the retrofit installation at the MAB level and 24 months of data after the retrofit installation at the MAB level, and that the autocorrelation of outcomes is 0.2. Finally, we assume that the intra-class correlation (ICC) is 0.01 and the number of clusters is capped at 25 MABs.

Assuming that of the 690 surveys attempted, we are able to obtain 483 completed surveys (and their associated KEDS bill numbers), we would be able to detect a decline in *monthly consumption of electricity* by about 2 percent to 11 percent, depending on the coefficient of variation. As shown, the minimum detectable effects can be quite large if the coefficient of variation (ratio of standard deviation to mean) is high. However, it drops as the standard deviation goes down relative to the mean. Our calculations suggest that if the coefficient of variation is 0.5, the autocorrelation is 0.2, the ICC is 0.01, and we can get data for 48 time periods, we could estimate impacts as small as 2 percent decline in electricity consumption. If the number of households for whom we obtain administrative data is higher than 483, then our statistical precision could improve.

[29] We conservatively estimate a response rate of 70 percent, given the low response rates for the baseline survey conducted by the IC for the SEEK activity in December 2019. Survey fatigue was a common reason for the low response rates.

Exhibit 9. Minimum Detectable Effects for AER Sub-activity

Outcome	Baseline Mean	Coefficient of Variation	Minimum Detectable Effect	% Change from Baseline	Number of Attempted Surveys	Projected Number of Completed Surveys	Time Periods (in terms of months)
AER sub-activity: Individual Apartment Owners							
Monthly Electricity Consumption (kWh)	312.5 kWh *(household level)*	0.5	6 kWh	1.92%	690 households	483 households	48 (pre: 24; post: 24)
		1	12 kWh	3.84%			
		2	24 kWh	7.68%			
		3	36 kWh	11.52%			

Note: We assumed a confidence level of 95 percent, two-tailed tests, 80 percent power, 30 percent sample attrition for the direct beneficiaries, intra-class correlation of 0.01, and an autocorrelation of 0.20. Baseline mean value at the household level has been assumed from Eurostat 2019. We assumed that for the ITS analysis, the period will be split in the following way: pre-period: minimum 24 months; and post-period: minimum 24 months.

4.2.3 Sampling Strategy

HER sub-activity. The entire beneficiary sample, 2,100 households, will be included in the study sample. If any of the households selected for the household survey in the HER sub-activity are also receiving benefits under the AER sub-activity, we will drop these households from the final sample.

AER sub-activity. For the individual apartment unit survey, we will survey all apartment units within each of the 25 MABs that opt-into the AER sub-activity.

4.2.4 Data Sources, Outcomes, and Anticipated Timing of Data Collection

The main quantitative data sources for the HER and AER sub-activities will include the components listed below.

HER survey (joint responsibility of IC and evaluator). Data will be collected using a household survey for households under the HER sub-activity. For the first iteration of the HER sub-activity, (i.e. 500 households), the primary quantitative data collected will have two main goals: (1) to develop an understanding of the households participating in the program thereby enabling the IC to make adjustments to programming before the second iteration; and (2) to conduct the CEA as described in **Section 4.4.1,** for iteration 1 to create a measure for the cost-effectiveness of the payment plan chosen for the first iteration. As per IMPAQ's understanding, there is no endline survey envisioned by the IC for the first iteration beyond the household satisfaction survey.

For our main sample for the HER sub-activity (i.e. 2,100 households in the second iteration for HER), we will use primary data collected by in-person household surveys at two points in time: baseline and endline to estimate mechanisms of change for our main outcome variables (i.e. electricity consumption and expenditure). Specifically, primary data collected through the baseline and endline survey will have two main goals: (1) collect information to understand the substitution between different energy sources, the mechanisms of change, and the impact of the BC&O campaign, and (2) collect information on the demographic and socio-economic characteristics of the household which will be important to use as control variables in the ITS analysis since the administrative data does not give us this information. We will interview multiple people in the household with the most knowledgeable person answering the questions, as needed. This is because, since the major proportion of household heads in Kosovo are male, interviewing only household heads will exclude women's perspectives. The goal of the baseline data

collection will be to obtain pre-intervention values of our main outcome variables and information on the demographics and socio-economic characteristics which will act as control variables in our ITS analysis. The goal of the endline data collection will be to quantify the change in energy usage behavior for each household. The survey will include detailed questions about the demographics, physical characteristics of the house, and different sources of energy and cost of each source, among other outcomes.

The IC will collect the baseline data for the HER survey, with IMPAQ's support. At baseline, the IC will procure the data collection firm and IMPAQ will lead the enumerator training, survey instrument development, pilot testing, and quality checks. At endline, IMPAQ will be responsible for procuring the data collection firm and leading the data collection efforts. **Section 4.2.6** details our approach for ensuring data quality.

AER survey (joint responsibility of IC and evaluator). Similar to the HER sub-activity, data will be collected using a survey for individual apartment units under the AER sub-activity. The AER survey will also be administered at two points in time: baseline and endline to estimate mechanisms of change for our main outcome variable (i.e. electricity consumption) at the individual apartment unit level and to understand MAB-level behavior change evolution. Primary data collected through the baseline and endline AER survey will be important to understand the substitution between different energy sources, the mechanisms of change, and the impact of the BC&O campaign for individual apartments and for the MAB as a whole. The goal of the baseline data collection will be to obtain pre-intervention values of our main outcome variables and data on demographic and socio-economic characteristics which will act as controls in the ITS analysis. The goal of the endline data collection will be to quantify the change in energy usage behavior for each apartment unit and to understand MAB-level changes in energy consumption in common areas and behavior change.

The AER survey will include two modules: (1) the first module will be similar to the HER survey and will be administered to all household heads in the individual apartment units; and (2) the second module will be administered to only those apartment units where the household head is also a HOA representative and will be skipped for all other apartment units. The first module will include questions similar to the HER sub-activity, which will try to understand individual decision making for each apartment unit and will include questions around demographics, decision to contribute money to the MAB level retrofits, energy consumption from different sources, and behavior change questions, among others. Baseline information from the first module will provide important information on individual apartment units which will be used as control variables in the ITS analysis. The second module which will be administered to HOA representatives only will include questions about the physical structure of the MAB, buy-in of individual apartments, and behavior change questions at the MAB level, among others.

The IC will collect the baseline data for the AER survey, with IMPAQ's support. At baseline, the IC will procure the data collection firm and IMPAQ will lead the enumerator training, survey instrument development, pilot testing, and quality checks. At endline, IMPAQ will be responsible for procuring the data collection firm and leading the data collection efforts. **Section 4.2.6** details our approach for ensuring data quality.

Administrative data (responsibility of evaluator). This data will be collected for beneficiary households in the HER and AER sub-activities. The administrative data will also allow us to measure outcomes on a continuous basis. KEDS data will also provide us reliable information about our main outcome variables:

monthly electricity consumption and expenditures.[30] Additionally, it will provide historical electricity consumption information for each household under HER and each apartment unit under AER. IMPAQ will be responsible for coordinating with KEDS to obtain necessary administrative data. During the baseline in-person surveys, we will ask households for their KEDS invoice number, so that we can track their electricity usage from the KEDS database. We will collect historical data for each household from January 2018 onward and until the end of the evaluation period in May 2024.[31] We will execute a data-sharing agreement with KEDS, under which they will transmit data to IMPAQ periodically.

The KEDS administrative data will be transmitted on a monthly basis and will include the following variables as shared by the Reporting and Analysis Manager of KEDS with IMPAQ on May 21, 2020: (a) customer ID, (b) customer name, (c) customer address, (d) customer municipality, (e) total electricity debt, (f) actual electricity debt, (g) last electricity payment in euros, (h) last electricity payment date, (i) bill number, (j) month/year, (k) season, (l) electricity consumption during high tariff times of day, (m) electricity consumption during low tariff times of day, (n) bill amount, and (o) reading date. Data collected be KEDS is the main source of household level electricity consumption for all official country-level estimates for Kosovo and is thus reliable. In discussions with KEDS in May 2020, they have in principle agreed to share the personally identified data of each beneficiary of the HER and AER sub-activities if consent is obtained from beneficiaries. IMPAQ has shared an official (MCC-approved) data sharing agreement with KEDS subsequently and this agreement is expected to be executed in the following months accounting for COVID-19 related administrative delays.

Monitoring data (responsibility of MFK and IC). We will use monitoring data from MFK and the IC for the SEEK activity, to understand the progress made toward implementation of the HER and AER sub-activities. We will track the "process" indicators in the RELP monitoring and evaluation plan. **Exhibit 10** provides details about the outcomes and timing of data collection for each data source.

Exhibit 10. Data Sources, Outcomes, and Timing of Data Collection

Data Source	Tentative Timing of Data Collection	Outcomes
HER survey ▪ HER sub-activity (500 households in iteration 1, 2,100 households in iteration 2)	▪ **Baseline** ***Iteration 1:*** Rolling from August 2020 – December 2020; ***Iteration 2:*** Rolling from March 2021 – September 2021 ▪ **Endline:** September – October 2023	**Baseline (500 – iteration 1 and 2,100 – iteration 2)** ▪ Electricity reliability (outages) ▪ Electricity consumption and expenditure (disaggregated by seasons as needed) ▪ Energy consumption and expenditure from all other sources ▪ Awareness of and willingness to pay for energy efficiency ▪ Knowledge of energy efficiency ▪ Investment in electrical appliances ▪ Household and housing unit characteristics ▪ Baseline level of insulation of home, including current levels of investments in energy efficiency retrofits **Additional outcomes at endline (only for iteration 2)** ▪ Retrofit decisions ▪ Level of satisfaction

[30] Since electricity consumption is the main outcome variable, access to administrative data on a regular basis will give us the highest levels of accuracy for this outcome variable.
[31] The threshold period ends in September 2021.

Data Source	Tentative Timing of Data Collection	Outcomes
		▪ Other benefits, apart from energy savings ▪ Likelihood of sustaining investments ▪ Quality and maintenance of energy efficiency retrofits ▪ Shifts in energy efficiency usage ▪ Impacts on peer networks ▪ Disaggregated impacts by gender, age of household head, ethnicity, and economic well-being
AER survey ▪ AER sub-activity (~690 apartment units in 25 MABs)	▪ **Baseline:** Rolling, August 2020 – September 2021 ▪ **Endline:** September – October 2023	**<u>Module 1 of AER survey (all apartment units)</u>** ***Baseline (Module 1 of AER survey)*** ▪ Electricity consumption and expenditure ▪ Energy consumption and expenditure from all other sources ▪ Awareness of and willingness to pay for energy efficiency ▪ Knowledge of energy efficiency ▪ Investment in electrical appliances ▪ Household demographic and socio-economic characteristics ***Endline (Module 1 of AER survey)*** ▪ Shifts in energy consumption ▪ Level of satisfaction ▪ Other benefits, apart from energy savings ▪ Disaggregated impacts by gender, age of household head, ethnicity, and economic well-being *(only in in apartment survey)* **<u>Module 2 of AER survey (only HOA representatives)</u>** ***Baseline (Module 2 of AER survey)*** ▪ Awareness of and willingness to invest in apartment-level energy efficiency ▪ Outages ▪ Level of insulation in common areas and building envelope ▪ Apartment building characteristics ▪ Financial status of HOA ▪ Characteristics of HOA, including numbers of members and meetings, and roles and responsibilities ▪ Relationship of HOA with municipality ▪ MAB composition (numbers of apartment owners, tenants, and vacant apartments) ***Additional outcomes at endline (Module 2 of AER survey)*** ▪ Retrofit decisions ▪ Level of satisfaction ▪ Likelihood of sustaining investments ▪ Shifts in energy efficiency usage ▪ Quality and maintenance of energy efficiency retrofits ▪ Changes in MAB composition
Administrative data ▪ HER sub-activity ▪ AER sub-activity*	▪ Monthly, January 2018– May 2024	▪ Electricity consumption and expenditure
Monitoring data	▪ Throughout program lifetime	▪ RELP monitoring and evaluation plan process indicators

**Administrative data for the AER sub-activity is contingent on an apartment owner survey being conducted in each apartment in a MAB and acquiring the KEDS bill number for each individual apartment.*

4.2.5 Analysis Plan

The analysis plan of the HER sub-activity will comprise of two components: (1) an ITS approach to understand the impact of the HER and AER sub-activities on the main outcomes of interest i.e. electricity consumption and expenditure; and (2) a pre-post design which will complement the ITS approach by providing nuanced information on demographic characteristics, the mechanisms of change including an understanding of substitution between different energy sources, and the underlying behavioral reasons for changes in electricity consumption identified by the ITS approach. We detail the ITS analysis and the pre-post analysis below.

1. ITS Analysis

We will estimate the causal impact of the HER and AER sub-activities on electricity consumption and expenditure using the ITS approach. For the HER sub-activity, since at this stage it is unclear how many intervention models are proposed for the second iteration, the main goal will be to examine the combined effects of all the intervention models under the HER sub-activity, pooling households from all different models that might eventually be implemented.

Households in our main sample can enroll in the HER and AER sub-activities at different points in time between March 2020 and September 2021. Thus, each household will have a different "treatment month." We will create "treatment bins" for each 4-week period from March 2020 to September 2021.[32] All households falling within a 4-week bin will be assigned that treatment bin, and the pre-intervention period and post-intervention period will be unique for each treatment bin. For instance, for all households retrofitted from March 1, 2021 through March 31, 2020, the *pre-intervention period* will be before March 2021 and the *post-intervention period* will be after March 31, 2020. For each treatment bin we will visually inspect for changes in the trend of electricity consumption shortly before and shortly after this time period (Bernal et al. 2017).[33] This will allow us to investigate whether there are any changes happening concurrently with SEEK that might influence the results.

We will use the following specification to estimate the causal impact of the HER and AER sub-activities on electricity consumption and expenditure, obtained using administrative data (Linden 2015). As a note, the analysis will be run separately for the HER and AER sub-activities:

$$y_{it} = \alpha + \beta_1 T_t + \beta_2 Post_t + \delta T_t * Post_t + \beta_3 X_t + \beta_4 Z_i + \mu_{it} \qquad (1)$$

where y_{it} is the electricity consumption for household *i* at time *t* where time is measured in months; $Post_t$ is a binary indicator equal to 1 for all months after a retrofit is installed in a household, and 0 otherwise. T_t is a variable representing the time since the start of the study period; Z_i includes baseline characteristics specific to the household, such as information about the physical structure of the house, assets/income, and demographic characteristics; and μ_{it} is a random error term that exhibits

[32] The final decision on the length of time for each treatment bin will be based on the implementation schedule of the HER retrofits by the IC.

[33] In case there is not sufficient variation in the timing of retrofit installation, or if there are some bins with very few households in them, then we will consider defining one uniform "post-intervention" period (October 2021) after all retrofits are completed.

autocorrelation. β_2 is the level change following the intervention and δ, is the slope change following the intervention. Finally, we will account for autocorrelation in the standard errors.

One of the assumptions of the ITS design is that the approach assumes away any impacts of seasonality. For instance, it is possible that any change in electricity consumption we observe in October 2021 is purely the result of seasonality effects. To account for this X_t includes variables to account for temporal patterns of electricity consumption, including month-of-year effects and time-of-day effects (if available from the KEDS data), and weather related variables such as outside temperature (Wagner et al. 2002). Finally, all outcomes under the ITS approach will be disaggregated by key groups including gender, age of household head, ethnicity, and household economic well-being.

2. Pre-post Analysis

We will complement the ITS analysis, using the pre-post analysis, which will provide more nuanced information on the mechanisms of change (including consumption of different energy sources and the substitution between them) over time for three groups: (1) households in the HER sub-activity; (2) individual apartment units in the AER sub-activity; and (3) HOA representatives in the AER sub-activity. The pre-post analysis will compare the level of an outcome before and after the intervention. We will use the following regression model:

$$y_{it} = \alpha + \beta_1 Post_t + \beta_3 X_{it} + \mu_{it} \qquad (2)$$

where y_{it} is the outcome of interest (e.g., energy consumption from other sources) for household/HOA representative *i* at time *t* (t = baseline or endline); and $Post_t$ is a binary indicator equal to 1 for endline and 0 for baseline. X_{it} includes characteristics specific to the household/MAB, such as information about the physical structure of the house, assets/income, and demographic characteristics; and μ_{it} is a random error term. The coefficient of interest, β_1, provides estimates of the impact of the HER/AER sub-activity on our outcomes of interest.

The primary analysis will be based on the responses of all household heads for the HER and AER sub-activities and HOA representatives for the AER sub-activity. We will also conduct subgroup analyses to inform the evaluation question related to differences in changes over time by income or gender. Findings from the pre-post analysis, while not causal since time-varying factors unrelated to the program could still be driving some of the observed changes, are critical since they will help in understanding the behavioral mechanisms at play as a result of an improvement in energy efficiency.

4.2.6 Data Collection, Data Processing, and Data Quality

The evaluation's success depends on the collection of high-quality data, particularly the accuracy, reliability, and timeliness of the data. In this context, to ensure high quality data collection, we will actively anticipate risks and minimize threats to quality that are inherent in the data collection process.

Instrument Development. IMPAQ will develop baseline and endline survey instruments for households for the HER and AER sub-activities and HOA representatives for the AER sub-activity. All guides will first be developed in English, then translated into Albanian. If other languages (such as Serbian) are needed, depending on the ethnicity of participants, we will ensure that a qualified translator completes this work. We anticipate that the household and the HOA representative/apartment owner survey will last about 40 minutes each. The survey instrument will be clearly formatted, and the questions properly ordered and easily understood, to ensure high and accurate item response.

Development of Survey Instruments

- Questions must be worded simply, clearly, and in an unbiased manner, so that respondents can readily understand what is expected of them.
- Question response categories must be appropriate, mutually exclusive, and as exhaustive as possible, given the intent of the question.
- Questions must be accompanied by clear, concise coding instructions and probes, so that interviewers know what is expected of them.
- Questions must be arranged by topic area, with transitional sentences to ensure proper flow between topics, to promote respondent understanding.
- Questions must be ordered to minimize question-order bias effects and item nonresponse, with sensitive questions placed toward the end of the instrument.

Programming. The survey instruments will be programmed in English, Albanian, and (if necessary) Serbian, in a computer-assisted personal interviewing instrument, by the data collection partner hired by the ICs. At baseline, IMPAQ will provide technical assistance to the IC by checking the translation and testing the programmed version of the instrument. At endline, IMPAQ will be responsible for the data collection and will lead this process.

Enumerator Training. Before the data collection starts, IMPAQ will hold intensive training sessions for the field team responsible for data collection (hired by IC at baseline and by IMPAQ at endline).[34] The IMPAQ team will train data collection supervisors and enumerators to effectively use tablets to implement the in-person surveys offline, without the need for a Wi-Fi connection. The "classroom" training will last 3-4 days. We will review the training guide developed for enumerator training by the data collection partner, which will outline survey administration, consent, and proper storage of data and tablets. This document will also include the rationale and background of the study and other items that allow enumerators to answer study questions. The enumerator training sessions will include training in general interviewing skills and program-specific training. The general interview skills training includes sessions on how to (a) avoid interview bias, (b) use probing techniques, (c) build the trust and cooperation of respondents, (d) read questions and record responses, and (e) adhere to the questionnaire to standardize responses and ensure proper data recording and confidentiality. The SEEK program-specific training, guided by an interviewer manual, will includes lectures, mock interviews, and in-class exercises.

Field Testing. After three days of "classroom" training, the IMPAQ team will provide an opportunity for the enumerator team to practice with real respondents for 1-2 days, before going into the field. The field team, in collaboration with IMPAQ and the IC, will conduct a pre-test with at least six to eight households

[34] As per IMPAQ's original statement of work, IMPAQ was expected to support enumerator training and field-testing at baseline. As per an agreed-upon contract modification by MCC in March 2020, IMPAQ will now lead the enumerator training and field testing at baseline and endline. The data collection firm will be hired by the IC at baseline and by IMPAQ at endline.

and one to two HOAs that will have not enrolled in the HER and AER sub-activities. After the pilot, all enumerators will regroup with the IMPAQ team in the field to debrief. The IMPAQ team will discuss and address any issues encountered during the pilot, to ensure smooth and consistent survey implementation. Final revisions may be made to the tools, based on enumerator feedback.

Human Subjects Protection Plan. We describe in detail the institutional review board requirements and process for application in **Section 10**. Additionally, during the enumerator training, IMPAQ will brief enumerators about procedures for interviewing respondents, protecting respondents' privacy and confidentiality, and securing data. Overall, our field team will also be trained in procedures for contacting respondents, protecting respondent privacy and confidentiality, and securing data, thus ensuring high compliance with ethical guidelines to conduct research. Furthermore, after data collection, the evaluation team will protect the privacy and confidentiality of respondents by storing the data on secure servers and separating personally identifiable information from the survey data.

Data Collection and Data Quality Checks. An IMPAQ data expert will test the instruments in the programmed tablets, as well as run quality checks (check for missing data, abnormal values, skip patterns, etc.) to ensure high quality output from the field. IMPAQ will provide oversight and support throughout the data collection process. Finally, the survey firm will be required to conduct callbacks with 10 percent to 15 percent of respondents at baseline and endline, to verify responses.

4.3 Qualitative Approach: HER and AER Sub-activities

In this section, we describe the qualitative design of the HER and AER sub-activities to track outcomes related to program implementation, mechanisms of change, and sustainability. IMPAQ's qualitative approach to evaluate the HER and AER sub-activities will use a variety of methods, including document review, interviews, and case studies to examine the implementation, mechanisms, lessons learned, and sustainability of the sub-activities. Throughout our qualitative data collection, we will emphasize developing a deeper understanding of the gender considerations, by conducing women-only FGDs and by qualitatively assessing the gender components of the activities as implemented. To complement the quantitative findings, our qualitative approach will look at the outputs and intermediate outcomes in the RELP theory of change, through a theory-testing process-tracing component, a method that examines the mechanisms through which any associated outcomes are achieved (Collier 2011). We plan to collect five rounds of qualitative data—at baseline, three rounds of interim data collection, and at endline.

In **Section 4.3.1**, we describe the qualitative methods and in **Section 4.3.2**, we define the outcomes. In **Section 4.3.3**, we outline the data collection and finally, **Section 4.3.4** provides an overview of the sample and data collection rounds.

4.3.1 Methods

HER sub-activity. Our data collection will include FGDs with both beneficiary and "non-beneficiary" households. We define non-beneficiary households as those households who reside within the same community and beneficiaries. From the non-beneficiary households, we can learn why households that were exposed to the same outreach campaigns decided not to participate, to determine whether any adjustments can be made to the HER sub-activity to encourage future enrollment. For example, it may emerge that lower-income households do not understand the benefits, including the ultimate cost savings, of participating in the program. Alternatively, we may find that these households are very aware of the benefits and would like to participate, but are unable to navigate the financing process and/or pay

the necessary costs up front. Once implementation for the HER sub-activity is completed and we have preliminary findings, we will hold FGDs with a mix of beneficiary households with both high and low energy consumption, to examine underlying behavior and contributing factors. We will also hold KIIs with associated stakeholders, including qualified installers. Finally, we will examine the extent to which gender was incorporated by the IC at various implementation stages (as laid out in the "HER Operations Manual"), including selection of beneficiaries, design of grant levels, procurement of qualified installers, BC&O campaign, capacity building, and installation of retrofits.

AER sub-activity. For the AER assessment, we propose a process-tracing method with seven single case studies. The unit for the case study is the MAB, and we will select one MAB per municipality. We will ensure that our cases represent special beneficiary target groups, including poor/low income, disabled/seriously ill, female-headed (including single mothers and elderly women), and minorities. Within each case study, we will hold an FGD or small group interview with apartment residents and HOA representatives (or alternative leadership). As part of the case studies, we will interview those who led the MAB retrofitting design and construction, which includes the corresponding designers, contractors, and supervisors. With guidance from MFK and MCC, we will interview any other ancillary stakeholders as appropriate and feasible, such as municipality leaders, if the municipalities invested in the MABs as planned. We will incorporate SEEK monitoring, evaluation, and learning data and our own quantitative data in the individual case studies. We will follow each case over the life of the AER sub-activity, plus 2.5 additional years after the end of the threshold, giving us longitudinal data to provide insights into the activity's theory of change. Finally, we will examine the extent to which gender was incorporated by the IC at various implementation stages of AER (as laid out in the "AER Operations Manual"), including selection of MABs, design of grant levels, establishment of HOAs, building audit and program design, procurement and contracting, BC&O campaign, capacity building, and installation of retrofits at the MAB level.

At the end, for both the HER and AER sub-activities, we will review program documents to gain information about the implementation of the SEEK activity. On an ongoing basis, we will review progress, annual, quarterly, and monitoring and evaluation reports, and any other relevant documents. The document review will also assess work plans, timelines, and schedules to gain a full understanding of the design, implementation, and progress of the HER and AER activities. We will also continue to collect documentation about legal, economic, and technical regulations, and laws, policies, and news and media reports about new developments in the Kosovo energy sector.

4.3.2 Outcomes

We will track the progress, successes, and challenges related to implementation **(EQ1**[35]**)**, including outreach campaigns, establishment of HOAs (AER), investment by municipalities, willingness to accept retrofits, loan/financing process, construction cost overruns, and the timeliness and quality of retrofits.

We will examine the assumptions (or mechanisms) conceptualized by the RELP theory of change as they relate to the impacts (or lack thereof) of the HER and AER sub-activities on reducing electrical outages, energy use, and energy expenditures **(EQ2**[36]**)**. By comparing households where there was the largest

[35] *EQ1: Were the activities implemented as designed and were the outputs outlined in the project logic achieved?*
[36] *EQ2: Did the SEEK activity result in reduced electricity consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (BC&O activities, installation of energy*

decrease in energy consumption with those with the smallest decrease (or even increase), we hope to identify the essential elements necessary for scaling up. A better understanding of the changes induced by SEEK that are related to "outcomes" and "impacts" will help explain the underlying mechanisms. This information will enable us to have greater insight when we interpret the results from the ITS approach. For instance, we expect that the findings could reveal that the HER and AER sub-activities were not successful in reducing electricity consumption. That is, it is possible that the BC&O campaign increased knowledge of and awareness about the benefits of energy efficiency, but beneficiaries use more electricity because they are heating their homes more and/or they purchase additional electrical appliances. It is also possible that there was no increase of knowledge or change of attitudes in the first place.

The evaluation will provide evidence to inform decision-making, identify lessons learned and good practices, and provide recommendations for the design and implementation of future energy efficiency programs. However, through our baseline and interim data collection, we can identify lessons learned from the early stages of the HER and AER sub-activities, which can inform scale-up or improve activities within the life of the program. For example, through exploring the experiences of both beneficiaries and installers, there may be opportunities to improve program efficiency, modify BC&O, etc.

Finally, to assess sustainability **(EQ7[37])**, we will collect qualitative data through KIIs with implementers at the end of the threshold program, and with households and MABs towards the end of the evaluation period. For HER, our purposive sample will qualitatively look at the diffusion of energy efficient behaviors, because we will be comparing households that chose to sign up for the activity against those that did not. Through KIIs with contractors and qualified installers, we can learn whether demand for energy efficiency retrofits has increased beyond the scope of the HER and AER sub-activities. Through AER case studies and FGDs with households that enrolled into the HER sub-activity relatively early in the enrollment cycle, we can examine whether and how attitudes and behaviors related to energy use shift or drop off over multiple years. We will also explore whether a realistic cost-efficient model that reduces energy usage was identified (based on interviews with MFK), and if so, whether it has the potential to be scaled up nationwide **(EQ3[38])**.

4.3.3 Data Collection

To evaluate both the HER and AER sub-activities, we will collect five rounds of qualitative data: baseline (round 1), interim (rounds 2, 3, and 4), and endline (round 5), described in more detail below. The first three rounds of data collection will focus on implementation and short-term outcomes, as the timing is concentrated during the threshold period and the first heating season post-retrofits. The last two rounds of data collection will occur later, and will focus on the long-term outcomes and sustainability of the SEEK activity, including providing contextualization for the findings from the ITS analysis and the endline survey. See **Exhibit 11** for a breakdown of the sample, timing, and outcomes covered in each round.

efficient equipment)? Were energy savings previously envisioned during preliminary baseline studies and energy audits achieved, and why?

[37] *EQ7: Did the program meet the stated objective of reducing the gap between supply and demand of electricity? How sustainable are critical outcomes of the program, and why?*

[38] *EQ3: Did the SEEK activity identify a cost-effective model for reducing household energy use that could be scaled up nationwide and what is the overall expected energy reduction?*

Baseline (Round 1). We propose collecting baseline data during the winter season of the first year of SEEK implementation , because this will allow us to interview households participating in the first iteration of the HER sub-activity and AER MABs that have already completed the retrofits and have experienced at least a few months of cold weather. From this, we can provide formative feedback to MFK and the IC on households' experience with and perceptions of the HER and AER sub-activities, particularly focusing on how different aspects of the model (such as the incentive amount) impact the decision to apply and the ultimate decision to adopt energy efficiency measures. In addition to examining current behavior during the heating season, we will ask households to project their anticipated behavior in the future, which may inform the IC about what type of BC&O is necessary during the second iteration of the HER sub-activity. In addition, we will ask households about the timeframe (outreach → application → energy audit → financing → installation → reimbursement) to see whether there are any weak points where there is potential to fall short during scale-up. This can provide preliminary insights into the main outcomes of interest (energy efficiency, consumption, and expenditure) for the second iteration. We will also hold two focus groups with non-beneficiaries, specifically for households in the intervention area who were exposed to the same BC&O campaigns. This can provide formative feedback for the second iteration on what worked and did not work during the BC&O.

Interim (Round 2). We suggest collecting interim data at the end of the threshold period, where we can learn from MFK if and how a cost-effective model has been selected. This will also be immediately after the completion of the HER and AER sub-activities so the ICs, designers, contractors, qualified installers, and energy auditors will still be available. They can provide information related to successes, challenges, and variation between models (if applicable). HER and AER beneficiaries can provide further information related to implementation, and at this stage, we can fully answer **EQ1**.[39]

Interim (Round 3). Our third round of data collection will occur after the first heating season. At this point from the ITS analysis, we will know whether or not there was a decrease in energy consumption immediately after the beneficiaries made energy efficient investments. Our qualitative data collection can explore the mechanisms of any change, including the effectiveness of BC&O.

Interim (Round 4). Our fourth round of data collection will occur one year later, after the second heating season. The RELP theory of change posits that increased awareness about the benefits of energy efficient measures and retrofitting households and apartments with energy efficient improvements will lead to reduced energy consumption. Again, from the ITS analysis, we will know the pattern of energy consumption in participating SEEK households and MABs. There are many possibilities, all of which we can help explain through qualitative data. For example, there may be no change or an increase in energy consumption the first year (if participants switch from non-electrical heating sources to electricity), followed by a decrease in the second year as participants become more familiar with energy efficient behaviors regarding electricity use. Or, there may be a decrease in the first year (if participants maintain consistent comfort levels) followed by an increase in the second year (if participants purchase other electrical appliances with little or minimal increases to their usual electricity bill). Or, there might not be any change at all, as we may find that participants are used to a certain bill, and are now able to heat more of their house for the same monthly cost.

[39] *EQ1: Were the activities implemented as designed and were the outputs outlined in the project logic achieved?*

Endline (Round 5). We suggest collecting qualitative endline data *after* the endline quantitative survey data collection. Doing so will allow us to tailor our questions to follow up on the results from this analysis, including any unexpected or unusual findings.

Data Processing and Data Quality

The IMPAQ team will be primarily responsible for leading the qualitative data collection efforts. We will hire local facilitators and translators to assist where needed. The IMPAQ team will provide training and oversight to those providing support to ensure that the data is of the highest quality. We will follow the Federal Policy for the Protection of Human Subjects and will inform respondents of the purpose of the research and their right to confidentiality. We also will seek the consent of each participant to take part in the assessment and to include their names in a list of interviewees. As we collect data, we update protocols based on lessons learned, to maximize data quality and account for any changes in program implementation.

All guides will first be developed in English, then translated by a member of our team into Albanian. If other languages (such as Serbian) are needed, depending on the ethnicity of FGD participants, we will ensure that a qualified translator completes this work. A minimum of two people will be present for each KII/FGD—one to lead facilitation and the other to take notes. With the consent of the participants, KIIs and FGDs will be recorded. We will take detailed notes of the KII/FGD to form the foundation for subsequent analysis, which will be finalized on the same day that data collection has occurred. We will develop a coding plan tied to the evaluation questions and to any key topics that emerge during data collection. After applying the codes to the FGD and KII summaries, we will analyze the data and identify themes that emerge, using NVivo or a similar qualitative data analysis software. We will compare responses from respondents based on role (such as implementer or beneficiary) and respondent type (such as gender, as women may have a different experience) within each data collection round and draw out areas of overlap, as well as divergence, to understand the implementation context more completely. These data will also inform our interpretation of the quantitative results, particularly factors that may contribute to expected or unexpected impacts and that reveal how different stakeholders experienced RELP activities and benefits.

4.3.4 Sample

HER sub-activity. For the HER baseline data collection, we will hold at least six FGDs representative of the HER target groups during the first iteration (low-income/vulnerable households, female-headed households, minority households, and households with people with disabilities or people who are chronically ill), including geographic representation. Two FGDs will be with households in targeted communities who did not participate in the program. For the second round of data collection, we will hold FGDs with similar households who participated in the second iteration of SEEK. For the subsequent rounds of data collection, including the endline, our sample will be contingent on findings from the quantitative analysis (both the ITS and the survey). We will select respondents based on variation in outcomes (for example, consumers with increased electricity consumption and consumers with decreased electricity consumption), as well as those who might help contextualize or explain any unusual quantitative findings. Again, we will consider representation in gender, income level, ethnicity, region, and other relevant characteristics, including holding several women-only FGDs to get the perspective of female-headed households concerning their experiences, as they are a key focus of the program. For the HER endline data collection, we will aim to have approximately six-eight participants in each FGD.

AER sub-activity. For the AER qualitative data collection, we will select seven MABs as cases to follow over the five rounds of data collection. Case selection will be purposive and will first be decided by location (one case per municipality). Using the SEEK quantitative data, we will ensure a diverse range of representation on characteristics including resident demographics (low-income, female-headed, and/or minority-owned); age and general condition of the MAB; incentive level; retrofit quality, cost, and/or timeframe; and satisfaction of apartment owners. We will select three cases that have started or completed retrofits before baseline data collection (for the first round) the remaining four cases towards the end of program implementation (for the second round).

Exhibit 11. Summary of Rounds of Qualitative Data Collection for HER and AER Sub-activities

Data Source/ Respondent	Evaluation Questions and Outcomes
Baseline (Round 1) – During SEEK Implementation ***Exposure Period:*** After the first 3-4 months of implementation, that is, after the first iteration of HER retrofits is complete and at least some AER retrofits installed ***Tentative Date: January/February 2021***	
HER Beneficiaries: up to 4 FGDs *(HH who completed retrofits in iteration 1)* **Non-participating Households:** up to 2 FGDs with non-participant HH **AER Beneficiaries:** 3 FGDs *(case study MABs where retrofits have started or completed)* **MFK:** 4-6 KIIs *(Energy Director, Energy Specialist, Monitoring and Evaluation Specialist, other relevant staff)* **GFA:** 4-6 KIIs *(HER and AER IC Leader, BC&O expert, other relevant staff)* **Retrofit Implementers:** up to 10 KIIs *(contractors, qualified installers, energy auditors)*	**EQ1:** Were the activities implemented as designed and were the outputs outlined in the project logic achieved? **Outcomes:** ▪ Quality and appropriateness of design ▪ Project milestones and benchmarks; delays and challenges in implementation ▪ Operational processes and systems; use of data management systems and communication procedures ▪ Beneficiary selection (households and MABs) ▪ Design of household incentives and BC&O campaigns ▪ Energy audit training ▪ Incorporation of gender-specific components ▪ Engagement with municipalities (AER) ▪ Experience with training/technical assistance ▪ Experience with financing options ▪ Experience with application, financing, and installation ▪ Perceptions of installation quality ▪ Problems with construction cost overruns ▪ Engagement with MFK/IC ▪ Experience with training/technical assistance (if relevant) ▪ Alignment with organization/government priorities **EQ2:** Did the SEEK activity result in reduced electricity consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (BC&O activities, installation of energy efficient equipment)? Were energy savings previously envisioned during preliminary baseline studies and energy audits achieved, and why? **Outcomes:** ▪ Reasons for applying (HER beneficiaries) and not applying (non-participant households) to SEEK ▪ HOA formation and engaging apartment owners (AER) ▪ Incentives, information, and guidance needed to make household/MAB energy efficiency retrofits and investments ▪ Perceptions of electricity quality, reliability, and affordability ▪ Awareness of energy saving measures and benefits ▪ Willingness to pay for energy efficiency

Data Source/ Respondent	Evaluation Questions and Outcomes
Ancillary Stakeholders: up to 10 KIIs *(municipality leadership, lending institutions, etc. as appropriate)*	▪ Retrofit decisions ▪ Projected energy consumption and energy cost savings ▪ Anticipated further investments in energy-efficient measures
Interim (Round 2) – End of Threshold ***Exposure Period:*** After all HER and AER activities are completed and the threshold program in Kosovo concludes ***Tentative Date: October/November 2021***	
HER Beneficiaries: up to 4 FGDs *(HH who completed retrofits in iteration 2)* **AER Beneficiaries:** up to 4 FGDs *(case study MABs not interviewed in Round 1)* **MFK:** 4-6 KIIs *(Energy Director, Energy Specialist, Monitoring and Evaluation Specialist, other relevant staff)* **GFA:** 4-6 KIIs *(HER and AER IC Leader, BC&O expert, other relevant staff)* **Retrofit Implementers:** up to 10 KIIs *(contractors, qualified installers, energy auditors)* **Ancillary Stakeholders:** up to 10 KIIs *(municipality leadership, lending institutions, etc. as appropriate)*	**EQ1:** Were the activities implemented as designed and were the outputs outlined in the project logic achieved? **Outcomes:** ▪ Project milestones and benchmarks; delays and challenges in implementation ▪ Beneficiary selection (households and MABs) ▪ Design of household incentives and BC&O campaigns ▪ Incorporation of gender-specific components ▪ Engagement with municipalities (AER) ▪ HOA formation and engaging apartment owners (AER) ▪ Experience with training/technical assistance ▪ Experience with financing options ▪ Experience with application, financing, and installation ▪ Perceptions of installation quality ▪ Problems with construction cost overruns ▪ Engagement with MFK/IC ▪ Experience with training/technical assistance (if relevant) ▪ Alignment with organization/government priorities **EQ2:** Did the SEEK activity result in reduced electricity consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (BC&O activities, installation of energy efficient equipment)? Were energy savings previously envisioned during preliminary baseline studies and energy audits achieved, and why? **Outcomes:** ▪ Reasons for applying (HER beneficiaries) and not applying (non-participant households) to SEEK ▪ Incentives, information, and guidance needed to make household/MAB energy efficiency retrofits and investments ▪ HOA formation and engaging apartment owners (AER) ▪ Perceptions of electricity quality, reliability, and affordability ▪ Awareness of energy saving measures and benefits ▪ Willingness to pay for energy efficiency ▪ Retrofit decisions ▪ Projected energy consumption and energy cost savings ▪ Anticipated further investments in energy-efficient measures ▪ Effectiveness of BC&O ▪ Shifts in energy usage ▪ Successes and challenges

Data Source/ Respondent	Evaluation Questions and Outcomes
	EQ3: Did the SEEK activity identify a cost-effective model for reducing household energy use that could be scaled up nationwide and what is the overall expected energy reduction? **Outcomes:** ▪ HER: Identification of cost-effective model ▪ Perceived scalability of model ▪ External factors and events affecting implementation ▪ Best practices and lessons learned ▪ Perceived program impacts ▪ Sustainability of program outcomes ▪ Spillover effects (influence on other households/MABs)
Interim (Round 3) – End of First Heating Season ***Exposure Period:*** After participants have experienced one full heating season post- HER and AER retrofit installation ***Tentative Date: April/May 2022*** ***Interim (Round 4) – End of Second Heating Season*** ***Exposure Period:*** After participants have experienced two full heating seasons post- HER and AER retrofit installation ***Tentative Date: April/May 2023***	
HER Beneficiaries: up to 4 FGDs **AER Beneficiaries:** up to 4 FGDs **Ancillary Stakeholders:** up to 10 KIIs	**EQ2:** Did the SEEK activity result in reduced electricity consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (BC&O activities, installation of energy efficient equipment)? Were energy savings previously envisioned during preliminary baseline studies and energy audits achieved, and why? **Outcomes:** ▪ Contextual information to explain results from the ITS analysis in the short-term (one heating season) and the long-term (two heating seasons) ▪ Experience with and perceptions of energy efficiency opportunities; behavior change ▪ Increased energy-related investments not related to SEEK program ▪ External factors and events affecting energy use ▪ Effectiveness of BC&O over multiple heating seasons ▪ Changes in energy efficiency ▪ Shifts in energy usage (sustained or drop-off over time) ▪ Realization of energy savings ▪ Changes in consumption uses; new purchases; changes in other household investments ▪ Condition of retrofits over multiple heating seasons, including any necessary maintenance/upkeep ▪ Successes and challenges
Endline (Round 5) – End of Evaluation Period/Post- Quantitative Data Collection ***Exposure Period:*** After the endline survey has been administered and analyzed, in the middle of the third heating season post- HER and AER retrofit installation ***Tentative Date: January/February 2024***	
HER Beneficiaries: up to 10 FGDs **AER Beneficiaries:** 7 FGDs	**EQ7:** Did the program meet the stated objective of reducing the gap between supply and demand of electricity? How sustainable are critical outcomes of the program, and why? **Outcomes:** ▪ Contextual information to explain endline survey findings

Data Source/ Respondent	Evaluation Questions and Outcomes
Ancillary Stakeholders: up to 10 KIIs	▪ Contextual information to explain results from the ITS analysis over 2.5 heating seasons ▪ Additional energy efficient investments or actions taken ▪ External factors and events affecting energy use ▪ Effectiveness of BC&O over multiple heating seasons ▪ Changes in energy efficiency ▪ Shifts in energy usage (sustained or drop-off over time) ▪ Realization of energy savings ▪ Changes in consumption uses; new purchases; changes in other household investments ▪ Condition of retrofits over multiple heating seasons, including sustainability of any necessary maintenance/upkeep ▪ Spillover effects ▪ Successes and challenges; how to scale up/improve

4.4 CEA

4.4.1 CEA: First Iteration of HER Sub-activity

As per our understanding, the first iteration of the HER sub-activity will be rolled out from August through December 2020, during which 500 households will sign up for the installation of energy efficient retrofits on a rolling basis. The CEA for the first iteration will estimate the cost-effectiveness of the implemented model. The period of analysis will match the implementation period, August–December 2020. For this analysis, we will rely on measures of electricity consumption obtained as part of baseline data collection and administrative data received from KEDS. We will start by calculating costs and benefits for the average household under this model. The income benefit from reduced electricity use will be the difference between household electricity use after the retrofit (in December) and before the retrofit (at baseline), using administrative data from KEDS. Financial benefits will be calculated starting with a measure of reduced electricity use:

$$\textit{Incremental Household Electricity Savings} = \textit{Electricity Use}_{\textit{baseline}} - \textit{Electricity Use}_{\textit{after retrofit}}$$

where,

$\textit{Electricity Use}_{\textit{baseline}}$ = *Household electricity consumption (kWh) at baseline (before the retrofit)*

$\textit{Electricity Use}_{\textit{after retrofit}}$ = *Household electricity consumption (kWh) (after the retrofit)*

Valuing the reduced household electricity use will be done using the energy tariff rate charged to households:

$$\textit{Electricity Cost Savings} = \textit{Residential Electricity Tariff} \times \textit{Incremental Household Electricity Savings}$$

Additionally, under the payment plan offered during the first iteration, households are expected to receive some kind of incentive, voucher, or subsidy. This will be treated as a cash inflow (or financial benefit) from the household CEA perspective. **Exhibit 12** and **Exhibit 13** detail the benefits and costs.

Exhibit 12. CEA Benefits for First Iteration: Household Level

Benefits	Data Source
Electricity cost savings	▪ **Quantity of electricity consumed (kWh/household)** will be collected from meter data, daytime and nighttime, provided by KEDS. We will have baseline data, which will be compared to data collected about electricity use after the retrofit (in December 2020). We will also try to collect data from a comparable time period (e.g., December 2019) to estimate electricity consumption before the retrofit, to minimize the influence of weather changes on our pre- and post- electricity consumption estimate. ▪ **Price per unit of electricity (€/kWh)** will be measured by the day/night tariff rate charged to households (provided by the utility company).
Incentive payments/ subsidies	▪ The baseline survey will provide information about the payment support at the household level, estimated on an average basis per participating household (measured in € per household). This information will also be validated with the IC.

Household costs during this time should be included in the model. These costs include household expenditures for all energy efficiency investments made during this period, including equipment and installation costs. We will rely on the baseline data to estimate these costs, on average, for the households in the iterative study.

Exhibit 13. CEA Costs for First Iteration: Household Level

Costs	Data Source
Energy efficiency equipment costs	▪ The baseline survey includes questions to collect actual household expenses associated with energy retrofit equipment purchased for the home. Data will be averaged to estimate equipment purchase costs per household (ideally, to account for different home sizes and investments), measured in € per household.
Installation costs	▪ The baseline survey includes questions to collect actual household expenses and time spent toward equipment installation in the home. This data will be averaged across all participating households (measured in € per household).

The cost-effectiveness ratio (CER) will result in the amount of savings per dollar invested at the household level for the selected payment plan, using the following equation:

$$CER_{\ household} = \frac{Electricity\ Cost\ Savings + Incentive/Subsidy}{Energy\ Efficiency\ Equipment\ Costs + Installation\ Costs}$$

The household-level CER will be useful for creating a measure for the cost-effectiveness of the payment plan chosen for the iterative study, from the household perspective. This is a useful perspective that informs the decisions that households make about whether an energy retrofit is appropriate for their homes and their budgets. Ideally, we will want to see a ratio above one, indicating that household savings are greater than costs.

It is important to note that this perspective does not capture all costs associated with any retrofit, since it deducts the incentive/subsidy that households receive from the overall costs they incur (the IC's costs in incentivizing households to make these investments). To account for the true cost of those retrofits measured against the energy savings, we will adjust the CER ratio by removing the incentive/subsidy cash inflow to the household and instead treat it as a cost, using the following equation:

$$CER_{\ all\ costs} = \frac{Energy\ Cost\ Savings}{Energy\ Efficiency\ Equipment\ Costs + Installation\ Costs\ + Incentive/Subsidy}$$

If possible, we would also include the IC's administration costs in the denominator. This would be much closer to reflecting the true cost of supporting and subsidizing energy retrofits at the household level (taking into account both the household and the IC's investments).

4.4.2 CEA: Second Iteration of HER Sub-activity

We will conduct a detailed CEA for the second iteration of the HER sub-activity and will include financial transactions (cash inflows and outflows) for different beneficiaries. Using a financial approach, rather than an economic approach, will allow us to focus on the financial viability of the program and the direct impact on the household income.

The CEA for the second iteration of the HER sub-activity will be conducted after the completion of endline data collection in September–October 2023. The theory of change for the CEA will follow the SEEK theory of change. Electricity efficiency programs are a demand-side investment that are intended to reduce costs for households and in turn, increase their disposable income. According to the SEEK theory of change, reduced costs will come from reduced energy consumption. Similarly, it is expected that some households may switch from using more expensive non-electric energy sources, such as firewood, as the per-unit cost of electricity decreases. This will also contribute to reduced total energy costs at the household level. Naturally, the extent to which energy efficient measures reduce consumption will depend on the elasticity of energy demand and whether households are under-heating their homes. As noted in the "Evaluability Assessment," energy efficient measures could actually lead to an increase in energy consumption or warmer homes, at the same expenditure levels. In effect, it is possible that a reduction in expected overall energy savings from an energy efficiency investment may cause a behavioral response leading to increased energy consumption (World Bank 2017b). To the extent that this is the case, results from the CEA will be discussed within the broader context, using qualitative data from the evaluation.

The CEA will include the perspectives of different beneficiaries and the MFK/IC. We will present the results as a cost-effectiveness ratio of the amount of costs saved per euro invested in each efficiency retrofit. All cost-effectiveness results will be conducted in real (constant) euros. All results will be in present value terms (using 2020 as the base year), using MCC's standard 10 percent social discount rate. For the HER sub-activity, the period of analysis will match the full lifecycle of each energy retrofit selected.

1. Methodological Approach for HER Sub-activity: Household Perspective

The household perspective will measure the average impact of the program on the participating customers, by measuring the change in their total annual energy expenditures and by adding applicable incentive payments and subtracting equipment and installation costs incurred by them (on average).

The cost-savings benefit will be quantified and monetized by calculating the per-unit cost and quantity consumed of each energy source used at the household level. We will use KEDS administrative data about electricity consumption and data from the baseline and endline surveys to compile a basket of energy sources used at the household level (e.g., firewood, coal). We will use pre- and post-data from the baseline and endline surveys to estimate incremental energy savings before and after the retrofit, as a direct benefit of the intervention.

In addition to the change in energy consumption, it could be that reducing the use of firewood in homes, for example, will lead to health benefits. The default approach is to exclude health benefit streams and to only include them when evidence suggests that the electrification investment is likely to improve health sufficiently to warrant inclusion in the analysis (MCC 2013). The baseline and endline surveys will also include questions about health outcomes that are tied to indoor air pollution and their associated medical costs. If the results from the endline survey suggest that there has been an impact on the health outcomes, we will also estimate the reduced health expenditures and include them as a direct financial benefit at the household level.

Energy consumption savings will be measured as:

$$Energy\ Savings = Energy\ Use_{baseline} - Energy\ Use_{endline}$$

where,

$Energy\ Use_{baseline}$ = *Energy consumption for all energy sources, baseline*
$Energy\ Use_{endline}$ = *Energy consumption for all energy sources, endline*

Income benefits from energy consumption savings will be measured as:

$$Energy\ Cost\ Savings_{household} = Pricing\ Structure \times (Energy\ Use_{baseline} - Energy\ Use_{endline})$$

This calculation will take into account independent changes in price risks, such as changes to energy contracts, structures, or tariff rates. Finally, if there are significant health impacts at endline, the final benefits will be measured as:

$$Household\ Benefits = Energy\ Cost\ Savings_{household} + (Medical\ Costs_{baseline} - Medical\ Costs_{endline}) + Incentives/Subsidy\ Payments$$

Exhibit 14 and Exhibit 15 present the detailed benefits and costs from the household perspective.

Exhibit 14. CEA Benefits for Second Iteration: Household Perspective

Benefits	Data Source
Cost savings from energy consumption reduction (across all sources of energy)	▪ **Incremental quantity of electricity consumed** (kWh/household) will be collected monthly from meter data, daytime and nighttime, provided by KEDS. This will be converted to annual equivalents. ▪ **Incremental quantity of off-grid energy (or any non-electric heating sources, such as firewood or coal) consumed** (unit per source of off-grid energy / household) will be calculated from the household survey at baseline and endline. Endline energy estimates will be collected in 2023, but we will assume that measurable changes in energy consumption occur immediately after the retrofit.[40] ▪ **Price per unit of electricity** (€/kWh) will be the day/night tariff rate charged to households (provided by the utility company). Real prices will be used and assumed to be constant throughout the period of analysis, unless data suggests that real prices will change during this period. ▪ **Price per unit of off-grid energy** (€/unit of energy consumed) will be provided by the baseline and endline surveys. An average per-unit price will be calculated across all

[40] In the final year of data collection, incremental energy saved (both electric and off-grid) will be assumed to continue for the remainder of the period of analysis for the CEA. For example, if the period of analysis is 10 years, the incremental energy saved measured in the final year of data collection (Year 3) will be expected to continue for the following 7 years of the period of analysis. If any qualitative information suggest that this is an inaccurate assumption, we will make adjustments and note this deviation in the report.

Benefits	Data Source
	households. Real prices will be used and assumed to be constant throughout the period of analysis, unless data suggests that real prices will change during this period.
Cost savings from health benefits (indoor air pollution)	▪ The baseline and endline surveys will create an approximate estimate of the incremental difference in medical costs attributable to the program. If there is any evidence of a significant impact on health, any measurable difference will be included in the benefit stream (measured in € per household).
Incentive payments/ subsidies	▪ This data, provided by the baseline and endline surveys, will be triangulated with the IC and estimated on an average basis per participating household (measured in € per household).

Only costs associated with the different energy retrofits will be included in the household CEA. Other energy retrofits that household may invest in independently will be excluded from the CEA.

Exhibit 15. CEA Costs for Second Iteration: Household level

Costs	Data Source
Energy efficiency equipment costs	▪ The endline survey will ask households about their direct expenses associated with energy retrofits purchased for their homes. Equipment costs will be averaged across all participating households to account for different home sizes and investments (measured in € per household).
Installation costs	▪ The endline survey will ask households about any installation costs associated with the energy retrofits. These costs will be averaged across all participating households (measured in € per household).

The cost-effectiveness ratio (CER) will result in the amount of savings per euro invested at the household level, using the following equation:

$$CER_{household} = \frac{Energy\ Cost\ Savings + Medical\ Cost\ Savings + Incentive/Subsidy}{(Energy\ Efficiency\ Equipment\ Costs + Installation\ Costs)_{household}}$$

For the households, we will also be able to calculate the payback period for the energy efficiency investments, or the time period when their energy savings (and other financial benefits) meet or exceed their investments in the retrofit.

2. Methodological Approach for HER Sub-activity: MFK/IC Perspective

The MFK financial perspective (Exhibit 16) will simply add all costs associated with administering the SEEK activity, since no cash inflows/benefits are expected for MFK. These costs include the incentives or subsidies that will be directly paid to the households, as well as the other SEEK implementation costs that are expected to influence household behavior in the HER activity (such as the BC&O activity) and other relevant costs, such as the MFK administration costs. These costs will be added together and divided by the total number of household beneficiaries, to obtain a per-household cost of implementing SEEK.

$$Costs_{MFK} = \frac{Incentive/\ Subsidy\ +\ Other\ implementation\ costs\ +MFK\ Administration}{Households\ participating\ in\ HER\ activity}$$

Exhibit 16. CEA Benefits for Second Iteration: MFK/IC Perspective

Costs	Approach
Incentive payments/ subsidies	▪ The baseline and endline survey data will be triangulated with the IC on an average basis for each participating household (measured in € per household).
Other implementation costs	▪ All direct SEEK program costs relevant to the HER activity will be provided by the IC (measured in € per household).
MFK administration, monitoring and evaluation costs	▪ Typically, these costs, accruing to MFK, are assumed to be 20 percent of the total direct program implementation costs in MCC economic rate of return analyses. This convention will be used for this analysis (measured in € per household).

3. Methodological Approach for HER Sub-activity: Utility Perspective

If households reduce their electricity consumption, this will allow KEDS to provide that unused energy (equivalent to the amount of energy that is reduced due to the HER activity) to other households, resulting in financial savings from KEDS' perspective. Given that demand exceeds supply for some months in the year in Kosovo, it is not assumed that this will lead to reduced revenues for the utility, but rather that KEDS will be able to provide the incremental energy saved from the HER activity to other non-HER households. This should also lead to import cost savings, because less electricity will need to be imported. Because the utility is not expected to incur any incremental costs as a result of the HER activity, this perspective will only measure incremental benefits (or cash inflows) from KEDS' perspective. These benefits are included in **Exhibit 17**.

Exhibit 17. CEA Benefits for Second Iteration: Utility Perspective

Benefits	Approach
Incremental energy supply cost savings (electricity only)	▪ **Incremental quantity of electricity consumed per household** (kWh/household) will be collected monthly from meter data provided by KEDS. This will be converted to annual equivalents. ▪ **Levelized cost to supply imported energy** (€ per kWh) will be provided by literature and verified with KEDS, to ensure that the incremental quantity of electricity saved for HER households displaces the need for additional energy, valued at the cost of importing energy.

This benefit will be calculated as a total benefit (not a per-household benefit, as in the other perspectives) to understand the total value to KEDS:

$$Benefit_{KEDS} = Incremental\ Energy\ Consumption\ Savings/HH * (Levelized\ Cost\ to\ Supply\ Energy) * HH\ participating\ in\ HER$$

4. Methodological Approach for HER Sub-activity: Integrated Perspective

The three perspectives above will be combined into one integrated perspective and CER for each of the retrofits supported in the full HER activity:

$$CER_{integrated} = \frac{Energy\ Cost\ Savings_{household} + Medical\ Cost\ Savings + Incentive + \frac{Benefit_{KEDS}}{HH\ participating\ in\ HER}}{(Energy\ Efficiency\ Equipment\ Costs + Installation\ Costs)_{household} + Costs_{MFK}}$$

This value will be presented in 2020 euro values on a per-household basis for all energy retrofits, discounted at 10 percent for the lifespan of the retrofit selected for the HER activity.

4.4.3 CEA for AER Sub-activity

This CEA for the AER sub-activity will take a similar approach as the one described above. In addition to the household-level perspective, the CEA will also examine the cost-effectiveness from the perspective of KEDS and the IC/MFK and will include the costs of the apartment building level investments, all of which will be summed into an integrated perspective. The CEA for AER will be completed following the endline survey (expected in September–October 2023). We will present the results as a cost-effectiveness ratio of the amount of costs saved per euro invested in household-level energy retrofits. All cost-effectiveness results will be measured in real (constant) euros. All results will be in present value terms (using 2020 as the base year), using MCC's standard 10 percent social discount rate. Since the apartment buildings will likely select energy retrofits with different lifespans, we will select a common period of analysis for all energy retrofits (e.g., 10 years) and any remaining value of the retrofit investment will be credited at the end of the 10-year period, allowing for comparability across different apartments and different energy retrofits.

1. Methodological Approach for AER Sub-activity: Household Perspective

The household perspective will measure the impact of the program on the households within the participating apartment buildings, by measuring the average change in their total annual energy expenditures and subtracting their average contributions towards the energy retrofit investments made by the apartment buildings.

We will quantity and monetize the cost-savings benefit by calculating the per-unit cost and quantity consumed of each energy source used at the household level. We will use KEDS administrative data on electricity consumption and data from the baseline and endline surveys to compile a basket of energy sources used at the household level (e.g., firewood, coal). We will use pre- and post-data from the baseline and endline surveys to estimate incremental energy savings before and after the retrofit, as a direct benefit of the intervention.

Energy consumption savings will be measured as:

$$Energy\ Savings = Energy\ Use_{baseline} - Energy\ Use_{endline}$$

where,

$Energy\ Use_{baseline}$ = *Energy consumption for all energy sources, baseline*
$Energy\ Use_{endline}$ = *Energy consumption for all energy sources, endline*

Income benefits from energy consumption savings will be measured as:

$$Energy\ Cost\ Savings_{household} = Pricing\ Structure \times (Energy\ Use_{baseline} - Energy\ Use_{endline})$$

This calculation will take into account independent changes in price risks, such as changes to energy contracts, structures, or tariff rates.

Exhibit 18 and **Exhibit** 19 present the detailed benefits and costs from the household perspective.

Exhibit 18. CEA Benefits for AER: Household Perspective

Benefits	Data Source
Cost savings from energy consumption reduction (across all sources of energy)	▪ **Incremental quantity of electricity consumed** (kWh/household) will be collected monthly from meter data, daytime and nighttime, provided by KEDS. This will be converted to annual equivalents. ▪ **Incremental quantity of off-grid energy (or any non-electric heating sources, such as firewood or coal) consumed** (unit per source of off-grid energy / household) will be calculated from the household survey at baseline and endline. Endline energy estimates will be collected in 2023, but we will assume that measurable changes in energy consumption occur immediately after the retrofit.[41] ▪ **Price per unit of electricity** (€/kWh) will be the day/night tariff rate charged to households (provided by the utility company). Real prices will be used and assumed to be constant throughout the period of analysis, unless data suggests that real prices will change during this period. ▪ **Price per unit of off-grid energy** (€/unit of energy consumed) will be provided by the baseline and endline surveys. An average per-unit price will be calculated across all households. Real prices will be used and assumed to be constant throughout the period of analysis, unless data suggests that real prices will change during this period.

Only costs that households contribute to the apartment building level energy retrofits will be included in the household CEA. Other energy retrofits that household may invest in independently will be excluded from the CEA.

Exhibit 19. CEA Costs for AER: Household level

Costs	Data Source
Energy efficiency financial contributions to apartment building	▪ The baseline survey will ask households about their direct contributions towards the energy retrofits chosen and installed by the apartment buildings. These contributions will be averaged across all participating households to account for different home sizes and investments (measured in € per household).

The cost-effectiveness ratio (CER) will result in the amount of savings per euro invested at the household level, using the following equation:

$$CER_{\,household} = \frac{Energy\ Cost\ Savings_{\,household}}{(Energy\ Efficiency\ Financial\ Contributions)}$$

For the households, we will also be able to calculate the payback period for the energy efficiency investments, or the time period when their energy savings (and other financial benefits) meet or exceed their investments in the retrofit.

2. Methodological Approach for AER Sub-activity: Apartment Building Perspective

The apartment building perspective will measure the impact of the sub-activity on the participating apartment buildings, by measuring the change in their total annual energy expenditures as a result of

[41] In the final year of data collection, incremental energy saved (both electric and off-grid) will be assumed to continue for the remainder of the period of analysis for the CEA. For example, if the period of analysis is 10 years, the incremental energy saved measured in the final year of data collection (Year 3) will be expected to continue for the following 7 years of the period of analysis. If any qualitative information suggest that this is an inaccurate assumption, we will make adjustments and note this deviation in the report.

the introduction of energy retrofits. This CEA will only estimate the benefits for participating apartment buildings and will not make assumptions about scaling this activity to other buildings.

The income benefit from reduced electricity use will be the difference between apartment electricity use after the retrofit (at endline) and before the retrofit (at baseline). Financial benefits will be calculated starting with a measure of reduced electricity use:

$$\textit{Incremental Apartment Building Electricity Savings} = \textit{Electricity Use}_{\textit{baseline}} - \textit{Electricity Use}_{\textit{after retrofit}}$$

where,

$\textit{Electricity Use}_{\textit{baseline}}$ = *Apartment building electricity consumption (kWh) measured during the baseline period (before the retrofit)*

$\textit{Electricity Use}_{\textit{after retrofit}}$ = *Apartment building electricity consumption (kWh) (after the retrofit)*

Valuing the reduced apartment electricity use will be done using the tariff rate charged to households:

$$\textit{Electricity Cost Savings}_{\textit{apartment}} = \textit{Residential Electricity Tariff} \times \textit{Incremental Apartment Building Electricity Savings}$$

Additionally, apartment buildings are expected to receive some kind of grant from the IC. This will be treated as a cash inflow (or financial benefit) from the CEA perspective. **Exhibit 20** and **Exhibit 21** detail the benefits and costs.

Exhibit 20. CEA Benefits for AER Sub-activity: Apartment Level

Benefits	Data Source
Electricity cost savings	▪ **Quantity of electricity consumed** (kWh/apartment) will be collected monthly from meter data, daytime and nighttime, provided by GFA Consulting Group. ▪ **Price per unit of electricity** (€/kWh) will be measured by the day/night tariff rate charged to households (provided by the utility company).
Grants	▪ The IC will provide information about the grants it has supported or promoted at the apartment building level, estimated on an average basis per participating apartment (measured in € per apartment building).
Energy efficiency financial contributions to apartment building	▪ The endline survey will ask households about their direct contributions towards the energy retrofits chosen and installed by the apartment buildings. These contributions will be summed across all participating households to account for an average total contribution by all households in each participating apartment building (measured in € per apartment building).

The model will also include costs incurred by the apartments for all energy efficiency retrofit investments, including equipment and installation costs. All these costs will be measured from the endline survey data and estimated on an average basis per participating apartment building. Grant support data will be provided by the IC.

Exhibit 21. CEA Costs for AER Sub-activity: Apartment Level

Costs	Data Source
Energy efficiency equipment costs	▪ The baseline survey will ask each apartment building about their direct expenses associated with their energy retrofits. Equipment costs will be averaged across all participating apartment buildings (measured in € per apartment building).

Costs	Data Source
Installation costs	▪ The baseline survey will ask each apartment building about any installation costs associated with the selected energy retrofit. These costs will be averaged across all participating apartment buildings (measured in € per apartment building).

The cost-effectiveness ratio (CER) will result in the amount of savings per dollar invested at the apartment building level, using the following equation:

$$CER_{\,apartment} = \frac{Electricity\ Cost\ Savings_{\,household} * households\ per\ apartment\ building + Grant + \ Energy\ Efficiency\ Financial\ Contributions}{(Energy\ Efficiency\ Equipment\ Costs + Installation\ Costs)_{\,apartment}}$$

3. Methodological Approach for AER Sub-activity: MFK/IC Perspective

The MFK/IC financial perspective will include all costs associated with administering the AER grants to apartments, since no cash inflows/benefits are expected for the MFK/IC (**Exhibit 22**).

$$Costs_{\,MFK} = \frac{Grant\ costs\ +\ Other\ implementation\ costs\ +MFK\ administration}{Apartments\ participating\ in\ AER\ activity}$$

Exhibit 22. CEA Costs for AER Sub-activity: MFK/IC perspective

Costs	Approach
Grant costs	▪ IC will provide information about the cost for providing grants to apartments. This data will be estimated on an average basis per apartment (measured in €).
Other implementation costs	▪ IC will provide information about all direct AER program costs that are relevant to the AER sub-activity (measured in € per apartment).
MFK administration, monitoring and evaluation costs	▪ Typically, these costs, accruing to MFK, are assumed to be 20 percent of the total direct program implementation costs in MCC economic rate of return analyses. This convention will be used for this analysis (measured in € per household).

4. Methodological Approach for AER Sub-activity: Utility Perspective

Because the utility is not expected to incur any incremental costs as a result of the AER sub-activity, this perspective will only measure incremental benefits (or cash inflows) from KEDS' perspective. These benefits are included in **Exhibit 23**.

Exhibit 23. CEA Benefits for the AER Sub-activity: Utility Perspective

Benefits	Approach
Incremental energy supply cost savings (electricity only)	▪ **Incremental quantity of energy consumed per apartment** (kWh/apartment) will be collected monthly from meter data provided by KEDS. This will be converted to annual equivalents. ▪ **Levelized cost to supply imported energy** (€ per kWh) will be provided by literature and verified with KEDS, to ensure that the incremental quantity of electricity saved for AER apartments displaces the need for additional energy, valued at the cost of importing energy.

This benefit will be calculated as outlined below:

$$Benefit_{\,KEDS} = Incremental\ Energy\ Consumption\ Savings/Apartment * (Levelized\ Cost\ to\ Supply\ Energy) * Apartments\ participating\ in\ AER$$

5. Methodological Approach for AER Sub-activity: Integrated Perspective

The four perspectives above will be combined into one integrated perspective and CER for each of the retrofits supported in the AER sub-activity at the household level. This calculation will be:

$$CER_{integrated} = \left[\frac{Grant + \frac{Benefit_{KEDS}}{Apartments\ participating\ in\ AER}}{(Energy\ Efficiency\ Equipment\ Costs + Installation\ Costs)_{apartment} + Costs_{MFK}}\right] * households\ per\ apartment\ building + Energy\ Cost\ Savings_{household}$$

Note, this equation does not take into account transfers of money between beneficiaries and focuses instead on total costs and benefits (for example, the contribution that households make to their apartment buildings for the energy retrofit is considered a transfer and therefore, not included in the integrated perspective.) This value will be presented in 2020 € values on a per-household basis, discounted at 10 percent.

4.4.4 Limitations of CEA Approach

The methodology for all three CEA analyses uses a financial perspective, not an economic CEA. Where a financial CEA focuses on actual cash transactions as outlined above, there are several welfare and broader economic improvements that are not captured in this approach. As such, there are several benefits that are not accounted for in this analysis, including:

- **Increased comfort for households that have warmer homes.** With the introduction of energy retrofits, households may benefit from warmer homes while remaining at the same level of energy consumption. While this would not result in a cash inflow or outflow for the household, it could lead to improved welfare for each member of the household.
- **The value of time spent collecting and burning firewood for lighting and heating purposes.** If households reduce their consumption of off-grid energy sources, they will likely experience time savings and the increased convenience associated with collecting or purchasing fewer units of these energy sources. These are nonmonetary benefits for households.
- **Reduced deforestation and associated environmental benefits.** The introduction of energy efficiency measures could play a significant role in reducing the consumption of firewood in Kosovo This would prevent deforestation and possibly provide other environmental benefits, such as reduced pollution or soil erosion. While this is a benefit to the economy, we will not estimate this benefit as part of the financial CEA, which focuses only on the financial transactions from the perspective of the various beneficiaries of this program.

Additionally, there may be costs that are not accounted for in this approach. These costs may include **maintenance of energy retrofits** at the household level, or any incremental costs for households if their energy retrofit needs upgrading or maintaining during the period of analysis.

5. DHM ACTIVITY: EVALUATION DESIGN

In this section, we provide a detailed overview of the mixed-methods design for the evaluation of the DHM activity of RELP. We start by outlining the implementation status of the DHM activity in **Section 5.1**. Next, in **Section 5.2** and **Section 5.3**, we outline the quantitative and qualitative approaches.

5.1 Implementation Status of DHM Activity

Termokos provides district heating service to approximately 12,000 apartment units in 300 apartment buildings in Pristina through distribution of piped hot water heat from a central plant.[42] The IC for the DHM activity is expected to field the baseline technical and household survey in September 2020. Metering of households is expected to start thereafter and continue for 12 months. However, there are implementation delays expected due to COVID-19. We have taken this timeline into consideration while proposing the evaluation design for the DHM activity.[43]

5.2 Quantitative Approach

In this section, we describe the quantitative approach for evaluating the DHM activity. In **Section 5.2.1**, we describe the quantitative methods and in **Section 5.2.2**, we outline the power calculations. **Sections 5.2.3**, **5.2.4**, and **5.2.5** present the sampling strategy, data sources, and analysis plan. Finally, **Section 5.2.6** summarizes the data collection.

5.2.1 Methods

The main outcomes of interest in the quantitative approach of the DHM activity described below will be district heating consumption and expenditure **(EQ4.1[44])**. We will examine additional outcomes, including electricity reliability; knowledge of consumption-based billing (CBB); perceptions of affordability of district heating; investment in electrical appliances; level of satisfaction with the activity; likelihood of sustaining investments; and quality and maintenance of meters and thermal valves. Similar to SEEK, all outcomes will be disaggregated by key socioeconomic characteristics, including gender, age of head of household head, ethnicity, and household's economic well-being.

IMPAQ will use a quantitative impact evaluation and a performance evaluation to assess the impacts of the DHM activity. The impact evaluation (ITS approach) will quantify the causal impact of the DHM activity on district heating consumption, while the performance evaluation (pre-post approach) will provide suggestive information about additional outcomes outlined above. Data collection under the DHM activity will include data collected from two sources: (1) administrative data to be used in the ITS approach, and (2) primary data collected in the field to be used both in the ITS approach and in the pre-post approach. Administrative data will provide us information on the monthly district heating consumption and primary data will provide us information on demographics, household characteristics, and behavioral patterns that will act as important control variables in the ITS analysis and will allow us to understand mechanisms of change. We describe these methods below.

[42] Termokos has 385 substations (200 residential; 185 commercial/institutional).

[43] Based on the "DHM Operations Manual," the IC is working with municipalities and/or consumers to co-finance the metering. As of the drafting of this report, it was unclear what the structure of this financing would be and if there would be a voluntary take-up component to the household-level meters and thermal valves.

[44] *EQ4.1: Did the DHM activity result in a change in energy consumption in the intervention areas?*

1. ITS Approach

Similar to the HER and AER sub-activities, we propose to estimate the causal impact of the DHM activity on district heating consumption and expenditure, using an ITS approach. As per IMPAQ's understanding, under the DHM activity, once meters and thermal valves are installed, households will begin receiving bills based on CBB. The ITS approach will estimate the causal impact of the installation of meters and thermal valves (and thereby CBB) by analyzing time-series data before and after CBB is started, and assess to what extent household-level outcomes changed immediately after the initiation of CBB relative to a possible preexisting trend.[45] The ITS analysis will rely on frequent observations before the initiation of CBB is completed to estimate a possible preexisting trend. It assumes that preexisting trends would have continued over time in the absence of CBB. Like in SEEK, when estimating the impacts of CBB using the ITS approach, we will account for month, weekday and time-of-day variations. To account for seasonal effects, we will collect at least 14 months of pre-intervention information and at least 21 months of post-intervention information. This is because Termokos only bills households for 7 months each year (October–April). Thus, in the pre-intervention period, we will receive data for the heating seasons of 2018/19 and 2019/20, or 14 months, and in the post-intervention period, we will receive data for 2021/22, 2022/23, and 2023/24, or 21 months. Finally, to account for heterogeneity between households, we will also use a household's demographic and socio-economic characteristics obtained from the primary quantitative baseline and endline surveys as control variables in the ITS approach.

We will use the ITS approach to estimate impacts for the main outcome variables of interest: district heating consumption and expenditure, as reported by Termokos. For the time period before the meters are installed, based on IMPAQ's discussions with Termokos in June 2020, Termokos has billing data available both, at an aggregate sub-station level (which can include many apartment buildings) and also for each individual apartment unit. Termokos has in principle agreed to share the data at the apartment unit level and IMPAQ is in the process of executing a data sharing agreement with Termokos.[46]

The IC for this activity is scheduled to conduct a baseline survey in September 2020, with approximately 30 percent of the beneficiary sample for the DHM activity. That is, approximately 3,600 households of a total 12,000 beneficiaries will be surveyed. Households surveyed by the IC through this baseline survey will also form our main sample for the ITS approach. To facilitate data access, we will ask households during the baseline survey to provide their Termokos bill number and consent for data access.

2. Pre-post Approach

Similar to **Section 4**, we will complement the ITS approach with a pre-post longitudinal analysis of the same households, to explore additional outcomes of interest such as consumption and expenditure on other sources of energy, substitution between energy sources, satisfaction with meters, and changes in the comfort level of homes. While the ITS approach will only use administrative data from Termokos, the pre-post approach will utilize primary survey data collected at baseline and endline. Specifically, IMPAQ

[45] Based on the "DHM Operations Manual," it is expected that all beneficiary households will start CBB only during the October 2021–April 2022 heating season. Therefore, there will be little variation in the *timing* of households' initiation of CBB.

[46] If Termokos does not agree to provide data at the apartment unit level for the time period before the meters are installed, then to ensure comparability in the unit of analysis, before and after meters are installed we will extrapolate values for the individual apartments. That is, our main unit of analysis at all points in time will the apartment unit. In the event that in the pre-period Termokos only provides aggregate monthly consumption for the whole apartment building, based on the number of apartments units in each apartment building (as quoted by Termokos), we will assign values to individual apartment units. This will ensure that the unit of analysis will stay the same in the pre- and post-periods.

will conduct an endline survey with the same set of households with whom the IC will conduct the baseline survey. The endline survey will be conducted in September–October 2023, at which point all households are expected to be metered. We will compare outcomes measured at endline, September–October 2023, with those measured at baseline, to provide information about the evolution of household-level outcomes. Findings from the pre-post analysis, while not causal (since factors unrelated to the DHM activity could be driving the results), are still critical to help us understand the underlying mechanisms of change driving district heating consumption changes identified in the ITS approach.

5.2.2 Statistical Power

To determine the sample size required for this study, we computed the minimum detectable effects for one key outcome: monthly district heating consumption. As outlined above, the IC will survey approximately 3,600 households at baseline. We plan to survey the same households at endline. We assume that at baseline and endline, we will have a completed survey from a total of 2,520 households, for a 70 percent response rate.[47] That is, we will be able to obtain Termokos bill numbers and link them to administrative data for the ITS analysis for only these 2,520 households.

We estimate minimum detectable effects for the key outcome variable, monthly district heating consumption. As the baseline value of this outcome, we will use estimates of district heating consumption from Termokos's data for the October 2018–April 2019 heating season. Based on this data, the average consumption for a household in Pristina was approximately 300 kWh per month.[48] However, we do not have estimates for the standard deviation of district heating consumption. We present results using 4 ratios of standard deviation to mean encompassing a broad range—0.5, 1, 2, and 3. As outlined above, we assume that through administrative data from Termokos, we will obtain 14 months of data before the meter and thermal valve installation (i.e. initiation of CBB) and 21 months of data after the installation/initiation of CBB, and that the autocorrelation of outcomes is 0.2. Using these parameters, we provide our minimum detectable effect estimates in **Exhibit 24**.

Assuming that of the 3,600 surveys attempted, we obtain 2,520 completed surveys (and their associated Termokos bill numbers)**,** we would be able to detect a decline in *monthly district heating consumption* by about 1 percent to 6 percent, depending on the coefficient of variation. As shown, the minimum detectable effects increase if the coefficient of variation (ratio of standard deviation to mean) is high. However, it drops as the standard deviation goes down relative to the mean. Our calculations suggest that, if the coefficient of variation is 0.5, the autocorrelation is 0.2, and we can get 35 time periods (14 pre-intervention rounds and 21 post-intervention rounds), we could estimate impacts as small as a 0.83 percent change in district heating consumption applying the ITS approach.

[47] Similar to SEEK, we assume a conservative response rate of 70 percent, since based on anecdotal evidence and the baseline survey conducted by the IC for SEEK, there is considerable survey fatigue among households in Kosovo.

[48] Average consumption in Pristina across all substations (1 substation roughly corresponds to 1-2 apartment buildings) for the heating season of October 2018–April 2019 was 100 MW per month, according to Termokos. The average number of apartments in an apartment building is 30, implying that average consumption per apartment unit is around 3 MW (300 kWh) per month for the heating season.

Exhibit 24. Proposed Sample Sizes for DHM Activity

Treatment Effect	COV	Minimum Detectable Effect	% Change from Baseline	Number of Attempted Surveys	Projected Number of Completed Surveys	Time Periods
ITS Analysis						
Impact of CBB on district heating consumption	0.5	2.5 kWh	0.83%	3,600 households	2,520 households	35 months (pre: 14; post (21))
	1	5 kWh	1.67%			
	2	9.5 kWh	3.17%			
	3	15 kWh	5%			

Notes:
COV = coefficient of variation.
The main outcome variable for the power calculations is district heating consumption with a baseline mean of 300 kWh per month. We assumed a confidence level of 95 percent, two-tailed tests, 80 percent power, 30 percent sample attrition for the direct beneficiaries, and an autocorrelation of 0.20. We assumed that for the ITS analysis, time periods will be split in the following way: pre-period: pre-period: 14 months; and post-period; 21 months.

Similar to SEEK, we also calculate the minimum threshold of beneficiaries that need to be reached for a viable design effect for the DHM activity. That is, in the event that the IC is unable to reach 3,600 households during baseline data collection, what would be the minimum number of beneficiaries (and corresponding minimum number of completed surveys) that would need to be reached to determine program impacts? As a note, this is a smaller concern for the DHM activity, since there is a larger pool of beneficiaries to choose from i.e. 12,000 beneficiaries. We present the minimum threshold number of surveys in **Exhibit 25**. Assuming a coefficient of variation of 0.5 and with all the other assumptions remaining similar to before, if the IC only surveys 1,000 households and the corresponding number of completed surveys is 700 (assuming 70 percent response rate), then we would be able to detect a 1.28 percent decline in district heating consumption. On the other hand, if 2,500 households are surveyed (corresponding number of completed surveys is 1,750) then we can detect a change as small as 0.8 percent decline in district heating consumption.

Exhibit 25. Minimum Threshold Number of Surveys for DHM Activity

Outcome	Baseline Mean	Coefficient of Variation	Minimum Detectable Effect	% Change from Baseline	Number of Attempted Surveys	Projected Number of Completed Surveys	Time Periods (in terms of months)
Monthly District Heating Consumption (kWh)	300 kWh *(household level)*	0.5	4 kWh	1.33%	1,000 households	700 households	35 (pre: 14; post: 21)
			3.5 kWh	1.17%	1,500 households	1,050 households	
			3 kWh	1%	2,000 households	1,400 households	
			2.5 kWh	0.83%	2,500 households	1,750 households	

Note: We assumed a coefficient of variation of 0.5, confidence level of 95 percent, two-tailed tests, 80 percent power, 30 percent sample attrition for the direct beneficiaries, and an autocorrelation of 0.20. We assumed that for the ITS analysis, the period will be split in the following way: pre-period: minimum 14 months; and post-period: minimum 21 months.

5.2.3 Sampling Strategy

As described above, at baseline, the IC will select 3,600 households for the baseline survey. This entire sample will be included in our study sample. If at endline, there is significant attrition among this sample, we will select a "replacement" sample, drawn from the beneficiary list for the DHM activity. This will turn our sample into a repeated cross-sectional sample, instead of the proposed longitudinal study. In selecting this replacement sample at endline, we will attempt to identify a sample that is closely matched with our baseline sample on characteristics including gender of head of household, approximate physical area of the house, and household size, among others.

5.2.4 Data Sources, Outcomes, and Anticipated Timing of Data Collection

The main quantitative data sources for the DHM activity will include the following:

Household survey (responsibility of IC and evaluator). We will use data collected by in-person household surveys at two points in time: baseline and endline. Primary survey data will serve two purposes: (1) to gather information on demographic and socio-economic characteristics of the household which will serve as control variables in the ITS analysis; and (2) to help understand the mechanisms of change and the driving force behind any changes to district heating consumption. The IC will be responsible for baseline data collection and IMPAQ will be responsible for endline data collection. The main respondent of this survey will be household heads. The goal of the baseline data collection will be to obtain pre-intervention values of our outcome variables including consumption of energy from other sources and information about demographic characteristics, and physical characteristics of the apartment. The goal of the endline data collection is to collect information similar to the baseline survey and also collect additional information about the experiences with CBB, new tariffs, and perceptions about its effectiveness.

Administrative data (responsibility of evaluator). The second data source will be administrative data collected for beneficiary households sampled under the IC's baseline survey. The goal of collecting administrative data will be to measure our main outcome variables (i.e. district heating consumption and expenditure) on a continuous basis. We will obtain administrative data from both Termokos and KEDS, since we expect that the households on the district heating network will also have a KEDS connection. The administrative data from Termokos will provide information about the monthly district heating consumption, and administrative data from KEDS will provide monthly data about electricity consumption, if any. IMPAQ will be responsible for coordinating with Termokos and KEDS to obtain necessary administrative data. During the baseline in-person surveys, we will ask households to provide their Termokos invoice number and their KEDS invoice number and will also obtain consent to access their data. We will initiate a data-sharing agreement with Termokos and KEDS to transmit data to IMPAQ on a monthly basis in addition to historical data as requested by IMPAQ.

The Termokos administrative data will be transmitted on a monthly basis and will be at the individual apartment level in both the pre- and post-periods. The data will potentially include the area of the household and the heating demand for the household in kilowatt hours monthly, from October to April of every year (i.e. the heating season when Termokos switches on its heating). In IMPAQ's initial discussions with Termokos in December 2019 – January 2020, Termokos indicated that in the pre-period only data at the substation/apartment building level would be available. However, subsequently in meetings held in March and June 2020, Termokos mentioned that they have access to individual

apartment level data for the pre-period and will seek permission to share that as well. IMPAQ has shared a data sharing agreement with Termokos and is awaiting approval by Termokos to share data at the individual apartment unit level. Data collected by Termokos is the main source of household level district heating consumption for Pristina and is thus reliable. In discussions with Termokos in June 2020, they have in principle agreed to share the personally identified data for each beneficiary of the DHM activity if consent is obtained from beneficiaries. IMPAQ has shared an official (MCC-approved) data sharing agreement with Termokos subsequently and this agreement is expected to be executed in the following months accounting for COVID-19 related administrative delays.

Monitoring data (responsibility of MFK and IC). We will use monitoring data from MFK and the IC for the DHM activity to understand the progress made toward implementation. We will track the "process" indicators in the RELP monitoring and evaluation plan.

Exhibit 26 provides details about the outcomes and timing of data collection for each data source.

Exhibit 26. Summary of Quantitative Data Collection for DHM Activity

Data Source	Tentative Timing of Data Collection	Outcomes by Data Collection Round
Household survey: ▪ 3,600 households	**Baseline** ▪ September 2020 **Endline** ▪ September-October 2023	**Baseline** ▪ Outages ▪ Energy use and expenditure from all sources ▪ Perceptions of affordability of district heating and level of current bill ▪ Current levels of comfort during winter ▪ Current tariff levels ▪ Investment in electrical appliances ▪ Household, housing unit, and apartment building characteristics ▪ Baseline level of insulation of home **Additional outcomes at endline** ▪ BC&O campaign ▪ Metering decisions ▪ Level of satisfaction with different aspects of the activity ▪ Likelihood of sustaining investments ▪ Quality and maintenance of meters and thermal valves ▪ Shifts in energy efficiency usage
Administrative data	At least 5 heating seasons (2018/19; 2019/20; 2021/22, 2022/23, and 2023/24)	**Monthly Basis (Baseline, midline, endline)** ▪ Consumption and expenditure from district heating *(from Termokos)* ▪ Electricity consumption and expenditure *(from KEDS, if applicable)*
Monitoring data	Throughout program lifetime	RELP monitoring and evaluation plan process indicators

5.2.5 Analysis Plan

1. ITS Approach Analysis

We will estimate the causal impact of the DHM activity by using Termokos data. Our analysis for this activity is similar to the one outlined for the HER and AER sub-activities in **Section 4.** We will account for confounding factors such as temperature, season-specific fixed effects, and month-of-year fixed effects, in addition to controlling for baseline characteristics obtained from the baseline survey—including household size, physical area, household income or asset index, ethnicity, energy consumption patterns, and distance from city center of Pristina—as control variables in our ITS specifications.[49]

2. Pre-post Approach Analysis

We will complement the ITS analysis with the pre-post analysis, and we will estimate changes in other outcomes (including consumption of different energy sources and the substitution between them) over time. We will use data from only two points in time, baseline (June 2020) and endline (September–October 2023). We will analyze data from the pre-post approach in a fashion similar to the one described in **Section 4.** Although our primary analysis will be based on the whole sample (all households), we will also conduct subgroup analyses by: (a) gender of the head of household, (b) ethnicity of the head of household, (c) age of household head, and (d) vulnerable or poor households, as defined by income.

5.2.6 Data Collection, Data Processing, and Data Quality

All the data processing and data quality checks outlined for the household surveys outlined in **Section 4.2.6** will also be followed for primary survey data collection for the DHM activity. Similarly, to acquire the administrative data from Termokos and KEDS, we will execute a data sharing agreement with Termokos and KEDS to transmit data to IMPAQ on a periodic basis. We will closely monitor the incoming Termokos data and administer similar data quality checks as described above for the household surveys.

5.3 Qualitative Approach

IMPAQ's qualitative approach will complement the quantitative methods described above for the proposed performance and impact evaluation of the DHM activity. We will use a variety of methods including FGDs, KIIs, and document reviews to examine implementation, lessons learned, and sustainability of the DHM activity. In **Section 5.3.1**, we describe the main methods for our qualitative approach, and in **Section 5.3.2**, we define the outcomes and their expected timing. **Section 5.3.3** provides an overview of the data collection, and **Section 5.3.4** summarizes the sample and rounds of data collection for the DHM activity.

5.3.1 Methods

We will use **FGDs** to get the perceptions and experiences of district heating customers about the transition toward a CBB system, as part of the DHM activity. We will include households with a mix of characteristics to develop insights into the contributing factors towards behavior change. To holistically assess the DHM activity, we will also hold **KIIs and small group interviews** with representatives from Termokos, the ICs, ERO, MFK, and any other associated stakeholders. The goal will be to assess and

[49] While the baseline and endline survey instruments will inquire about household income ranges and building age, there remains a concern that households will not report income accurately and might not know how old their buildings are. Thus, distance to city center, will serve as proxies for household economic well-being and state of buildings in Pristina.

contextualize program activities, including those related to preparing institutional, organizational, and regulatory measures; setting end-user tariffs; and designing new billing and collection procedures. Finally, we will conduct a thorough **document review** of work plans, timelines, schedules, annual and quarterly progress reports, and monitoring and evaluation plans, to assess the design and implementation status of the DHM activity. As with the SEEK activity, we will also focus on evaluating the gender components of this activity. We will examine the extent to which the program has incorporated gender mainstreaming in the implementation of the DHM activity, as defined in the "DHM Inception Report." Specifically, we will assess the program's technical assistance to ERO on strengthening gender mainstreaming and the level of awareness among implementation teams and program beneficiaries.

5.3.2 Outcomes

We will track the progress, successes, and challenges related to implementation **(EQ1**[50]**)**, including willingness to accept meters/thermostatic valves, the timeliness and quality of installation, BC&O campaigns, and the transition to CBB.

We will also examine the contribution of various components of the program (technical assistance to key institutions, BC&O activities, installation of energy efficient equipment, and implementation of CBB) toward any changes in district heating consumption **(EQ4.2**[51]**)**. To explore the causal mechanisms behind any outcomes and the theory of change, we will ask beneficiaries about their knowledge, attitudes, and behaviors regarding energy efficiency, usage, and expenditure. At endline, we will probe to understand the type of energy behavior before CBB (including being able to control one's own heat usage) and current behavior. This information will give us greater insight when we interpret the findings from the quantitative analyses. If the quantitative findings reveal that the DHM activity was successful in reducing district heating consumption, we will focus on probing which specific components helped district heating consumers achieve these reductions. For instance, it is possible that there may not be an increase in awareness about the benefits of energy efficiency, but district heating consumption goes down to compensate for higher tariffs and subsequent increases in monthly bills. However, it may be that CBB will not lead to an increase in consumer bills in a meaningful way, but rather the thermostatic valves allow customers to lower the ambient temperatures more comfortably and reliably than by opening windows.

Finally, to assess sustainability **(EQ7**[52]**)**, we will collect qualitative data about the plans for sustainability through KIIs with implementers at the end of the threshold program and with district heating customers (households and businesses), Termokos, and ERO at the end of our evaluation period. We will triangulate the sustainability plans developed during the threshold period with results through the end of the evaluation period to see which elements were and were not sustainable. Through FGDs with district heating customers, particularly those metered at the beginning, we can examine whether and how attitudes and behaviors related to district heating use and expenditure change over multiple years. Through KIIs with Termokos and ERO, we can learn about the financial stability of Termokos, whether

[50] *EQ1: Were the activities implemented as designed and were the outputs outlined in the project logic achieved?*
[51] *EQ4.2: What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and implementation of consumption-based billing)?*
[52] *EQ7: Did the program meet the stated objective of reducing the gap between supply and demand of electricity? How sustainable are critical outcomes of the program, and why?*

the tariffs were calculated appropriately, and whether plans for network expansion in Pristina were realized **(EQ4[53])**.

5.3.3 Data Collection

As with the HER and AER sub-activities, we will collect five rounds of qualitative data: baseline (round 1), interim (rounds 2, 3, and 4), and endline (round 5), described in more detail below. The first three rounds of data collection will focus on implementation and short-term outcomes, as the timing is concentrated during the threshold period and the first heating season post- meter and thermostatic valve installation. The last two rounds of data collection will occur later, and will focus on the long-term outcomes and sustainability of the DHM activity, including providing contextualization for the findings from the ITS analysis and the endline survey. See **Exhibit 27** for a breakdown of the sample, timing, and outcomes covered in each round.

Baseline (Round 1). We propose collecting baseline data during the winter season of the first year of implementation, as this will allow us to interview households and businesses with installed meters and thermostatic valves and that should have experienced at least a few months of CBB and/or BC&O. At baseline, we will also hold FGDs with district heating customers and businesses who have not yet been metered. This can provide preliminary insights into acceptance of installation and consumption and expenditure, as we will ask households to program their anticipated behavior in the future. We will interview MFK and the ICs to learn about implementation progress, including tariff design and BC&O. We will interview Termokos, ERO representatives, and other stakeholders if appropriate to provide additional context on tariff design, engagement with MFK and the ICs, an alignment with government and business priorities.

Interim (Round 2). To provide additional information related to successes, challenges, and short-and long-term effects after the program has been implemented, interim monitoring data will be collected from both beneficiaries and implementers. As with SEEK, will conduct interviews with MFK, the ICs, and contractors immediately after completion of the implementation period, while staff from implementing agencies are still available. DHM beneficiaries can provide further information related to implementation, and at this stage, we can fully answer **EQ1**[54].

Interim (Round 3). Our third round of data collection will occur after the first heating season. At this point from the ITS analysis, we will be able to ascertain whether or not there was a decrease in energy consumption and expenditure immediately after the installation of meters and thermostatic valves. Our qualitative data collection will explore the mechanisms of any change, including the effectiveness of BC&O.

Interim (Round 4). Our fourth round of data collection will occur one year later, after the second heating season. The RELP theory of change posits that installing meters and thermostatic valves along with the BC&O initiatives – that is, paying for energy used and being able to control the temperature within an individual household or business – will lead to reduced energy consumption. Again, from the ITS analysis,

[53] *EQ4: Did the DHM activity result in a change in energy consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and implementation of consumption-based billing)? How many new households shifted to district heating as a result of the DHM activity?*

[54] *EQ1: Were the activities implemented as designed and were the outputs outlined in the project logic achieved?*

we will know the pattern of energy consumption. This interim data will shed light on the longer-term effectiveness of the BC&O, as well as any shifts in energy usage.

Endline (Round 5). For both SEEK and DHM, we will time our qualitative endline data collection *after* the endline survey data collection. Doing so will allow us to tailor our questions to follow up on the results from this analysis, including any unexpected or unusual findings.

Data Processing and Data Quality

Our qualitative data collection process for the DHM will be similar to the process for the SEEK. We will develop FGD and KII guides based on the topics outlined in **Exhibit 27**. All guides will be developed in English, then translated by a member of our team into Albanian. We will hire local facilitators and translators to assist where needed. With the consent of participants, KIIs and FGDs will be recorded. We will develop a coding plan tied to the evaluation questions and any key topics that emerge during data collection. After applying the codes to the FGD and KII transcriptions, we will analyze the data and identify themes that emerge, using NVivo or a similar qualitative data analysis software. We will analyze the data by respondent role and type, including by gender. For additional details about our qualitative data collection process, see **Section 4**.

5.3.4 Sample

For the DHM baseline data collection, we will hold one KII with a business and three FGDs with households that have installed meters and thermostatic valves. We will also hold one KII with a business and three FGDs with district heating customers that have yet to install meters and valves. We will ensure representation across demographic and other relevant characteristics, including holding a women-only FGD to get the perspective of female-headed households or household members. For the subsequent rounds of data collection, including the endline, we will purposively select respondents based on variation in outcomes from the quantitative analysis (both the ITS and the survey), specifically differences in district heating consumption over multiple seasons (customers with increased and decreased energy expenditure). Again, we will consider representation in gender, income level, ethnicity, and other relevant characteristics. We will have approximately 6-8 participants in each FGD. During all round of data collection, we will interview a small number of businesses that use district heating, and any other relevant stakeholders identified in conjunction with MCC and MFK (such as municipality leadership, if the municipality co-finances installation).

Exhibit 27. Summary of Qualitative Data Collection for DHM Activity

Data Source/ Respondent	Evaluation Questions and Outcomes
Baseline (Round 1) – During DHM Implementation ***Exposure Period:*** After the first 3-4 months of implementation, where some district heating customers have installed meters and thermostatic valves and some have not. ***Tentative Date: January/February 2021***	
District Heating Customers – Households: 3 FGDs *(HH who completed installation)* 3 FGDs *(HH who have not completed installation)*	**EQ1:** Were the activities implemented as designed and were the outputs outlined in the project logic achieved? **Outcomes:** ▪ Quality and appropriateness of design ▪ Project milestones and benchmarks; delays and challenges in implementation ▪ Implementation and progress of activities

Data Source/ Respondent	Evaluation Questions and Outcomes
District Heating Customers – Businesses: 1 KII or small group interview *(business who has completed installation)* 1 KII or small group interview *(business who has not completed installation)* **MFK:** 4-6 KIIs *(Energy Director, Energy Specialist, Monitoring and Evaluation Specialist, other relevant staff)* **ICs:** 4-6 KIIs *(Decon International, iC consulenten, CES Clean Energy Solutions, Optima Energy Consulting, Installation Contractor)* **Termokos:** 1-2 KIIs or small group interviews **ERO:** 1-2 KIIs or small group interviews *(Director of District Heating Department, Thermal Energy Analyst)* **Ancillary Stakeholders:** 1-2 KIIs *(such as Municipality of Pristina and others as appropriate)*	▪ Operational processes and systems; use of data management systems and communication procedures ▪ Design of BC&O campaigns ▪ Incorporation of gender-specific components ▪ Engagement with MFK/ICs ▪ Experience with training/TA, financing options and applicants ▪ Perceptions of installation quality ▪ Tariff design for CBB ▪ Regulatory framework ▪ Alignment with government priorities ▪ Perceptions of district heating quality, reliability, comfort levels, and affordability ▪ Perceptions of and concerns regarding CBB **EQ4:** Did the DHM activity result in a change in energy consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and implementation of consumption-based billing)? How many new households shifted to district heating as a result of the DHM activity? **Outcomes:** ▪ Incentives, information, and guidance needed to accept meter and thermostatic valve installation ▪ Incentives, information, and guidance needed to make changes in energy consumption ▪ Awareness of energy saving measures and benefits ▪ Willingness to pay for district heating investments (thermostatic valves) ▪ Projected energy consumption and energy cost savings
Interim (Round 2) – End of Threshold ***Exposure Period:*** After all DHM activities are completed and the threshold program in Kosovo concludes ***Tentative Date: October/November 2021***	
District Heating Customers – Households: 1-2 FGDs **District Heating Customers – Businesses:** 1-2 KIIs or small group interviews **MFK:** 4-6 KIIs *(Energy Director, Energy Specialist, Monitoring and Evaluation Specialist, other relevant staff)* **ICs:** 4-6 KIIs	**EQ1:** Were the activities implemented as designed and were the outputs outlined in the project logic achieved? **Outcomes:** ▪ Project milestones and benchmarks; delays and challenges in implementation ▪ External factors and events affecting implementation ▪ Design of BC&O campaigns ▪ Incorporation of gender-specific components ▪ Engagement with MFK/ICs ▪ Experience with training/TA, financing options and applicants ▪ Perceptions of installation quality ▪ Tariff design for CBB ▪ Regulatory framework ▪ Alignment with government priorities ▪ Perceptions of district heating quality, reliability, comfort levels, and affordability

Data Source/ Respondent	Evaluation Questions and Outcomes
(Decon International, iC consulenten, CES Clean Energy Solutions, Optima Energy Consulting, Installation Contractor) **Termokos:** 1-2 KIIs or small group interviews **ERO:** 1-2 KIIs or small group interviews *(Director of District Heating Department, Thermal Energy Analyst)* **Ancillary Stakeholders:** 1-2 KIIs *(as appropriate)*	▪ Perceptions of and concerns regarding CBB ▪ Best practices and lessons learned ▪ Perceived program impacts ▪ Sustainability of program outcomes **EQ4:** Did the DHM activity result in a change in energy consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and implementation of consumption-based billing)? How many new households shifted to district heating as a result of the DHM activity? **Outcomes:** ▪ Incentives, information, and guidance needed to accept meter and thermostatic valve installation ▪ Incentives, information, and guidance needed to make changes in energy consumption ▪ Effectiveness of BC&O ▪ Awareness of energy saving measures and benefits ▪ Willingness to pay for district heating investments (thermostatic valves) ▪ Projected energy consumption and energy cost savings ▪ Shifts in energy usage ▪ Successes and challenges
Interim (Round 3) – End of First Heating Season ***Exposure Period:*** After participants have experienced one full heating season after consumption-based billing ***Tentative Date: April/May 2022*** ***Interim (Round 4) – End of Second Heating Season*** ***Exposure Period:*** After participants have experienced two full heating seasons after consumption-based billing ***Tentative Date: April/May 2023***	
District Heating Customers – Households: 1-2 FGDs **District Heating Customers – Businesses:** 1-2 KIIs or small group interview **Termokos:** 1-2 KIIs or small group interviews **ERO:** 1-2 KIIs or small group interviews **Ancillary Stakeholders:** 1-2 KIIs *(as appropriate)*	**EQ4:** Did the DHM activity result in a change in energy consumption in the intervention areas? What was the contribution of various components of the program toward any changes in energy consumption (technical assistance to key institutions, BC&O activities, and implementation of consumption-based billing)? How many new households shifted to district heating as a result of the DHM activity? **Outcomes:** ▪ Contextual information to explain results from the ITS analysis in the short-term (one heating season) and the long-term (two heating seasons) ▪ Experience with and perceptions of energy efficiency opportunities; behavior change ▪ External factors and events affecting energy use ▪ Effectiveness of BC&O over multiple heating seasons ▪ Changes in energy efficiency ▪ Shifts in energy usage (sustained or drop-off over time) ▪ Realization of energy savings ▪ Changes in consumption uses; new purchases; changes in other household investments ▪ Successes and challenges
Endline (Round 5) – End of Evaluation Period/Post- Quantitative Data Collection	

Data Source/ Respondent	Evaluation Questions and Outcomes
Exposure Period: After the endline survey has been administered and analyzed, in the middle of the third heating season after switch to consumption-based billing ***Tentative Date: January/February 2024***	
District Heating Customers – Households: 6 FGDs **District Heating Customers – Businesses:** 1-2 KIIs or small group interviews **Termokos:** 1-2 KIIs or small group interviews **ERO:** 1-2 KIIs or small group interviews **Ancillary Stakeholders:** 1-2 KIIs *(as appropriate)*	**EQ7:** Did the program meet the stated objective of reducing the gap between supply and demand of electricity? How sustainable are critical outcomes of the program, and why? **Outcomes:** ▪ Contextual information to explain endline survey findings ▪ Contextual information to explain results from the ITS analysis over 2.5 heating seasons ▪ External factors and events affecting energy use ▪ Effectiveness of BC&O over multiple heating seasons ▪ Changes in energy efficiency ▪ Shifts in energy usage (sustained or drop-off over time) ▪ Realization of energy savings ▪ Changes in consumption uses; new purchases; changes in other household investments ▪ Successes and challenges ▪ Perceived program impacts ▪ Best practices and lessons learned

6. IPP ACTIVITY: EVALUATION DESIGN

IMPAQ's approach to evaluate the IPP activity will be to track and analyze the entire implementation process of this activity, starting from facilitation of funding to the IPPs, to the eventual launch of the IPP projects. Since this activity aims to work with approximately five IPPs, our approach will be primarily qualitative. **Section 6.1** describes the implementation status of the IPP activity, and **Section 6.2** details the evaluation methodology.

6.1 Implementation Status of IPP Activity

The IPP activity has two complementary components. The first is to support the KCGF through technical assistance and direct funding, to expand its mandate to include offering specific guarantees for renewable energy. The second is to develop a standardized framework for financing renewable energy programs and to provide complementary technical assistance to targeted IPPs who have already received pre-authorization from the ERO. Technical assistance will also be provided to banks to make lending for renewable energy an attractive venture. Based on IMPAQ's understanding, the IPP activity has also been expanded to support businesses in Kosovo in purchasing solar panels to reduce or eliminate their reliance on the national grid.[55]

In February/March 2020, MFK procured the services of the project management consultant, who oversees and coordinates the technical capacity-building services of KCGF, partner institutions, and IPP developers. The project management consultant will also assist in raising and deploying needed capital to support the renewable energy guarantee window. As of May 2020, MFK has yet to procure all of the technical assistance consultants. Training and capacity building to KCGF will begin in the spring of 2020 and will be ongoing through the end of the threshold period in September 2021. The renewable energy IPP guarantee window is slated to open in the second quarter of 2021, with subsequent submissions to KCGF.

Based on the timeline and information gathered during our scoping trip and subsequent meetings, the timeline is not yet clear for when IPPs are expected to reach financial close and build their projects/connect to the grid. As per IMPAQ's current understanding, the targeted IPPs have pre-authorization from the ERO and are part of a select group that benefits from feed-in tariffs and a power purchasing agreement with the Kosovo Energy Corporation. They are also subject to a tight timeline dictated by the ERO. This timeline, according to our recent meetings with MFK, might not allow enough time for the IPPs to obtain sufficient capital. The individual IPPs will be working on different timetables, depending on where they are in the authorization process, how quickly they meet the requirements of the newly established framework, and when they obtain financing. As we also expect that this activity could continue beyond the evaluation period, we propose to primarily evaluate the IPP activity using a process-tracing approach, described below.

[55] A recent call with MFK (4/28/2020) revealed that the scope of the IPP activity will be expanded to include financing for large businesses to install solar panels. These panels would not be connected to the grid, but rather would be for private consumption to power day-to-day operations of the individual businesses.

6.2 Evaluation Methodology

In this section, we describe the evaluation methodology for the IPP activity. In Section 6.2.1 and Section 6.2.2, we describe the main evaluation methods and outcomes, respectively. In Section 6.2.3, we discuss the data collection process, and Section 6.2.4 details the sample for the IPP activity.

6.2.1 Methods

The number of beneficiaries targeted and the timeframe of the IPP activity do not allow us to conduct an impact evaluation or examine the end outcomes, so we propose using a process-tracing approach. This approach will help us determine qualitatively whether the main objectives—including the creation of a standardized framework for renewable energy financing and the establishment of a sustainable renewable energy guarantee window—are in the process of being met. This approach will also allow us to assess whether the activity implementation is following the theory of change. A process evaluation tracks cases throughout implementation and focuses specifically on the inputs and activities. Process tracing emphasizes the causal sequencing of activities, and descriptively links factors (including activities) to outcomes (Collier 2011). While we are unable to link specific factors to the ultimate outcome of the IPP activity (increased electricity supply in Kosovo), as this hypothesized outcome will not occur until much later than the evaluation period, we will examine the sequence of events that are related to the intermediate outcomes (e.g., creating the guarantee window, establishing a pipeline of bankable projects, reaching financial close, connecting to the grid) to see whether the theory of change holds. We will also examine the timeline for ERO authorization and lender project financing, to see whether the IPP activity leads to processes that are more efficient. The process-tracing approach will include information from KIIs and site visits, as well as document review, including any project-monitoring data from MFK and KCGF.

We will conduct KIIs with a wide variety of stakeholders at five points during the evaluation. During baseline data collection (round 1), we will conduct site visits to interview IPPs that plan to apply for financing under this activity. We will also hold KIIs with businesses that plan to use the project to buy solar panels for private consumption. Because KCGF and other lending institutions (banks) are critical for program success, we will interview representatives from both regarding their experience with providing and receiving technical assistance. We will also hold KIIs with MFK, the project management committee, and technical assistance providers about program implementation, including delays and challenges. During the scoping mission, we learned that the ERO can be slow to process applications from and grant approvals to IPPs. In addition, uncertainty about whether IPPs can meet the required timeline to benefit from feed-in tariffs and the Kosovo Energy Corporation about purchasing may make this an unattractive venture for potential IPPs and lenders. Because government support from the ERO related to feed-in tariffs and electricity purchasing is crucial for making these projects appealing investments, we will interview representatives from the ERO. With guidance from MFK and MCC, we also will interview other ancillary stakeholders as appropriate, such as representatives from Kosovo Energy Corporation/KEDS, who should eventually buy electricity produced by the IPPs through the government's power purchase agreement.

During our subsequent rounds of data collection, we will include participating IPPs and any IPPs who originally participated but have dropped out. Including nonparticipating IPPs will be useful to get the perspective of those who were unsuccessful in becoming bankable. This will allow us to determine how the renewable energy project in Kosovo is progressing after the completion of capacity-building

activities. This may also provide enough time to examine early outcomes, such as the status of construction undertaken by some IPPs. We will also gather information about the experience of IPPs during the processes of preparing their application and loan approval, and about the time required to receive the final approval of ERO to begin construction. We will also gather insights into external contextual factors, for example, the support provided by ERO and Kosovo Energy Corporation. While the timing of our evaluation is too early to examine energy production, at endline, we will ask KCGF, banks, and IPPs to share their predictions for future activities. For KCGF, we will examine whether the training and technical assistance is sufficient not only for it to raise capital and open the renewable energy credit guarantee window, but also to give it the capacity to leverage other guarantee opportunities. We will also look at the amount of funds raised by KCGF in this time period and whether there is any change in the guaranteed amount and the type of projects receiving credit guarantees. For banks, we will examine whether difficulties in financing renewable energy projects have been alleviated, and whether the status of the regulatory environment is conducive for development of renewable energy projects and for project financing in general.

Finally, the process-tracing approach will incorporate quantitative data gathered from the interview, document review, and program-monitoring data. Progress reports, design documents, financial documents, monitoring and evaluation data, and first-hand accounts from the IPPs will provide key outcomes such as access to and amount of financing, licensing status, and construction and connection status (if applicable and available).

6.2.2 Outcomes

We will track the progress, successes, and challenges related to implementation **(EQ1**[56]**)**, including: the training and capacity building of IPP developers, banks, and KCGF; establishing the Kosovo Project Acceleration Advisory Committee; developing a standardized framework for renewable energy financing; the timing of the ERO licensing and approval process; and opening the renewable energy credit guarantee window. Given the timeframe of our evaluation, it is not possible to answer whether the IPP activity ultimately contributes to increasing the electricity supply in Kosovo. If any IPPs are connected to the grid by the end of the evaluation period, we will measure total electricity generated in megawatt hours (MWh) to estimate the program's impact. We will also explore whether the IPP activity was successful in reducing the perception of financial risk for investors and commercial lenders, both in the energy sector, and for program funding in Kosovo more generally **(EQ5**[57]**)**. Through KIIs with KCGF, banks, IPPs, and the ERO, we can learn whether the barriers identified to financing and constructing renewable energy have been reduced. Finally, we will also look at intermediate outcomes, such as the development of new financial instruments, or if more applications for renewable energy projects are approved, the long-term sustainability of project financing for IPPs in Kosovo **(EQ7**[58]**)**.

[56] *EQ1: Were the activities implemented as designed and were the outputs outlined in the project logic achieved?*

[57] *EQ5: Was the IPP activity successful in reducing the perception of financial risk for investors and commercial lenders in the energy sector, and why? Did the IPP activity contribute to increasing the electricity supply in Kosovo? If so, by how much and why?*

[58] *EQ7: Did the program meet the stated objective of reducing the gap between supply and demand of electricity? How sustainable are critical outcomes of the program, and why?*

If private businesses purchase solar panels for personal consumption, as recently planned through this activity, we can also potentially measure the reduction in their electricity consumption through KEDs data, if feasible.

6.2.3 Data Collection

We will collect five rounds of qualitative data, following the same data collection schedule as SEEK and DHM. Note that while we are using baseline (round 1), interim (rounds 2, 3, and 4), and endline (round 5) to define the rounds of data collection, the progress of each IPP will be different, and therefore each round will track the implementation progress of each IPP, as well as the implementation benchmarks of MFK and KCGF.

Baseline (Round 1). Our first round of in-person data collection will occur approximately 3-4 months into implementation and will provide preliminary insights about whether the activity is on track to build a bankable renewable energy program pipeline. Many of the interviews can be conducted virtually, and if possible, our country expert and field work manager can visit identified IPPs in the pipeline.[59]

Interim (Rounds 2, 3, and 4). We will collect three rounds of interim data, timed to correspond with the SEEK and DHM qualitative data collection: at the end of the threshold period, 6 months post-threshold, and 1.5 years' post-threshold. Data collection will be both qualitative and qualitative in nature. The interim data collection is crucial to assess the ongoing progress of each IPP against the theoretical assumptions of the program We will hold interviews with MFK, the project management consultant, and technical assistance consultants immediately after the threshold period ends, while staff are still available.

Endline (Round 5). Endline for IPPs, KCGF, banks, and other relevant stakeholders (such as the ERO) will occur two and half years post-threshold. With the process-tracing approach, we can evaluate (a) whether the program successfully completed the activities (established the Kosovo Project Acceleration committee, provided training, opened the IPP credit guarantee window); and (b) whether these activities led to intermediate outcomes for the IPPs (licensing, reaching financial close, construction, connection to the grid).

Data Processing and Data Quality

Our qualitative data collection process will be similar to what we propose for the SEEK and DHM activities (in **Sections 4** and **5**). As with our proposed assessments of SEEK and DHM, to ensure that qualitative data provides insights into program implementation, successes, and challenges, we will work closely with the ICs, MCC, and MFK to identify the appropriate respondents from stakeholder groups to participate in semi-structured KIIs or small group interviews during all rounds of data collection

6.2.4 Sample

For the IPPs, our sample will be the entire population of five IPPs who have already received pre-authorization from ERO, and that KCGF and MFK have identified as viable candidates for this activity. We will interview two to three banks targeted by this program. For the subsequent rounds of data collection, we will select the same solar IPPs, this time based on their status. If any IPPs have dropped out of the program (decided to sell their authorization or did not receive approval for financing, licensing, or

[59] If possible, given the progress of the IPP activity to date, particularly given the COVID-19 situation in Kosovo at the time.

construction), we will include them in our data collection, to learn about any weaknesses in the program, including whether the technical assistance provided was sufficient. We will also interview up to two private businesses who have decided to purchase solar panels for private consumption during each round of data collection. **Exhibit 28** summarizes the data sources, timing, sample sizes, and outcomes.

Exhibit 28. Summary of Data Collection for IPP Activity

Data Source/ Respondent	Outcomes
IPPs: 5 KIIs and site visits in each round of data collection *(entire population of IPPs targeted by KCGF/MCC as viable candidates)*	**Qualitative Outcomes:** ▪ Experience with training and technical assistance ▪ Readiness to apply for loans ▪ Experience with financing options ▪ Successes and challenges ▪ Progress toward intermediary outcomes **Quantitative Outcomes:** ▪ Amount of financing ▪ Licensing status ▪ Construction status, connection status ▪ Total MWh produced (if applicable)
Businesses: 1-2 KIIs or small group interviews in each round of data collection *(businesses who have or are planning to install solar panels as part of this activity)*	**Qualitative Outcomes:** ▪ Reasons for purchasing solar panels ▪ Perceptions of electricity quality, reliability, and affordability ▪ Experience with financing and installation ▪ Perceptions of installation quality ▪ Changes in energy expenditures ▪ Actual time required to break even on investment in solar panels ▪ Successes and challenges ▪ Influence on other businesses ▪ Sustainability, such as maintenance required for solar panel upkeep ▪ Additional energy efficient actions taken **Quantitative Outcomes:** ▪ Electricity consumption and expenditure
Only Rounds 1 and 2 **MFK:** 4-6 KIIs **ICs:** 2-4 KIIs *(FMI staff, technical assistance consultants)*	**Qualitative Outcomes:** ▪ Program milestones and benchmarks; delays and challenges in implementation, including ○ Establishing Kosovo Project Acceleration Advisory Committee ○ Conducting training and capacity building for IPP developers and banks ○ Developing standardized frameworks and templates ▪ Perceived program impacts ▪ Best practices and lessons learned ▪ Successes and challenges ▪ Progress toward intermediary outcomes ▪ Sustainability of program outcomes **Quantitative Outcomes:** ▪ Number of IPPs with access to financing ▪ Amount of financing ▪ Licensing status ▪ Construction status, connection status ▪ Total MWh produced (if applicable)

Data Source/ Respondent	Outcomes
KCGF: 1-2 KIIs or small group interviews in each round of data collection	**Qualitative Outcomes:** ▪ Experience with training and capacity building ▪ Amount of capital raised ▪ Number and amount of renewable energy guarantees offered in the renewable energy guarantee window ▪ Percentage of guarantees offered to lenders ▪ Spillover to other program guarantees (beyond solar/renewable energy) ▪ Best practices and lessons learned ▪ Successes and challenges **Quantitative Outcomes:** ▪ Number of loans guaranteed ▪ Amount of guarantee
Banks: 2-3 KIIs in each round of data collection *(as appropriate)*	**Qualitative Outcomes:** ▪ Engagement with technical assistance providers ▪ Experience with training and technical assistance ▪ Alignment with priorities ▪ Spillover to other forms of program financing ▪ External factors and events effecting implementation **Quantitative Outcomes:** ▪ Number of loans provided ▪ Amount of loans provided
ERO: 1-2 KIIs or small group interviews in each round of data collection	**Qualitative Outcomes:** ▪ Engagement with MFK / project management committee ▪ Alignment with organization/government priorities ▪ External factors and events effecting implementation (such as changes to feed-in tariffs, timing of approval process)

7. WEE ACTIVITY: EVALUATION DESIGN

In this section we present the evaluation design of the WEE activity. The WEE activity is designed to ensure equal economic opportunities for women in the energy sector with the long-term objective of increasing women's employment in the energy and other STEM sectors, because women's low labor force participation in Kosovo is a barrier to equitable and inclusive growth. The approaches to achieve this include four sub-activities: (1) technical assistance and grants for women entrepreneurs; (2) internships and apprenticeships for young women interested in entering the energy sector; (3) scholarships for women studying energy-related fields; and (4) summer camps for teenage girls to foster their interest in STEM fields including energy. The WEE activity focuses on increasing women's employment in the energy sector and while reducing electricity consumption is one goal of these activities, it is not the main aim. In the original RELP design, the WEE activity was a sub-activity of the SEEK activity. However, based on MCC's decision in July 2020, since the WEE activities are not directly linked into the HER or the AER sub-activities under SEEK, they will be evaluated separately as a stand-alone group of activities.

7.1 Implementation Status of WEE Activity

We present below the implementation status of the four sub-activities of the WEE activity as of submission of this report in September 2020:

- *Scholarship Program.* The scholarship program was launched in February 2019. Through a selection process 28 scholarship recipients were selected to attend a 2-year associate's degree program at Des Moines Area Community College in Iowa (2019-2021). This activity was put on hold due to COVID-19 and the cancellation of classes. Some students remained in Iowa, and some returned to Kosovo. Now that the summer break has completed, classes are scheduled to begin again in Fall 2020, and students are expected to return from Kosovo to resume the program. In addition, some internships, which are part of the study program, have also resumed and those participants will return to those internships.
- *Internship Program*. MFK will implement the internship program. In January 2020, the internship program was advertised in local media outlets to disseminate information about the internship program. A total of 26 energy companies have expressed interest in offering approximately 100 internships to women. Plans for opening applications in April 2020 were suspended due to COVID-19, and the internship program was put on hold. Current plans hope to re-initiate the program in Spring 2021.
- *WiSci Camp.* The first year of camp was held in August 2019. This was preceded by efforts to train counselors, finalize camp logistics, and secure funding. The camp included leadership training and engagement in pre-camp programming opportunities through the web and newsletters. Year 2020 of the WiSci camp was cancelled due to COVID-19, and the second version of the WiSci Camp (year 2021) is under evaluation (as of August 2020).
- *Entrepreneurship Grants and Technical Assistance.* This activity is at the design and conception stage. Women-headed businesses (not necessarily in the energy sector) from all sectors will be eligible to apply for grants under this sub-activity. This activity was initiated with the launch of applications; however, due to COVID-19 restrictions, the activity was put on hold in March of 2020. It is expected to resume in Fall 2020.

7.2 Methods

The evaluation of the WEE activity will use multiple methods related to the two main evaluation questions outlined in Exhibit 29, to analyze the *implementation* and *effectiveness* of project activities.

Exhibit 29. WEE Activity Evaluation Design

Evaluation Question	Methods and Rationale
EQ1 Were the WEE activities implemented as designed and were the outputs outlined in the project logic achieved?	▪ **Document review and monitoring and evaluation data** will be used to analyze whether project targets and outcomes have been achieved as per the initial plan. ▪ **Semi-structured KIIs and FGDs** will be used to understand challenges and barriers encountered during project implementation. Semi-structured KIIs will include qualitative and quantitative questions. ▪ **Site visits** will be used to interview girls/women and facilitating agents, such as camp counselors and supervisors.
EQ6.1 How effective were program activities at increasing women's employment in the energy sector, and at increasing investments in energy efficiency among women entrepreneurs?	▪ **Semi-structured KIIs and FGDs** with project beneficiaries will serve to assess project results directly. Semi-structured KIIs will include qualitative and quantitative questions. ▪ **Secondary data** will be analyzed to assess project influence on women's academic and employment goals and access to job opportunities in the energy sector. In the case of entrepreneur grants, data will be analyzed to determine whether grant subsidies did in fact lead to more efficient use of energy, lower energy costs, and improved profit margins. ▪ **Site visits** will qualitatively evaluate young women's involvement in companies, including whether they are doing administrative or technical jobs.

7.3 Outcomes

Various qualitative and quantitative indicators will be used to evaluate project implementation and effectiveness, described in Exhibit 30.

Exhibit 30. WEE Activity Outcomes to Assess Project Implementation and Effectiveness

Component	Quantitative Outcomes	Qualitative Outcomes
Technical assistance and grants for women energy entrepreneurs	▪ Average percent decrease in energy consumption among grant recipients from initiation to end of activity ▪ Average percent decrease in energy costs among grant recipients from initiation to end of activity ▪ Average increase of profit margin due to decrease of energy use / energy costs among grant recipients from installation of grant to end of grant activity ▪ Amount of funds that women's enterprises invest in energy efficiency ▪ Percentage of women entrepreneurs investing in energy efficiency Percent increase in energy efficient investments	▪ Women entrepreneurs' perceptions about investing in energy efficiency

Internships and apprenticeships for young women in the energy sector	▪ Percentage of women finding employment with host company upon completion of WEE internship / apprenticeship ▪ Percentage/number of participating companies that have added technical specifications in hiring guidelines to hire women ▪ Percentage of participating companies that employ intern as staff upon completion of WEE internship / apprenticeship Percentage of women finding full-time employment in the STEM field sector within two years after completing WEE internships/ apprenticeships	▪ Women's satisfaction with experience ▪ Knowledge acquired during internships ▪ Skills acquired during internships ▪ Supervisors' satisfaction with experience ▪ Improved capacity in company's management and/or management of women ▪ Company's willingness to hire more women in the future
Scholarships for women interested in joining the energy sector	▪ Percentage/number of scholarships granted to young women from among those who apply ▪ Percentage and number of scholarship recipients finding full-time employment in a STEM sector field within two years after scholarships	▪ Satisfaction with experience ▪ Perceived knowledge acquired during scholarship ▪ Increased confidence among young women about their plans to pursue a STEM sector field
Summer camp for teenage girls	▪ Percentage/number of summer camp attendees for young women from among those who apply ▪ Level of interest in pursuing STEM-related fields of study due to summer camp experience (1-5) Percentage/number increase in young girls who attend summer camp each year ▪ Number/percent increase in young girls who apply to attend the camp each year ▪ Percentage/number increase in young girls who attend summer camp each year	▪ Satisfaction with experience ▪ Perceived knowledge acquired ▪ Mentors' satisfaction with camp experience ▪ shifted perceptions of women in STEM sectors

7.4 Data Collection

Data collection for the WEE activity will be aligned with the qualitative data collection proposed for the HER and AER sub-activities in **Section 4.3**, but not necessarily collected simultaneously. That is, wherever possible we will leverage our local presence in Kosovo to track participants under the different sub-activities of WEE and collect data at appropriate times (e.g. before the start and end of internships or before the start and end of summer camp).

Data for the WEE activity will be collected using semi-structured KIIs and FGDs, as well as through review of existing records and documentation and post-activity participant surveys. The KIIs will be semi-structured and will include *both* qualitative and quantitative questions. The main questions for the qualitative data collected will cover perception change, project success, sustainability, experience with the project, and project outcomes at the sample unit level. To evaluate project effectiveness and sustainability, a number of records will be collected and reviewed at baseline and endline. For example, for the entrepreneurial grants activity, cash flow and profit and loss statements from entrepreneurial

entities will be collected and analyzed for evidence on increase investments in energy efficiency measures or increased profits related to decreased energy expenses. For the internship activity, participating company human resource records will be analyzed at baseline and endline to know whether or not proportional hiring of women has increased, if women serving internships have been hired, or if guidelines for hiring have been adapted to facilitate increased hiring of women. We will also analyze summer camp applications and attendances to determine an increase over time, which will suggest sustainability of this effort. Finally, post-activity surveys will be used to solicit information as to whether women using scholarships or participating in internships have gained employment in STEM sector fields within two years after their WEE experience. This evidence will provide insights as to the outcome of these interventions. Likewise, follow up with summer camp participants may shed light on the continued interest and pursuit of STEM-sector fields after summer camp activities. Finally, secondary data will be analyzed to assess project influence on women's academic and employment goals and access to job opportunities in the energy sector.

7.5 Sample

Sampling for the WEE evaluation will vary depending on the sub-activity under analysis. For the women energy entrepreneur technical assistance and grants, the sample will include women entrepreneurs; for the internship sub-activity, the sample will consist of young women and their supervisors; for scholarships, the sample units will be scholarship recipients; and for the summer camp, the sample will comprise both young women participants and mentors. See **Exhibit 31** for a more detailed breakdown of sample sizes during each round of data collection.

Exhibit 31. Summary of Data Collection for WEE Activity

Data Source	Sample Sizes per Round	Sampling Strategy	Timing of Data Collection
FGDs and KIIs ▪ Women Entrepreneurs	▪ 2 FGDs with women entrepreneurs ▪ 2 semi-structured KIIs with entrepreneurs who have invested in energy efficiency	▪ Randomly select women entrepreneurs ▪ Purposefully select women entrepreneurs who have invested in energy efficiency	**Baseline** January–February 2021 **Interim** 1. October–November 2021 2. April–May 2022 3. April–May 2023 **Endline** January–February 2024
FGDs and KIIs ▪ Young women, supervisors (*internship component*)	▪ 4 FGDs (2 with women interns/apprentices; 2 with their supervisors; 6-10 individuals in each) ▪ 4 semi-structured KIIs (2 with interns: 1 with woman who secured a position after the internship; 1 with a woman who was unable to secure a position after the internship; 1 with each of their supervisors)	▪ Randomly select interns/apprentices and supervisors ▪ Purposefully select from successful beneficiaries[60]	

[60] We consider beneficiaries "successful" when they have found full-time employment as a result of their internships.

FGDs and KIIs ▪ Young women (*scholarship component*)	▪ 2 FGDs (6-10 individuals) ▪ 2 KIIs (case studies)	▪ Randomly select scholarship recipients ▪ Purposefully select from successful beneficiaries	
FGDs and KIIs ▪ Young women, Camp Mentors (*summer camp component*)	▪ 2 FGDs (1 with mentors; 1 with students; 6-10 individuals) ▪ 2 semi-structured KIIs with students ▪ 2 semi-structured KIIs with mentors	▪ Randomly select mentors and camp participants	
KIIs ▪ MFK staff ▪ ICs	▪ 2 semi-structured KIIs with RELP staff responsible for implementation of the 4 different WEE components	▪ Use purposeful sampling	**<u>Baseline</u>** January-February 2021 **<u>Endline</u>** October–November 2021
Document review; RELP monitoring and evaluation data; Secondary data	▪ Data about the change in percentage/number of women who apply for STEM sector fields in Kosovo universities will be obtained by the Ministry of Education, Science, and Technology. ▪ Data on business energy usage of grant recipients during activity period ▪ Data on business energy costs of grant recipients during activity period. ▪ Financial data on businesses (income/expenditures) of grant recipients that covers activity period ▪ Data from hiring guidelines of participating companies and technical specifications for their contractors that promote or facilitate the hiring of women. ▪ Quantitative data about the change in percentage/number of women who graduate in STEM-related fields will be obtained by the Ministry of Education, Science, and Technology.		Ongoing throughout project lifespan

8. EVALUATION OF BROADER SECTORAL OUTCOMES

In this section, we describe our evaluation design to examine the impact of RELP on broader sector-level outcomes, using a mix of quantitative and qualitative methods described in **Section 8.1**. We will draw on a number of data sources (described in **Section 8.2**) to answer questions about the extent to which RELP reduced the gap between supply and demand of electricity.

8.1 Methods

To determine whether the RELP activities have reduced the demand-supply gap **(EQ7.1**[61]**)**, we propose to conduct a mixed-methods analysis, as described below.

1. **Simulations.** We will use simulations to gauge the *projected* impacts on the demand-supply gap **(EQ7.1)** if SEEK were to be scaled up all over Kosovo. We will focus only on the SEEK activity, since the DHM activity is restricted to only one municipality, Pristina. We will apply the approach described below and the causal quantitative estimates of the SEEK activity obtained from beneficiary-level data described in detail in **Section 4**. We will only be able to model the demand side of the demand-supply gap, since none of the activities of RELP impact electricity supply in any meaningful way. To the extent feasible, we will use estimates of changes in energy supply as a result of an increase in IPPs producing renewable energy to examine changes in supply, however, as described in **Section 6**, changes in electricity supply are not expected to happen during the lifetime of the project. Simulations of broader sectoral impacts will provide suggestive evidence about the *expected reductions in the demand-supply gap*, after the end of the evaluation period from 2025–2030. We will use the following general approach:

    - *Calculate the change in total electricity consumption for HER and AER beneficiaries.* We will use the ITS approach described in detail in **Section 4** to derive estimates of the change in electricity consumption because of the HER and AER sub-activities.[62]
    - *Empirically derive short-run income and price elasticities of electricity demand at the household and MAB levels.* Following Houthakker (1951) and Rapanos and Polemis (2006), we will model the electricity consumption decision as a result of installation of retrofits under RELP.[63] We will first specify a theoretical model of consumer utility maximization from electricity consumed and other goods, subject to income constraints. The model will illustrate how income and cost of electricity affects the private optimal choice of electricity consumption *after* the installation of retrofits. From this, we will derive estimating equations for the income and price elasticities of electricity demand. We will fit several specifications, using the data from the HER and AER sub-activities, and derive empirically both income and price elasticity estimates. This will allow us to understand the responsiveness of electricity demand to price and income. This will then allow us to forecast the changes in electricity

[61] *EQ7.1: Did the project meet the stated objective of reducing the gap between supply and demand of electricity?*

[62] As pointed out by Greenstone and Fowlie (2015), gains from an investment in energy efficiency are realized through two main channels: reduced energy consumption and increased consumption of energy services (e.g. space heating) due to reductions in the price of energy services (i.e. it costs less to heat the same space).

[63] The most popular approach has been the estimation of a single equation demand function expressed in a log linear form, where total electricity consumption is a function of income, price of electricity, and the number of heating degree days.

consumption induced by scaling up SEEK all over Kosovo, as well as the associated energy rebound effects.

- *Forecast the savings trajectory for HER and AER beneficiaries.* We will use time-series models (e.g., ARIMA) to forecast the electricity consumption of our sample of 2,600 HER beneficiaries for the 5-year period once the project ends, 2025–2030 (e.g., see Jain et al. 2018).
- *Simulate estimates under different scenarios.* Using the forecasted trajectory and the elasticities described above, we will provide estimates for different time periods and different subgroups. That is, we will provide expected energy reduction projections as a result of SEEK expansion for the 5-year time window 2025–2030.[64] We will account for several variations in our simulations: anticipated rebound effect, anticipated/assumed energy efficiency adoption rates, anticipated/assumed changes in building stock, programmed population rates, programmed income, and programmed number of households (urban and rural).

2. **Administrative data.** While the simulation exercise described above will estimate projected impacts *after* the evaluation period, we will also use administrative data to estimate actual sectoral impacts due to RELP *during* the evaluation period i.e. up until 2024. Specifically, we will analyze changes in the demand-supply gap **(EQ7.1**[65]**)** due to RELP. We will use administrative data from Termokos, KEDS, and the Kosovo Transmission System and Market Operator to assess key indicators of electricity consumption (*demand)* and electricity generation, transmission, and distribution *(supply)* before and after the introduction of RELP, to examine whether there has been any actual impact on the demand-supply gap **(EQ7.1)**. We will examine changes in levels and trends of these key indicators before and after the introduction of RELP. We will examine aggregate monthly and yearly changes in the main outcomes at the municipality level, if feasible. These aggregate estimates will provide a broad picture about changes in the demand-supply gap. However, they will not enable us to assign attribution to RELP or understand the mechanisms at play.
3. **Qualitative analysis.** We will complement the administrative data with qualitative data, to understand demand-side behavioral changes at the household level and supply-side changes that could be affected by other policies and unrelated factors in Kosovo that could also lead to broader sectoral-level changes. We will conduct KIIs, posing questions focused on understanding the impacts on electricity consumption if RELP were to be scaled up more widely in Kosovo. We envision interviewing respondents from Termokos, KEDS, ERO, MFK, MED, and other relevant stakeholders. We will track the development, passing, and implementation of policies, laws, and regulations throughout the energy sector and incorporate these policies into our simulation estimates.

8.2 Outcomes and Data Collection Method

Exhibit 32 summarizes the outcomes and the data sources / data collection methods. We propose to collect and analyze data about measures related to the demand-supply gap.

[64] We will use secondary data sources, including government reports and other external reports, to obtain information about population growth rates, income growth rates, and the rate of urbanization for Kosovo.

[65] *EQ7.1: Did the project meet the stated objective of reducing the gap between the supply and demand of electricity?*

Exhibit 32. Data Collection Method and Key Outcomes for Broader Sectoral-Level Changes

Data Source	Anticipated Time Frame	Key Outcome
Administrative data: ▪ Monthly from KEDS, Termokos, ERO/ Kosovo Transmission System and Market Operator	▪ Monthly, January 2018–May 2024 ▪ Beneficiary level and municipal level	**<u>Demand-supply gap</u>** ▪ Maximum load/peak load ▪ Load factor ▪ Hourly consumption in a 24-hour period ▪ Maximum, minimum, monthly demand ▪ Average price of electricity **<u>Other programs, policies, and regulations affecting grid-level outcomes</u>** ▪ Tracking development, passage, and implementation of policies and laws
Document review from relevant stakeholders: ▪ News and media ▪ Legal, economic, and technical regulations, ▪ Documents of energy programs in the region	▪ Ongoing on a monthly basis, December 2019–May 2024	
KIIs with relevant stakeholders including Termokos, KEDS, ERO, MED, MFK, and other relevant stakeholders	▪ See **Section 4.3, 5.3** and **Section 6** for timing of qualitative data collection.	

9. OVERALL EVALUATION CHALLENGES AND LIMITATIONS

In this section, we outline the critical risks to the implementation of RELP more broadly and the potential impacts on the project's evaluation. We divide this into two domains, with **Section 9.1** outlining challenges to project implementation and **Section 9.2** detailing evaluation limitations and mitigation strategies. Finally, in **Section 9.3**, we outline the assumptions and data needs for the evaluation of each activity under RELP.

9.1 Challenges to Project Implementation

We present below the main foreseeable challenges to the implementation of RELP.

1. **Implementation delays.** It is possible that the implementation of the SEEK, DHM, and IPP activities may be delayed, due to various challenges, including: (a) U.S./MCC sanctions related to import tariffs for goods from Serbia and a temporary suspension of funding; and (b) the COVID-19 pandemic. These project delays create the risk that the implementation of activities will be rushed to meet target deadlines or will not be completed by the end of the threshold period, currently scheduled for September 2021. With respect to the COVID-19 pandemic, there may be an added risk of a lack of sufficient number of beneficiaries for the HER, AER, and DHM activities since households may be reluctant to allow installations in their homes during a pandemic. These risks may have implications for our evaluation design (especially for DHM, which relies crucially on the timing of the installation of meters) and for our proposed data collection and interim report submission dates which are crucially linked to the implementation timeline. We will document any substantive modifications to implementation plans and work with MCC to update the evaluation design, to the extent possible.

2. **Complexity of project logic.** The project logic is quite complex in terms of the number of nodes, interdependencies, and outcomes required to reach the project's primary objective. Ideally, the project logic should be divided into one per activity, providing a clearer view of each activity. The complexity of the project logic necessitates a complex evaluation design, involving many stakeholders. This introduces challenges for attributing project impacts to any one project component or stakeholder. This also introduces challenges for ascertaining impacts, since the inclusion of so many stakeholders can lead to conflicting findings. Additionally, significant attrition of any of these stakeholders will hinder our ability to definitively attribute impacts.

3. **Risk mitigation plans.** MFK's risk mitigation plan does not provide sufficient detail and labels many risks as "low." The majority of the mitigation measures rely heavily on MFK involvement and its ability to coordinate with many different stakeholders, including the ICs, ERO, Termokos, KCGF, and the municipal governments. We consider this an overall risk, as this assumes that MFK has the capacity to continuously monitor and engage with the ICs and other project stakeholders and to take action to counter these risks.

9.2 Limitation and Mitigation Strategies of RELP Evaluation

Next, we present below the main limitations of the RELP evaluation and proposed mitigation strategies for each activity.

9.2.1 SEEK and DHM Activities

1. **Disentangling impacts between AER and HER sub-activities.** There are potentially some households that will receive both the HER and AER sub-activities at the same time. We will disentangle impacts by removing the overlap sample. If the extent of overlap exceeds the sample required to detect meaningful effects, we will include these "overlap" households and estimate impacts with and without them, to understand the independent impact of the HER sub-activity and the combined effect of the HER and AER sub-activities.

2. **Lack of consent for access to KEDS and Termokos data.** Obtaining consent from households to allow us to link their apartment-level electricity consumption/district heating consumption to the KEDS/Termokos database is a crucial aspect of using administrative data for the ITS analysis to analyze our main outcome variable of interest, electricity consumption/district heating consumption for the SEEK and DHM activities, respectively. To ensure that we can get this information, the consent form will clearly lay out that no personally identifiable information will be shared with anyone except the IC, the evaluator, MCC, and MFK, and that the data will only be used for the purpose of implementation and evaluation of RELP. Furthermore, we will assure respondents that there are no risks to them from participating in the data collection.

3. **Attrition of respondents.** Since both SEEK and DHM are longitudinal studies, attrition is a significant concern. We have accounted for the challenge of attrition in our evaluation design and calculated sample sizes for minimum detectable effects, based on a high nonresponse rate of 30 percent. Additionally, at baseline, we will implement methods to reduce attrition, such as obtaining comprehensive contact information from respondents. If attrition from the sample is high, we will discuss with MCC the option of sampling the new residents in the homes that received retrofits under SEEK. Additionally, for DHM we will potentially update our design and implement a repeated cross-sectional design. That is, at endline, instead of surveying the same households from baseline, we will sample a new set of households.

4. **Lack of access to household-level energy consumption data from KEDS and Termokos.** During IMPAQ's scoping trip both KEDS and Termokos were amenable to sharing household level data for their customers. Subsequently, the lowest level of aggregation we have obtained from Termokos is only at the building level. We will continue to work with MFK and Termokos to obtain household level data. In the event that we have access to only building level data from Termokos, we will revise our evaluation approach to conduct analysis only at the building level. However, this would reduce the sample size and the associated power to detect meaningful effects, since the number of buildings in Pristina under the district heating network is only ~300. Similarly, in the event that KEDS is unwilling to share their data, implementation of the ITS approach will become challenging and we will instead focus on the pre-post approach using primary survey data.

9.2.2 IPP Activity

1. **Timing of project implementation.** As documented in Section 9.1, it is possible that various components of RELP may be delayed. The technical assistance to banks, IPPs, and other stakeholders is not scheduled to occur until the beginning of 2021, and the renewable energy guarantee window is not scheduled to open until immediately before the end of the threshold

period. This already tight timeline makes it difficult to assess any outcomes, and further delays will influence our ability to do a full process-tracing evaluation. In addition, delays on the side of the ERO in processing applications and granting approvals may extend the timeline beyond the scope of the evaluation. We will work with MCC to modify the design of our evaluation (including the timing of data collection), as necessary, if large project deviations occur.

9.2.3 WEE Activity

1. **Attributing broader sectoral impacts of the WEE activity.** It will be challenging to directly analyze the impact of the WEE activity components (internships, scholarships, summer camp, and grants) on women's employment. Offering internships, scholarships, and the women in science summer camp only addresses the demand side of the employment problem, by providing women with sector-related job skills and by creating interest in women and encouraging them to apply for jobs in utilities. On the supply side, employers must be willing to hire more women. Additionally, even on the demand side, the sample sizes targeted by the WEE activity are quite small. Furthermore, the sample suffers from sampling bias and information gathered from this sample should be interpreted with caution. To assess project implementation, we will collect data only from those women who are already interested in the energy sector, i.e., women attending internships, scholarships, summer camps, etc. This may lead to skewed data about women being interested in the energy field, with an overall lack of data about why women are not interested in the field in the first place. Thus, this evaluation will only be able to provide suggestive evidence about the impact of the WEE activity on broader sectoral-level changes in women's employment, from both demand and supply sides.

9.2.4 Overall Challenges

1. **Similar programs during the threshold period.** Another limitation arises from the fact that during the threshold period, there may be other Government of Kosovo and/or donor-funded programs, which would make it difficult to attribute changes in energy consumption to MCC's investments. To mitigate this, IMPAQ will keep track of outside investments in the energy sector and monitor any programs that may affect RELP beneficiaries. We will work with MFK to establish procedures for sharing this information. We will adapt the ITS regression equation to include additional indicators for other programs. However, if a program is switched on very close to the time that MCC RELP activities are taken online, we will not be able to disentangle the two and separately identify the impact of MCC's investment.

2. **Small sample sizes for qualitative data collection.** The qualitative approach also has several limitations. The data collected will be from a small sample, and as is the nature of qualitative data, the results are not necessarily generalizable, but rather show the broad spectrum of types of perspectives that may be encountered across beneficiaries and stakeholders. We will mitigate this by continuous monitoring, using qualitative data collection, throughout the evaluation lifespan. That is, in addition to baseline and endline, we will have three rounds of interim qualitative data collection.

3. **Lack of experimental approach in the evaluation design.** The implementation design of all three activities under RELP do not include counterfactual groups, which makes a purely experimental evaluation design challenging. That is, none of the activities are randomly assigned to a

"treatment" group. For instance, under SEEK, households will voluntarily enroll in the program and in DHM, all Termokos customers are included in the activity. Finally, in IPP, the program will purposively select IPPs to be included as beneficiaries. Given the lack of randomization, we have proposed rigorous quasi-experimental approaches, where feasible, and pre-post approaches to complement them.

9.3 Assumptions and Data Needs

We conclude this section by presenting the assumptions and data needs that underlie the evaluation design proposed for all three activities.

9.3.1 SEEK, DHM, and WEE Activities

The following assumptions and data needs underlie the evaluation of the SEEK, DHM, and WEE activities:

- We have access to MFK and the IC's monitoring data in a timely manner.
- Households consent to our access to their KEDS and Termokos records.
- KEDS and Termokos provide access to administrative data.
- The IC procures a subcontractor for baseline data collection in a timely manner.
- There is low attrition of households from baseline to endline.
- We have access to documents including progress, annual, quarterly, and monitoring and evaluation reports, and any other relevant documents.
- Direct and indirect beneficiaries and implementers consent to participate in KIIs and FGDs.

9.3.2 IPP Activity

The following assumptions and data needs underlie the evaluation of the IPP activity:

- We have access to documents, including progress, annual, quarterly, and monitoring and evaluation reports, and any other relevant documents.
- Banks, IPPs (including those who have dropped out of the project), implementers, and government stakeholders consent to participation in KIIs.
- Project management committees, KCGF, and technical assistance consultants are able to provide administrative and project data.

10. ADMINISTRATIVE ISSUES

In this section, we describe the administrative issues relevant to conducting the evaluation of RELP. In Section 10.1, we present a summary of the institutional review board requirements. In Section 10.2 and Section 10.3, we detail the process for preparation of public-use data files and steps taken to ensure data protection, respectively. In Section 10.4, we discuss our dissemination plan and in Section 10.5, we detail the evaluation team's roles and responsibilities. Finally, Section 10.6 presents the evaluation budget.

10.1 Summary of Institutional Review Board Requirements and Clearances

We will ensure that our team members, including enumerators and contractors working on the project, adhere to the ethical guidelines outlined in the American Evaluation Association's "Guiding Principles for Evaluators," in conjunction with MCC's commitment to, where feasible, respect and follow the Common Federal Policy for the Protection of Human Subjects. For the submission of the institutional review board (IRB) package, we will prepare a detailed research protocol which will include: (a) project summary, rationale, and background information; (b) study goals and objectives; (c) study duration, schedule, design, and methodology; (d) procedure for recruiting participants; (e) safety considerations; (f) data management and analysis; (g) quality assurance; (h) expected outcomes of the study; (i) dissemination plan; (j) anticipated problems and their management; (k) informed consent; and (l) data privacy. We will also include copies of all quantitative and qualitative instruments, in addition to the completed IRB questionnaire. Since IRB approvals are valid for one year, we will apply for renewals for additional years, as needed. Our team has extensive experience in preparing for the IRB in many sectors, including energy. For more than a decade, IMPAQ has successfully obtained IRB clearances on hundreds of domestic and international studies. We have a record of granted exemptions for research conducted in low-income settings. We will apply for IRB approval in the United States and in Kosovo. We will revise survey instruments and guides, based on feedback from the U.S. and/or local IRB in Kosovo.

10.2 Preparing Data Files for Access and Documentation

We will prepare all public-use documentation for the quantitative data in accordance with MCC's most recent guidelines. To protect the identity of individuals, all personally identifiable information will be stripped from the public-use dataset. Variables with fewer than five observations in a given category will be masked by aggregating their categories into broader ones, to minimize the risk of identification (Seastrom 2017). Continuous variables such as age may be aggregated to hide values at the low and high ends, which could potentially identify an individual. In these instances, if necessary, variables will be placed in groupings that preserve cutoffs needed for constructing outcome variables. We will include a list of all variables that will be excluded from or will be masked in the public-use data file (and therefore from the data dictionary), along with the rationale for excluding or masking each variable. The public-use data will be accompanied by documentation, which will contain a data dictionary providing details about each of the variables included in the public-use dataset. Each variable will be presented, along with the summary of the data, including the frequency of each response, the proportion of respondents who chose it, and when applicable, the value label. Finally, the public-use data documentation will also contain a detailed crosswalk between all outcome variables and the variables used to construct them.

10.3 Data Protection

Personally identifiable information. IMPAQ's process for handling personally identifiable information is designed to reduce the exposure of personal identifiers to an absolute minimum. One of the first steps taken when a dataset arrives at IMPAQ from an outside source is to identify the personally identifiable information, such as names, addresses, and phone numbers. From the first transmission point, all data are encrypted with SSL encryption to minimize exposure during the transit. After transmission, data rests in a special storage system that is encrypted with a FIPS 140-2 encryption compliance system. Only designated people within IMPAQ have access to the data, to move it to a segregated PCI-compliant network. This process, along with many other data-security procedures and the hardware and software infrastructure in place at IMPAQ, minimizes the chance of unauthorized access to personal identifiers. Once the datasets containing personally identifiable information are successfully transferred from the source to the final system, a comprehensive process begins to clean the data. As we have done in previous studies, IMPAQ will document the entire process of cleaning the data, keeping written records of every decision made during the process and the reason for each decision, so that, if necessary, it would be possible to exactly replicate the process at a later date.

Data security. A state-of-the-art router, a firewall, and intrusion detection protect the IMPAQ computing network and prevention system monitored 24-7 by the security operations center of our internet service provider. All IMPAQ executives annually complete the federal information system security awareness training that is compliant with Federal Information Security Management Act (FISMA) and provided online by the Department of Defense at http://iase.disa.mil/eta/iss_icv5/. IMPAQ holds credit monitoring and privacy notification insurance with the Chubb Insurance Group. Database servers are maintained in a secure server room, with physical access restricted to authorized IMPAQ information technology staff. Data are protected using a layered firewall infrastructure, local network DMZs, active port analysis and monitoring, regular password reassignment, server login access control, application of latest security patches to operating systems, and network monitoring for suspicious activities. IMPAQ information technology security personnel routinely review these logs for inappropriate activities, and take action as needed.

10.4 Dissemination Plan

One of the primary purposes of an evaluation is to learn from the evidence about what works and what does not. The IMPAQ team will prepare a presentation within a month of approval of the interim and endline reports. We will present these findings at MCC headquarters and in Kosovo. Both presentations will be in English; however, we will translate the presentation into local languages upon MCC's request. At the interim report stage, a policy brief (in English and local languages) of approximately four pages will target MCC decision makers and government officials. The brief will focus on key findings and lessons learned at the interim stage. At endline, IMPAQ will undertake several other learning and dissemination activities. For example, the team may review any evaluation-related material prepared by MCC's media team, to ensure quality. Finally, we will seek to disseminate evaluation findings to the larger policy audience at both interim and endline, through publications, workshops, and conferences.

10.5 Evaluation Team Roles and Responsibilities

With our partners, Women Influencing Health Education and Rule of Law (WI-HER), Ideas42, and Causal Design, the IMPAQ team is uniquely qualified to implement the RELP Evaluation. **Exhibit 33** outlines the

roles and responsibilities of the RELP team and reflects the latest proposed staffing structure for the RELP evaluation.

Exhibit 33. Evaluation Team Roles and Responsibilities

Team Member	Role	Responsibility
Sonam Gupta	Project Director	Oversee overall project management and provide supervision of team members on all evaluation activities and deliverables, including document review, evaluation design, fact-finding field visits, data collection, meetings, workshops, data analysis, and writing.
Nils Junge	Energy Expert	Co-lead with project director, the development of a rigorous evaluation design, using the appropriate combination of quantitative and qualitative methodologies, to answer the evaluation questions.
Daniel Elliot	Program Manager	Serve as primary point of contact for all contractual and operational matters between the evaluation team and MCC; provide contractual, operational, and logistical support.
Vanessa Hoffman	Qualitative Expert	Lead the qualitative component of the evaluation, in coordination with the project director and team leader; develop evaluation methodologies and data collection instruments, as well as sampling methodologies for qualitative data collection.
Uttara Balakrishnan	Quantitative Associate	Support the project director and the team leader in the review of quantitative instruments, sampling, evaluation design, data analysis, report writing, and development of deliverables and other tasks.
TBD	Qualitative Associate	Support the qualitative expert in the design of the qualitative approach for the evaluation, data collection instruments, and sampling methodologies; support in the analysis of the qualitative data, report writing, and dissemination of evaluation findings.
Daniel Zaas and Connor Harrison	Research Analysts	Participate in the design of the evaluation and data collection protocols, data analyses, and report writing; support the overall implementation and management of the evaluation.
Visar Zekaj	Country Expert and Field Work Manager	Manage in-country administrative and logistical aspects of the evaluation.
Leszek Kasprowicz	Energy Expert	Provide technical expertise about energy related aspects of the evaluation from an engineering perspective; provide in-depth input and feedback about evaluation instruments.
Zija Kamberi	Legal and Regulatory Expert	Provide legal and regulatory insight about all aspects of the evaluation.
Causal Design	CEA Expert	Provide CEA-related technical input about all evaluation activities; lead analysis of the findings in comparison with preexisting CEA analyses, and how these analyses can be updated with the new data.
Ideas42	Behavioral Economics Expert	Contribute to the preparation of research instruments to conduct qualitative and quantitative research to identify behavioral determinants that may be adopted by various audiences in Kosovo.
WI-HER	Gender Expert	Work with IMPAQ to coordinate and align all gender-related work and collaborate on relevant evaluation activities and deliverables.

REFERENCES

Adan, Hassan, and Franz Fuerst. 2016. "Do Energy Efficiency Measures Really Reduce Household Energy Consumption? A Difference-in-Difference Analysis." *Energy Efficiency* 9 (5): 1207–19. https://doi.org/10.1007/s12053-015-9418-3.

Akpandjar, George, and Carl Kitchens. 2017. "From Darkness to Light: The Effect of Electrification in Ghana, 2000–2010." *Economic Development and Cultural Change* 66 (1): 31–54. https://doi.org/10.1086/693707.

Albu, J. B., N. Sohler, R. Li, X. Li, E. Young, E. W. Gregg, and D. Ross-Degnan. 2017. "An Interrupted Time Series Analysis to Determine the Effect of an Electronic Health Record–Based Intervention on Appropriate Screening for Type 2 Diabetes in Urban Primary Care Clinics in New York City." *Diabetes care* 40 (8): 1058-1064.

Allcott, Hunt, and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives* 26 (1): 3–28. https://doi.org/10.1257/jep.26.1.3.

Ameli, Nadia, and Nicola Brandt. 2015a. "Determinants of Households' Investment in Energy Efficiency and Renewables: Evidence from the OECD Survey on Household Environmental Behaviour and Attitudes." *Environmental Research Letters* 10 (4): 044015. https://doi.org/10.1088/1748-9326/10/4/044015.

Ameli, Nadia, and Nicola Brandt. 2015b. "What Impedes Household Investment in Energy Efficiency and Renewable Energy?" *OECD Economics Department Working Papers*. Vol. 1222. https://doi.org/10.1787/5js1j15g2f8n-en.

Angelucci, M., and V. Di Maro. 2010. *Program evaluation and spillover effects*. Impact-Evaluation Guidelines No. IDB-TN-136. Washington, DC: Inter-American Development Bank.

Ashraf, Nava, James Berry, and Jesse Shapiro. 2007. *Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia*. w13247. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w13247.

Aydin, Erdal, Nils Kok, and Dirk Brounen. 2017. "Energy Efficiency and Household Behavior: The Rebound Effect in the Residential Sector." *The RAND Journal of Economics* 48 (3): 749–82. https://doi.org/10.1111/1756-2171.12190.

Barron, Manuel, and Maximo Torero. 2017. "Household Electrification and Indoor Air Pollution." *Journal of Environmental Economics and Management* 86 (November): 81–92. https://doi.org/10.1016/j.jeem.2017.07.007.

Bayer, Patrick, Ryan Kennedy, Joonseok Yang, and Johannes Urpelainen. 2020. "The Need for Impact Evaluation in Electricity Access Research." *Energy Policy* 137 (February): 111099. https://doi.org/10.1016/j.enpol.2019.111099.

Beér, J., W. Mielczarski, and D. Taylor. 2012. *Kosovo: Kosovo Power Project: Report of the SFDCC External Expert Panel to the World Bank.*

http://siteresources.worldbank.org/INTENERGY2/Resources/Kosovo_Power_Project_Expert_Panel_Report.pdf.

Bernal, J. L., S. Cummins, and A. Gasparrini. 2017. "Interrupted time series regression for the evaluation of public health interventions: a tutorial." *International journal of epidemiology*, 46 (1), 348-355.

Bowen, Brian H., James A. Myers, Arzana Myderrizi, Blendi Hasaj, and Blerina Halili. n.d. *Kosovo Household Energy Consumption Facts and Figures*. 82.

Buckheit, Bruce C. 2012. *Re-Evaluation Kosovo's Least Cost Electricity Optio*n. The Sierra Club, The Kosovar Institute for Development Policy (INDEP).

Bytyci, Fatos. 2019. "GE to Build Kosovo's New 500 MW Coal Power Plant." Edited by Kirsten Donovan. *Reuters*, May 3. www.reuters.com/article/us-contourglobal-kosovo-ge/ge-to-build-kosovos-new-500-mw-coal-power-plant-idUSKCN1S917R.

Chai, Kah-Hin, and Catrina Yeo. 2012. "Overcoming Energy Efficiency Barriers through Systems Approach—A Conceptual Framework." *Energy Policy* 46 (July): 460–72. https://doi.org/10.1016/j.enpol.2012.04.012.

Cohen, Jessica, and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment ." *Quarterly Journal of Economics* 125 (1): 1–45. https://doi.org/10.1162/qjec.2010.125.1.1.

Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science & Politics* 44 (4): 823–30. https://doi.org/10.1017/S1049096511001429.

Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35–69. https://doi.org/10.1257/aer.100.1.35.

Davis, Lucas W., Sebastian Martinez, and Bibiana Taboada. 2020. "How Effective Is Energy-Efficient Housing? Evidence from a Field Trial in Mexico." *Journal of Development Economics* 143 (March): 102390. https://doi.org/10.1016/j.jdeveco.2019.102390.

Decon International. 2020a. *Operations Manuals for District Heating and Metering*. Reliable Energy Landscape Project, funded by Millennium Foundation Kosovo.

———. 2020b. *DHM Inception Report* (Draft). Reliable Energy Landscape Project, funded by Millennium Foundation Kosovo.

———. 2020a. *District Heating Metering Implementing Contractor—Inception Report*. Funded by Millennium Challenge Corporation.

Dhami, Sanjit S. 2016. *The Foundations of Behavioral Economic Analysis*. First edition. Oxford: Oxford University Press. Chapter 22.

Dowson, Mark, Adam Poole, David Harrison, and Gideon Susman. 2012. "Domestic U.K. Retrofit Challenge: Barriers, Incentives, and Current Performance Leading into the Green Deal." *Energy Policy* 50 (November): 294–305. https://doi.org/10.1016/j.enpol.2012.07.019.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2008. "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya." *American Economic Review* 98 (2): 482–88. https://doi.org/10.1257/aer.98.2.482.

EBRD (European Bank for Reconstruction and Development). 201). *EBRD invests in green economy program in Kosovo*. Retrieved from http://www.ebrd.com/news/2017/ebrd-invests-in-green-economy-program-in-kosovo.html

English, Alicia, Anila Qehaja, Flamur Breznica, Arif Hoti, Michael Waschak, and Jim Myers. 2015. *Energy Consumption and Potentials for Energy Efficiency Implementation: Analyzing Low-Income, Low-Service Areas of Kosovo—Final Report*. RIT Center for Energy and Natural Resources. https://doi.org/10.13140/RG.2.1.4816.1368.

European Commission Liaison Office, Kosovo. 2010. *Study for Decommissioning of Kosovo-A Power Plant*. European Union-funded program.

Farnsworth, Nicole, Donjeta Morina, David JJ Ryan, Gresa Rrahmani, Vanessa Robinson-Conlon, and Iliriana Banjska. 2018. *Kosovo Gender Analysis*. Kosovo Women's Network. https://womensnetwork.org/wp-content/uploads/2018/10/womens-network.pdf

Farnsvede, Ulf, Ariana Qosaj-Mustafa, and Nicole Farnsworth. 2014. *Country Gender Profile*. https://abgj.rks-gov.net/assets/cms/uploads/files/Publikimet%20ABGJ/Profili%20Gjinor%20i%20Vendit_2014%20-ALB.pdf

Fleiter, Tobias, Ernst Worrell, and Wolfgang Eichhammer. 2011. "Barriers to Energy Efficiency in Industrial Bottom-up Energy Demand Models—A Review." *Renewable and Sustainable Energy Reviews* 15 (6): 3099–3111. https://doi.org/10.1016/j.rser.2011.03.025.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2018. "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." *The Quarterly Journal of Economics* 133 (3): 1597–1644. https://doi.org/10.1093/qje/qjy005.

Fretheim, A., S. B. Soumerai, F. Zhang, A. D. Oxman, and D. Ross-Degnan. 2013. "Interrupted time-series analysis yielded an effect estimate concordant with the cluster-randomized controlled trial result." *Journal of clinical epidemiology* 66 (8): 883-887.

Galvin, Ray. 2015. "Integrating the Rebound Effect: Accurate Predictors for Upgrading Domestic Heating." *Building Research & Information* 43 (6): 710–22. https://doi.org/10.1080/09613218.2014.988439.

Gerarden, Todd D., Richard G. Newell, and Robert N. Stavins. 2017. "Assessing the Energy-Efficiency Gap." *Journal of Economic Literature* 55 (4): 1486–1525. https://doi.org/10.1257/jel.20161360.

Gerőházi, Éva, and Hanna Szemző. 2016. *Recommendations for Armenia and Bosnia and Herzegovina on Encouraging Energy Efficient Renovation on the Multi-Family Housing Stock*. Metropolitan Research Institute (MRI.

GFA Consulting Group. 2020a. *Pilot Incentives for Household Investments in Energy Efficiency Implementing Contractor—Subsidies in Energy Efficiency in Kosovo, Inception Report*. Funded by Millennium Challenge Corporation.

———. 2020b. *HER Operations Manual* (Draft). Reliable Energy Landscape Project, funded by Millennium Foundation Kosovo.

———. 2020c. *AER Operations Manual* (Draft). Reliable Energy Landscape Project, funded by Millennium Foundation Kosovo.

Ghesla, Claus, Manuel Grieder, and Renate Schubert. 2018. "Nudging the Poor and the Rich—A Field Study on the Distributional Effects of Green Electricity Defaults." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3147028.

Gillingham, K., and K. Palmer. 2014. "Bridging the Energy Efficiency Gap: Policy Insights from Economic Theory and Empirical Evidence." *Review of Environmental Economics and Policy* 8 (1): 18–38. https://doi.org/10.1093/reep/ret021.

Giraudet, Louis-Gaëtan. 2020. "Energy Efficiency as a Credence Good: A Review of Informational Barriers to Energy Savings in the Building Sector." *Energy Economics*, February, 104698. https://doi.org/10.1016/j.eneco.2020.104698.

Granade, Hannah Choi, John Creyts, Anton Derkach, Philip Farese, Scott Nyquist, and Ken Ostrowiski. 2009. *Unlocking Energy Efficiency in the U.S Economy*. McKinsey & Company. https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/epng/pdfs/unlocking%20energy%20efficiency/us_energy_efficiency_exc_summary.ashx

Grogan, Louise, and Asha Sadanand. 2013. "Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua." *World Development* 43 (March): 252–65. https://doi.org/10.1016/j.worlddev.2012.09.002.

Hausmann, Ricardo, Dani Rodrik, and Andres Velasco. 2005. *Growth Diagnostics*. John F. Kennedy School of Government, Harvard University. https://growthlab.cid.harvard.edu/files/growthlab/files/growth-diagnostics.pdf

Houthakker, H. S. 1951. "Electricity Tariffs in Theory and Practice." *The Economic Journal* 61 (241): 1. https://doi.org/10.2307/2226608.

Howarth, Richard B., and Bo Andersson. 1993. "Market Barriers to Energy Efficiency." *Energy Economics* 15 (4): 262–72. https://doi.org/10.1016/0140-9883(93)90016-K.

Huntington, H., L. Schipper, and A.H. Sanstad. 1994. "Markets for energy efficiency." Editor's introduction. *Energy Policy* 22: 795-797.

IEA (International Energy Agency). 2019. *SDG7: Data and Projections*. Paris: IEA. https://www.iea.org/reports/sdg7-data-and-programions.

IMPAQ International. 2018. *Gender, LGBTI, and persons with disabilities assessment*. USAID/Kosovo. https://www.edu-links.org/sites/default/files/media/file/Gender-LGBTI-PWD%20assessment%20Kosovo%202018.pdf.

———. 2020. *Evaluability Assessment*. MCC Kosovo. https://www.edu-links.org/sites/default/files/media/file/Gender-LGBTI-PWD%20assessment%20Kosovo%202018.pdf.

Loftus, P. J., A. M. Cohen, J. C. S. Long, and J. D. Jenkins. 2015. "A critical review of global decarbonization scenarios: what do they tell us about feasibility?" *Wiley Interdisciplinary Reviews: Climate Change* 6 (1): 93–112.

Jacob, Robin, Marie-Andree Somers, Pei Zhu, and Howard Bloom. 2016. "The Validity of the Comparative Interrupted Time Series Design for Evaluating the Effect of School-Level Interventions." *Evaluation Review* 40 (3): 167–98. https://doi.org/10.1177/0193841X16663414.

Jain, Vipul, Arun Kumar Sangaiah, Sumit Sakhuja, Nittin Thoduka, and Rahul Aggarwal. 2018. "Supplier Selection Using Fuzzy AHP and TOPSIS: A Case Study in the Indian Automotive Industry." *Neural Computing and Applications* 29 (7): 555–64. https://doi.org/10.1007/s00521-016-2533-z.

Kammen, D. M., M. Mozafari, and D. Prull. 2012. *RAEL Report. Sustainable Energy Options for Kosovo: An Analysis of Resource Availability and Cost*. University of California, Berkeley. http://rael.berkeley.edu/kosovoenergy.

Kakalejčíková, Zita. 2017. *How to improve residential energy efficiency in South Eastern Europe and CIS*. Habitat for Humanity International Europe, Middle East and Africa. https://www.habitat.org/sites/default/files/How%20to%20improve%20REE_policy%20brief_FINAL_0.pdf.

Kitchens, Carl, and Price Fishback. 2015. "Flip the Switch: The Impact of the Rural Electrification Administration 1935–1940." *The Journal of Economic History* 75 (4): 1161–95. https://doi.org/10.1017/S0022050715001540.

Kittner, Noah, Hilda Dimco, Visar Azemi, Evgenia Tairyan, and Daniel M Kammen. 2016. "An Analytic Framework to Assess Future Electricity Options in Kosovo." *Environmental Research Letters* 11 (10): 104013. https://doi.org/10.1088/1748-9326/11/10/104013.

Kline, Patrick, and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129 (1): 275-331.

Kokx, L. 2017. *National Building Energy Efficiency Study for Kosovo*. World Bank. https://www.worldbank.org/content/dam/Worldbank/Feature%20Story/ECA/kosovo/Kosovo%20Eptisa%20Final%20Report_2013.04.13.pdf.

Kontopantelis, E., T. Doran, D. A. Springate, I. Buchan, and D. Reeves. 2015. "Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis." *BMJ* 350, h2750.

Kosovo Agency of Statistics. 2015. *Annual Energy Balance in the Republic of Kosovo for 2015*. Republic of Kosovo. https://ask.rks-gov.net/media/2412/annual-energy-balance-in-the-republic-of-kosovo-for-2015.pdf.

———. 2019a. *Economic Statistics: Statistical Repertoire of Enterprises in Kosovo, Q4 2019*. Republic of Kosovo. https://ask.rks-gov.net/media/5283/rep-stat-mbi-ndër-ekono-në-kosovë-tm1-2019-shqip.pdf.

———. 2019b. *Labor Force Survey Q3 2019*. Republic of Kosovo. https://ask.rks-gov.net/media/5250/louboure-force-survey-q3-2019.pdf.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2019. *Does Household Electrification Supercharge Economic Development?* Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w26528.

Matkovic Puljic, Vlatka. 2016. *The Unpaid Health Bill—How Coal Power Plants in the Western Balkans Make Us Sick*. Health and Energy Alliance (HEAL). https://www.researchgate.net/publication/309615389_THE_UNPAID_HEALTH_BILL_-_How_coal_power_plants_in_the_WESTERN_BALKANS_make_us_sick.

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1): 159–217. https://doi.org/10.1111/j.1468-0262.2004.00481.x.

Millennium Challenge Corporation. 2013. *Guidance for MCC Economic Analysis of Power Sector Projects*. Version date: November 18.

———. 2017. *Threshold Program Grant Agreement between the United States of America and the Republic of Kosovo*. https://assets.mcc.gov/content/uploads/tpaa-kosovo.pdf.

———. 2018a. *Reliable Energy Landscape Project (RELP) Final Design Report*. 18.

———. 2018b. *Kosovo Labor Force and Time Use Study*. 18. https://millenniumkosovo.org/wp-content/uploads/2018/11/MCC-Kosovo-Labor-Force-and-Time-Use-Study-Final-Research-Report-1.pdf.

Millennium Foundation of Kosovo. 2018. *Monitoring and Evaluation Plan*. Version 2. Funded by Millennium Challenge Corporation.

———. 2017. *Kosovo Constraints Analysis*. https://assets.mcc.gov/content/uploads/Constraints-Analysis-Kosovo-1.pdf.

Ministry of Economic Development. 2011. *Republic of Kosovo Heating Strategy 2011–2018*. Republic of Kosovo.

———. 2012. *Mining Strategy of Republic of Kosovo 2012–2015*. Republic of Kosovo.

———. 2016. *Third National Plan of Action for Energy Efficiency (NEEAP) in Kosovo*. Republic of Kosovo.

———. 2017. *Republic of Kosovo Energy Strategy 2017–2026*. Republic of Kosovo.

———. 2018. *Renewable Energy Progress Report of the Republic of Kosovo 2016–2017*. Republic of Kosovo.

Ministry of Education, Science, and Technology. 2018. *Education Statistics 2017-2018*. Republic of Kosovo. https://ask.rks-gov.net/media/4145/stat-e-arsimit-2017-2018-shqip.pdf.

Ministry of Energy and Mining. 2011. *Kosovo Efficient Energy Action Plan 2010–2018*. The Republic of Kosovo. http://mzhe-ks.net/repository/docs/Kosovo_Energy_Strategy.pdf.

Munshi, Kaivan, and Jacques Myaux. 2006. "Social Norms and the Fertility Transition." *Journal of Development Economics* 80 (1): 1–38. https://doi.org/10.1016/j.jdeveco.2005.01.002.

Mockrin, S. C., L. D. Byers, and D. E. Koshland. 1975. "Subunit Interactions in Yeast Glyceraldehyde-3-Phosphate Dehydrogenase." *Biochemistry* 14 (25): 5428–37. https://doi.org/10.1021/bi00696a008.

Penfold, Robert B., and Fang Zhang. 2013. "Use of Interrupted Time Series Analysis in Evaluating Health Care Quality Improvements." *Academic Pediatrics* 13 (6): S38–44. https://doi.org/10.1016/j.acap.2013.08.002.

Qehaja, Anila. 2015. *Cost Benefit Analysis of Residential Insulation in Kosovo*. Thesis. Rochester Institute of Technology (RIT) Scholar Works 105.

Rewald, Rebecca. 2017. *Energy and Women and Girls: Analyzing the needs, uses, and impacts of energy on women and girls in the developing world*. Oxfam. https://www.oxfamamerica.org/static/media/files/energy-women-girls.pdf.

Reyna, Janet L., and Mikhail V. Chester. 2017. "Energy Efficiency to Reduce Residential Electricity and Natural Gas Use under Climate Change." *Nature Communications* 8 (1): 14916. https://doi.org/10.1038/ncomms14916.

Rapanos, Vassilis T., and Michael L. Polemis. 2005. "Energy Demand and Environmental Taxes: The Case of Greece." *Energy Policy* 33 (14): 1781–88. https://doi.org/10.1016/j.enpol.2004.02.013.

Rosenzweig, Mark R., and T. Paul Schultz. 1989. "Schooling, Information and Nonmarket Productivity: Contraceptive Use and Its Effectiveness." *International Economic Review* 30 (2): 457-477.

Rud, Juan Pablo. 2012. "Electricity Provision and Industrial Development: Evidence from India." *Journal of Development Economics* 97 (2): 352–67. https://doi.org/10.1016/j.jdeveco.2011.06.010.

Sahiti, Naser. 2012. *National Background Report on Energy for Kosovo*. 70. https://danube-inco.net/object/document/10210/attach/0_National_Background_Report_Energy_Kosovo_2012.pdf.

Sanzillo, Tom, and David Sclissel. 2016. *The Proposed New Kosovo Power Plant*. Institute of Energy Economics and Financial Analysis (IEEFA).

Schleich, Joachim, and Edelgard Gruber. 2008. "Beyond Case Studies: Barriers to Energy Efficiency in Commerce and the Services Sector." *Energy Economics* 30 (2): 449–64. https://doi.org/10.1016/j.eneco.2006.08.004.

Seastrom, M. 2017. *Best practices for determining subgroup size in accountability systems while protecting personally identifiable student information*. Washington, DC: U.S. Department of Education, Institute of Education Sciences. http://ies.ed.gov/pubsearch

Sorrell, Steve. 2015. "Reducing Energy Demand: A Review of Issues, Challenges and Approaches." *Renewable and Sustainable Energy Reviews* 47 (July): 74–82. https://doi.org/10.1016/j.rser.2015.03.002.

Ulriksen, Marianne S., and Nina Dadalauri. 2016. "Single Case Studies and Theory-Testing: The Knots and Dots of the Process-Tracing Method." *International Journal of Social Research Methodology* 19 (2): 223–39. https://doi.org/10.1080/13645579.2014.979718.

Vasquez, Claudia, Rhedon Begolli, Linda Van Gelder, and Sameer Shukla. 2018. *Western Balkans: Directions for the Energy Sector*. Washington, DC: World Bank Group. http://documents.worldbank.org/curated/en/201391544823541838/Western-Balkans-Directions-for-the-Energy-Sector.

Wagner, A. K., S. B. Soumerai, F. Zhang, and D. Ross-Degnan. 2002. "Segmented Regression Analysis of Interrupted Time Series Studies in Medication Use Research." *Journal of Clinical Pharmacy and Therapeutics* 27 (4): 299–309. https://doi.org/10.1046/j.1365-2710.2002.00430.x.

Walle, Dominique van de, Martin Ravallion, Vibhuti Mendiratta, and Gayatri Koolwal. 2015. "Long-Term Gains from Electrification in Rural India." *The World Bank Economic Review*. https://doi.org/10.1093/wber/lhv057.

Warrander, Gail, Verena Knaus, Ivan S. Abrams, Maria Randell, and David Priestley. 2014. *Kosovo: The Bradt Travel Guide*. Bradt Travel Guides. Guilford, Connecticut: The Globe Pequot Press.

Wilson, Alex. 2015. *Energy Community: Prospects and Challenges*. European Parliamentary Research Service.

World Bank. 2011. Review of *Background Paper: Development and Evaluation of Power Supply Options for Kosovo*. Prepared by DHInfrastructure.

———. 2015. *The World Bank in Kosovo: country snapshot 2015*. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/476691486988988059/The-World-Bank-Group-in-Kosovo-country-snapshot-201.

———. 2016. *Doing Business 2017: Equal Opportunity for All*. Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-0948-4.

———. 2017a. *Republic of Kosovo Systematic Country Diagnostic*. Washington, DC: World Bank. https://doi.org/10.1596/26573.

———. 2017b. *Energy Sector Management Assistance Program. Assessing and Measuring the Performance of Energy Efficiency Projects*. Technical Report 011/17. Washington, DC: World Bank.

———. 2018. *Energy in Kosovo*. Washington, DC: World Bank. Last updated October 2018. https://www.worldbank.org/en/country/kosovo/brief/energy-in-kosovo.

Appendix A: HER Iterative Study Design

In this section, we detail the design of the HER iterative study, based on information reported in the IC monitoring, evaluation, and learning plan for the SEEK activity. Section A.1 details the incentive models originally proposed, and Section A.2 outlines the study sample for the iterative study. Finally, Section A.3 outlines the selection of the optimal incentive model.

A.1 Incentive Models

The IC originally proposed the following six incentive models for the HER sub-activity:

- **Model 1.** A grant is given in the form of *in-kind* products and materials, which are purchased at wholesale. The household members participate with their *own work* and by paying for supervision and part installation to the qualified installers (QI) they selected. The key variables available for iterative testing are the incentive level (proportion of products/materials to be provided), and the specific marketing and/or knowledge dissemination campaign about do-it-yourself renovation of houses/apartments.
- **Model 2.** A grant is given in the form of *in-kind* products and materials purchased at wholesale. The household participates by paying for installation *through the QI*. The key variables available for iterative testing are the incentive level (proportion of products/materials to be provided), and the specific "Best Price" marketing campaign.
- **Model 3.** *Cash* reimbursement is provided for the approved part of total program costs (products, materials, and works) implemented by the QI selected by the household. The remaining part of the total cost of installation may be co-financed by the related municipality. The key variables available for iterative testing are the cash incentive level (proportion of total retrofit-program costs to be reimbursed), as well as the general SEEK marketing campaign.
- **Model 4.** *Cash* reimbursement is provided for the approved part of total program costs (products, materials, and works) implemented by the QI selected by the household. The remaining part of the total cost of installation may be financed through a custom loan from a partner financing institution. The key variables available for iterative testing are the cash incentive level (proportion of total retrofit-program costs to be reimbursed), as well as the specific marketing campaign on lending options from the partner financing institution.
- **Model 5.** *Cash* reimbursement is provided for the approved part of total program costs (products, materials, and works) implemented by the QI selected by the household. The remaining part of the total cost of installation will be financed through a custom loan from a partner microfinance institution. Similar to Model 4, the key variables available for iterative testing are the cash incentive level (proportion of total retrofit-program costs to be reimbursed), as well as the specific marketing campaign of the partner microfinance institution.
- **Model 6.** Under this vendor incentives program, a *grant* is given to households that implement energy efficiency measures through selected vendors, with installation done by a QI also selected by the household. The key variables available for iterative testing are the cash incentive level provided to the vendor (proportion of total retrofit-program costs to be reimbursed), as well as the specific "Best Quality" marketing campaign.

According to the "HER Operations Manual," of the six models identified above, only one (similar to Model 3) will be tested in the first iteration of the HER sub-activity. According to this model, MFK will co-finance part of the total program cost by providing a nonreturnable grant. In the second iteration, one to three models will be tested. It has not been decided yet which models will be tested in the second iteration.

A.2 Study Sample

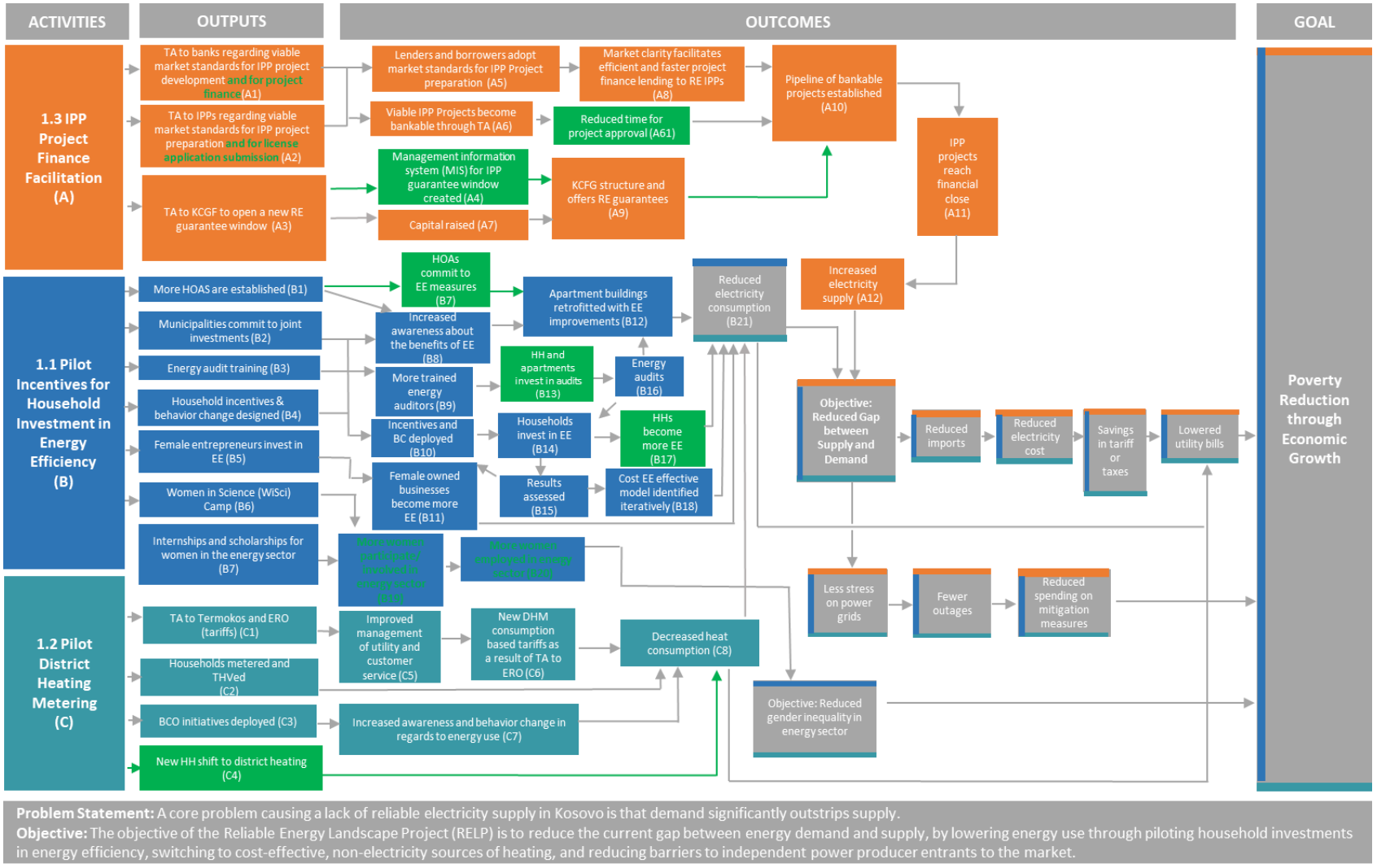
1. The first iteration of HER will target at least 500 house or apartment retrofits throughout Kosovo. The households will be selected to include the following factors:
    - Participation of homes with number of rooms representative of the Kosovo population (71 percent with as many as four rooms and 29 percent with more rooms)
    - Distribution of homes in accordance with the Kosovo population (about 57 percent rural and 43 percent urban)
    - Inclusion of poor, vulnerable, female-headed, and minority households in shares that are representative of the Kosovo population (about 7 percent, 9 percent, 11 percent, and 7 percent, respectively; some are included in more than one of these categories)
    - One hundred homes from each of the five regions defined by the SEEK program (Pristina, Ferizaj-Gjilan, Prizren, Peja-Gjakova, Mitrovica)
    - Identification of a counterfactual control group of at least 10 homes (10 percent of the iteration's retrofit targets) in each of the 5 defined regions that will not benefit from the HER sub-activity. The numbers of poor, vulnerable, female-headed, and minority households, as well as the rural/urban distribution and the range of number of rooms, will be representative of the Kosovo population.
2. The second iteration of HER will target 2,100 households throughout Kosovo.

A.3 Selection of Optimal Incentive Model

At the end of the iterative study, the IC will evaluate the overall performance and lessons learned, and then will propose the optimum model— or combination of incentives—for energy efficiency renovations of houses and apartments in Kosovo, according to various segments of the residential sector. Since the first iteration will involve only one model, the optimal model selection criteria will most likely be applied for the second iteration, when one to three models will be tested. The overarching criteria for this final selection will be:

- **Cost-effectiveness of models.** This is the key supply-side consideration. The limited available government budget must be used effectively to maximize the results and impact, not only for beneficiaries, but also for the national economy.
- **Demand for application of specific model designs.** This is the key demand-side consideration. The results and impact for the national economy will multiply if the manner in which the available government budget is used stimulates the maximum acceptance and interest of the beneficiaries.

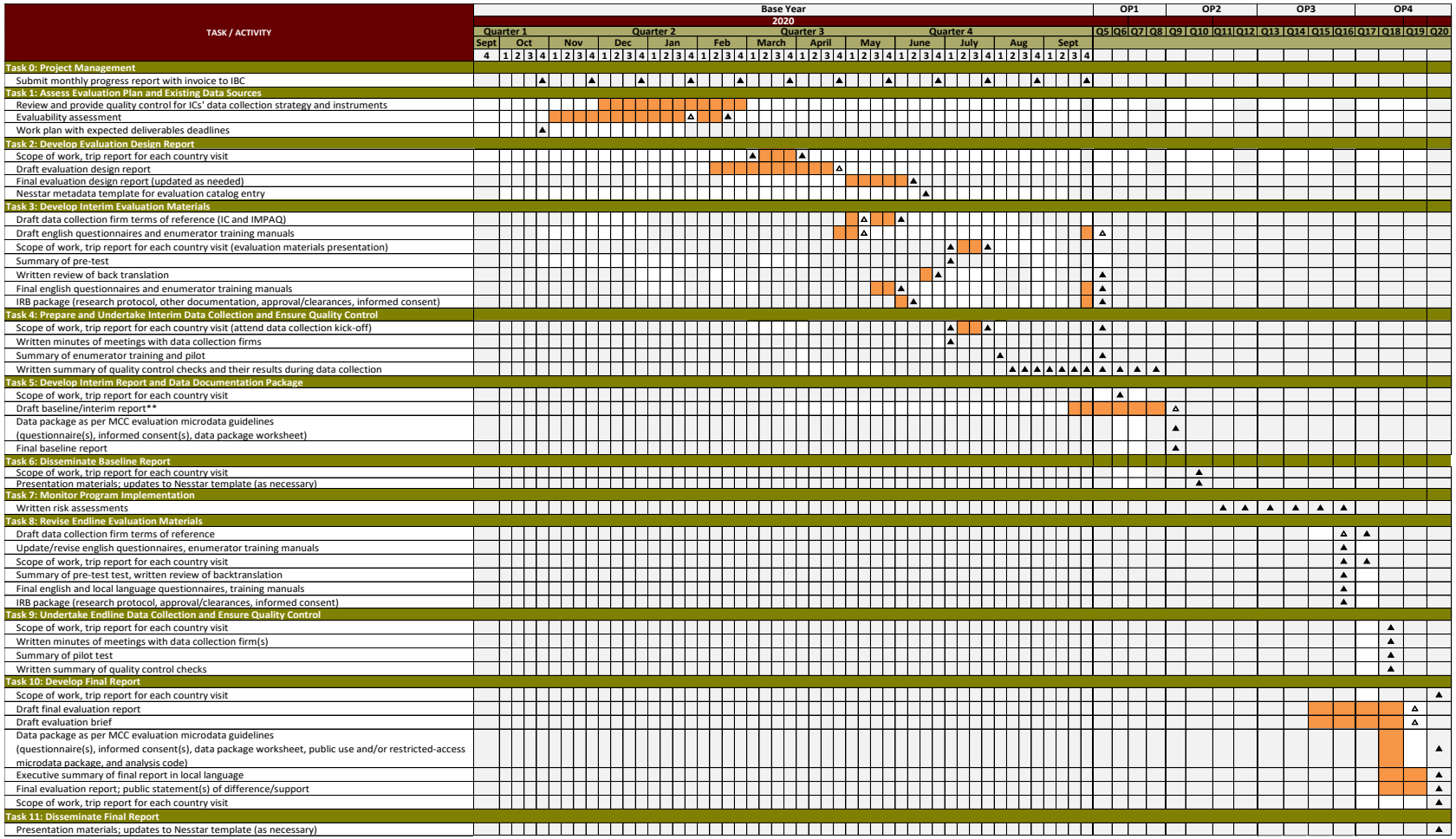
Appendix B: Suggested Changes to RELP Logic Model from IMPAQ's 'Evaluability Assessment'[66]



[66] BC: behavior change, BC&O: behavior change and outreach, EE: energy efficiency, RE: renewable energy, TA: technical assistance and THV: thermal heat valves.

Appendix C: Evaluation Work Plan

Exhibit 34. Evaluation Work Plan



Key: Draft △ Final ▲ Prep Work (IMPAQ)
Technical assistance to IC

***Submission of interim report is tied to the conclusion of baseline quantitative data collection*

Exhibit 35. Activity Timelines and Data Collection

Activity/Timeline	Ideal Exposure Period	Data Collection Source and Timing
HER, AER: ***Retrofits installation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)*** **DHM:** ***Meter and thermal valve installation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)***	▪ **Energy Consumption Outcomes** ○ Short-run impacts can be observed in heating season right after retrofit/meter installation ○ Medium-run impacts can be observed after at least 1 full heating season ○ Long-run impacts can be observed after at least 2 full heating seasons ▪ **Behavior Change Outcomes** ○ Short-run impacts can be observed after at least 1 full heating season ○ Medium-run impacts can be observed after at least 2 full heating seasons	▪ **Baseline** ○ Quantitative data: Month X – September 2021 ○ Qualitative data: After the first iteration of HER and before the beginning of the second iteration. Qualitative for DHM and AER will happen at the same time ▪ **Interim** ○ Qualitative data: 1. After the end of the threshold period 2. At the end of 1 heating season 3. At the end of 2 heating seasons ▪ **Endline** ○ Quantitative data: After 2 heating seasons and before the start of the third heating season ○ Qualitative data: After the endline quantitative data collection to put results in perspective
IPP ***Implementation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)***	▪ **Readiness of IPPs** ○ Observed at 6 month intervals to ascertain progress ▪ **Progress of Renewable Energy Projects** ○ Observed at 6 month intervals to ascertain progress ▪ **Longer-term Outcomes** ○ Observed approximately 2-3 years after threshold implementation	▪ **Baseline** 3-4 months after start of implementation ▪ **Interim** 1. After the end of threshold period 2. 6 months after the end of threshold 3. 1.5 years after the end of threshold ▪ **Endline** ○ Approximately 2 years after end of threshold
WEE 1. **Internships** 2. **Scholarships** 3. **WiSci Camp** 4. **Grants** ***Implementation in a phased manner from Month X – September 2021 (i.e. till end of Threshold)***	**Internships, Scholarships, WiSci Camp, Grants** ▪ Perception shift about women, satisfaction with internships/scholarship/camp/grants, knowledge acquired, employer perceptions, sources and consumption of energy, energy efficient measures invested etc. will be observed every 6 months to determine gradual shift in behavior	▪ **Baseline** 3-4 months after start of implementation ▪ **Interim** 1. After the end of threshold period 2. 6 months after the end of threshold 3. 1.5 years after the end of threshold ▪ **Endline** Approximately 2 years after end of threshold

Kosovo Reliable Energy Landscape Project Evaluation

Contract No. 95332419C0104

Evaluability Assessment

Submitted To:

Rabia Chaudhry, Project Officer
MCC - Headquarters
Millennium Challenge Corporation

Submitted By:

IMPAQ International, LLC
10420 Little Patuxent Parkway, Suite 300
Columbia, MD 21044

Date: January 24, 2020

TABLE OF CONTENTS

TABLE OF EXHIBITS

LIST OF ACRONYMS

BC&O	Behavior Change and Outreach
EBRD	European Bank for Reconstruction and Development
EU	European Union
HOA	Homeowners Associations
IC	Implementing Contractor
IPP	Independent Power Producer
KCGF	Kosovo Credit Guarantee Fund
MAB	Multi-apartment Buildings
MCC	Millennium Challenge Corporation
M&E	Monitoring and Evaluation
MFK	Millennium Foundation of Kosovo
PMC	Project Management Committee
RELP	Reliable Energy Landscape Project
SEEK	Subsidies for Energy Efficiency in Kosovo
STEM	Science, Technology, Engineering, and Mathematics
TA	Technical Assistance
THV	Thermal Heat Valve

SECTION 1. INTRODUCTION

On September 12, 2017, the U.S. Government's Millennium Challenge Corporation (MCC) and the Government of the Republic of Kosovo signed a $49 million program to spur economic growth and reduce poverty in Kosovo by September 2021.[67] To address one of the main constraints to Kosovo's economic growth—unreliable electricity supply—the MCC Kosovo Threshold Program designed the Reliable Energy Landscape Project (RELP). The objective of RELP is to reduce the current gap between energy demand and supply. On the demand side, RELP will address two barriers to investments in energy efficiency: the general lack of consumer awareness of energy saving measures and their benefits, including the benefits of metering; and for poor households, the lack of ability to pay for these measures. This will be achieved by lowering energy use through piloting household investments in energy efficiency, metering existing district heating consumers, and switching new consumers to cost-effective district heating. On the supply side, RELP will aim to reduce barriers to renewable energy entrants to the market, by stimulating the business environment for renewable energy production. There are three main activities under RELP.[68]

Activity 1.1 Subsidies for Energy Efficiency in Kosovo

The primary technical focus of the Subsidies for Energy Efficiency in Kosovo (SEEK) intervention packages are energy efficient measures (including improving thermal insulation in walls and roofs, energy efficient windows, weather sealing, energy efficient water heaters, and energy efficient biomass stoves and furnaces) that will reduce the consumption of electricity for heating and promote the use of energy efficient heating devices. SEEK aims to provide incentives for households and apartment buildings to invest in retrofits to reduce household energy consumption, in addition to incentives aimed at increasing the involvement of women in the energy sector.[69] Behavior change and outreach (BC&O) is a crucial component to achieve the intended objectives of this activity. It consists of three sub-activities:

1. **Household energy retrofits**. This will provide incentives to eligible residential consumers to invest in energy efficient retrofits. One of the key objectives is to test and evaluate the most cost-effective delivery approaches for incentives for different beneficiary groups. Another objective is to deliver the best behavior change approaches, which can be scaled up beyond the Threshold Program. Consequently, there is an iterative component to design the optimal intervention packages.[70]
2. **Apartment energy retrofits**. SEEK will provide grants to residential apartment communities in selected municipalities to implement energy efficient upgrades to common areas and the thermal envelope. To facilitate implementation, homeowners associations (HOA) will be established, and municipalities will co-invest in retrofits in apartment buildings.
3. **Women energy entrepreneur technical assistance**. This is designed to ensure equal economic opportunities for women in the energy sector. Approaches to achieve this include: (a) women energy entrepreneur technical assistance (TA) and grants, (b) paid internships and apprenticeships for young women in the energy sector, (c) academic scholarships for women interested in joining the energy sector, and (d) summer camps for teenage girls to foster their interest in the energy sector.

[67] MCC Kosovo Threshold Program Grant Agreement: https://assets.mcc.gov/content/uploads/tpaa-kosovo.pdf

[68] The project description in Section 1 is based on the latest *RELP M&E Plan* (2018). We received an updated theory of change in December 2019, which is used in Sections 3 and 4.

[69] Approximately 2,600 households and 25 apartment buildings will receive benefits under the SEEK activity.

[70] Based on conversations with MCC and the implementing contractor for household energy retrofits during the scoping mission, it is not clear if the iterative component will be retained.

Activity 1.2 Pilot District Heating Metering

This activity aims to support the transition of district heating residential customers on the Termokos network in Pristina from area-based billing to quantity-based billing through the installation of household meters. The energy saved due to this intervention is likely to improve the service for more than 12,000 existing consumers who already live in buildings connected to district heating services in Pristina (based on the RELP terms of reference). Further, the energy saved will enable expansion of district heating to new consumers. The number of new consumers reached will depend on the energy savings realized from installing meters in residences of existing district heating customers. Termokos and the Energy Regulatory Office will work collaboratively to design tariffs. Additionally, the project will test different modalities and strategies for packaging district heating metering and energy efficiency services to maximize customer value, reliability, and sector cost efficiency. This will entail feasibility work, including formative research to design a social and behavior change campaign.

Activity 1.3 Independent Power Producer Project Finance Facilitation

On the supply side, this activity addresses the problem of unreliable electricity supply by facilitating entry into the renewable energy market. This aim is to expand the Kosovo Credit Guarantee Fund's (KCGF's) mandate of establishing a new renewable energy independent power producer (IPP) guarantee window, and also to create a Kosovo Project Acceleration Program within the Millennium Foundation of Kosovo (MFK). In particular, TA will be provided to KCGF for fundraising, to banks for designing project finance products, and to IPPs for submitting applications. The activity foresees the tasks illustrated in **Exhibit 1**.

Exhibit 1. Independent Power Producer Activity Tasks

Developing a standardized framework for renewable energy project financing meeting international best practices in the context of Kosovo's energy market

Providing training to IPP developers, banks, KCGF, and local consultants on deliverable management of documents and fund-raising activities

Facilitating the development of a standardized format for bankable renewable energy IPP transactions for submission to partner banks

Establishing a Kosovo Project Acceleration Advisory Committee, consisting of key renewable energy stakeholders to develop the standardized framework for renewable energy financing

The purpose of this evaluability assessment is to lay the foundation for the evaluation design report by assessing five key dimensions of RELP: (a) problem diagnostic, (b) project objectives and logic, (c) risks and assumptions, (d) project participants and beneficiaries, and (e) accountability and learning metrics. For this assessment, we have relied on the following documents: (a) MCC's *Kosovo Constraints Analysis*, (b) *RELP Final Design Report*, (c) *RELP Monitoring and Evaluation (M&E) Plan* from 2018 (project logic and indicator definition table from December 2019), and (d) additional literature, including academic articles.

The rest of this report is organized as follows. **Section 2** examines if there is evidence to support the problem diagnostic. **Section 3** evaluates whether the project logic is clearly defined with adequate assumptions. **Section 4** examines the risks, assumptions, and risk mitigation strategies. **Section 5** evaluates whether participants are clearly defined and justified in terms of geographic scope and eligibility criteria. Finally, **Section 6** assesses if metrics for monitoring and measuring results are adequately defined.

SECTION 2. PROBLEM DEFINITION AND DIAGNOSTIC

In 2017, MFK conducted a comprehensive analysis of constraints that hinder sustainable economic growth and poverty alleviation in Kosovo. The constraints analysis examined issues related to historic institutions, geographic position of the economy, human capital and health, infrastructural challenges, financial access, and macroeconomic risks. The analysis identified three main binding constraints to economic growth and poverty alleviation: (a) unreliable supply of energy; (b) weak rule of law and perceptions regarding the rule of law; and (c) poor environmental services (water, environment, and health).[71] We evaluate whether the key problem that led to the conceptualization of RELP, i.e., that the electricity supply is inadequate to meet end-user demand, is supported by concrete evidence. We evaluate the supporting evidence in the following three pillars:

Pillar 1: Evidence of a binding constraint. The constraints analysis clearly outlines that a key contributor to poor economic growth in Kosovo is a large gap between electricity supply and demand. Based on the work by Haussmann, Rodrik, and Velasco (2005), there are four broad tests identified in the growth diagnostics methodology, for a constraint to qualify as a "binding" constraint to growth.[72,73]

The constraints analysis comprehensively shows that the first test is met, i.e., that the shadow price of the constraint should be high in comparison with other comparable countries. Specifically, while electricity prices and industrial energy prices in Kosovo are similar to those in comparable countries, Kosovo lags behind in energy efficiency.[74] Electricity consumption levels and the share of electricity bills in household expenditures is higher than in other comparable countries. Furthermore, using outages (a more reliable measure of the shadow cost of electricity), Kosovo ranks the worst. However, the constraints analysis does not mention Kosovo's absence of natural gas supply and its inability to import the commodity (due to lack of infrastructure) as a binding constraint (Kosovo Energy Strategy, 2017-2026).[75] Kosovo's inability to import gas, and thus diversify its energy supply, can potentially be a significant constraint: it limits the country's energy options and reduces the country's ability to keep costs down.

The second test is that changes in the constraint should be correlated with growth and investments. Here, the constraints analysis provides evidence about the value of gross domestic product (GDP) per capita lost due to electrical outages per capita. According to the constraints analysis, outages cost around 5% of GDP annually. The constraints analysis highlights losses to both businesses and households due to unreliable energy supply. Specifically, Kosovo businesses incur EUR 296 million in *additional* costs annually, due to irregular energy supply. Low voltage also jeopardizes household well-being, including spoiling goods and

[71] Binding constraints are those that, if relieved, would produce the largest gains in growth and entrepreneurship of any potential constraint areas.

[72] Haussmann, Rodrik, and Velasco (2005) lay out a theoretical framework wherein they divide the factors affecting growth into two categories: high cost of financing and low return to domestic investment. They suggest that identifying binding constraints to growth is important, since governments face administrative and political limitations and hence, must prioritize strategies.

[73] The four tests include: (a) the shadow price of the constraint should be high with respect to comparator countries; (b) changes in the constraint should be correlated with growth and investment; (c) economic agents should be working around the constraint, and (d) sectors or industries that are dependent on the constraint should have lower growth.

[74] Electricity prices are usually below cost recovery, since governments explicitly or implicitly subsidize prices. Thus, prices in Kosovo and other comparable countries are similar.

[75] The Trans Adriatic Pipeline, connecting Greece, Albania, and Italy to supplies from Azerbaijan, is due to be completed in 2020. In theory, a spur could have been added to link up Kosovo, but for various reasons, was not feasible.

harming costly electronic items, and the constraints analysis clearly articulates this with quantitative evidence (World Bank, 2017).

The third test to qualify as a binding constraint is that economic agents should be working around the constraint. The constraints analysis provides evidence that this test too is satisfied, by showing that Kosovo solves its energy generation gap by importing electricity. Importing electricity is costly and price per megawatt hour averages around EUR 55 (compared with the in-country price per megawatt hour of EUR 28).

The fourth test to identify a binding constraint is that sectors or industries dependent on the constraint should have lower growth. On this front, the constraints analysis does not provide sufficient evidence about the sectors and industries most affected by an unreliable energy supply and the resultant energy costs of these industries. Using the World Bank *Doing Business* report (2013) as a benchmark, the constraints analysis shows that smaller firms are worse off due to high energy costs. However, a more detailed outlining of which sectors and industries depend most on energy and their respective growth rates (compared with sectors that do not depend as much on energy), is needed. Additionally, it is essential to discuss how the current district heating tariff structure (including fixed and variable components) affects consumption levels and expenditures of economic agents. This will allow us to understand the extent to which changes in these tariffs through RELP will affect residents.

Pillar 2: Institutional context and local political economy. The constraints analysis comprehensively defines the problem (an unreliable electricity supply) and provides sufficient evidence to support the problem diagnostic. However, there is limited discussion about how, given the constraints imposed by the unreliable supply, an exclusive focus on renewable energy as proposed under RELP is the correct approach. Specifically, political will is required to transition completely to renewable energy.

On the one hand, a background paper by the World Bank (2011) highlights that Kosovo's growing energy needs can only be met by building new coal plants that leverage the country's large reserves of lignite coal.[76,77,78] Additionally, the potential for wind energy appears limited. A 2010 study, funded by Switzerland's Renewable Energy and Energy Efficiency Promotion in International Cooperation platform, found that there were very few areas with wind speeds exceeding 6 meters per second, a minimum needed for commercial potential in the region (World Bank, 2011). Other studies have estimated a solar power potential of 77 megawatts (MW) (at high costs), a small hydropower potential of 64MW, and a large hydropower potential of 305MW. As highlighted in the constraints analysis, Kosovo's power demand is expected to reach 1500MW by 2025, only a fraction of which can be met with renewable sources.

On the other hand, Kammen et al. (2012) highlight that a coal-dominated future is not sustainable for Kosovo. The authors also recognize that the economic costs of frequent power outages are equivalent to the pollution risks faced by Kosovo residents if more coal fired plants were to be opened. A low carbon path that integrates aggressive deployment of energy efficiency and uses large and small-scale hydropower, solar, and biomass power will deliver 34% of total energy demand and 60% more jobs by

[76] Kosovo has the fifth-highest reserves of lignite coal in the world (World Bank, 2011). Domestic lignite reserves are estimated to amount to 12.5 billion tons, of which 10.9 billion tons are exploitable (Kammen et al., 2012).

[77] Even assuming all required renewable energy capacity is built by 2017, there would still be a supply-demand gap of almost 1,000MW by 2025 (World Bank, 2011). A mix of renewable and thermal energy would be needed to meet Kosovo's long-term energy demand.

[78] While the World Bank did not move forward with building new coal plants, in the Kosovo Energy Efficiency and Renewable Energy Project funded by the Bank, one key activity is rehabilitation of the Kosovo B plant.

2025. According to the authors, this option will result in costs savings of 5% to 50% relative to a baseline scenario that includes a new coal power plant. However, these estimates assume a high level of consumer willingness and ability to invest in energy efficient measures, evidence for which is not available.

Based on available evidence, these two studies recommend vastly different energy choices for Kosovo. These critical energy and developmental tradeoffs facing Kosovo have not been sufficiently highlighted in the constraints analysis. The choice of renewable energy is not straightforward, and more evidence is needed about the political will to adopt renewable energy, as outlined in RELP. From the constraints analysis and the *RELP Final Design Report*, it is unclear where RELP sits in relation to other initiatives by the government, the government's overall energy strategy, and the strategies of other donor agencies.[79] To meet its energy strategy, the Government of Kosovo has requested financing support from the World Bank for expansion of renewable energy sources in public buildings, after the decommissioning of the Kosovo A coal-fired power plant.[80,81] Additionally, the European Bank for Reconstruction and Development (EBRD) has funded two large sustainable energy projects in Kosovo: Green Economy Financing Facility and Kosovo Sustainable Energy Project (KoSEP). The EBRD projects finance energy efficiency measures in apartment buildings, by collaborating with local financial institutions, including TEB Bank, to offer subsidized credit.[82] In this context, there are two key questions: (a) will RELP's focus on energy efficiency in apartment buildings complement or supplement the World Bank and EBRD projects, and (b) will there be coordination between RELP and the EBRD project, to avoid overlap?

Finally, the *RELP Final Design Report* and *M&E Plan* do not adequately address several political economy concerns that RELP could potentially confront. The first concern is Kosovo's inability to participate in the European Network of Transmission System Operators for Electricity (ENTSO-E), which is held up by disagreements over Kosovo's official name (as mentioned in the constraints analysis). This increases the length of time between an unexpected outage and transmission system response. RELP as designed does not foresee big shifts in renewable energy supply. This implies that Kosovo will continue to depend on its own coal-fired power plants and/or electricity imports.[83] For the latter, entry to ENTSO-E is vital, and if that doesn't happen, the problem of unreliable electricity supply could persist. Second, Kosovo continues to be plagued with corruption, and perceptions of corruption, which deter foreign direct investment in vital sectors, particularly electricity. Finally, many funders (including the World Bank) have pulled out of

[79] The Government of Kosovo is pursuing an energy strategy that is a combination of: (a) developing a modern and highly efficient thermal generation facility, using best available techniques and meeting EU environmental standards; (b) decommissioning the 50-year old Kosovo A power plant; (c) rehabilitating the Kosovo B power plant to meet the requirements of the EU's Industrial Emissions Directive; (d) developing renewable energy sources, including small hydro, wind, and biogas; (e) improving energy efficiency in public buildings and incentivizing efficiency in the private sector; (f) promoting integration of regional markets; (g) enhancing efficiency of electricity distribution; and (h) building institutional capacity within the energy sector.

[80] Specifically, the World Bank has the Kosovo Energy Efficiency and Renewable Energy Project, which has been active since 2014 and is set to end in 2020.

[81] A third 500-MW coal-fired power plant, the New Kosovo Plant, is being planned. ContourGlobal, a United Kingdom-based power generator, will build it. Funding for the plant was secured from General Electric after the World Bank pulled out. Construction was expected to begin in late 2019.

[82] Additional information on the Green Economy Financing Facility can be found at https://ebrdgeff.com/kosovo/ebrd-launches-energy-efficiency-framework-for-homes-in-kosovo/

[83] https://www.forbes.com/sites/davekeating/2018/03/13/how-an-electricity-dispute-in-kosovo-made-people-across-europe-late-to-work/#4db94cbe19c9

investing in Kosovo's new lignite powered coal plant, citing environmental concerns.[84] This could have implications for future relationships with these funders for renewable energy investments in Kosovo.

Pillar 3: Proposed intervention and resolution of root causes. The constraints analysis identifies "unreliable electricity supply" as the binding constraint for economic growth and the *Final Design Report* identifies "excessive household energy use" as the root cause of this problem. The *Final Design Report* provides evidence about the disparity in energy intensity between Kosovo and other comparable countries. Specifically, if households in Kosovo were heated to standard indoor conditions, energy intensity would be 300 kilowatt-hours per square meter per year (kWh/m^2/year), almost double the European Union (EU) average of 174 kWh/m^2/year (Kokx, 2017). Since higher energy intensities indicate a higher price or cost of converting energy into GDP, this implies that Kosovan households are heating inefficiently. In 2016, Kosovo imported 555 gigawatt hours of electricity at a cost of EUR 47 million—more than 40% higher than the cost of domestic electricity consumption (MCC, 2018a). The *Final Design Report* comprehensively documents the extent to which current energy consumption costs have a negative impact on household budgets. While the report also outlines the extent to which RELP activities would result in utility cost savings because of the retrofits, utility cost savings in the report are based on full uptake and hence, there is a need to provide more conservative estimates that assume lower uptake rates. The *Final Design Report* also does not provide sufficient evidence about how energy efficient measures would reduce household energy consumption in Kosovo. That is, the extent to which energy efficient measures would reduce consumption would depend on the elasticity of energy demand (which correlates with income) and whether households are under-heating their homes. Hence, the energy efficient measures could actually lead to an increase in energy consumption. Alternatively, the energy efficient measures could simply lead to warmer apartments at the *same* expenditure levels. Further, even if households were aware of the benefits of energy efficiency, it is not straightforward that they would be willing to pay for it. Additionally, there is limited focus on the longer-term sustainability of the energy efficient measures by households and apartments, either in the form of mandates or in the form of medium- to long-term repair/maintenance of energy efficient equipment.

Finally, while both the constraints analysis and the *Final Design Report* identify insufficient electricity supply as the key problem, it is not clear how inefficient electricity consumption is the "root cause" of this problem, especially since only 5% to 10% of households use electricity for heating in Kosovo and the majority use firewood. The main causes of insufficient electricity supply as highlighted above include: (a) the institutional context and political economy; (b) Kosovo's inability to import gas; and (c) the continual reliance on coal-fired power plants, which has led to funding limitations because of environmental concerns. Looking ahead, from the perspective of RELP, the constraints analysis developed to conceptualize RELP comprehensively outlines that reduction of the gap between electricity demand and supply will be one of the key levers to kick-start economic growth. However, it is unclear whether the focus on renewable energy and energy efficiency under RELP would be sufficient to address an important driver of poverty in Kosovo: an unreliable electricity supply and uncertainty about whether there is sufficient political will to focus exclusively on renewable energy.[85]

[84] See for instance, https://www.reuters.com/article/us-contourglobal-kosovo-ge/ge-to-build-kosovos-new-500-mw-coal-power-plant-idUSKCN1S917R

[85] Even wealthier countries like Germany have not been able to transition completely to a low-carbon energy system and coal still accounts for 40% of Germany's electricity supply (World Bank, 2015).

SECTION 3. PROJECT OBJECTIVES AND THEORY OF CHANGE

In this section, we present and critically evaluate RELP's objectives and theory of change. We will work backwards to identify the extent to which the current program logic clearly links activities and short-term outputs to RELP's overall goals. The RELP theory of change lays out the causal pathways through which short-term outputs, long-term outcomes, and eventual project goals are achieved. Theories of change are important for M&E processes, and their critical assessment is important to understand the extent to which project activities are appropriate to achieve the project's key objectives (Corlazzoli & White, 2013).

3.1 Assessing Project Objectives

The overarching goal of the Kosovo MCC Threshold Program is to achieve poverty reduction through economic growth. According the *M&E Plan*, RELP specifically, has four broad goals: (a) reduce the overall cost of electricity supply in Kosovo through reduction of expensive imports and reduction of stress on distribution systems; (b) reduce gender inequity in the energy sector; (c) improve overall economic growth in Kosovo; and (d) put in place district heating metering in the Pristina district heating network as a prerequisite to significantly scaling up district heating networks in Kosovo and other benefits such as reducing emissions from dirty fuels, and improving power sector efficiency through waste heat use. RELP aims to achieve these objectives by lowering energy use through piloting household investments in energy efficiency (household efficiency retrofits and apartment building efficiency retrofits), switching to cost-effective district heating, and reducing barriers to IPP entrants to the market.

The main objectives of RELP stem directly from the identified problem diagnostic in the RELP *Final Design Report*, wherein the demand for electricity significantly outstrips supply of electricity in Kosovo. Thus, the project objectives flow coherently from the problem diagnostic. Residential electricity use accounts for around 60% of total electricity used in Kosovo. Household electricity demand is mainly driven by household demand for outputs like heating, water heating, lighting, cooling, and cooking. As highlighted in the previous section, Energy efficient investments are expected to produce the same level of output from a reduced number of units of energy input. RELP aims to address two root causes and barriers to investments in energy efficiency: the general lack of consumer awareness of energy saving measures and their benefits, and the lack of ability or willingness to pay for them. The proposed approach to address these root causes involves a mix of awareness raising, incentive piloting (to make investments affordable), regulatory support, and technical assistance/capacity building.

3.2 Assessing Project Theory of Change

RELP's theory of change focuses on a combination of renewable energy investments on the supply side and energy efficiency incentives on the demand side to achieve the project's main goals. It is increasingly recognized that both renewable energy and energy efficiency play a vital role in the rapid transformation of the energy sectors of industrialized, emerging, and developing countries (Hsu, Rosengartin, Furter, & Xie, 2017). In this regard, the *RELP Final Design Report* clearly lays out the *potential* for energy efficient investments in Kosovo in the residential sector to lower energy use. Specifically, according to a 2015 World Bank report titled "National Building Energy Efficiency Study for Kosovo," the investment potential for feasible residential energy efficiency measures was over EUR 500 million (Kokx, 2017).

However, there is little *empirical evidence* outlined in the *M&E Plan* and the *RELP Final Design Report* on how well energy efficiency measures actually work in terms of reducing energy use. A review of the literature confirms that while there is widespread consensus in the literature that the residential sector is

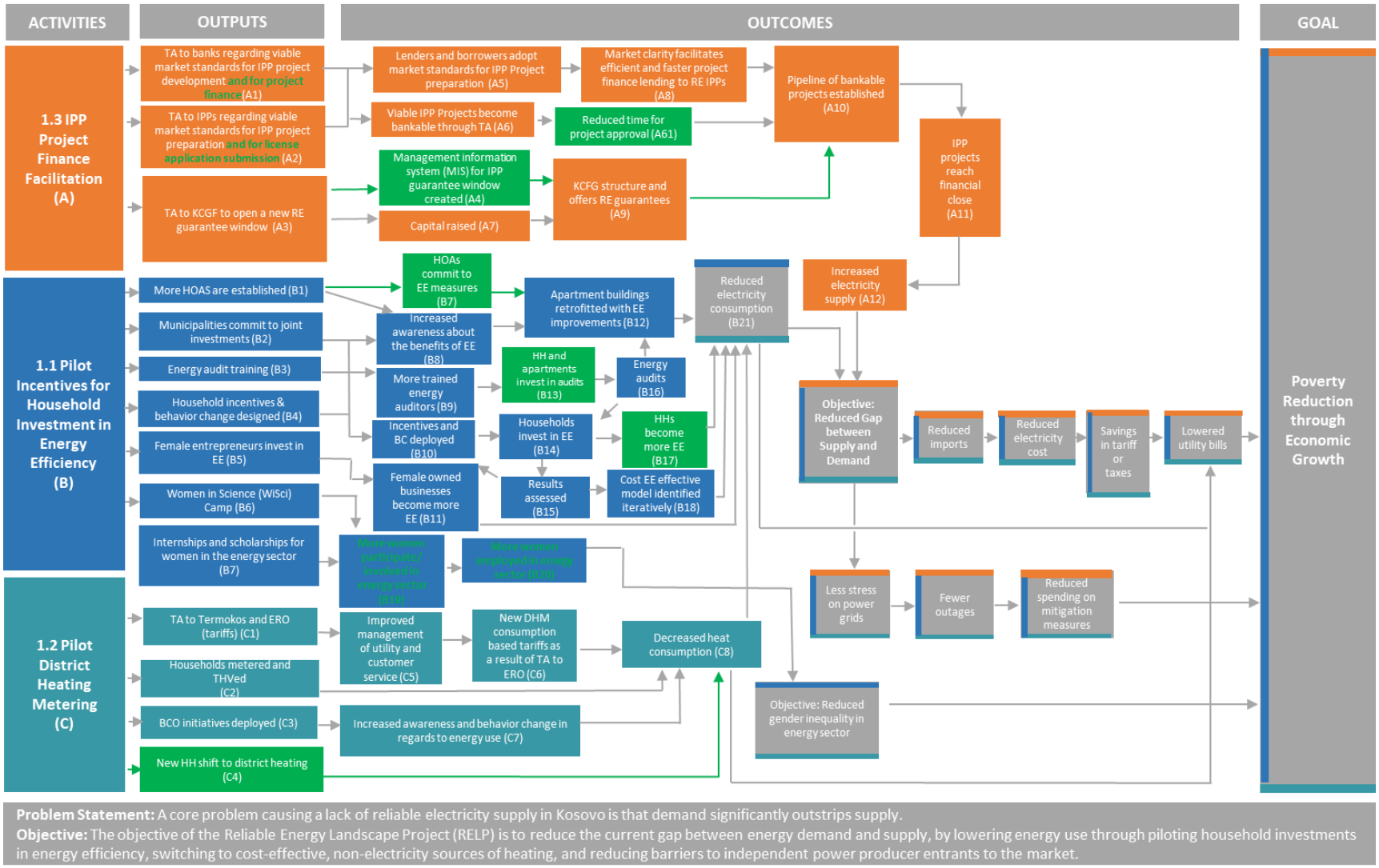
one of the most significant single sectors for energy consumption presenting high cost-efficient potentials for mitigation, recent years' experience has shown that there are considerable barriers to fully realize economically effective and technically feasible energy savings opportunities (Gillingham & Palmer, 2014; Frederiks et al., 2015a; Knoop & Lechtenböhmer, 2017). Further, improvements in energy efficiency do not regularly lead to one-to-one reductions in energy consumption (Galvin, 2014), as energy efficiency gains alter the perceived cost of comfort and may thereby generate shifts in consumption patterns—a 'rebound effect' (Aydin et al., 2017). This discrepancy between the expected/realized energy savings and the optimal/actual investments in energy efficient technologies is often referred to as the 'energy efficiency gap' or 'energy efficiency paradox', which has been illustrated and examined in multiple studies (Schleich & Gruber, 2008; Chai & Yeo, 2012; Allcott & Greenstone, 2012; Ameli & Brandt, 2015; Gerarden et al., 2017). That is, households and apartment buildings in the SEEK sub-activity might actually increase their energy consumption after the installation of retrofits. Similarly, households in the district heating metering sub-activity who receive new meters and thermal valves might also start heating larger areas of their households because of meter induced energy savings. Thus, RELP needs to elucidate clearly, how the specific sub-activities of SEEK will lead to unequivocal energy savings and reduced electricity consumption (leading to the objective of reduced gap between supply and demand).[86]

In this context, as outlined in the *M&E Plan*, there is commitment to iterate and understand what works to reduce electricity consumption through an iterative study design. The *M&E Plan* stresses the use of data collection in the iterative design cycle of the RELP to systematically analyze and revise pilot models for packaging incentives for energy efficiency project investments in the residential sector. The engine driving the experimentation cycle is a unique intra-project monitoring, evaluation, and learning framework within the RELP. We use the latest version, from December 2019, of RELP's logic model. The project logic provides a useful framework for understanding how the project's three main activities are expected to lead jointly to the desired impacts. However, some aspects of the logic model lack clarity, and some important assumptions are missing. Further, the latest version of the logic model is very complex in terms of a number of nodes and interdepencies between them. Further, for each goal, a number of outcomes need to be achieved before achieving that goal.

Suggested revisions to the theory of change are marked in green. In **Exhibit 2**, we describe the revisions we recommend to the logic model for the evaluation and justifications for each revision below. Our revisions center on clarifying the links between incentives to households, apartment buildings, and women entrepreneurs and reduced electricity consumption. The theory of change is least developed for the sub-activities centered around reducing gender inequality in the energy sector since it is unclear how the separate activities will lead to broader sectoral level changes. We present revisions separately for each of the three main activities of RELP.

[86] Additionally, it is not immediately clear how RELP situates itself among the other projects in the country. For instance, the World Bank and EU are working with Termokos to expand their capacity and meter new households. These investments will happen either prior to or concurrently with RELP. Thus, changes in energy consumption will not be directly attributable to RELP in a causal framework.

Exhibit 2. Reliable Energy Landscape Project Logic, With Suggested Revisions[87]



[87] BC: behavior change, BC&O: behavior change and outreach, EE: energy efficiency, RE: renewable energy, TA: technical assistance and THV: thermal heat valves.

Exhibit 3. Suggested Revisions to Reliable Energy Landscape Project Theory of Change

Revision	Explanation
Activity 1.3 IPP Project Finance Facilitation	
TA to IPPs	During the scoping mission it was emphasized that extensive TA will be provided to IPPs to help them prepare the license application to submit to the Energy Regulatory Office, to reduce the processing time for applications. We have made this more explicit by adding this to **Box A2**. Additionally, it is not clear what "viable market standards" means. There needs to be more clarity on what specific market standards are being referred to in particular and to what degree are IPPs unaware of them.
TA to banks	During the scoping mission, it was emphasized that extensive TA will be provided to banks including development of financial products focused on project finance. We have made this explicit in **Box A1**. Additionally, similar to IPPs, more clarity is needed on what market standards are being referred to for banks.
Reduced time for project approval	One of the main goals of the IPP sub-activity as emphasized during the scoping mission was to reduce the time it takes to submit a licensing application to the Energy Regulatory Office for approval and get it approved. We have made this more explicit by adding "Reduced time for project approval" in **Box A61.**
Renewable energy guarantee catalyzes investments in IPP projects and establishes a pipeline of bankable IPP projects	The latest logic model linked "KCGF structures and offers renewable energy guarantees" (**Box A9**) only to "IPP projects reach financial close" (**Box A11**). This implicitly assumes the renewable energy guarantees will catalyze investments in IPP projects and form a pipeline of bankable IPP projects. To make the link more explicit we have linked it (**Box A9)** to "Pipeline of bankable projects established" (**Box A10**), instead.
Management Information System (MIS) for IPP guarantee window	The latest logic model listed MIS for IPP guarantee window as a short-term output. However, setting up of the MIS is contingent upon TA provided to KCGF. To make this explicit we have moved the "MIS for IPP guarantee window" (**Box A4**) as an outcome to be measured after TA has been provided to KCGF.
Activity 1.1 Subsidies for Energy Efficiency in Kosovo	
Women in Science Camp leading to more women participating in the energy sector	In the logic model, participation in the WiSci camp (**Box B6**) will lead to more women participating in/involved in the energy sector (**Box B19**). It is unclear how this will be achieved. During the scoping mission, we found that WiSci camps would include sessions on all science, technology, engineering, and mathematics (STEM) fields, only one of which is energy. Given this, we recommend adding intermediate outcomes such as (a) stated level (1-5) of interest in the energy sector; and (b) change in percentage of women who enroll in energy related fields at various levels of schooling, to measure the extent to which the camps generated interest in the energy sector.
More women employed in the energy sector	In the latest logic model, internships, scholarships, and science camps will lead to more women being employed in the energy sector (**Box B20**). Higher employment levels of women in the energy sector is a long-term outcome, unlikely to be measurable in the next four years of RELP. Instead we recommend replacing this outcome in the logic model with more measurable short to medium-term outcomes such as: (a) stated level (1-5) of interest in the energy sector; (b) change in percentage of women who changed plans for sectors of study when leaving college (toward energy related fields); (c) change in percentage of women who enroll in energy related fields at various levels of schooling; (d) change in percentage of women who graduate in energy-related fields; and (e) change in percentage of women who apply for energy-related jobs.
More women participate/are involved in the energy sector	The current logic model has internships, scholarships, and science camps leading to "more women participate/involved in the energy sector" (**Box B19**). It is unclear how this will be monitored/measured. A clearer definition of what 'involved in the

	energy sector' implies is important. In particular, higher involvement could imply higher employment rates/higher applications for employment/higher number of teachers and counsellors willing to engage in the energy sector.
Households invest in energy efficiency, making them more energy efficient and reducing electricity consumption	**Box B14** "Households invest in energy efficiency" is linked to "Reduced electricity consumption" (**Box 21**). This implicitly assumes that investing in energy efficient retrofits will make households reduce their energy consumption. However, energy consumption might stay the same or increase if households start heating more rooms in their homes. Further, there are important heterogeneities involved. Specifically, because of the rebound effect, electricity consumption may not fall to the same degree that energy efficiency increases, especially for relatively better off households. We have added, "Households become more energy efficient" (**Box B17**) to make this more explicit.
Apartment buildings retrofitted with energy efficient improvements if HOAs are willing to invest in them	In the latest logic model, the short-term output "More HOAs established" (**Box B1**) implicitly assumes that HOAs that are established are well-functioning and are willing to consider energy efficient retrofits in the apartment buildings. This potentially depends on an internal agreement of individual apartment owners. We have made this more explicit by including a short-term output "HOAs commit to energy efficient investments" (**Box B7**).
More trained energy auditors will lead to more energy audits, if households are willing to pay for energy audits	In the latest logic model, "More trained energy auditors" (**Box B9**) was linked to "Energy audits" (**Box B16**). This implicitly implies that households would understand the value of and be willing to pay for energy audits. Further, for poorer households, the willingness to pay for audits will be quite low. To make this explicit we have added, "Households and apartments invest in energy audits" (**Box B13**).
Activity 1.2 District Heating Metering	
New households shift to district heating.	The latest logic model directly links "Decreased heat consumption" (**Box C8**) to "New households shift to district heating" (**Box C4**). However, this chain in the logic model assumes that Termekos is at capacity and the only way for it to expand its customer base is by a reduction in heat consumption from the existing households. However, during the scoping mission we determined that Termokos can easily expand from 70MW (current average consumption) to 140MW (Termokos's capacity). Further, Termokos is involved with projects with other donors like the World Bank and EU who are looking to expand the capacity of Termokos to 280MW. Thus, expanding district heating to new households need not be contingent on decreased heat consumption from existing households. Consequently, we have moved it as a main short-term output.

In **Exhibit 4,** we review key project assumptions that link project activities and outputs to RELP outcomes. Assumptions in green are part of the latest logic model (December 2019 version), while those in red are missing assumptions that we recommend adding. While the theory of change and the *M&E Plan* addressed some of the key assumptions, some important assumptions were missed, as explained in **Exhibit 4**. Broadly, assumptions for the IPP, SEEK, and district heating metering activities lack a focus on political will, long-term funding, and sustainability.

Exhibit 4. Review of Key Project Assumptions

Assumption	Assessment
Activity 1.3 IPP Project Finance Facilitation	
Assumption: Lack of expertise is the only reason that IPP projects	There is an implicit assumption that lack of TA is the only constraint that IPPs face in terms of having long processing times for their license applications. IPPs could have other constraints including delays on the side of the Energy Regulatory Office in processing applications and granting approvals.

Assumption	Assessment
applications have long processing times	
Assumption: Lack of market clarity is the only constraint to establishment of bankable projects	While weak market standards and lack of understanding of them are constraints for both borrowers and lenders, this might not be the only constraint to setting up viable IPP projects. IPP projects should be viable for both borrowers to invest in and for lenders to lend to. For instance, on the lender side, during the scoping mission, KCGF highlighted that financing IPP projects offers a lesser return for banks than other investments. This is coupled with an uncertain return given the long license processing times and uncertainty over who will buy the electricity generated by IPPs. On the borrower side, other important constraints for IPPs (apart from project financing) include long licensing approval times and land acquisition.
Assumption: The number of IPP projects reaching financial close is sufficient in number to increase overall electricity supply	The logic model assumes that enough IPP projects of sufficient scale would have been created by the end of the project period to enable overall increases in electricity supplied. This depends crucially on the robustness of the pipeline of bankable projects. The scale of these IPP projects is also critical. If these projects are not scaled up sufficiently, then Kosovo might continue to rely on electricity imports to meet the electricity shortfall.
Assumption: KCGF will be ready in time to provide guarantees to IPPs	For the IPP sub-activity to achieve its goals of ensuring IPPs reach financial close and increase electricity supply there is an implicit assumption that the renewable energy guarantee window will be available in time. During the scoping mission, we learned that KCGF would not be ready to provide guarantees for a while since they are still in the fund-raising phase. More broadly, given the complexity of the logic model, many different pieces need to align in a timely manner for the goals of the project to be achieved by the end of the project period.
Assumption: There will be sufficient buyers for the renewable energy produced by IPPs	There is an implicit assumption that there will be sufficient buyers for the electricity produced by the IPPs. Specifically, during the scoping mission, IPPs mentioned that it is not always a given that the Kosovo Energy Corporation will buy the renewable energy produced by them, even if purchase agreements are in place.
Assumption: KCGF continues to offer renewable energy guarantees in the medium to long-term	Renewable energy projects have a long gestation period before becoming commercially viable. The long-term incentives for KCGF to continue providing renewable energy guarantees is unclear. Further, KCGF should have the capacity to meet increased demand in the end for renewable energy guarantees from more small and medium-sized enterprises.
Assumption: The Energy Regulatory Office approves projects, grants licenses and sets the price of electricity in a timely fashion	During the scoping mission, we learned that the Energy Regulatory Office provides licenses for IPPs to start operations. The licensing period can range from 3 to 4 years. Additionally, we learned that the office may not be interested in investing in solar for political reasons, which may delay the licensing process further. In order for IPP projects to reach financial close, there is an implicit assumption that licenses are granted in a timely manner and IPPs don't drop off midway through the licensing process, which as we learned can be time and resource intensive involving a fair amount of paperwork. Finally, since, the initial price acts as an incentive for IPPs to be engaged in the solar energy production, lack of clarity can create market uncertainty and discourage IPPs from entering the market. That is, the potential for the feed-in tariffs to change during the course of the application period, makes the process for IPPs inherently risky.
Assumption: Land acquisition for setting up solar/wind/thermal power plants is straightforward	While attracting project financing is a binding constraint for IPPs, during the scoping mission we learned that land acquisition for setting up solar/wind/thermal power plants is also a challenge. The implicit assumption for a pipeline of viable IPP projects to be established is that land for these projects can be easily acquired.

Assumption	Assessment
Assumption: A large number of donors are identified and are willing to provide capital	In order for KCGF to raise enough capital to structure and offer renewable energy guarantees, there should be a sufficient number of donors identified. That is, TA to KCGF on how to open a new renewable energy guarantee window might not be sufficient to raise capital. KCGF needs to actively engage with different types of stakeholders and identify potential sources of funding.
Activity 1.1 Subsidies for Energy Efficiency in Kosovo	
Assumption: HOA's are well functioning and willing to pay for and invest in energy efficient measures	In the logic model, it is implicitly assumed that once HOA's are established, they will be well functioning. The HOA's will require BC&O to understand the importance of energy efficient measures. The willingness also stems from individual apartment owners agreeing to the measures. It needs to be elucidated how this will be established, e.g., through voting and/or some minimum percentage of apartment owners agreeing. Finally, even if HOAs are willing to invest, there is an implicit assumption that credit constraints (if any) of HOAs will be eased.
Assumption: Retrofitting apartment buildings will lead to reduced electricity consumption	This assumes that individual apartments within these apartment buildings do not purchase more electrical equipment with their cost savings. This also assumes that each of these apartment buildings does not witness an increase in the number of residents, assuming apartment buildings have vacancies. If newly retrofitted apartment buildings become more valuable for families to rent out, then apartment buildings might see an overall increase in energy consumption, or the energy consumption might stay the same.
Assumption: Households positively respond by investing energy efficiency	In the logic model, for the household energy retrofits activity, it is implicitly assumed that once incentives are deployed households will invest in energy efficient retrofits i.e., there will be an increase in the willingness-to-pay for energy efficient measures. This also assumes that a one-time BC&O will be sufficient for households to change their behavior. However, this critically depends on the effectiveness of the BC&O. Further, even if the BC&O is sufficient, households may still face other binding constraints such as credit constraints.
Assumption: Households investing in energy efficient retrofits will lead to reduced electricity consumption	There is an implicit assumption that no other changes will take place such as replacing wood with electricity as a heating source, heating more parts of one's home, expansion of the square footage of homes, and purchase of newer appliances, which will simultaneously increase electricity consumption even if households have adopted the energy efficient retrofits.
Assumption: Households and apartment buildings are willing to invest in energy audits	While ensuring trained auditors is critical, it is equally important that households and apartment buildings are willing to invest in energy audits in the first place. That is, they understand the relevance and importance of conducting the audit *and* are willing to invest money in it.
Assumption: Households and apartment buildings understand the findings from energy audits and have the capacity to implement changes	While households and apartment buildings might be aware of the need for energy audits because of effective BC&O, they might not act on the audit findings, for several reasons. Households may not be literate enough to understand the findings. Further, even if the audits reveal several ways of improving energy savings in one's house, a household will only undertake these investments if it has the financial resources to do so. Else, they will delay such expenditure. Thus, electricity consumption will stay the same even after an energy audit.
Assumption: Female owned businesses become more energy efficient leading to reduced electricity consumption if no other changes happen	While female owned businesses might become energy efficient and witness more energy savings, similar to households, they may choose to expand their businesses and/or hire more people, leading to an overall decline or no change in electricity consumption.

Assumption	Assessment
Assumption: TA given to female entrepreneurs is sufficient to invest in energy efficiency	For female entrepreneurs to invest in energy efficiency there should be a sufficient number of such firms who are willing and able to invest in energy efficiency from a cost perspective. Further, the TA and BC&O given to female entrepreneurs should be effective and sufficient. From a longer-term perspective, this also assumes that there is institutionalization of BC&O wherein female entrepreneurs in the future too will be motivated and incentivized to invest in energy efficiency. Finally, it is not clear how other women-led businesses will reduce energy consumption through a demonstration effect given that the targeted number of businesses under RELP is not large.
Assumption: There is sufficient demand for women by employers in energy efficient utilities and district heating jobs	Offering internships, scholarships, and the women in science summer camp only solves the demand side of the employment problem by creating interest and encouraging women to apply for jobs in utilities. On the supply side, however, employers must be willing to hire more women in their firms. This involves conducting BC&O with firms, encouraging them to hire women. Even in the case of internships, female interns can only stay on as permanent employees if the firms are willing to make them full-time employment offers.
Assumption: Science camps will generate an interest in energy among teenage girls	Based on information gathered during the scoping trip, the WiSci camp is targeted to all STEM fields. That is, there is not an exclusive focus on generating interest in the energy sector. Given this, there is an implicit assumption in the logic model that the WiSci camp will lead to reduced gender inequality in the energy sector.
Assumption: Women energy entrepreneur TA activities can achieve objective of reducing gender inequality in the energy sector	An overarching objective of the women energy entrepreneur TA activity is that the different sub-activities will have demonstration effects and/or motivate other females. It is not clear how that will be accomplished if the number of participants is limited. Based on information from the scoping trip, scholarships are provided to around 28 women, and these women will be provided internships as well. The WiSci camp covered 100 teenage girls and the sample size for the women's entrepreneur grant is still to be decided.
Activity 1.2 District Heating Metering	
Assumption: New households shift to district heating	There is an implicit assumption that new households will know how to connect to district heating. Will households be able to choose whether to switch or not? Alternatively, will district heating simply be connected to a building and all households will automatically be connected?
Assumption: Households are willing and available to accept installment of equipment. .	A fundamental assumption underlying the BC&O strategy is that the communication delivered by the Implementing Contractor (IC) will be sufficient and effective. This relies on households understanding how quantity based-district heating metering will help them reduce utility bills. Given the diversity within the country, MFK and the IC will need to ensure that the BC&O is tailored for households with low literacy, socially vulnerable households including single mother households, and poor to very poor households.
Assumption: Through meters and valves, households are incentivized to decrease heat consumption	A crucial requirement for households to decrease heat consumption in the medium to long-run is that BC&O activities should be institutionalized and the TA given to Termokos and HOAs is sufficient for them to continue working with households even after the end of the project. It is not clear whether the BC&O will be one-time or will be continuous throughout the life of the project e.g., through monthly BC&O reminders on a household's energy bill. Further, even if BC&O is adequate and awareness levels are high, households might still not reduce heat consumption if there are other costs to metering. Additionally, because of the anticipated energy savings because of the meters, households might heat more parts of their dwelling. This in turn, might not lead to energy savings and might in fact, lead to either a similar bill or a higher bill for households.

Assumption	Assessment
Assumption: Municipality will co-finance THVs	This activity relies on municipalities co-financing thermal heat valves (THVs). A crucial requirement for this is not only engaging them for the duration of the project but also incentivizing them to continue to do so in the future.
Assumption: Ownership of meters beyond the substation is adequately defined	During the scoping mission, we found that the ownership of the meters beyond the substation level is not adequately defined i.e., Termokos is responsible for the meters only up to the substation level. An additional concern not highlighted in the *M&E Plan* is that this lack of clarity of ownership implies that service, mantainence, and repair of meters is unclear and can lead to households not using the meters and/or THVs. This in turn, would not lead to reduced heat consumption. Further, there is no TA directed toward this. The project is working with the Energy Regulatory Office to define ownership of meters and the responsible parties for repair and/or maintenance.
Assumption: MFK secures continued support from the national and local governments for the district heating metering	It is unclear to what extent the TA to Termokos addresses the relationship between Termokos, the national government, and the municipal governments who have to co-finance the THVs. While the project logic addresses technical capacity, it focuses less on political will of various levels of government in continuing district heating metering in the future. Additionally, it is important to acknowledge that MFK will be conducting advocacy not only at the municipal level but also at the national level, in addition to providing TA to Termokos, and devise a plan around this. Finally, as highlighted by stakeholders in the scoping mission, they would value frequent communication and building of a relationship with the funder and ICs.

SECTION 4. RISKS AND ASSUMPTIONS

Based on our review of the *SEEK Draft Inception Report*, *M&E Plan*, and the MFK RELP Assumptions, Risks, and Mitigations Measures (December 2019 version), the project team has clearly defined the risks to achieving project results, and has defined the mitigation strategies. At this stage, the project has not indicated how they will monitor risks, nor how they will alter design/implementation based on the realization of risks. We suggest learning from the EBRD projects (as well as similar projects) on these types of risks faced and mitigation strategies adopted.

While we do not assess every risk identified by the project, we present key risks in **Exhibit 5** below. Furthermore, we note four overarching risks that are not specifically identified by the project. These risks are not specific to a particular activity and have implications for the entire project.

1. The theory of change is quite complex in terms of the number of nodes, interdependence, and the number of outcomes required to reach the project's goal, demonstrating that the project is inherently risky.
2. The risk mitigation plan does not provide sufficient detail, and labels many risks as "low." The majority of the mitigation measures rely heavily on MFK involvement. We consider this an overall risk, as this assumes that MFK has the capacity to continuously monitor and engage with the ICs and other project stakeholders and take action to counter these risks.
3. Several activities are dependent on different government bodies (for example, the Energy Regulatory Office), which have frequent turnover and potentially different priorities.
4. Slight delays in the project lead to a risk that the implementation of activities will be rushed to meet target deadlines, or will not be completed by the end of the Threshold period.

Each assumption presented in **Exhibit 4** in the previous section faces the risk of not holding. In addition, in **Exhibit 5**, we review key project risks by activity. Across all activities, ineffective BC&O appears to be the greatest risk, and therefore the project should particularly focus on this component.

Exhibit 5. Review of Key Project Risks

Risk	Mitigation Strategy	Assessment
Activity 1.3 IPP Project Finance Facilitation		
Lenders and borrowers do not adopt market standards	MFK will coordinate efforts with KCGF and main lenders and borrowers in order for market standards to be adopted by all parties.	While poor existence and understanding of market standards is a constraint for both borrowers and lenders, this might not be the only constraint to setting up viable IPP projects. IPP projects should be viable for both borrowers to invest in and for lenders to be willing to lend money.
There is no pipeline of bankable projects	MFK will coordinate efforts with KCGF and main lenders and borrowers in order for a pipeline of bankable projects to be established.	The project team should also consider coordinating efforts with the Energy Regulatory Office as their support related to feed-in tariffs and electricity purchasing is crucial for making these projects attractive investments.
IPPs fail to implement (build) their project/s, as a result	MFK and project management committee (PMC) will incentivize	Incentives from MFK and PMC may not be sufficient. Again, if the Energy

Risk	Mitigation Strategy	Assessment
electricity supply will not increase	IPPs to implement this project after financial close.	Regulatory Office and Kosovo Electricity Distribution Company are not supportive, IPPs will be disincentivized from implementing their projects.
TA to banks, IPPs, and KCGF is not appropriate and sufficient	MFK will manage the contract with PMC and will make sure that the TA is at the required levels.	Again, IPP projects should be viable for both borrowers to invest in it and for lenders to be willing to lend money. As mentioned in the assumptions, uncertainty from the Energy Regulatory Office about price and Kosovo Energy Corporation about purchasing may make this not an attractive venture for potential IPPs. Viability may not be achievable even with sufficient TA.
Activity 1.1 Subsidies for Energy Efficiency in Kosovo		
Insufficient participants or benefits may be awarded to non-targeted groups (higher income, etc.)	Excessive and targeted campaigning. For women's sub-activities, detailed market research and concept, testing will be carried out. For apartment energy retrofits, the IC will contact early Urban and Social Department of the targeted municipalities to identify possible multi-apartment buildings (MABs) with the highest needs.	This critically depends on the effectiveness of the outreach campaign, and on the appropriateness of the incentive plans. Low returns on investment, insufficient subsidies, and ineffective BC&O will disincentivize participation. Mitigation strategy should include more details about what the outreach and campaigning will entail to attract a sufficient number of targeted program participants for all three sub-activities.
Misunderstanding of incentive plans/complicated application process	The application and incentives will be clearly explained. Information will be provided in multiple languages and will clearly indicate all needed documents for obtaining the incentives. IC field staff and installers will be well prepared on how to explain properly the incentive plan to all interested participants.	Both the incentive plan and application must be clearly designed and user-friendly. A confusing or burdensome incentive plan and application will disincentivize participants from signing up.
Non-qualified materials are installed (or installed incorrectly); energy efficient measures are implemented partially or poorly; Implementation of additional works (not related directly with energy efficient measures) may have possible adverse impacts on energy savings achievements	The qualified installers will be solely responsible for final functionality of the implemented energy efficient measures. Misconduct will be penalized (loss of business opportunities under MFK-supported projects). Contract with the household/HOA will define that work not related directly to energy efficient measures should be financially borne by the household/HOA.	Project team should consider quality-monitoring plan for installers. Mitigation strategy does not include how to disincentivize work that may have adverse impacts on energy savings achievements.

Risk	Mitigation Strategy	Assessment
Women owned businesses do not invest in energy efficiency and do not become more energy efficient	MFK and IC will continuously engage with private sector to promote this activity	The mitigation strategy lacks detail, so after the concept note has been developed, the project team should consider clarifying the risks and mitigation strategies specifically related to the planned activities.
Lack of interest by interns and host companies to participate in internship program. Very low employment opportunities for women recipients of internship and scholarship program.	MFK will continuously adopt outreach campaigns and will design incentives to reach as many women as possible and energy host companies which will offer internship opportunities. MFK will adopt a track records system to track employment opportunities for scholarship and internship recipients.	Outreach may not be sufficient if there are limited employment opportunities in Kosovo. In addition, to lead to sectoral level changes, the planned number of women participating would have to be significant and continuous.
Activity 1.2 District Heating Metering		
BC&O is not well developed and not effectively implemented	Termokos must have a good relationship with the IC and work together to understand BC&O needs and key risks.	Given the diversity within the country, MFK and the IC will need to ensure that the BC&O is tailored for households with low literacy, socially vulnerable households including single mother households, and poor to very poor households, and implemented by qualified field staff who can effectively connect with the targeted groups. MFK and the IC will also need to ensure that the BC&O is offered for an appropriate time period, which may be longer if households are resistant to change.
Incentives (meters, THVs, BC&O, tariff rates) and mechanisms (establishing HOAs) do not work and there is no decrease in heat consumption	MFK and IC will adapt BC&O continuously. In addition, MFK and IC will coordinate with the Energy Regulatory Office for tariff review and design. TA to Termokos, HOAs to continuously work with households.	A crucial requirement for a decrease in heat consumption is that BC&O activities should be effective and institutionalized, and appropriate support is given to Termokos and HOAs.

SECTION 5. PROJECT PARTICIPANTS

In this section, we review the project participants as defined by the ICs, and discuss whether the selection is based on credible, quantifiable criteria, and how selection can be linked to the evaluation. **Exhibit 6** describes the number of participants, the selection criteria, and the associated design and implementation of RELP's activities and sub-activities. For many of the activities, both the selection criteria and the variation in design/implementation have yet to be determined.

Exhibit 6. Project Participants

Activity	Number of Participants	Selection Criteria	Design and Implementation
SEEK—household energy retrofits sub-activity	▪ 500 households (1st iteration) ▪ 2100 households (2nd iteration)	▪ 1st iteration – no target percentages defined, but objective is to reach low-income, female-headed, minority and vulnerable households ▪ 2nd iteration – target group will be defined (TBD based on monitoring, evaluation, and learning data)	▪ 1st iteration—income and gender will be explicitly considered when offering incentives ▪ 2nd iteration—program implementation, including incentive levels, will be updated for different sub-groups (TBD based on monitoring, evaluation, and learning data)
SEEK—apartment energy retrofits sub-activity	Up to 25 MABs in 5-7 preselected municipalities (Pristina, Mitrovica, Gjakova, Lipjan, Gracanica, Viti, and Novobrdo)	Not fully established, but will include factors such as: ▪ potential of reducing energy demand ▪ electricity used for space heating and air conditioning ▪ availability of energy consumption proofs ▪ socio-economic background (inclusion of low-income, female-headed, minority and vulnerable families), ▪ established HOA (or alternative authorized entity or representative)	▪ At least 10 MABs will be renovated within the first year (by end of September 2020) ▪ Whole-building approach ▪ Incentive levels will range between 35% and 95%, and will consider SES and apartment size
SEEK—women energy entrepreneur TA sub-activity, female entrepreneurs	No target number of participants – potential participants will be selected from over the country, including Pristina, Mitrovica, Gjakova, Gracanica, Peja, Gjilan, Ferizaj, and Prizren	▪ Women owned/managed energy service providers and individual energy service providers ▪ Women owned/managed companies and individual entrepreneurs and/or groups of women engaging in crafts and income generating activities ▪ Non-governmental organizations aligned with women may also be considered	▪ Incentive plan TBD, but will vary by applicant type (individual women entrepreneurs will receive higher grants than established women-owned businesses) ▪ Training and capacity building activities will vary based on needs of participants

Activity	Number of Participants	Selection Criteria	Design and Implementation
		▪ Special attention will be given to minority groups (Roma, Ashkali and Egyptians) and those with disabilities	
SEEK—women energy entrepreneur TA sub-activity, WiSci Camp, Scholarship Program, and Internship Program	▪ WiSci: 100 teenage girls (ages 15-18) and 12 counselors from Kosovo, Albania, Serbia, N. Macedonia, Montenegro, and the United States ▪ Scholarship Program: 28 women ▪ Internship Program: 200 women	▪ Scholarship Program: Highly selective application process ▪ [WiSci and Internship Program – selection criteria not specified	▪ WiSci: Annual two-week summer camp, with curriculum designed to encourage adolescent girls to pursue further education and potential careers in the STEAM fields ▪ Scholarship Program: Full scholarships for two-year Associate Degrees from the University of Iowa, with study areas covering the energy field ▪ Internship Program: MFK will pair recent college graduates with host companies in the energy sector and subsidize their salaries
IPP	No target or expected number of IPP developers determined yet, however, the feed-in tariff supports up to 30 MW of electricity generation	▪ Potential IPPs will be targeted from the solar IPPs who have already applied to the Energy Regulatory Office for licensing	▪ TBD, but will include training and TA to support viable IPP projects, including on how to apply for loans and navigate licensing
District Heating Metering	Termokos district heat supply network in Pristina, with approximately 12,000 connected households	▪ All connected households will be considered	▪ TBD, but will include capacity building in addition to installation of metering systems

SECTION 6. METRICS FOR MEASURING RESULTS

In this section, we will examine if the metrics proposed by MCC for measuring accountability and learning in RELP are clearly defined.[88] We will examine these metrics along three separate dimensions: (a) indicators and data sources for monitoring project implementation; (b) indicators and data sources for monitoring project results; and (c) evaluation potential to maximize learning and accountability.

6.1 Monitoring Project Implementation

Assessing accountability and learning indicators related to project implementation includes determining whether project activities are being implemented as planned, reaching the desired populations, and producing the outputs required to improve short- and medium-term outcomes. To measure indicators related to project implementation, we examined the ***process*** and ***output*** indicators of the *M&E Plan*. Process indicators include comprehensive indicators for measuring the value of implementation contracts signed, and the percent of money. They measure progress toward the completion of project activities to assess whether the project is on schedule to meet its primary objective. Output indicators measure immediate results arising from project implementation. In general, the process and output indicators in the *M&E Plan* are comprehensive for each of the three activities of RELP. We assess in **Exhibit 7**, any process or output indicators that are missing or incomplete.

Exhibit 7. Assessing Process and Output Metrics for Monitoring Project Implementation

Result in Project Logic	Indicator	Explanation
Activity 1.3 IPP Project Finance Facilitation		
TA to banks for project finance	None	The *M&E Plan* should include monitoring around the number of banks that started developing new financial products based on project finance because of the TA provided to them.
TA to IPPs for license application submission	None	Given the time lag in this particular sub-activity (i.e., there will be a gestation period of a couple of years from when IPPs submit their license applications, acquire funding, and start producing electricity), it is important to include short and medium-term process indicators for the TA to IPPs for license application submission. These could include: the final number of projects funded, proportion of IPPs getting faster license processing times, and processing time for applications.
Lenders and borrowers adopt market standards for IPP project preparation	Number of banks and IPPs that adopt renewable energy financing market standards	The *M&E Plan* should include a plan for measuring immediate outputs including the development of market standards, measurement of delays, and bureaucratic hurdles. Further, there should be outputs, which measure the willingness of lenders and borrowers to adopt these standards once they are developed.

[88] "Learning" as defined by MCC includes the commitment to understand the causal relationships and effects of the intervention and to facilitate the integration of M&E findings in the design, implementation, analysis, and measurement of current and future interventions. "Accountability" as defined by MCC includes the obligation to report on and accept responsibility for all funded activities and attributable outcomes.

Result in Project Logic	Indicator	Explanation
KCGF structures and offers market guarantees	Amount of R renewable energy guarantees offered by KCGF	While the amount of renewable energy guarantees offered is important, the monitoring plan should include more immediate outputs to measure TA to KCGF and their willingness to provide renewable energy guarantees in the first place.
Viable IPP projects become bankable through TA	Number of IPP projects becoming bankable through TA	The *M&E Plan* should include immediate outputs of TA to IPPs since the TA to IPPs will include a number of components from informing IPPs about project standards, to training IPPs to use templates, and helping IPPs access financing. It is critical to monitor each step of the TA provided to IPPs. For instance, indicators could include number of templates created for different processes, number of IPPs using the templates, number of loans submitted with new templates, and finally the number of projects that receive funding from banks.
Activity 1.1 Subsidies for Energy Efficiency in Kosovo		
Apartment buildings retrofitted	Number of apartment buildings entrances retrofitted	While the number of apartment buildings retrofitted is important, the project should also monitor the extent to which energy efficient behavior is institutionalized within these apartment buildings. In addition to the number of apartments retrofitted, it is important to include a plan for monitoring continued energy efficiency in the future. Further, since a proportion of these apartments will be retrofitted with co-financing from municipalities, the *M&E Plan* should monitor that as well.
Female entrepreneurs invest in energy efficiency	Amount that female entrepreneurs invest in energy efficiency	We recommend adding several indicators for monitoring this sub-activity: (a) an immediate output of TA provided to female entrepreneurs would be to monitor the proportion of entrepreneurs who agree to invest in energy efficiency from the total number who were given TA, (b) since the extent to which women entrepreneurs change their behavior depends critically on the TA being provided, there should be a plan to monitor the TA as well, (c) types of energy efficient measures invested in should also be monitored, since these can range from minimal (e.g., light bulbs) to maximal (e.g., changing the building envelope), and (d) indicators around sector level changes which capture change beyond the entrepreneurs chosen for RELP and which can give an indication of the demonstration effects.
HOAs are well-functioning and willing to commit to energy efficient improvements	None	A crucial component of apartment buildings being retrofitted is that HOAs are established and are willing to have their apartment buildings retrofitted. The *M&E Plan* should thus, include a plan for monitoring both, the functioning of HOAs and the BC&O they receive which will influence their willingness to commit to energy efficient investments. There should also be a plan for measuring their commitment to invest since even after receiving BC&O they might still not commit to invest, or commit too much and actually invest only a proportion of it.
Types of energy efficient retrofits adopted by households and apartment buildings	None	According to the latest design of the SEEK activity, households and apartment buildings will jointly decide on the type of retrofit they will install with the qualified installer. This decision can be based on both, actual need of a particular retrofit and cost considerations. It is important that the project monitors if households actually install the retrofits they have been advised on by the qualified installer, or do

Result in Project Logic	Indicator	Explanation
		they only install a subset of these retrofits. This will give an indication of the main motivations (awareness and/or cost) that drives these energy efficient installations.
Process	Value of signed implementation contracts Percent and value Disbursed	These indicators are clear and easy to measure. It might also be helpful to include more nuanced process indicators, which track the percentage of milestones achieved/deliverables, met within each of the proposals funded by the contracts.
Female youth attending science camps	None	We recommend adding process indicators to track the number of teenage girls who sign up for the WiSci camp. In addition, we recommend also monitoring the proportion of girls who sign up (as a proportion of those who were given information about the WiSci camp). Finally, we also recommend adding process indicators around the number of counsellors trained to continue camp leadership in the WiSci camps.
Female interns in the energy sector	None	We recommend adding process indicators to track the number of women who get internships in energy related firms. In addition, we recommend tracking the proportion of women who get internships (as a proportion of those who are targeted). Finally, we recommend adding process indicators around the number of employers who are willing to hire female interns in their firms.
Women receiving scholarships	None	We recommend adding process indicators around the types of sectors in which women receive their degree to assess the extent to which they are relevant to the energy sector. We also recommend adding indicators around the proportion of women who receive scholarships who come back to Kosovo to work in the energy sector.
Activity 1.2 District Heating Metering		
Household metered	Number of district heating meters installed in households	While measuring the number of district heating meters installed and comparing against targets is important, there should also be outputs related to the immediate project activity of working with Termokos and TA. Specifically, there should be outputs around installation of meters in a timely and efficient manner, development of BC&O strategy, receptiveness of Termokos to organizational assessment, and infrastructural issues encountered while installing meters.
Process	Value of signed implementation contracts	Similar to SEEK, the *M&E Plan* should include an approach to measure specific milestones related to installation of district meters in individual houses including a plan to measure delays in installation.

6.2 Monitoring Project Results

Assessing accountability and learning indicators related to project results includes determining whether project activities led to the desired short- and medium-term outcomes. To measure indicators related to project results we examine the ***outcome*** indicators of the *M&E Plan*. Outcome indicators measure intermediate effects of project activities, which is extremely useful in assessing if the project is meeting its objective. In general, the outcome indicators in the *M&E Plan* are comprehensive for each of the three activities of RELP. All quantifiable outcomes to measure project impact and performance have been captured. One major lacuna in the *M&E Plan* is that the outcome indicators do not measure the extent to which project incentives, behaviors have been adopted, and frameworks, mechanisms, and good

governance practices have been institutionalized to ensure long-term sustainability of outcomes. We specify below six main areas in which sustainability and institutionalization should be measured over the long-term:

1. **Financial stability and long-term viability of IPP projects**. While it is important to examine if IPP projects reach financial close, it is also important to examine the potential for long-term financial stability of the IPP projects. This could be done by including indicators, which examine the types of systems put in place within the IPP projects to make them financially sound.
2. **Recurrent capital expenditure required by female owned businesses**. To understand the extent to which female owned businesses are likely to make continuous investments in energy efficiency including recurring expenditures like energy audits, it is necessary to understand if and to what extent systems have been institutionalized within the business to support on-going/continued energy savings behaviors and energy efficient practices.
3. **Recurrent expenditures required by households and apartment buildings.** Similar to women entrepreneurs, it is also important to track recurrent expenditures made by households and apartment buildings on maintenance and repair of energy efficient retrofits.
4. **Financial status of households**. While lack of information about energy efficient investments is an important barrier to the adoption of these measures, financial capacity and willingness to pay/invest are also major constraints. The *M&E Plan* needs to include indicators that track the likelihood that households will continue to invest in energy efficiency in the medium to long-term. That is, would households be able to repay the loans taken for the energy efficient investments? Will they continue to invest in energy audits? Will they spend resources to maintain and repair energy efficient equipment? Finally, will the Government of Kosovo be able and willing to sustain incentives to households and apartments that adopt energy efficient measures and that demonstrate energy reductions.
5. **Sustainability of WiSci camps and internships.** For the WiSci camps and internships for women, it is important to track ownership of these activities in the medium to long-term. In particular, are there planned training activities for camp counsellors and mentors for women selected for internships? How often will recruitment happen both for the camps and for the internships? Are relevant stakeholders available to take over the funding of these activities once RELP ends? Finally, is there a plan for long-term continuity of these activities through issuance of regular advertisements, procurements, and trainings?
6. **Sustainability of TA provided at various levels.** It is important to understand the extent to which the TA provided at various levels in the project logic will be sustained after RELP ends. This is important since it will also provide an indication of stakeholder commitment to energy efficiency in the end. In particular, we recommend measuring the establishment of ongoing TA made available – through establishment of energy associations and government public education programs around energy efficiency.

6.3 Evaluation Potential for Maximizing Learning and Accountability

We will conduct the following main activities as part of the RELP evaluation: 1) provide TA to ICs to finalize the overall research design including project communications and outreach materials, develop survey instruments, administer surveys to collect household level data for the project, and provide quality control

to ensure data is collected on all the outcomes of interest to facilitate comparison across different groups and time; 2) collect qualitative data to assess the project's progress and fidelity of implementation and to monitor trends in stakeholder engagement, beneficiary perceptions, adherence to evaluation design, and project challenges and opportunities; 3) analyze qualitative and quantitative data (survey, administrative, and program) to measure the impact of different components on the project outcomes; and 4) assess the cost-effectiveness of different program activities with focus on informing program sustainability.

Using the findings from the scoping trip, we will work with MCC, MFK, and the ICS to design a robust and feasible evaluation design for RELP. The evaluation design report will detail the key evaluation questions and the quantitative and qualitative methodology used to answer each question. The RELP evaluation will assess the extent to which the project activities reduced electricity demand in targeted areas in Kosovo. The evaluation will assess the implementation of each of the activities under RELP and verify the validity of program logic across all three activities. Further, the evaluation will also generate evidence for lessons learned from implementation for MCC and MFK and provide recommendations for program sustainability. In particular, we will examine the extent to which there is a plan to institutionalize energy efficient behaviors through long-term tax incentives and/or mandates to shift energy efficient practices over time, and finally incentives for broader and long-term demand-creation. Adhering to MCC's M&E policy, this evaluation will help MCC, MFK, and its partners, including the Government of Kosovo, understand the effects of MCC's RELP on the energy sector in Kosovo and will provide evidence to improve performance of government and private stakeholders in designing and sustainably enforcing energy efficient behaviors.

REFERENCES

English, A., Qehaja, A., Breznica., Hoti, A., M. W. & J. M. (2015). *Energy consumption and Potentials for Energy Efficiency Implementation: Analyzing Low Energy Consumption and Potentials for Energy Efficiency Implementation: Analyzing Low Income Low Service Areas of Kosovo*.
Retrieved from https://doi.org/10.13140/RG.2.1.4816.1368

Allcott, H., & Greenstone, M. (2012). "Is there an energy efficiency gap?" *Journal of Economic Perspectives*, *26*(1), 3-28.

Ameli, N., & Brandt, N. (2015). "Determinants of households' investment in energy efficiency and renewables: Evidence from the OECD survey on household environmental behavior and attitudes. *Environmental Research Letters*, *10*(4), 044015.

Aydin, E., Kok, N., & Brounen, D. (2017). "Energy efficiency and household behavior: the rebound effect in the residential sector. *The RAND Journal of Economics*, *48*(3), 749-782.

Chai, K. H., & Yeo, C. (2012). "Overcoming energy efficiency barriers through systems approach—a conceptual framework. *Energy Policy*, *46*, 460-472.

Corlazzoli, V., & White, J. (2013). *Practical Approaches to Theories of Change in Conflict, Security and Justice Programmes: Part 2: Using Theories of Change in Monitoring and Evaluation*. London: DFID / Search for Common Ground

EBRD. (2018). *EBRD invests in green economy project in Kosovo*. Retrieved from http://www.ebrd.com/news/2017/ebrd-invests-in-green-economy-project-in-kosovo.html

Frederiks, E. R., Stenner, K., & Hobman, E. V. (2015). "Household energy use: Applying behavioral economics to understand consumer decision-making and behavior. *Renewable and Sustainable Energy Reviews*, *41*, 1385-1394.

Galvin, R. (2014). "Making the 'rebound effect' more useful for performance evaluation of thermal retrofits of existing homes: defining the 'energy savings deficit' and the 'energy performance gap'. *Energy and buildings*, *69*, 515-524

Gerarden, T. D., Newell, R. G., & Stavins, R. N. (2017). "Assessing the energy-efficiency gap. *Journal of Economic Literature*, *55*(4), 1486-1525.

Gillingham, K., & Palmer, K. (2014). "Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy*, *8*(1), 18-38.

Gulati, M. (2013). "Development and Evaluation of Power Supply Options for Kosovo: A Background Paper. In *Conference on Sustainable Energy for Kosovo, Pristina*.

Hausmann, R., & Rodrik, D. A. Velasco. (2005). *Growth diagnostics." Mimeo, Harvard University*.

Hsu, A., Rosengartin, C., Furter, A. W., & Xie, Y. (2017). "Renewable Energy and Energy Efficiency in Developing Countries: Contribution to Reducing Global Emission." *United Nations Environment Programme, Third Report*.

Kammen, D. M., Mozafari, M., & Prull, D. (2012). *Sustainable Energy Options for Kosovo* (pp. 1–56). University of California, Berkeley.

Knoop, K., & Lechtenböhmer, S. (2017). "The potential for energy efficiency in the EU Member States–A comparison of studies. *Renewable and Sustainable Energy Reviews*, *68*, 1097-1105.

Kokx, L. (2017). *National Building Energy Efficiency Study for Kosovo*. World Bank.

Ministry of Economic Development. (2017). *Energy Strategy of the Republic of Kosovo 2017-2026*. Retrieved from http://mzhe-ks.net/repository/docs/Kosovo_Energy_Strategy.pdf

MCC. (2017). *Kosovo Constraints Analysis*.
Retrieved from https://assets.mcc.gov/content/uploads/Constraints-Analysis-Kosovo-1.pdf

MCC. (2018a). *RELP Final Design Report*.

MCC. (2018b). *RELP Monitoring and Evaluation Plan*.

MCC. (2019). SEEK *Draft Inception Report*.

Schleich, J., & Gruber, E. (2008). "Beyond case studies: Barriers to energy efficiency in commerce and the services sector. Energy Economics, 30(2), 449-464.

World Bank. (2017). *Doing Business 2017: Equal Opportunity for All*. Washington, DC: World Bank.

World Bank. (2015). *National Building Energy Efficiency Study for Kosovo*. Washington, DC: World Bank.

Appendix E: MCC Comments and Response Matrix

Exhibit 1. First Round Comments form MCC

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	overall comment	Disaggregation by gender are proposed for various responses. Will this be disaggregated by gender of respondent or gender of household head? Both? For any given surveyed household under both PIEE and DHM, it would be useful to know who within the household will be surveyed, and how you will approach asking this. For example, it would not be good to survey the HH head primarily, as 90% are men. In addition, in the section where approaches for data collection are summarized for both DHM and HER/AER, it would be helpful to specify interviews with female business owners, and FGDs with women in HHs/women businesses. I suggest you make sure to do some targeted FGDs with women only to get at the issue of gender differences in usage/experiences, and reflect that in your plan.	This is a valid concern and at a minimum the evaluation will disaggregate all relevant outcomes by gender of household head. We will also rethink the survey respondent decision to make it more inclusive within the constraints of time and budget including surveying multiple people in household to ensure adequate gender representation. Additionally, in our qualitative data collection, we will also conduct women-only focus groups for both SEEK and DHM. See Sections 4.3.4 (HER/AER) and 5.3.4 (DHM) for description of targeted women-only focus groups.
MCC	2	There is a reference to metering in the first paragraph. As there are other devices to be installed (e.g. thermostatic heat valves and heat allocators) that facilitate energy savings, you may want to mention something like "metering and other energy saving devices"	This has been updated on Page 2 to include metering and other energy saving devices.
MCC	3	It was mentioned that the study was designed to employ a second iteration that would implement the model identified in the first. I'm not sure this was necessarily the design - there was the general mention of iterative design and an iterative approach as well as the reference to evaluations of one phase that could lead to changes to design variables in the next phase. Thus, as opposed to testing-implementation, it seemed to be testing-testing, with learning in between and the cost-effective model(s) developed by the end of the Program. Possibly the two are indeed the same thing but I wanted to note this nuance in case it made any difference.	This has been updated on Page 3 and the first paragraph of Page 4. We have removed the reference to “testing-implementation” and have updated that the study will include "testing-testing" with learning in between.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	16	In reference to outcomes table for DHM: Gender and social inclusion is integrated into the DHM behavior change component, and into technical assistance for doing disaggregated analysis of consumer behavior. It would be useful to know what impact this integration had on outcomes of consumer behavior, especially if this is to be scaled up. Please ensure disaggregation of DHM outcomes by women-headed households, age groups, income levels, and ethnicities and add the same question as posed for SEEK on seeking to understand experiences of different groups, including women.	This has been updated in Exhibit 3 to include disaggregation of DHM outcomes by women-headed households, age groups, income levels, and ethnicities.
MCC	16	In order to answer the question "Were the activities implemented as designed?", what will be used to represent "design"? I imagine it will be the pre-implementation documents such as the contractually approved Operations Manual?	Yes, this has been updated in Exhibit 3 for EQ1.
MCC	17	RQ2 questions: could you please add a question to assess if energy consumption or any savings differed by gender, income of household? This is presumably what you are trying to get at in your earlier statement where you note results will be disaggregated by key variables. But it would be good to capture this explicitly with a specific question.	Based on the EMC meeting on July 16, 2020, we have included the finalized evaluation questions in Exhibit 3.
MCC	17	RQ6: outcome column. One of the items here you have is 'proportion of interns offered jobs by host company.' this should be broadened to either host company or another relevant employer. As long as the internship helps them obtain a job, we will reach our objective, even if it's not a job with that one particular internship host company.	We have updated this in Exhibit 3 and Section 7.3 (Exhibit 30). Tracking the girls' employment after their internship may be useful. Though it will be impossible to attribute STEM-sector field employment to the internship, tracking trends may lend insight. Therefore, we have maintained and broadened this indicator to track whether internship/apprenticeship participants found employment in related fields within two years. Further, we have added an indicator to track whether and what percent of participating companies employed program intern(s), which is meant to provide some insight as to whether or not the internship program provides a pathway into employment and whether or not the internship

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
			influences the host company in its openness to hiring women.
MCC	19	Evaluation timeline: is this timeline still on track given COVID related travel risks? It seems to me it may have to be modified.	We have updated Section 3.2 with Exhibits 4 and 5 detailing the exposure periods and linking them to the data collection timing. Data collection timing is crucially linked to implementation. In the event implementation is delayed due to COVID, data collection timelines will shift as well.
MCC	20	Status update on projects: these have shifted now that we have been delayed due to COVID and the hold. Wisci 2020 is cancelled. The internship program is on hold. The women entrepreneur applications were launched but also had to be put on hold in mid-March.	We have updated the document to reflect the postponements and suspensions. Language has been added in Section 7.1 to note the aim to reinitiate in 2021.
MCC	23	Might there be any seasonal sensitivities around the average consumption that could be worth exploring in order to answer the research questions? E.g. knowing about the average during the summer or winter season and the subsequent reduction in energy during following seasons where the 500 or 2100 retrofits are made.	We have clarified this in Section 4.2.1. We will control for month effects in our ITS regressions. Additionally, since we will have monthly administrative data, we will also examine disaggregated impacts by season. We have updated this in Exhibit 10.
MCC	23	Does the statistical power related to 70% still hold the further away from 2100 GFA is able to achieve? For example, in a worst case scenario where only 1000 beneficiaries are reached, are 700 surveys sufficient? Is there a minimum threshold of surveys such that we might want to make sure there is a corresponding minimum number of beneficiaries reached? If so, communicating this to GFA could be valuable as they go about attempting to reach the highest number possible. This is also relevant for Section 5 and Decon.	In Section 4.2.2 (Exhibit 8) we present the different design effects that we would be able to identify if the IC is unable to reach 2,100 beneficiaries. We present effect sizes for scenarios with 400, 800, 1200, and 1600 beneficiaries. Similarly, for DHM, in Section 5.2.2 we present in Exhibit 25, effect sizes for different scenarios of number of households reached. As a note, for DHM this might not be an issue since the projected number of beneficiaries is approximately 12,000 of which the IC will survey approximately 3,600 at baseline. That is, there is a larger pool of households to choose from.
MCC	31	As it relates to outcomes, has IMPAQ been able to confirm whether some of the households retrofitted could not be using electricity? If this is so, do we believe there is a possibility to learn about models of realizing efficient energy consumption for those who may reach a time where they begin consuming electricity and can choose to take	As per the details provided in the IC's Operations Manual for the HER sub-activity, households should be currently using electricity in order to be eligible for the retrofit. We have clarified this on the top of Page 4.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
		advantage of certain incentives (or not) so to reduce their consumption?	
MCC	36	Missing from the WEE description is the overall focus of the project on female employment, through inclusion of requirements in the technical specs for contractors to hire women. Please add this throughout.	As discussed with the MCC GSI, this comment is related to all activities, and is not specific to the WEE activity. We have addressed this at two places: (1) we have updated the activity descriptions for SEEK and DHM on Pages 3 and 4 to reflect the project's requirements for contractors to hire women, and (2) in relationship to this recommendation, in Section 7.3 we have added an indicator in the internship/apprenticeship activity to determine if the hosting companies add guidance at some point during participation in the WEE activity that promotes the hiring of more women.
MCC	37	In reference to outcomes table row, "Technical Assistance and Grants": Will these outcomes cover efficient use of energy, lower energy costs, and improved profit margins?	We have added three quantitative indicators in Section 7.3 (Exhibit 30 outcomes table) and corresponding data collection information in the Section 7.4 (Exhibit 31 data collection table).
MCC	37	Under scholarship you list under outcomes "Percent increase in number of women who graduate in energy-related fields." It is not clear what this relates to. Also, it would be helpful to specify what will be classified as an energy related field. For example, computer technologies could be relevant but not immediately obvious as falling under an energy related field. Do you mean to focus on a % increase nationwide which could be problematic, or from our baseline of MFK beneficiaries, which is 0? I am not sure this indicator makes sense given these issues. In the WiSci section, rather than focusing on the % of women who show interest/study energy, you should focus on STEM more generally, given the WiSci focus was broader than just energy and given that a non-energy but still a STEM field could still lead to energy sector employment (e.g. IT). Finally, for the scholarship recipients, some may not have a job right after graduating, but might	The following changes were made Section 7.3 and 7.4. (1) Deleted the last bullet in the scholarships section measuring women graduating in energy-related fields; (2) Changed 'energy' or 'energy-related fields' to 'STEM' or ' STEM sector fields' and did the same throughout; (3) Added that a job would be pursued within two years of completion of internships and scholarships. Note that this will require a follow-up survey to track scholarship alumnae, so language to this data collection process has been added to Section 7.4 (Exhibit 31 data collection table).

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
		secure one within a year or so after graduating. Does that count? There may be in other words a time lag.	
MCC	44	Referring to the first paragraph and the note of unused energy and demand exceeding supply, this may not be relevant for most of the year, where Kosovo sees energy consumption under its peak capacity. Moreover, less consumption could actually lead to higher tariffs, though at a very very marginal and levelized rate, even if bills are lower because the MWh energy purchased is less. But regardless it does seem to be true that KEDS would not necessarily have any direct financial savings or incur incremental costs.	We have clarified this in Section 4.4.2 (subsection 3) that demand exceeding supply is only a feature in certain months of the year.
MCC	64	Take care to discern among SAIDI and SAIFI unplanned vs. planned and SAIDI and SAIFI for transmission vs distribution.	In the new list of evaluation questions in Exhibit 3, the evaluation question assessing the impact of RELP on overall sectoral reliability has been removed. While we will still include questions on household level electricity reliability in the SEEK and DHM baseline and endline quantitative survey instruments, we will no longer estimate impacts on SAIDI and SAIFI. This has been modified in Section 8.
MCC	66	Noting COVID related challenges could be valuable as well.	We have updated this in Section 9.1
MCC	67	WEE sub-activities also provide sector-relevant job skills in addition to generating interest.	We have added sector-related job skills in Section 9.2.3. In Section 7.3 (Exhibit 30), we suggest that we keep 'knowledge gained' as a qualitative question to reveal participant increased confidence in understanding of the sector - to that end we have qualified this indicator to note 'perceived knowledge gained.' For the internship, questions have been added to both the intern and the supervisor end-of-internship assessment to inquire whether or not sector-related knowledge and skills were gained, and to request specific examples.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	5	Throughout the report, please be careful to use the terms goal, objective, and outcomes in line with MCC's terminology. Our goal is always to reduce poverty through economic growth. Each project has an Objective noted in the THP agreement. Measuring that objective should be the anchor of the evaluation (e.g. the objective-level outcomes guide sampling, timing, etc. above all other outcomes). All other results that the project is trying to achieve should be referred to as outcomes. At the start of the TOC section, please state the RELP objective as stated in the THP agreement.	We have added the RELP objective as stated in the THP agreement to the start of the TOC section (Section 2.2) and have clarified goals, objectives, and outcomes in the rest of the document.
MCC	6	I think it's worth noting at the end of the logic section that because this is part of a THP, no cost-benefit analysis was conducted and so none is discussed in this EDR.	We have updated this on Page 6 in Section 2.2.
MCC	7,13	To M&E Lead: why is the gender inequality outcome listed as an objective in the logic. Is it noted as such in the THP agreement? All: Our evaluations are designed first and foremost to assess the achievement of the project objective and, in doing so, to validate the theory of change/program logic. The gender component of the project should be part of the logic, not an objective in and of itself.	Based on the MCC EMC meeting on July 16, 2020, the WEE activity will now be treated as a standalone evaluation and this has now been moved to Section 7 of the EDR.
MCC	16	Minor point, but please refer to the questions guiding the study as "Evaluation" rather than "Research" questions. That conveys that this study focuses on evaluating the hypotheses that the project was testing, versus conducting research on outcomes of interest	We have updated this throughout the report.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	16	It's important to consolidate evaluation questions to the extent possible. This becomes particularly important when it comes to summarizing the overarching narrative of the evaluation. Some suggestions: - 1.1 and 1.2 can be combined; 1.2 is needed to answer 1.1. - 2.2 and 2.3 could be a follow-on Q to 2.1; they all seem to be asking about different aspects of the same outcome - 2.4 isn't an evaluation question. It should be answered in the conclusion of the analysis reports or a lessons section. Drop - How is 3.1 different from 2.1? Can those two be combined? - 3.3 isn't really an EQ either. Would it not be reported as part of the CEA for 3.2? Drop - 3.4 isn't an EQ. put it as a sub-Q if anything. - 4.2 should be a sub-Q of 4.1 - 4.3 is this a precursor to the outcome in 4.1? Can it be reported as part of answering 4.1? Drop	The updated evaluation questions decided as part of MCC EMC's meeting on July 16, 2020 are reflected in Exhibit 3.
MCC	16	If you're proposing different methodologies for different outcomes, that needs to be made explicit in the table, e.g. which of the outcomes will be pre-post vs. ITS? I think this mainly applies to quantitative outcomes. Same goes for data sources. You may want to break out the table into more rows.	We have updated Exhibit 3 for SEEK and DHM and noted the approach (ITS versus pre-post) for the quantitative outcomes.
MCC		This statement is confusing: while primary quantitative data will be collected at two points in time: baseline and endline. Qualitative data will also be collected at baseline and endline. Additionally, there will be three rounds of interim qualitative data collection for continuous monitoring and to examine changes that occur between baseline and endline. Is the interim qual data not also primary data? It's important to make explicit how many rounds of data collection MCC will be paying for through this contract.	This has been updated at the beginning of Section 3.2. The interim qualitative data is also primary data. There will be five rounds of primary qualitative data collection and two rounds of primary quantitative data collection.
MCC	19	Is there a typo here? Given the rolling nature of baseline quantitative data collection for SEEK, the submission of the draft **interim** report is tentatively scheduled for December 2021, approximately 1-2 months after the **baseline** quantitative data collection is expected to finish.	As per IMPAQ's SOW, we will submit two reports: (1) *interim report* after baseline data collection; and (2) *final report* after endline data collection.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	19	The rounds of data being collected by the implementers should be clearly noted / distinguished from the IMPAQ rounds	We have updated this at the beginning of Section 3.2.
MCC	17	At this stage, we have a reasonably clear idea of what data collection will be conducted by the implementing contractors. Please include that in the design description - e.g. for HER activity, the IC will collect baseline quantitative data as part of a "customer satisfaction survey". Include what types of variables are expected to be collected through this mechanism and how they will fit into the analysis expected to answer the evaluation questions. Similarly for DHM, include any information on discussions that already occurred with Decon.	We have updated the description of the design in Section 3.1 for SEEK and DHM.
MCC	22	Please describe the administrative data that will be used to estimate electricity consumption/expenditure. Who collects it and how? What's the quality of the data? Where is the KEDS data described? Is it certain that we will gain access to it? This can be addressed on page 25 under data sources. I see there is a reference to access being a risk at the end (p 67) - can MFK give us some indication of the likelihood of getting data before we finalize this EDR? Have preliminary discussions taken place?	We have provided details in Section 4.2.4 under "Data Sources" of the meeting between KEDS and IMPAQ on May 21, 2020, wherein preliminary discussions took place regarding access to the administrative data. KEDS shared the list of variables included in their dataset and we shared a data sharing agreement with them for initial review and questions.
MCC	25	It's not clear what the value of the baseline/endline survey will be to evaluate the HER activity, when the main methodology is ITS and is based on the KEDS data. Please explain/clarify. If you're only going to be able to speak to energy use through primary data collection, then how will you answer EQs	We have updated this in Sections 4.2.1 and 4.2.4. The main aim of the primary data collection is twofold: (1) to understand the mechanisms of change including the substitution between different energy sources which cannot be informed by utilizing only the administrative data, and (2) to provide information on the demographic and socio-economic characteristics which will used as control variables in the ITS analysis since the administrative data does not provide any of those details.
MCC	25	Footnote 36 (and perhaps more about data quality) should be in the text so that it's clear that the baseline data is a potential risk for the eval, since it is not in the control of the evaluator	This has been updated in Section 4.2.4.
MCC	26	Rabia/IMPAQ: Generally, the design is hard to follow all the different activities across each sub-activity and activity. I'm wondering if there is overlapping data collection and analysis to answer the same evaluation	We have updated this in Section 3.1 and added Exhibit 2. We have also added a short note below Exhibit 2 to make clear how the evaluation of the

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
		question, and whether that is critical. Is there a way to simplify this design to focus on what's critical? I'd like this EDR to make clearer how each sub-level analysis is going to build up to answering the objective-level question. Otherwise, I'm concerned about going so deep on every single activity.	different sub-activities will answer the broader objective-level question.
MCC	51	Please provide more information on Termokos data. How is it defined, collected? Do we know the likelihood of obtaining access?	We have updated this under "Data Sources" in Section 5.2.4.
MCC	General	The EDR is missing discussions about the timeframe of exposure for key outcomes. These should be used to justify the timing of data collection. The objective level outcomes should be the main driver of what the exposure period for interim and especially endline should be.	We have updated Section 3.2 with Exhibits 4 and 5 detailing the exposure periods and linking them to the data collection timing.
MCC	19	For each activity, include a column for the ideal exposure period required to capture changes in key outcomes being measured by each data collection activity. Some of these are immediate (e.g. energy consumption due to retrofits in HER), while others will require perhaps 6 months or so (e.g. behavior changes in homes with retrofits in HER). The timeline should reflect the needed exposure period and how it ties to implementation timelines -- the specific timeline of the evaluation data collection should be listed as notional contingent upon the implementation timeline and needed exposures. Right now, the EDR reads as having fixed timelines without dependencies on implementation or exposure periods. This approach will alleviate the challenge of trying to predict implementation timelines, given all the delays with the pandemic and uncertainties about infrastructure works going forward.	We have updated Section 3.2 with Exhibits 4 and 5 detailing the exposure periods and linking them to the data collection timing.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	13-14, 22, 49	For both HER and DHM, the ITS method will depend heavily on the availability of the energy consumption/ billing data from KEDS and Thermokos respectively. It's prudent to clearly specify which variables are expected to be used from the billing data (I presume it will only be energy consumption?) and which ones are expected to be collected from the pre-post surveys. We want to be mindful of not duplicating data collection for variables that may already be collected elsewhere (e.g. by the implementing contractor) and ensure that we are not spending level of effort on extraneous data that will not ultimately be used for any meaningful analyses.	For HER and AER we have updated this in Sections 4.2.1 and 4.2.4 and for DHM in Section 5.2.1 and 5.2.4 The main aim of the administrative data will be to obtain information on monthly energy consumption. On the other hand, the main aim of the primary data collection will be: (1) to obtain information on the demographic and socio-economic characteristics of the households which will act as control variables in the ITS analysis which will use administrative data and (2) to understand the mechanisms of change including the substitution between different energy sources which we will not be able to get from the administrative data.
MCC	21	*"For the HER sub-activity, direct beneficiaries include only the 2,100 households scheduled to be enrolled in the second iteration, and these will constitute our "main sample"* -- I do not quite understand the rationale of not using the 500 households during the first iteration but am interested to discuss more. For the 2100, can you address how you will disentangle the potentially differing effects on energy consumption of the different intervention models (1-3, unclear at this time) on these 2100 households? The ITS evaluation design essentially characterizes all 2100 households as part of one sample receiving the same intervention, while we may be looking at 3 sub-groups receiving different interventions. If this understanding is correct, the statistical power to detect minimum effects in Exhibit 5 would be decreased for **each intervention model.** If the goal is to identify the combined effects on energy consumption for **all intervention models** within HER, then this approach makes sense. Recommend clarifying in the EDR.	1) For iteration 1, we have clarified in Section 4.2.1, Section 4.2.4, and footnote 2w that the main goal will be to understand how the model implemented in iteration 1 impacts electricity consumption for different demographics and groups. This will then be useful for improving programming for iteration 2. We will not able to use them as the "main sample" since there is no endline survey proposed for them. 2) For iteration 2, since at this stage it is unclear how many models are going to be implemented, the goal is to identify the combined effects on energy consumption for all intervention models within HER. We have clarified this in Section 4.2.1 and in Section 4.2.5 (subsection 1).

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	26	What is the purpose of Baseline iteration 1 quantitative data (500 households)? Is this strictly necessary or could that information be gleaned simply from the 2100 iteration 2? Based on the outcomes in Exhibit 6, the 500 iteration 1 seems to be extraneous.	This has been updated in Section 4.2.1 and Section 4.2.4. The baseline data for iteration 1 has two goals: (a) to conduct a CEA of the model chosen for the first iteration and, (b) to provide learning for the second iteration to help the IC determine the impact of the model implemented in the first iteration and adjust the models implemented in the second iteration.
MCC	22	How critical is the input from the 690 spillover sample toward the simulation needed to answer the objective level outcome? The level of effort described for the spillover sample seems rather high. I'm foreseeing two major challenges: (1) Level of effort required in the development of the sampling frame using the 3 indirect beneficiaries/ direct beneficiary, and (2) Attrition of the indirect beneficiaries in the sample. Between the development of the sampling frame and the actual endline data collection, there will likely be a 2-year gap and we know that there is significant movement of people from rural to urban in Kosovo. If the spillover data won't significantly alter the objective outcome simulation results, it might not be worth the effort.	Based on the EMC meeting on July 16, 2020, the evaluation question related to spillovers has now been removed. However, we will collect qualitative information on non-participating households for the HER sub-activity to understand both their decision to not enroll in the program and if their views towards energy efficiency investments change over time (see Section 4.3.1).
MCC	19	The table in Exhibit 3 is a useful summary. Please also include the rough sample sizes of the data collection activities. Based on my reading of the report, all samples for the data collection activities identified are mutually exclusive across project activities (i.e. SEEK, DHM, and IPP samples are distinct and independent populations). Is there a way to economize data collection across the different RELP activities? The quant data for SEEK and DHM will be panels.	We have updated new Exhibit 5 with sample sizes. Since different activities have different beneficiaries, unfortunately, there is no way to economize the data collection. To the extent possible, qualitative data collection for all activities will take place at the same time as will endline quantitative data collection.
MCC	28	HER analysis plan: What is the complementarity between the ITS and pre-post methodologies? Is the idea that the ITS will only be useful for energy consumption because that is the only data available from KEDS bills? And the pre-post data will be more nuanced because it will have additional demographic and behavior data? How will these two analyses be combined to yield a cohesive narrative?	Yes, the pre-post analysis will provide more nuanced information than is possible to glean from the KEDS bills. We have updated this at the beginning of Section 4.2.5.
MCC	24	*"If any of the households selected for the household survey in the HER sub-activity are also receiving benefits under the AER sub-activity, we will drop these households from the final sample.*" That should not be the case because by definition the HER (single family households/	As per latest information from the IC (Meeting on July 30, 2020), there can be overlap between HER and AER households. That is, some households could receive both interventions.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
		standalone homes) are different from the beneficiaries or AER (apartment buildings). These two populations are mutually exclusive.	
MCC	34-35	Exhibit 7 is a useful summary. What is the justification for so many qual data rounds? For the HER and AER beneficiary households, there needs to be some additional explanation about geographical spread. What considerations will be undertaken to ensure that a relatively diverse group of beneficiaries from municipalities across Kosovo are included. Will the multiple qual data rounds target the same locations? Different locations? Why? How does the qual data tie into the quant data? There should be a clear plan where the qual data complements quant data, all in support of the eval questions and build toward answering the objective outcome. It can be easy for the qual data to meander and portions of it to be extraneous to the overall focus of the evaluation. Ensure that does not happen here.	Added clear rationale for the qualitative data rounds in Section 3.2 (Evaluation Timeline), Section 4.3.3 (Data Collection), and the revised Exhibit 11 (Summary of Rounds of Qualitative Data Collection for HER and AER Sub-activities). Section 4.3.4 (Sample) describes consideration for selection.
MCC	Overall for AER	Do we have clarity on who exactly will be sampled for the quant data (HOA representative or apartment owners)? What is the argument for either? What implications does that unit have for the sampling cost and the work planning? I recommend putting some thought toward the pros and cons of sampling each toward the evaluation's ability to speak to the key outcome (decrease in energy consumption). The unit of analysis will likely need to be different for behavior change outcomes (apartment) versus building level energy consumption (whole building) since the engineering retrofits are intended to address the building envelope. This applies to the CEA section as well, since the unit of analysis there is apartment, rather than the whole building.	As discussed in the MCC EMC meeting on July 16, 2020, we will survey both, individual apartment owners and HOA representatives. We will administer one survey with two modules, one module for all individual apartments and the second module only for those apartments which have an HOA representative. We have clarified this in Section 4.2.1 and Section 4.2.4. The CEA analysis for the AER sub-activity in Section 4.4.3 has also been modified accordingly to take into account the two different units of analyses.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	37	WEE: How are you proposing to obtain the quantitative data described in Exhibit 9? It seems that the primary information is coming from the MFK M&E Plan/ monitoring indicators. "*The KIIs will be semi-structured since they will also have a quantitative component. Specifically, we will include questions to track the quantitative indicators outlined in Exhibit 9 to assess program implementation and effectiveness.*" If this is referring to obtaining MFK monitoring data, that is not a viable strategy. The MFK M&E Plan will only have indicators directly tied to beneficiary level outputs within responsibility of the MFK (e.g. number of women receiving internships, scholarships, or entrepreneur assistance). All outcome level indicators (e.g. percent increase in number of women graduating in energy-related fields; percent increase in energy efficiency investments) that might require a survey will not be reported in the MFK M&E Plan. This section needs to be fleshed out a bit more to clearly identify the mechanism of obtaining the quant data.	We have updated Section 7.2 (Exhibit 29) to clarify that the KIIs will be semi-structured, that is, they will include qualitative and quantitative questions which will allow us to capture quantitative outcomes as outlined in Section 7.3 (Exhibit 30). Further, language has been inserted in Section 7.4 to clarify that data will not only be collected by FGDs and KIIs, but also by the analysis of existing reports at baseline and endline as well as some post-activity surveys to capture evidence of potential outcomes of activity or sustainability of activities that advance women.
MCC	38	Exhibit 10 - how are the KIIs and FGDs over the baseline, interim, and endline related? Will there be some attempt to follow up on the same individuals? Will individuals who have participated in previous rounds be excluded in subsequent rounds? What is the rationale for 5 rounds? What value do we hope to gain from each round? Include exposure period as part of rationale.	Added rationale for the qualitative data rounds in the Exhibit 5 in Section 3.2 (Evaluation Timeline), Section 4.3.3 (Data Collection), and the revised Exhibit 11 (Summary of Rounds of Qualitative Data Collection for HER and AER Sub-activities). See 4.3.1 (Methods) and 4.3.4 (Sample) for how respondents will be selected. For HER we plan to use different individuals per round (selection will be driven by energy habits). For AER, we will select 7 MABs as case studies to follow up with during subsequent rounds.
MCC	39	The Cost-Effectiveness Analysis (CEA) section is well done. Should be a very interesting analysis and directly applicable to evaluation questions. The limitation described in section 4.5.4 will need to be included while interpreting the results.	We will keep that in mind while interpreting the results in the interim and final reports.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	49	Include information about how area based Thermokos data will be handled to identify changes in energy consumption. As we know already. Thermokos only has consumption data at the sub-station level, which can include anywhere from an entire apartment building to multiple apartments within the same building. The consumption-based billing will change the unit of analysis from multiple apartments to single apartments. How will this change be accounted for in the ITS methodology?	We have updated this in Section 5.2.1 and 5.2.4. Based on IMPAQ's discussion with Termokos in June 2020, Termokos also has consumption data at the apartment unit level. Thus, in the "pre-period" i.e. before the meters are installed, we will have data at the apartment unit level and thus the unit of analysis will remain the same in the ITS analysis.
MCC	55	Does the second column in Exhibit 23 "Timing and Sample Size" refer to the number of KIIs and FGDs that will be conducted for **each round**? What is the rationale for 5 rounds? How do the rounds align with exposure period of the interventions? What is the geographical breakdown of the customers to be included in the qual sample? This table does not include information on how the qual samples will be selected. Justify.	Added rationale for the qualitative data rounds in the Exhibit 5 in Section 3.2 (Evaluation Timeline), Section 5.3.3 (Data Collection), and the revised Exhibit 27 (Summary of Qualitative Data Collection for DHM Activity). See 5.3.4 (Sample) for how respondents will be selected.
MCC	General for IPP	The approach seems reasonable, but I propose including some quantitative data similar to assessment of the WEE component to get an overarching view of the IPP activity. This should not be too difficult and using KIIs and document review to gather key quant information should work well. Recommend outlining the data gathering process for key quantitative indicators that will feed into the evaluation questions (e.g. additional energy kWh supply in the pipeline toward the Objective outcome "supply-demand gap). This could then be included in the simulation to answer the objective level outcome.	Added quantitative outcomes, see Section 6.2.1 (Methods) and Section 6.2.4 and Section 6.2.5 (Summary of Data Collection for IPP Activity). We have also updated the language in Section 8.1 for using electricity supply to address the objective level outcome.
MCC	overall	This is a solid design. The presentation of the design report by RELP activity is useful, but for work planning purposes it is difficult to get a quick summary of the timing of the various data collection activities. It would be helpful to include a table in Appendix C (Evaluation Work Plan) that describes all data collection activities over time, including RELP activity implementation milestones, exposure periods and the data collection activity.	This has been updated and we have added Exhibit 4 and 5 to Section 3.2. We have also added Exhibit 37 to Appendix C detailing exposure periods and data collection activities over time.

Exhibit 2. Second Round Comments from MCC

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	5	We also have some specific requirements for HER/AER contractors regarding hiring of women in the context of these contracts. This too contributes to our female employment in the energy sector focus, and should be noted here.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will take this into account while developing qualitative guides for the implementing contractors for the HER and AER activities.
MCC	18	For WEE – we are interested in energy efficiency, but we are also interested in supporting women-owned businesses to grow and increase revenue and employment. If a particular woman-owned business grows, if the investment creates business growth, then her business might end up consuming more energy overall, even if more efficiently. So lower energy costs overall may not be achieved and should not be the right focus, but it should be more efficient use of energy. We also never explicitly thought about 'improved profit margins" though we did think about and hope that the investment would create increased employment within these firms, especially female employment, in light of the low LFPR among women in Kosovo. Suggest therefore replacing profit margins with employment.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. Through the instruments, we will capture expansion of these firms, increased employment opportunities created during the project, and proportion of women who fill new positions.
MCC	20	Apartment owner is likely to be predominantly male, even if he is not the only user of energy inside that apartment. Please consider how to make this sample less biased towards male respondents.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. Based on discussions with MCC we will take this into consideration while designing the SEEK baseline quantitative instruments.
MCC	32	As you noted above, asking only the HH head could lead to answers representative of the primarily male respondents who will be the HH heads. Why not instead adopt a different approach and in some sample of cases interview the HH head, in others the spouse/another appointed HH member.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. Based on discussions with MCC we will take this into consideration while designing the SEEK baseline quantitative instruments.
MCC	84	This is not accurate. The activity was not put on hold, and the students remained in their program throughout. Classes were not cancelled, but there was modified instruction and increased virtual instruction. Some students went back to Kosovo, but this was just for a few weeks over the summer break, and all are back in Iowa now from their summer break.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	84	All internships have resumed now, but some have different jobs than originally envisaged because of COVID-19 and overall changed landscape of what jobs are still available. DMACC had to lay off a lot of people on campus to save on costs because of budget issues since COVID-19.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.
MCC	84	Program is in implementation. 10 interns are already in jobs; another 20 positions advertised, and applications are being accepted.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.
MCC	84	This activity was launched a few weeks ago. The mentoring/coaching component will start first for the first 20 most promising applicants. 70 applications were received for 20 spots. The rest of the program is also operational, and applications are being invited. Please see MFK website for more details.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.
MCC	85	See my earlier comment on this in particular: For WEE – we are interested in energy efficiency, but we are also interested in supporting women-owned businesses to grow and increase revenue and employment. If a particular woman-owned business grows, if the investment creates business growth, then her business might end up consuming more energy overall, even if more efficiently. So lower energy costs overall may not be achieved and should not be the right focus. We also never explicitly thought about 'improved profit margins" though we did think about and hope that the investment would create increased employment within these firms, especially female employment, in light of the low LFPR among women in Kosovo.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will use secondary data to determine whether grant subsidies did in fact lead to more efficient use of energy, contribute to business growth, and correlate with increased employment of women.
MCC	85	This still does not include our effort to promote female hiring in the context of different contracts. We have a female employment target and requirements for contractors on this front. Might be good to include this here as it really is also part of the overall effort.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will take this into account while developing qualitative guides for the implementing contractors for the HER and AER activities.
MCC	85	Please see my comment above. I don't think this would be the right metric to use.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will update the instruments to measure whether grant subsidies did in fact lead to more efficient use of energy, contribute to business growth, and correlate with increased employment of women.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	85	Same as above – if her business grows as a result of the program, overall energy costs may increase, and this would not be contrary to our objective. We just want them to use the energy more efficiently. This should therefore be reframed as a goal towards more efficient energy use, not less energy overall.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. In the instruments, we will explore what energy-saving mechanisms have been installed to understand the change in energy efficiency from initiation to conclusion of the grant program.
MCC	85	See above. I disagree with this metric as well. Suggest instead adding one on employment, as that was a specific area of interest and goal.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will include a focus on growth of women's businesses in the instruments. We will also include a focus on increased employment in the STEM sector or related field.
MCC	86	As noted during our discussions, whether they find employment with the host company or any other company is equally valuable. Therefore, we should not limit this just to the host company as they want to help give them good experience to be more employable more generally, with that one particular company or another one.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will include questions around host companies' hiring interns in addition to recording employment in any company in the STEM sector fields.
MCC	86	What does this mean? I do not understand what this seeks to capture. You are referring to any future job descriptions at the companies? We are also not encouraging the companies to do this, and there are other metrics you can focus on if you want to see if they are willing to hire more women – e.g. by looking at % changes in male/female employees at company, though that could take likely some years to shift meaningfully.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.
MCC	86	If some of the 28 were to decide to pursue more schooling, how would you count that?	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. In our instruments we will assess academic pursuits/intent to pursue further education in the STEM sector fields as expressed at the conclusion of their scholarship experience.
MCC	86	As noted in our calls, the camps are not an annual thing, even if MFK had hoped to host a 2nd one. So, I am not sure this metric makes sense.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.
MCC	86	See comment above. This metric does not make sense as the camp is not set up to be annual.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.

Reviewer Name/ Institution	Page Number	Comment	Consultant Response
MCC	87	Internships are already starting so if this is the intention you need to get in touch with MFK asap to start any data collection	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. We will take this into consideration while deciding dates for baseline data collection.
MCC	87	As noted above, getting a meaningful answer on this can be tricky. For example, a company may change their job ads to say something like 'women especially encouraged to apply' but this may not change the overall numbers. Also, the internship component is not helping the companies revisit or adopt any guidelines, so I do not think it makes sense to tie this directly to that activity. It's too much of a stretch.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time.
MCC	88	Shouldn't you get this data before activity start to compare against baseline? Please note this is one of 3 evaluation factors for the program, and the applicant will have to submit to the WEE program some data on this as part of their application.	Noted. Per MCC's recommendation we will not make changes to the EDR at this time. In our instruments, we will shift this metric to focus on the shift in energy source and to record commitment to energy efficiency through investments in energy-saving mechanisms.