Evaluation of the MCC Lesotho Compact Metolong Program (MP) and Urban and Peri-urban Water (UPUW) Activity

Evaluation Design Report

June 2018

This report was prepared independently by Danae Roumis, Miguel Albornoz, and Robin Clanahan of Social Impact, Inc. at the request of the Millennium Challenge Corporation (MCC).

EVALUATION DESIGN REPORT

MCC Lesotho Water Sector Project

Metolong Program & Urban and Peri-Urban Water Activity

Summative Evaluation

June 2018

Submitted to:
Millennium Challenge Corporation
875 Fifteenth Street, NW
Washington, DC 20005-2221
Contracting Officer's Representative: Algerlynn Gill
Contract MCC-13-BPA-0017, Task Order MCC-17-CL-0005

Submitted by:
Social Impact, Inc.
2300 Clarendon Blvd., Suite 1000
Arlington, VA 22201
703.465.1884
www.socialimpact.com

CONTENTS

TABLES & FIGURES

TABLES

FIGURES

ACRONYMS

AGOA	African Growth and Opportunity Act
ATT	Average Treatment Effect on the Treated
BOS	Bureau of Statistics
CAD	Computer-Aided Drafting
CBA	Cost-Benefit Analysis
CEO	Chief Executive Officer
CGM	China Garment Manufacturers
DHS	Demographic and Health Surveys
DRB	Disclosure Review Board
DRWS	Department of Rural Water Supply
EA	Enumeration Area
EDR	Evaluation Design Report
EMC	Evaluation Management Committee
EOC	End of Compact
ERR	Economic Rate of Return
ES	Effect Size
FDI	Foreign Direct Investment
FGD	Focus Group Discussion
GIS	Geographic Information System
GPS	Global Positioning System
HMIS	Health Management Information System
HQ	Headquarters
IE	Impact Evaluation
IEMS	Impact Evaluation Multipurpose Survey
IPTW	Inverse Probability of Treatment Weighting
IRB	Institutional Review Board
ITSA	Interrupted Time Series Analysis
ITT	Indicator Tracking Table
IWS	Intermittent Water Supplies
JMP	Joint Monitoring Programme
KII	Key Informant Interview
LHLDC	Lesotho Housing and Land Development Corporation
LMDA	Lesotho Millennium Development Agency
LNDC	Lesotho National Development Corporation
M&E	Monitoring and Evaluation
MA	Metolong Authority
MCA-L	Millennium Challenge Account - Lesotho
MCC	Millennium Challenge Corporation
MDES	Minimum Detectable Effect Size
MDG	Millennium Development Goal
ML	Megaliters

MOHSW	Ministry of Health and Social Welfare
MP	Metolong Program
MPMU	Metolong Program Management Unit
MSME	Micro, Small, and Medium Enterprises
MTI	Ministry of Trade and Industry
NRW	Non-Revenue Water
O&M	Operations and Maintenance
ODC	Other Direct Cost
OSM	Open Street Map
PSM	Propensity Score Matching
QED	Quasi-Experimental Design
SD	Standard Deviation
SE	Standard Error
SI	Social Impact
SME	Small and Medium Enterprises
SOW	Scope of Work
TY	Teyateyaneng
QED	Quasi-Experimental Design
UPUW	Urban and Peri-Urban Water
UTM	Universal Transverse Mercator
WASA	Water and Sewerage Authority
WASCO	Water and Sewerage Company
WB	World Bank
WTP	Willingness to Pay
WTW	Water Treatment Works

1 INTRODUCTION

1.1 Background & Country Context

Lesotho's urban population has grown rapidly in recent decades, due in large part to job opportunities in the water-intensive textile and garment industry based in the capital of Maseru, as well as population growth and expansion of peri-urban areas. By 2008, domestic and industrial demand for water in urban areas was rising faster than the available supply. The combination of urbanization, growing demand, and aging infrastructure put strain on water networks in urban areas, resulting in declining reliability of piped water supply. Under these conditions, the utility (WASCO) was not able to expand to underserved urban and peri-urban areas. Further, a secure water supply was needed to attract new foreign direct investment (FDI) in the textile and garment industry.

In a Compact with the Government of Lesotho (GoL) implemented between 2008 and 2013, the Millennium Challenge Corporation (MCC) aimed to address these challenges through two major investments, including the Urban and Peri-Urban Water (UPUW) Activity and the Metolong Dam Program (MP). Taken together, the objective of the UPUW Activity & Metolong Program was to ensure an adequate and reliable supply of high-quality water in urban areas of Lesotho, for domestic, commercial, and industrial use. The UPUW Activity was implemented in ten urban and peri-urban locations, each comprised of a tailored set of new or rehabilitated infrastructure. MCC co-financed the Metolong Program (MP) along with a consortium of other donors, to increase and provide a long-term, reliable bulk water supply to Maseru and surrounding areas.

1.2 Evaluation Purpose

Social Impact (SI) was contracted by MCC to conduct an evaluation of the MCC Lesotho urban water programs, including the Metolong Program and the UPUW Activity. The evaluation is broadly concerned with the evaluability, implementation, impact, and sustainability of the projects along with lessons that can be generated from the evaluation findings for future MCC Compacts implementing similar programming. This evaluation serves an accountability purpose by systematically assessing actual project implementation and evaluating impact on beneficiaries relative to the project's stated goals, while also serving a learning purpose by highlighting ways that MCC can improve the design and implementation of other urban water projects implemented in similar contexts.

1.3 Design Report Objectives

This design report presents SI's approach to evaluating the impact of the UPUW Activity and Metolong Program, including design and methodology, data collection, analysis, reporting, and dissemination. The primary focus of this summative impact evaluation is on household-level beneficiaries, though the evaluation also includes elements targeted toward industrial firms and small and medium enterprises (SMEs) in Maseru.

The structure of this report is as follows. We review the Compact interventions, beneficiaries, geographic coverage, and program theory of change in Section 2, followed by a review of relevant literature in Section 3. In Sections 4 and 5 we present the evaluation questions, along with a summary of work completed to date on this study. In Section 6, we elaborate on the proposed methodology, including design and identification strategy, sampling, data collection, and analysis plans; this section also includes a discussion of anticipated limitations and risks. We describe plans for reporting and dissemination in Section 7 and conclude with administrative information in Section 8.

2 LESOTHO COMPACT & INTERVENTIONS

2.1 Overview of Metolong Program & UPUW Activity

MCC entered into a Compact with the GoL between 2008 and 2013. This 362.5 million-dollar Compact included activities in the water sector, the health sector, and private sector development. The Water Sector Project aimed to increase access to improved water supply and sanitation for rural and urban communities, and separate interventions were carried out for urban and rural areas.

SI's evaluation is focused on MCC-funded interventions in Lesotho's urban water sector, including the Metolong Program and UPUW Activity. The UPUW interventions varied by site, and included intake works, pipeline extensions, water treatment works (WTW), reservoirs, reticulation, provisions for new household connections, and other complementary components. The MP included a new dam, WTW, pumping stations, storage reservoirs, and downstream conveyance. MCC funded the WTW, pumping stations, downstream conveyance and the Metolong Program Management Unit (MPMU). The objective of the MP was to increase the bulk water supply available for Maseru and other lowland urban areas and transport the water to nodal reservoirs serving these areas. The construction and transfer of ownership of all infrastructure under both projects was completed by the Compact end date on September 17, 2013, though the defects and liability period extended well beyond this date. A summary of the MP and UPUW Activity is provided in Table 1.

Table 1. Summary of MCC Lesotho urban water interventions

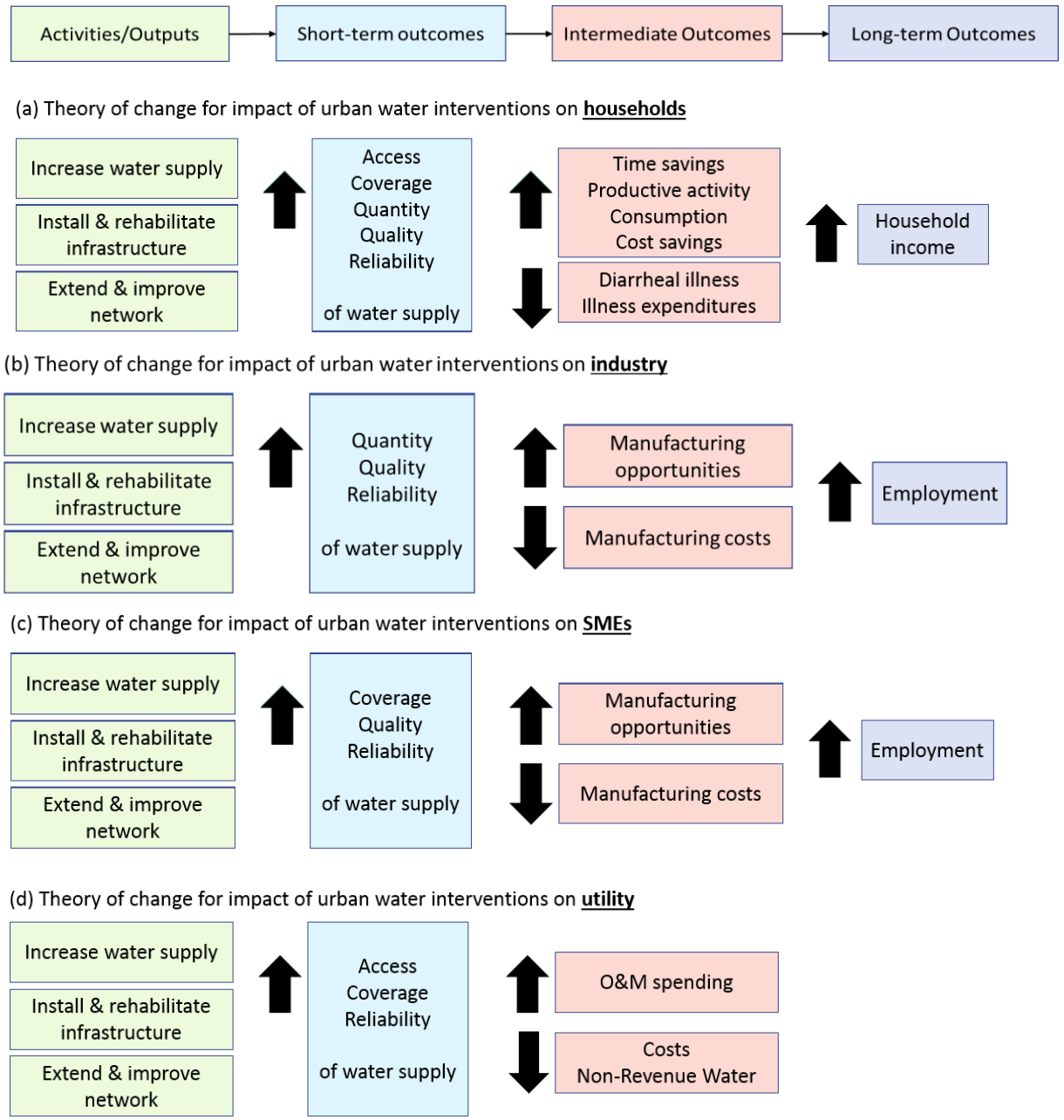
ACTIVITY	COMPONENTS
Metolong Program	**Construction of downstream water treatment works for the supply of water from the Metolong Dam to Maseru and the neighboring towns of Mazenod, Roma, Morija, and Teyateyaneng, and the establishment of the Metolong Program Management Unit ("MPMU").**
UPUW Activity	**Extension and rehabilitation of the urban and peri-urban water network.** ***Package 1* – Maseru, Mazenod, Roma, Morija, Teyateyaneng:** Rehabilitated reservoirs and pipelines; new reticulation, new household provisions, new public water points. Extended transmission to Teyateyaneng. Linked to the Metolong Program. ***Package 2* – Semonkong:** Water Treatment Works; water intake, new mains; new reticulation, household provisions, public water points; new reservoirs. ***Package 3* – Mafeteng, Mohale's Hoek, Quthing, Qacha's Nek:** Rehabilitate water treatment works; new water intake, mains; new reticulation, new household provisions, new public water points; new reservoirs; rehabilitated reservoirs, mains. ***Package 4* – Mokhotlong, Butha-Buthe, Leribe:** Upgrade Water Treatment Works; New and rehabilitated water intake, new mains; New reticulation, new household provisions, new public water points; New reservoirs; Rehabilitated reservoirs, mains. ***Package 5* – Mapoteng:** Chlorination, river source and water tank; Rehabilitate water intake, main; new community draw-off points; new reservoir.

Our impact evaluation design will primarily focus on the UPUW Activity. Households and businesses in Maseru and surrounding areas that were expected to benefit from the improved bulk water supply through the Metolong Program are supplied by the same water network which was targeted by UPUW Package 1. To the extent that the project logic of the two interventions is distinct and requires separate means of verifying project impacts, we will highlight how our evaluation design captures impacts of the MP and UPUW separately. Otherwise, our evaluation will measure outcomes of interest according to the UPUW site to which a given beneficiary or set of beneficiaries corresponds.

2.2 Project Logic and Economic Analysis

Figure 1 displays the theory of change for the Water Sector Project's urban water interventions. Each of the component theories of change, corresponding to the different intended beneficiaries, are described in detail in the following subsections along with corresponding economic analyses from MCC economic rate of return (ERR) calculations.

Figure 1. Theory of change for urban water interventions



2.2.1 Household Theory of Change

The household theory of change asserts that increasing the amount of water in an urban network, upgrading infrastructure, and improving and extending the network would lead to increases in access/coverage, reliability, and quality of water for households, which would result in time savings and reduced diarrheal illness, ultimately increasing the time and resources available for generating income. In general, as reported in our evaluability assessment, the project logic is plausible with some notable weaknesses, described below.

The logic posits that once infrastructure is in place, households will be willing and able to connect. It may be more plausible to assume willingness to pay in areas with existing networks, compared to Semonkong where water was previously provided free of charge. Similarly, the extent to which households could realize cost savings rests heavily on affordability, including connection fees and tariffs, and trade-offs in terms of coping costs (e.g. alternative sources, water treatment and storage, etc.). In Semonkong, this assumption may not apply given that residents were previously not paying for water. The project logic also assumes that increasing access to quality water would lead to increases in consumption of quality water. This may be complicated by multiple source use and household water storage in urban areas; it is also not clear from project documentation the extent to which quality in the system was expected to improve. Lastly, while one of the central assumptions of the project is that households will experience time savings, urban households in these contexts reportedly spend little time collecting water, especially relative to their rural counterparts, and it is not clear whether this time saved would have a meaningful effect on income generation in the household. Time savings may be more likely realized in previously unserved peri-urban areas, or Semonkong where the network is new.

Economic model: The three household-level benefit streams considered in MCC's economic rate of return (ERR) calculation for the UPUW activity include: (i) time saved in water collection, (ii) time saved in obtaining medical treatment for water-related illnesses, and (iii) reduced mortality for children under 5 due to water-related illnesses. The economic value of these benefits is basically assumed to be either the value of incremental labor availability from time not spent on water collection or obtaining medical treatment or an incremental increase in future income from a reduction in mortality for children under 5. The ERR assumes that households will experience a 50% reduction in time spent on collecting water and obtaining medical treatment for water-related illnesses, as well as a 30% reduction in mortality for children under five due to diarrheal illness. These three streams combine to contribute an 8.5% economic rate of return (ERR) on the UPUW investment, independent of any enterprise-level impacts.

2.2.2 Industry Theory of Change

MCC hypothesized that an increased supply of quality, reliable water would result in expansion and growth of industrial firms, ultimately leading to more employment opportunities and greater production. In practice, this hypothesis is specifically directed at firms in the "wet" textile and garment industries. These firms rely on water for productive uses in value-adding processes such as washing denim garments, laundering knit or woven garments, and dyeing textiles. This logic rests on two key assumptions: (i) an insufficient or unreliable water supply was a productive constraint to these firms at the outset of the project, and (ii) wastewater requirements also needed for industry expansion would be undertaken by firms themselves or other investors, given that water access was considered necessary but not sufficient for industry growth.

Economic model: The ERR for the Metolong Program is entirely directed at industry beneficiaries, identifying preserved and additional income resulting from the preservation and expansion of wet industry employment in Lesotho as its exclusive economic benefit streams. The ERR calculation assumes "that the termination of

AGOA[1] third country fabric provision will mean that a local milling facility is an absolute requirement for the continuation of the existing employment in the knitted fabric sector. Such a facility could not be developed without the increase in water supply afforded by the scheme, which is also a requirement for further expansion of the textile industry at Tikoe." The preservation of wet industry employment is the dominant benefit stream in the first four years of the ERR and is calculated as a function of the avoidance of reduced income in the knitted and woven fabric sector presumed to be enabled by a local milling facility. For the remaining sixteen years in the ERR, additional income is the dominant benefit stream, constructed by using factory space at the newly constructed Tikoe Industrial Estate as a proxy for total employment at the estate. It is assumed that each 4.00 m^2 of factory space will correspond to one new factory worker, and each five new factory workers will correspond to one new non-factory (commercial or office) worker. As the total factory space at Tikoe increases to 76,000 m^2, the additional income benefit stream eventually dwarfs the preserved income benefit stream by a margin of almost 4:1. Together, their estimated value leads to an ERR of 24.1% for MP.

2.2.3 Enterprise Theory of Change

Our review of due diligence and project documentation indicate that early planning documents did not differentiate between intended commercial- and industrial-level outcomes of the urban water programming. The two are often conflated in decisional materials, although the use of water by small and medium enterprises (SME) and large firms in Lesotho's wet industries, like garments and textiles, are reflected by different benefit streams in ERR calculations. For the purposes of our evaluation, SI assumes that the final benefit stream in the UPUW ERR is indicative of an SME-specific theory of change, while the benefit streams in the MP ERR are indicative of an industry-specific theory of change. Namely, the UPUW activity implicitly posited that increasing the amount of water in an urban network, upgrading infrastructure, and improving and extending the network would lead to increases in the supply of quality, reliable water for urban SMEs, who would benefit in the form of decreased manufacturing cost and/or increased manufacturing opportunities. These SMEs would, in theory, invest in productive capabilities because of improved access, reliability, and water quality.

Economic model: The final benefit stream in the UPUW Activity's ERR was an increase in private investment due to greater water availability. Given that Lesotho's garment and textile industry is centered nearly exclusively in Maseru, the situation of this benefit stream in the UPUW activity implies that this private investment would occur largely in commercial SMEs. It is assumed that the ratio of incremental private investment to infrastructure investment is 33% and that the share of value added in incremental investment is 0.86. The complementary private investment is assumed to phase in at a pace tracking the MCC investment in water infrastructure until it ultimately becomes a 63.6 million Maloti per year benefit stream from the 8th year of implementation forward. This benefit stream alone, if it were to be fully realized, would contribute an additional 14% ERR to the first three UPUW activity benefit streams, although the post-compact ERR calculation prorates this benefit according to a 50% likelihood that such complementary private investment will occur, making for a total UPUW Activity ERR of 15.5%.

2.2.4 Utility Theory of Change

For the utility (WASCO), MCC hypothesized that increased coverage and reliability would result in greater cost recovery, which could be allocated to operations and maintenance (O&M) to maintain the new infrastructure. This logic assumes that the new infrastructure would not introduce new costs to WASCO, and that they would solely benefit from the new revenue stream. To the extent that this infrastructure requires

[1] AGOA, or the African Growth and Opportunity Act, is a trade act enhancing market access to the U.S. for sub-Saharan African countries including Lesotho. The prime benefit of this act to Lesotho is that textiles and garments can be exported to the U.S. duty-free, allowing them to compete with cheaper firms from Asia.

increased cost to operate and maintain, tariffs must be high enough to cover these costs. Since there is an expected trade-off between the tariff and households' willingness to pay, increasing tariffs may result in slower coverage increases, and it is plausible that utility costs may rise faster than the customer base. So, the net revenue neutrality of the new infrastructure is essential at minimum for this benefit to be realized.

Economic model: Although increased incremental water revenue is included as a line-item in the MP ERR, it is "switched off," such that no allowance of it is taken in the calculation. This may be due to MCC's updated assumptions during the Compact that firms would not necessarily expand in response to changes in water supply due to the lack of milling in Lesotho.[2] Although other indicators relevant to WASCO were tracked during and following the Compact, including non-revenue water in M&E plans during implementation and amount of money budgeted for O&M in the post-Compact M&E plan, this specific indicator was never tracked and as such there is no explanation in M&E documentation of why the ERR calculation expected that it would not be relevant. Inclusion of this benefit stream would have increased the MP estimated ERR from 24.1%, accounting for industry-level benefit alone, to 29.5%.

2.3 Program Beneficiaries & Participants

The definition of beneficiaries is not consistent across all project documentation. Early project documentation defined beneficiaries as the full population of each urban area targeted by the MP and UPUW activity. Elsewhere, MCC posits that the project would benefit "urban and peri-urban people who are currently not being served and who do not have access to reliable/consistent supply." If this is assumed to be consistent with the full urban populations, it implies that all existing customers experience unreliable supply. In a later revision of the M&E plan, MCC references a new beneficiary analysis methodology, which results in a new estimate of 50% of the urban populations. The economic model focuses on the former, and estimates benefits for *previously unserved households*, which is assumed to be 40% of the urban populations. The UPUW Activity ERR also incorporated benefits to SMEs, which were removed from the M&E plan but retained in the ERR.

For the Metolong Program, the problem diagnostic acknowledged that rural poverty may be reduced through increased urban employment but was not formally accounted for as part of the beneficiary analysis. Likewise, the socioeconomic status of textile and garment industry workers who would experience the benefits of preserved or new employment was not explicitly integrated into the beneficiary analysis for the MP.

The following entities have the potential to benefit from the MP and UPUW Activity:

1. **Household beneficiaries:** Any household with a connection to WASCO, whether pre-existing or new in the ten UPUW sites (Maseru, Semonkong, Mafeteng, Mohale's Hoek, Qacha's Nek, Quthing, Butha-Buthe, Leribe, Mokhotlong, and Mapoteng)
2. **SME beneficiaries:** Owners and employees of any formal or informal SME with a connection to WASCO, whether pre-existing or new, in Maseru and surrounding areas.
3. **Industry beneficiaries:** Owners and employees of firms located on the premises of the Tikoe or Thetsane industrial estates. Targeted beneficiaries were textile and garment firms. However, we know from the scoping trip that Tikoe houses mostly non-wet industry firms at present.
4. **Utility:** The urban water utility, WASCO, which manages and operates networks across all of Lesotho, including in all ten UPUW sites. According to MCC's definitions, the utilities are considered 'participants' in the interventions, rather than direct beneficiaries.

[2] Based on communication from MCC Economic Analysis team in April 2018.

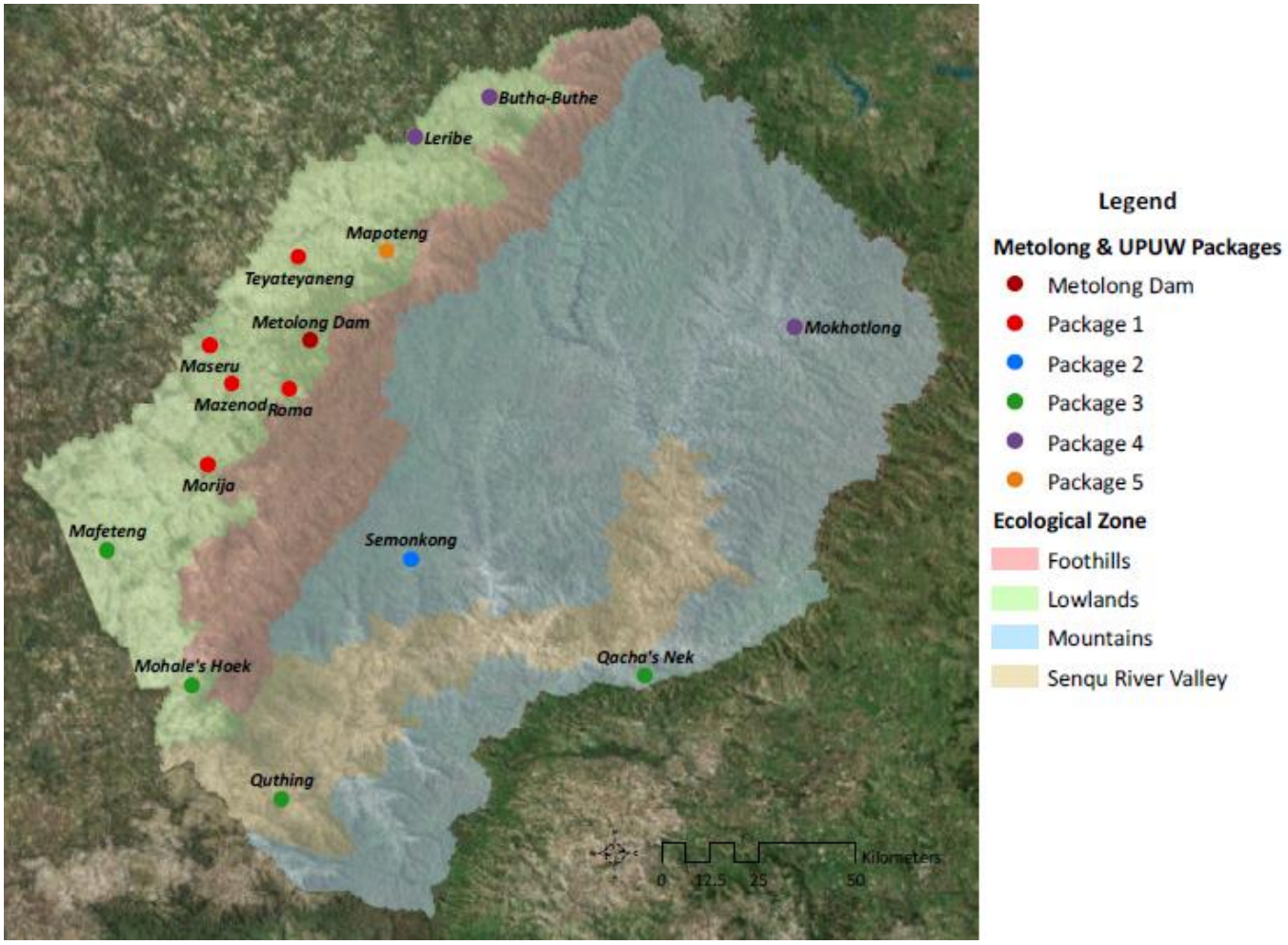
At the household-level, the original M&E Plan estimates that the UPUW activity had the potential to benefit up to 304,000 people (or 50,700 households), reflecting the entire population of the treatment areas. The final Compact M&E Plan revised this estimate downward to 124,248 people (around 20,700 households), including domestic and industrial consumers. Based on the closeout ERR, this reduction is likely to reflect MCC's own redefinition of beneficiaries to include only the portion of the urban population that did not have existing "piped, in-house water supplies." We cannot estimate what the number of households in the beneficiary group defined as such are with the information currently available. Borrowing the ERR figure that 60% of the urban population had a pre-existing connection, our beneficiary group is at least 1.5 times as large as the group defined by the ERR and the final Compact M&E plan, plus whichever proportion of the population has newly connected to WASCO networks since the closing of the Compact.

At industrial factories, improved water supply could also be expected to improve the sanitation and hygiene conditions in factories (partly depending on pre-intervention conditions), though as mentioned above employees themselves were not necessarily seen as beneficiaries, other than through increased employment at industrial firms. For SMEs, there is no estimate of how many or which businesses could be expected to benefit from the projects in original project documentation. For industry, however, the original M&E plan estimates that the MP would create employment opportunities for 39,750 people at the Tikoe industrial estate. No M&E plans cite employees in the knitted fabric sector as beneficiaries, although the MP ERR calculation assumes that 28,000 of these will have their employment preserved due to the bulk water supply. Increased employment at other industrial areas in Maseru, such as the Thetsane industrial area, was not considered although they are connected to the same network as Tikoe. Subsequent M&E Plans retreat from specifying the number of people for whom employment opportunities will be created, limiting the beneficiary analysis to a statement that the MP will provide 75 Ml/d of bulk water to the Thetsane/Tikoe industrial area and that industrial consumers are among the 124,428 individuals who will benefit from the increased urban supply.

2.4 Geographic Coverage

The urban water interventions were targeted primarily to urban areas in the lowlands of Lesotho, including the capital of Maseru and surrounding towns, as well as the majority of other urban areas around the country. Some of the targeted UPUW sites, including Mokhotlong, Semonkong, and Qacha's Nek, are located in the mountainous center and east of the country, with just one site located in the Senqu River Valley (see Figure 2). Some of these sites are considered peri-urban, since they are the most significant urban agglomerations in their respective surroundings but still have populations below 10,000 individuals with a mix of urban and rural livelihoods. These include Mapoteng, Mokhotlong, and Semonkong, as well as areas surrounding Maseru including Roma, Morija, Mazenod, and Teyateyaneng. In Semonkong, community members received water for free prior to the UPUW activity as part of the GoL's policy on rural water supply.

Figure 2. Map of MCC Lesotho urban water interventions



2.5 Key Program Indicators

In this section we summarize key indicators used in the Compact M&E Plans and Indicator Tracking Tables (ITT) to monitor the progress of the MP and UPUW Activity. We separate these out by targeted beneficiary group and identify how they are linked to the project logic. Where relevant, we include indicators that were dropped from the Compact due to feasibility, with the intention of inclusion during an evaluation.

2.5.1 Household

The household theory of change asserts that increasing the amount of water in an urban network, upgrading infrastructure, and improving and extending the network would lead to increases in access/coverage, reliability, and quality of water for households, which would result in time savings and reduced diarrheal illness, ultimately increasing the time and resources available for generating income.

Based on findings from SI's evaluability assessment, many indicators for intermediate and long-term outcomes including time savings, illness expenditures, and household productivity were removed from or never included in the M&E Plan during the Compact with the understanding that they would be included in an independent evaluation of the UPUW Activity. For example, the indicator for average time saved per household was removed from the M&E framework, because it was deemed not to be measurable within the Compact period (M&E Plan Amendment 2, p. 104). The only household-level indicator included in the fully amended (2014) M&E Plan was the number of households with provisions to connect to water networks. This was an output indicator with an End of Compact (EOC) target of 2,454 households. By EOC 2,312 households had been provided with such provisions, and two years following the Compact this figure had increased to 4,021 households, according to the Post-Compact ITT.

Many measures of short-term outcomes were included among the Compact's "goal" indicators, although these did not disaggregate between or attribute changes to the rural and urban water programs. Specifically, the Compact targeted an increase of the portion of the population with access to potable water from 71.5% at baseline to 92% at EOC. This figure was measured at 88.2% at EOC by the Lesotho Bureau of Statistics Continuous Multi-Purpose Household Survey. Earlier M&E plans included an indicator for new cases of diarrhea per 1,000 outpatient visits, with the goal of reducing these from 30.6 per 1,000 visits at baseline to 29 at 2013. However, the 2014 amendment of the M&E plan removed this indicator, citing that it was not feasible to measure in the absence of an independent evaluation.

2.5.2 Industry

For industry, MCC hypothesized that an increased supply of quality, reliable water would result in expansion and growth of industrial firms, ultimately leading to more employment opportunities and greater production. These desired long-term outcomes were originally reflected in the Compact M&E Plan with an impact-level indicator measuring the total factory workers employed by water-related industries in the Thetsane and Tikoe industrial parks. This figure was estimated as 22,700 workers at baseline with an EOC target of 40,000 workers. MCA-Lesotho planned to monitor this figure using LNDC reports. However, the indicator was stricken in the 2012 amendment to the M&E Plan on the basis that it was "no longer applicable as the result of de-scoping. The Urban and Peri-Urban Water Activity will no longer result in water connections to industries, hence no impact on business activity and employment is expected within industries during the Compact period." Nevertheless, the EOC target had actually been exceeded by 2010, with 42,000 workers employed at the two industrial parks. No indicators at the output, outcome, or impact level were included in the Post-Compact M&E Plan and ITT for industrial-level beneficiaries.

2.5.3 Small & Medium Enterprises

The implicit project logic for SMEs posits that increasing the amount of water in an urban network, upgrading infrastructure, and improving and extending the network would lead to increases in access/coverage, reliability, and quality of water for urban SMEs, who would benefit in the form of decreased manufacturing cost and/or increased manufacturing opportunities. These SMEs would, in theory, invest in productive capabilities because of the improved access, reliability, and quality of water.

At the outset of the project, MCA-Lesotho intended to use an enterprise survey to measure new enterprises opened due to the availability of water and new enterprises connected to the rehabilitated water network, with EOC targets of five and ten thousand enterprises, respectively. These indicators were revised in the 2010 Amendment to the M&E Plan to measure the total number of enterprises with an average water supply over 18 hours per day and average commercial water consumed at the business unit in cubic meters per month, still using an enterprise survey as a source. However, these indicators were removed in the 2012 amendment to the M&E Plan on the basis that the UPUW Activity would not count connections as a program output, and thus commercial water supply and consumption was no longer considered a direct outcome of the project.

The Post-Compact M&E Plan and corresponding ITT includes the only formally tracked program indicator for non-industrial enterprises: new commercial connections in the UPUW intervention areas. This indicator is characterized as an outcome indicator, which is a consistent reflection of earlier M&E Plans backing away from the assumption that actual water supply constituted an enterprise-level outcome. The indicator was not captured at baseline, but 2014 and 2015 measurements were reported as 719 and 601 new commercial connections, respectively. This metric does not constitute a true data point—it is measured as a simple 10% calculation of all new connections in the year reported by the WASCO Annual Report for the entirety of Lesotho. Hence, these figures are not only approximate but also not constrained to the UPUW intervention areas, as the definition of the indicator intends.

2.5.4 Utility

For the utility (WASCO), MCC hypothesized that increased coverage and reliability would result in greater cost recovery, which could be allocated to operations and maintenance (O&M) to maintain the new infrastructure. In the 2010 Amendment to the Compact M&E Plan, MCA-L introduced an output indicator particular to the MP measuring the annual megaliters of treated water delivered to WASCO from the Metolong WTW. The EOC target for this indicator was 27,375, with a Year 2 measurement of 0. It also introduced an outcome indicator for unaccounted for water, later renamed non-revenue water (NRW), which ultimately sought an EOC target of 25%, down from 34% at baseline. Several output indicators were added to these in the 2012 Amendment to the same plan, including: water pipes coverage, reservoirs constructed, reservoirs rehabilitated, and upgraded pumping stations. The EOC targets for these variables were 173.67 kilometers, eight constructed reservoirs, four rehabilitated reservoirs, and three upgraded pumping stations respectively. In the Post-Compact M&E Plan, an outcome indicator was added to track the amount of money budgeted for O&M by WASCO, mirroring the final intermediate outcome for this beneficiary. SI will mainly use administrative data and qualitative data from the process study to assess the way the projects affected the utility, including historical NRW by service center from 2010 to the present. SI will plot this data longitudinally and assess any discrete changes in the trend that may be associated with upgrades and additions to the infrastructure.

3 LITERATURE REVIEW

A review of relevant literature grounds the evaluation in the context of existing evidence, from Lesotho where available and from the literature on water and sanitation interventions more broadly. We present our review separately for expected impacts on household, small and medium enterprises, and industry.

We conducted our review by searching databases of peer-reviewed academic journal articles including RePEc (IDEAS, EconPapers), ResearchGate, PubMed, Web of Science, PLoS Medicine, ScienceDirect, as well as for academic, gray literature, and other research papers through Google Scholar.[3] To conduct our literature review, we used various combinations of the following search terms: *Lesotho, large-scale, urban, peri-urban, water, infrastructure, impact, evaluation, intermittency, reliability, quality, diarrhea, under five, time savings, wet industry, small medium enterprises, garment, textile, industry*. We also back-searched reference lists, and forward-searched for literature that referenced relevant peer-reviewed articles.

3.1 Summary of existing evidence

MCC's urban water activities in Lesotho were designed to support and extend urban water networks to improve and expand access to quality water for households, businesses, and industrial customers. Literature regarding the situation in the urban water sector in Lesotho prior to the Compact is somewhat mixed, with respect to how the sector's performance is presented. By 2002, it was apparent that the Water and Sewerage Authority (WASA) would not be able to meet future demand given existing resources for the critical industrial sector in Lesotho, nor for growing peri-urban populations relying on private vendors (Wason and Hall 2004). This analysis aligns with the problem diagnostic forming the basis for the Lesotho Compact investments. A World Bank study published before the Compact noted that while Lesotho performed better than average in the region with respect to access to improved water sources (at 88% of urban population), infrastructure to meet all domestic water needs was still lacking (Bogetic 2006). This same report noted higher than average access to improved sanitation (65% in Maseru), and suggested waste collection and treatment in urban areas was a larger issue than either of those above.

3.1.1 Households

The project logic underpinning these activities posits that upgrades to urban water networks will improve the production, quality, and reliability of urban water supply. This was expected to lead to greater coverage and access to improved water, resulting in benefits to the community including reduced water-borne illness, time savings, a reduction in defensive expenditures, and an increase in productive activity. Ultimately, this would lead to an increase in household incomes.

While there is a robust literature studying the link between various water, sanitation, and hygiene interventions on outcomes like those described above, most studies pertain to rural settings, with interventions implemented at the household level. Many studies focus on smaller-scale projects, including pilot programs with intensive, short-term, and closely monitored implementation. These interventions commonly introduce new technology or information that was not readily available previously, often at reduced or no cost, for the purposes of the research. In these cases, beneficiaries can usually be identified with precision and the 'treatment' they undergo

[3] RePEc (http://repec.org/); ResearchGate (http://www.researchgate.net); PLoS Medicine (http://collections.plos.org/water-and-sanitation); Google Scholar (http://scholar.google.com) . Search includes direct searches from relevant journals including Environmental Science & Technology, Journal of Water and Health, Water Resources, Sci Total Environment, WHO Bulletin, J Tropical Medicine & International Health.

is clearly defined. While the rural literature can provide some relevant evidence, context must be considered carefully when attempting to transport findings to the urban setting.

In contrast, MCC's investments take place in dynamic urban contexts, where outcomes of interest are subject to influence by a wide range of factors different from those in rural areas. MCC's investments are large-scale and system-wide, implemented through existing institutions and processes, focused on infrastructure and institutions more often than directly targeted to household beneficiaries. As a result, the populations impacted tend to be more numerous, geographically dispersed, and socioeconomically diverse. It can be challenging to pinpoint *a priori* which households will benefit and to what degree. Further, urban households are more likely to have access and use multiple sources of water already, varying in terms of reliability, quality, and convenience (Cairncross and Kolsky 1997; Stoler et al. 2015; Bello et al. 2010), meaning that substitutions rather than discrete switches to improved sources are more relevant in urban contexts. Indeed, a feasibility study carried out in advance of the MCC Compact in peri-urban areas around Maseru (Mazenod, Roma, and Teyateyaneng) found multiple source use reported by households in these areas varying by domestic activity (study described in MWH Due Diligence Study Project B final report). Even less has been written specifically about peri-urban settings, which may resemble urban areas with respect to economic access (e.g. to urban jobs or markets) but on the other hand are marked by a substantially lower level of access to public service provision (Swedish Water House 2007). A 2010 study estimating water demand in peri-urban areas of Lesotho found that the vast majority used piped water in some way, even though a small minority had access to their own tap (Bello et al. 2010); about half of the households in the sample used a public tap, and a third used water from a neighbor's tap. That study also notes the importance of reliable water to peri-urban households in Lesotho for gardening purposes as a source of income for providing fresh produce to nearby urban centers.

Rural literature: Overall the rural literature points to a causal link between improvements in water supply, improved health outcomes, time savings, reductions in coping expenditures, and productive activity (Whittington et al. 1990). Time savings resulting from a reduction in water collection time has been associated with an increase in productive activity (Galiani et al. 2009), as well as improved health outcomes for children (Pickering and Davis 2012). However, even within the context of these findings, meta-analyses and systematic reviews consistently find that source quality improvements are less effective in reducing diarrheal illness than point-of-use interventions, due to contamination during storage or food preparation. Moreover, sanitation and hygiene interventions are more effective in achieving improved health outcomes, as they target the direct cause of the disease transmission pathway (Waddington 2009; Fewtrell and Colford 2004; Falconi et al. 2017; Fan and Mahal 2011).

In some cases, quantity of per capita water consumption has been linked to reductions in diarrheal illness which is thought to be primarily related to the fact that increased water consumption is allocated toward sanitation and improving environmental health conditions in the household, whereas drinking water consumption has been shown to remain relatively stable regardless of total per capita water consumption (Thompson et al. 2001). A recent impact evaluation for MCC's rural water investments in Lesotho under the same Compact found that the interventions resulted in time savings of an average of 44 minutes per day, a significant reduction from 105 at baseline. The rural evaluation did not find statistically significant impacts of that program on diarrheal illness among children under five, illness-related expenditures, school absence, or the total number of hours worked by women in beneficiary households. An older study of diarrheal illness in the rural Lesotho Highlands found that households using improved water sources (protected springs) were less exposed to *E. coli* (Kravitz et al. 1999). Cooper-Vince et al. (2017) find, in a study from rural Uganda, significant association between water insecurity and school absenteeism.

Urban literature: In Lesotho, as is the case globally, the combination of continued urbanization, environmental changes, and aging infrastructure are increasingly putting enormous stress on public service provision in urban centers, including in water and sanitation (Bello et al. 2010; McDonald et al. 2014; J-PAL 2012; Gwimbi et al. 2011; Thompson et al. 2001; White et al. 1972). Increases in intermittency lead households to coping behaviors such as household water storage and use of alternative sources. Expansion in service has generally lagged the pace of urban population growth, and utilities face institutional challenges managing urban water supplies (Molapo 2005; Hunter et al. 2010). Despite this, the literature pertaining specifically to the impact of large-scale urban water infrastructure projects is limited. In fact, evaluations of MCC urban water investments in other countries are some of the more relevant examples of such studies. Below we summarize literature pertinent to the hypothesized impacts of the Lesotho urban water projects, including underlying assumptions in the project theory of change.

The project logic hypothesizes that conditional on the interventions' completion according to plan, an increased and reliable supply of quality water will lead to expanded coverage or access to improved water. In the context of the MCC Lesotho urban water interventions, this can be conceptualized in two ways: (1) improved service for those with existing connections, and (2) new connections for those previously unconnected. The populations most likely to benefit from improved service in Lesotho's urban areas are higher-income households, since they were more likely to have existing connections, while – if achieved by the project – an expanded coverage of new connections would be more likely to benefit poorer households (Molapo 2005). An increasing amount of research points to the link between reduced intermittency and improved water quality (Ercumen et al. 2015; Nelson and Erickson 2017; Adane 2017; Jeandron 2015; Kumpel and Nelson 2013). Recent research points to the large burden of disease attributable to intermittent water supplies (IWS), emphasizing that improved water sources do not uniformly offer safety from microbial contamination (Bivins et al. 2017; Shaheed et al. 2014). Therefore, with respect to both existing and new customers, the question is not only one of access to an improved or piped source, but also reliability and safety of the supply.

There is limited evidence regarding the impact of improvements in urban piped network infrastructure on take-up rates for new household connections. However, there are several studies assessing willingness to pay for new connections. Up-front costs remain a key barrier for many households even where network expansion is completed. Evidence from studies in Zambia and Kenya suggest low willingness to pay for improved quality (Ashraf et al. 2010; Kremer et al. 2011; Null et al. 2012; Blum 2014), while a study from Morocco suggests high willingness to pay for the convenience of a tap on premises relative to a public standpipe (Devoto et al. 2011). Further, some studies show household preferences for water source correlate with aesthetic quality often more than microbiological quality or even price (Kulinkina et al. 2016; Spencer 2008). Citing WTP studies conducted in Lesotho in 1996 and 2002, the feasibility report for the UPUW Activity reports that unconnected households exhibit a high willingness to pay specifically for the improved convenience and reliability of a connection but were likely to be severely constrained by steep connection costs (MWH Project B final due diligence report). It is not clear at this time whether WASCO has instituted any means to reduce this barrier for potential new customers (e.g. through connection fees paid in installments). A willingness to pay study conducted in 1996 in Lesotho found that in peri-urban areas, willingness to pay was positively associated with service quality, especially with regard to reliability (Sechaba Consultants 1996, cited in MWH Project B final due diligence report).

Households who gain new connections are likely to experience improvements in the reliability and quality of their water. There is wide consensus in the literature that water quality from household taps is an improvement upon public taps or other unimproved sources. One study published during the MCC Compact studied water

quality in a peri-urban area of Maseru district (Manonyane community) found that protected sources were significantly less likely to contain *E. coli* (7%) compared to unprotected sources (40%) (Gwimbi 2011). Still, it is important to bear in mind two important caveats with respect to these MCC interventions. First, the designation of a given water source as improved does not guarantee that it is free from contamination with fecal bacteria. Until recently, this designation was based only on the type of drinking water source, but it is increasingly recognized that source quality can vary substantially, sometimes with high levels of contamination in improved sources (Bain et al. 2014; Onda et al. 2012; Bartram Cairncross; Wright et al. 2004; Martinez-Santos 2017). As described earlier, water collection, storage, environmental conditions in the household, and multiple sources use in urban areas further complicate this relationship.

The project logic also assumes that improved access to quality water will lead to an increase in consumption of quality water, resulting in improved health outcomes, especially diarrheal illness among children. In general, the literature bears this hypothesis out, with piped households generally able to consume orders of magnitude more per capita per day, compared to households without a connection on premises. Molapo (2005) reports from a survey in Maseru a household average of about 500 liters for piped households, compared to 330 with a yard tap and 43 for unconnected households – dividing by the average household members, this provides comparable estimates to those reported in the baseline report by the previous evaluator (NORC), SI's recent Tanzania work, and other estimates in the literature. In urban contexts where multiple source use is common, households could potentially substitute lower quality sources which they treat, to higher quality sources that they do not treat (Onjala 2013). Bello et al. (2010) confirm that peri-urban areas had per capita consumption of water pre-intervention similar to that of rural areas (they estimate about 20 lpcd on average).

Improved access to quality water is hypothesized to result in time savings, due to a reduction in the amount of time spent collecting water, which is, in turn, expected to be re-allocated to productive activities. Baseline findings from Lesotho from 2010 (NORC report) report on the results of a household survey asking respondents their perceptions of how additional time savings might be allocated. Many households noted they would like to use additional time for income-generating activities, including starting a business. Nonetheless, there is limited evidence in the literature substantiating this hypothesized relationship in urban areas to the same degree as is documented in rural areas. One study from Morocco found that even significant time savings gained from new household connections (an estimated 82 minutes in the three days preceding the survey) were not re-allocated to productive activity for either men or women, but rather to leisure and social activities (Devoto et al. 2011). On the other hand, a retrospective study published recently from Kenya found time savings from water collection as a result of gaining access to piped water, which was allocated to income generating activities (Bisung and Elliot 2018). Bello's 2010 study in peri-urban Lesotho found that about half of households spent up to 10 minutes per round trip to collect water, with another third between 11-30 minutes. Only five percent of households reported spending more than an hour. While this doesn't directly address total hauling time on a weekly basis per household, it does give a sense of the situation pre-MCC interventions. While there is evidence that the greatest burden of water collection and water shortage still falls on women and girls even in the urban context, baseline qualitative findings from Tanzania offer additional nuance demonstrating that urban households are less likely to keep children home from school for water collection chores, recognizing the importance of school attendance. School absence, in that context, was ascribed more often to water shortages resulting in the inability to bathe or wash clothes. Likewise, there was very low reported absence or caregiving due to diarrheal illness in Tanzania at baseline.

The project logic assumes that households will experience cost savings as a result of lower expenditures on water and on treating water-borne diseases. A 2002 update to a previous willingness to pay study in Lesotho found that about half of unconnected households paid for water from other sources, at rates up to two times

as much as the lowest band of connected households (Sechaba Consultants 1996, cited in MWH Project B final due diligence report). This suggests the potential for cost savings among those who became connected as a result of the MCC Compact interventions, though the extent to which this might have occurred in reality also depends on price changes, tariff changes, household preference, source availability, and other potential WASA investments in the several years between that study and the MCC Compact. A related and important finding from that study is that willingness to pay for unconnected customers decreased dramatically as tariff bands increased. There was virtually no willingness to pay in the highest tariff band – the only one offering cost recovery to the utility. SI's recent analysis of short-term results from the Jordan Compact found self-reported improvements in supply reliability and increased water consumption, but no significant reductions in water expenditures on expensive alternative water sources. Bisung and Elliot (2018) also found cost savings for households due to access to improved sources, spent primarily on food.

It is important to note that the project logic simultaneously assumes an increase in consumption of quality water and a reduction in water expenditures. While in some cases this may be plausible (e.g. for households paying extremely large mark-ups for water from private vendors), it is not necessarily the case that households will save money as a result of the interventions. MCC's assumption of cost savings also includes a reduction in spending on diarrheal illness. Also of note is that MCC's problem diagnostic does not address other coping costs or defensive expenditures, such as treating drinking water, storing water for domestic purposes in the household, and reliance on multiple alternative sources. These represent other types of expenditures that could potentially decrease as a result of a household's gaining improved access (whether through an existing or new tap), but as above it remains to be seen whether any potential reductions for these costs exceed potential increases that may come simply through a substantial increase in consumption. It might also be the case that households set a certain water budget, given income and other constraints, and adjust their demand and prioritize within that budget. For example, a peri-urban household may switch from an expensive private vendor to a tap on premises and consume within their existing budget (assuming the vendor was more expensive than the tap), with a net zero impact on water expenditures.

3.1.2 Industry

The ERR calculation for MP states: "The key assumption, which has in due course proved valid, is that the termination of AGOA third country fabric provision will mean that a local milling facility is an absolute requirement for the continuation of the existing employment in the knitted fabric sector. Such a facility could not be developed without the increase in water supply afforded by the scheme, which is also a requirement for further expansion of the textile industry at Tikoe." In our evaluability assessment, we further clarified that this logic rests on an assumption that water access is a necessary but not sufficient condition for industrial growth, and that the wastewater requirements also needed for industry expansion would be undertaken by firms or other investors.

Due diligence documentation and concurrent WASO reports from before the Compact provide ample evidence that industrial and consumer demand combined to overburden the water supply available to urban areas before the provision of MP infrastructure. Specifically, WASCO's 2007/2008 report states, "The demand for water for domestic and industrial consumption, particularly in the designated urban centres, has increased tremendously and currently the Authority can partly meet the actual demand. It is observed that the water shortages can have serious repercussions on both economic development and public health, hence a need to expand the availability of potable water." While our review of the literature gives little indication that this constraint persisted in the years during and following the contract, it does find some evidence that a lack of wastewater treatment facilities may be an ongoing constraint to the local textile and garment industry.

We conducted a review of academic and gray literature to verify the extent to which water supply was, or still is, considered a constraint to industrial firms in Lesotho. As such, we reviewed literature specific to Lesotho. Most of the available literature on industrial water usage in Lesotho is focused on wastewater processes, including what environmental and health effects these processes can have on downstream communities.

A 2007 dissertation is frequently cited in papers concerned with industrial water usage in Lesotho since it uses a case study approach to intensively describe water management at the China Garment Manufacturers (CGM) Industrial Ltd. textile factories in Maseru (Masupha 2007). Although the paper does not describe water as a constraint to production at the plant visited in 2005, it does provide a granular view of how water is consumed and discharged at a typical industrial factory. Masupha describes four wet "process combinations" employed by the firm that each comprise unique combinations of eleven possible wet "operations." Assuming a 70kg load of garments, the paper finds that the process-combinations (including stone-washing/washing, stone-washing/bleaching, stone-washing/dyeing, and stone-washing/ bleaching/dyeing) require between 6,000 and 10,000 liters of water per load, plus an additional amount on the margin for quality-assurance checking of the garments, cleaning the factory floor, or generating steam to press the garments. Besides these process-combinations, the factory also used water for employee consumption either for drinking (4 liters/employee/day) or for using the restroom (30 liters/employee/day).

In all, the main CGM factory typically consumed 1.15 million liters of water per day for the wet process combinations, 10,000 liters of water per day for employee drinking water, and 75,000 liters of water per day for employee restroom use. Because the effluent water discharged from use for the wet processes contains chemicals that interfere with municipal wastewater processes, only the water used for employee consumption can be discharged into the municipal wastewater treatment system. The remainder must be treated in an on-site wastewater treatment plant specifically designed to remove pollutants from the manufacturing process before it can be recycled or discharged into local rivers or streams. In the whole of Maseru, the paper estimates that 12 Ml of the 28 Ml per day supplied to the city in 2002 was consumed by textile industries. 10.2 Ml/day of this was discharged as effluent, with the remainder lost due to evaporation and spillage. At least at the CGM Industrial plants, in instances that factories were discharging effluent in excess of what their wastewater treatment plants could accommodate, the effluent was being discharged directly into local streams and rivers.

Given the process described by Masupha, most academic work on water use by industrial firms in Lesotho is actually much more concerned with the end-state and consequent health and environmental effects of industrial wastewater than the constraint that water supply places on industry. Most papers that do address constraints on industry in Lesotho are more concerned with macroeconomic factors—particularly Lesotho's status as a privileged exporter to the U.S. market in the context of the African Growth and Opportunity Act (see Gibbon 2003; Bennett 2006; Morris, Staritz, and Barnes 2011; and Staritz and Morris 2013). One paper does corroborate the due diligence documentation's assertion that water shortage was a constraint to these industries prior to the Compact, and that any additional water diverted from a local river would have exacerbated water shortages for the local population at the time (Lall 2005). There are also mentions of water constraints in Lesotho's garment and textile industries in reports by the Overseas Development Institute (ODI), summarizing a project that ended in 2009, and in a 2004 IMF Country Report. The ODI report listed water (and wastewater) infrastructure among the list of constraints to the industry, which also included, "lack of physical infrastructure such as factory shells; the HIV/AIDS pandemic; inadequate training of supervisors and labour; and the high cost nature of the industry owing to low productivity" (ODI, n.d.). Similarly, the IMF country report listed the lack of water and factory shells as major constraints to growth (IMF 2004).

Another paper concludes that in industrialized areas of Lesotho, surface water quality has been significantly affected by industrial effluents being released into them, finding that water quality downstream of industrial

areas exceeded almost all South African thresholds for domestic and irrigation uses and was significantly worse than control areas sampled for the purposes of comparison (Pullanikkatil and Urama 2011). A separate dissertation found that as recently as 2013, the area around the Thetsane industrial area was "plagued by untreated waste water discharged into nearby rivers (Kamlana 2014).

Gaps in literature: Although the textile and garment industry in Lesotho has been heavily studied as an engine of growth and job creation, the intersection between the growth and expansion of this industry and water is not heavily explored. From a technical perspective it is clear that water supply is a key condition for firm growth and industry expansion, but recent literature does not address whether the increased bulk water supply made available by the MP and other concurrent programs such as the Tikoe-Thetsane Industrial Water Supply Project would have resolved water-related constraints to growth or if a lack of on-site wastewater treatment facilities has prevented this increased bulk supply from being fully utilized. The literature suggests, at least, that the industrial wastewater currently being produced by factories in the Thetsane industrial area is not adequately treated before it is discharged into local rivers and streams. Although the importance of water-related constraints for industrial growth and expansion may not be as severe as others, e.g. macroeconomic concerns, our research could help inform the extent to which these constraints are being considered by industry-leaders who may drive the industry's growth in Lesotho.

3.1.3 Small & Medium Enterprises

In general, the UPUW Activity's logic for improving the productivity of SMEs is poorly defined. Due diligence documentation does not establish the extent to which reliable water supply is a constraint to urban SMEs in treatment areas. The logic is perhaps best summarized in the ERR calculation as an expected private investment in response to the availability of water. A review of peer-reviewed and gray literature on the linkages between water supply and SME growth and prosperity in the Sub-Saharan Africa region yields few insights on the subject, indicating it is not a heavily trafficked area of academic concern.

Lesotho's official poverty reduction strategy prior to and towards the end of the Compact (2006 and 2012) emphasized embracing "micro, small, and medium enterprises" as engines for poverty reduction, given that an estimated 100,000 of these existed in the country in formal and informal capacities as of 2012 (IMF 2006). In 2006, constructing market centers with adequate access to infrastructure, including sanitation and water systems, was part of a strategy to support SMEs that was the highest ranked strategy in the "employment creation and income generation" priority (IMF 2012). This strategy of providing property solutions to SMEs persisted in the 2012 Poverty Reduction Strategy. This could potentially be related to a certain constraint pointed out by a 2004 World Bank report that, "SMEs are subject to the same corporate tax, electricity, water, and telecommunications rates as large corporations" (Farole and Winkler 2004).

Nevertheless, we find very little academic discussion of water as a constraint to SMEs, whether in Lesotho or in the Sub-Saharan Africa region. A 2011 study comparing water abundant "river towns" and arid "Karoo" towns in rural areas of South Africa in terms of enterprise development finds no major impacts of water abundance on the quantity or quality of enterprise assemblages. Where differences in these outcomes do exist, they are characterized by more typical economic thinking such as money circulation through a settlement aggregation either by wealthy inhabitants or by some commercial export base that attracts money from non-inhabitants (Toerien and Seaman 2011).

There are two studies, both somewhat dated, that specifically focus on the enterprise response to deficient infrastructure. The first, focused on SMEs in Kenya, finds that deficient water supply is most constraining to enterprises in the food sector, and generally was reported as a constraint more frequently by micro to medium-sized enterprises than by large or very large enterprises. Even then, only 5.4%-12.5% of surveyed enterprises

this size indicated water infrastructure was their greatest problem, compared to much larger figures for electricity, roads, and waste disposal (Kimuyu and Kayizzi-Mugerwa 1998). The second study, focusing on Ugandan enterprises, finds that economic benefits to SMEs of water supply improvements may be limited (Davis et al. 2001). These studies do little to dispel the notion presented by other studies that larger macro-economic conditions are more predictive of SME growth and profit.

While literature pertaining specifically to Lesotho is limited, we have some insights from a recent report by the Ministry of Trade and Industry (MTI), and another study conducted by Finscope on financial inclusion. A 2008 study conducted by the Ministry of Trade and Industry (MTI) in Lesotho stated that infrastructure needs, including water, were a larger constraint for those located in informal areas; according to their report, about 20% of enterprises located in the Lowlands ecological zones lacked access to adequate water infrastructure (MTI 2008). A more recent Finscope study looked at financial inclusion among a sample comprised largely of informal micro-enterprises (Finscope 2016). Water was not mentioned as a constraint to start a business by most respondents. The most binding constraint to starting a business was sourcing money, mentioned by 49% of respondents, with not enough customers (19%) and cash flow (19%) as the next most cited constraints. Connecting to water service was mentioned as an operational constraint by just 2% of respondents, with other constraints seemingly overshadowing this completely (major constraints instead were listed as sourcing money (35%), cash flow (26%), being owed money (22%), not enough customers (16%), too many competitors (13%), and several others above water). It was not mentioned at all as an obstacle to growing the business, with major constraints there instead listed as access to finance (20%), space to operate (16%), competition (15%), and others. Disaggregation of these data specifically for SMEs are not reported.

Gaps in literature: The literature reviewed does not present any quantitative estimates of the effect of water supply on SME growth, profit, or assets. While a gap in terms of our evaluation, it is not clear that this constitutes a gap more broadly in the sense that no other studies concerned with SME growth are calling for clarity on the issue. Nevertheless, to the extent that our evaluation finds effects of water access or reliability on SME investment or growth, it appears this information would be new to the literature.

3.2 Contribution of evaluation

To our knowledge, this would be the first evaluation assessing the effect of large-scale urban water infrastructure interventions on urban households across Lesotho. It will add to the currently scant body of literature documenting the impact of such interventions on economic and health outcomes for urban and peri-urban households. It will also add to the currently available literature pertaining to water as a constraint for industry and SMEs in Lesotho. For MCC specifically, this evaluation contributes both to accountability and learning objectives. In sites where the interventions were implemented successfully, the results would provide an estimate of the benefits provided by these interventions. Results from sites where new or upgraded works are not functioning as expected provide an opportunity to gather information about how challenges can be avoided in the future, as well as document any unintended consequences on the surrounding communities. Ultimately, this evaluation will produce lessons learned that can be directly applied to future large-scale water sector interventions.

4 EVALUATION QUESTIONS

As a whole, SI's evaluation is designed to answer the evaluation questions below. Q1 was addressed by the evaluability assessment, while Q2 through 6 and 7d were addressed through the process evaluation completed at the end of 2017. **The primary focus of the impact evaluation is on Q7a through 7c.**

Guided by the evaluation questions, the impact evaluation is focused on measuring change among a sub-set of the short-term and intermediate outcomes listed in the project logic, including access to quality water, water consumption, water-related illness, time savings, and cost savings. To these outcomes, we also add consumption of quality water, given its pivotal location in the hypothesized causal pathway between access and water-related illness, as well as its direct relevance to Q7b.

1. Is the program evaluable?
2. Was the program implemented according to plan? Are interventions operating according to plan? If not, what are the major issues, and to what extent were they affected by implementation fidelity?
3. What is the current functionality, use, and plan for managing and maintaining the infrastructure under the Metolong Program and UPUW Activity?
4. To what extent has a management unit been established for the Semonkong water system? To what extent has WASCO HQ provided support to those managing the new system in Semonkong?
5. To what extent has support been provided to the Water and Sewerage Company (WASCO) for the management of Metolong Dam, Water Treatment Works, and Pump Stations? If provided, who provided it, when was it provided, and how effective has this support and dam management been? Does a staffing plan exist for Metolong Dam? To what extent are positions occupied and what has turnover been to-date?
6. Do Operations and Maintenance plans exist for the Metolong Program and UPUW assets? How are these plans budgeted and funded? Are these O&M plans being observed and carried out?
7. **What were program results on key short-term and intermediate outcomes?**
    a. **To what extent has access to quality water increased? What activities, if any, has WASCO conducted to encourage households to connect to the network?**
    b. **To what extent are community members (including businesses such as manufacturing firms) using water from the urban water network and how has this changed since the Lesotho Compact started?**
    c. **To what extent are community members experiencing cost and time savings, or reductions in water-related illness?**
    d. How have the MP and UPUW programs impacted WASCO's income and costs? Has additional income been generated that can be directed to maintaining the new infrastructure?
8. What lessons can MCC or the Government of Lesotho apply to future programs related to program design, implementation, and sustaining results?

5 EVALUATION PROGRESS TO DATE

5.1 Evaluability Assessment

In responding to evaluation question 1 ("Is the program evaluable?"), SI conducted an evaluability assessment in March 2017. SI conducted this assessment using a version of MCC's five dimensions of evaluability framework that was adapted for appropriateness to an *ex post* evaluation context. This framework emphasized the technical feasibility of a potential evaluation alongside the value of undertaking an evaluation. The assessment investigated the extent to which the MP and UPUW activity had a well-defined problem diagnostic, a plausible and clear theory of change, an anticipation and understanding of realized risks, clearly identifiable beneficiaries, and a documented monitoring and evaluation framework with relevant indicators.

The assessment found the problem diagnostic to be mostly well defined for both MP and UPUW, particularly with respect to identified water shortages in the urban areas of Lesotho. It was weaker for the portion of UPUW that was justified by unreliable and low-quality water, with little quantitative evidence supporting that such a problem existed. Although most causal links for both activities were logically plausible, they were in general more typical of rural water interventions. Several realized risks were uncovered during the scoping trip for both projects that could jeopardize intended benefits for key groups of beneficiaries. Some UPUW infrastructure appeared to be non-functional, and the lack of wastewater treatment at Tikoe industrial estate may have prevented intended benefits from being fully realized, despite an improved water supply. While beneficiaries had been clearly defined in a broad sense, they had not been precisely identified or located, which appeared contingent on high-quality GIS files or other data. Finally, although the framework for measuring and evaluating results was clearly documented, many key indicators were ultimately dropped during the Compact.

The sum of these findings was that, although it was technically feasible to evaluate the projects, we would need to sequence our process evaluation ahead of our summative evaluation to first establish a clear understanding of project implementation and current operational status of the infrastructure, to inform our plans for evaluating the impact of the MP and UPUW Activity, since poor performance was likely to prevent some intended downstream impacts from materializing.

5.2 Process Evaluation

An SI team conducted data collection for the process evaluation in September of 2017. The purpose of the process evaluation was to assess how the project was implemented and managed, report on current functionality and use, and derive lessons learned that can be applied to future Compacts. It included two components: an implementation fidelity assessment and a performance evaluation. The process evaluation included both the MP and the UPUW Activity.

The implementation fidelity assessment primarily addressed evaluation questions 2, 3, and 6. The purpose of the assessment was to determine whether the program was implemented according to plan, as well as document current functionality and use of the works, and where applicable, to assess where observed problems in functionality are originating. We conducted structured observations at each project site, as well as key informant interviews with relevant stakeholders. Information from the site visits and key informant interviews was used to develop implementation fidelity scores for each installation. In order to adequately compare intended design to realized design, SI's Water Supply Expert populated the design and function requirements of relevant infrastructure in the structured observation protocols for each site in advance to the greatest extent possible using existing documentation. The performance evaluation component of the process

evaluation addressed evaluation questions 4, 5, and 7d through a combination of document review, site visits, and key informant interviews. In total, the SI team conducted thirty-one key informant interviews that included forty-two key informants.

We assessed implementation fidelity based on the following four elements: (i) Design, (ii) Installation, (iii) Management/O&M, and (iv) Funding. Collectively, these elements account for the most important factors contributing to whether the infrastructure is fully functioning and achieving intended results. We implemented a unique scoring scale for each of the four elements, which was summed to arrive at an overall fidelity index score. A single score allows comparison between different installations and types of infrastructure that might be quite different in nature. A justification for each score with supporting information, including photographs where applicable, was also included in a narrative for each installation.

The detailed implementation fidelity methodology and scoring criteria are outlined in SI's Process Evaluation EDR , while preliminary results were reported through a detailed trip report annex and subsequent presentation to MCC in November 2017; these documents are not public, but the results will be fully integrated into the final, public evaluation report. Below, we provide a high-level summary of the results of the implementation fidelity component of the process evaluation, to contextualize the discussion of the IE design in this report.[4]

In total, SI observed Compact works at eleven different sites. These site visits confirmed that many of the UPUW works were non-functional, requiring remediation from WASCO since the end of the Compact. While conducting the structured observations, SI's Water Supply Expert crafted a parallel performance ranking methodology meant to characterize the perceived functional status of the installations prior to any WASCO remedy, such that the implementation fidelity score could communicate actual implementation against planned implementation and the ranking system could communicate a forward-looking distance from adequate functioning for the infrastructure.

The perceived status ranking is a combination of a letter and number, such that the letter represents a category of functionality and the number represents the number of discrete issues requiring resolution for an installation to operate to design expectations. An "A" signifies performance up to design expectations. A "B" signifies failure to meet design expectations that does not require extensive or costly work for remediation. A "C" represents an adequate design that fails to meet expectations due to installation errors that will require costly and extensive work to remedy. Finally, a "D" represents major design faults that inhibit water delivery.

In Table 2, we display the results of the fidelity scoring and required remedy ranking for the MP and UPUW infrastructure. We present both the perceived status ranking and overall implementation fidelity scores by site in below as context for the value of evaluating outcomes of interest at each site. The implementation fidelity assessment found that there are problems of water supply at the majority of the UPUW sites, although new reservoirs and reticulation were reported by key informants to have increased the reliability of the water supply and new customer connections, compared to before the intervention. Despite the persistent snags at the new Semonkong plant, the plant there as well as the one at Mapoteng are both successfully supplying water for the first time to newly connected customers.

[4] The performance evaluation findings are not as relevant to the summative evaluation design. However, in the Final Evaluation Report, we will make note of any cases where performance evaluation findings yield important insights for summative evaluation results.

Table 2. Preliminary Implementation Fidelity and Perceived Status Findings

Ranking (required remediation)	Score (implement. fidelity 0-10)	Site and Major Issues Prior to WASCO Remedial Intervention
A	10	**Metolong Dam and complementary works** No major issues – all snags resolved
A	7.33	**Mapoteng** No major issues
B1	8.17	**Maseru and Surrounding Area** There is no reduced pressure connection from the new delivery main/high-level reservoirs to the old low-level distribution areas.
B1	6.83	**Qacha's Nek** Chemical dosing equipment supplied under the Compact has had to be replaced.
C3	8.17	**Semonkong** Filter beds are not delivering efficiently, resulting in low backwash efficiency. Telemetry and remote sensing are inoperative.
D2	6.83	**Butha-Buthe** Installation of delivery pump set prevents function, requiring use of pre-Compact pumps.
D2	5.00	**Quthing** River bed intake clogged, requiring installation of a temporary intake by WASCO.
D3	4.50	**Mokhotlong** New river and stream intakes both failed shortly after commissioning, requiring interim remedy by WASCO.
D5	4.50	**Leribe** Design of river bed sand intakes and blower prevent function.
D5	3.17	**Mafeteng** Clariflocculator, booster pumps, and generator design inhibit function.
D7	2.67	**Mohale's Hoek** Intake, pump sump, pumps, desludging pump, rising main design prevent function. Clariflocculator function inhibited.

As far as our impact evaluation methodology is concerned, it is important to understand first whether the infrastructure is operating as it was planned and second whether this operation can be attributed entirely to the program or to some other intervening actor. From KIIs and site visits during our process evaluation, we can verify the current functioning of the sites and also to whom this functioning can be attributed. In what we now call a "success" case, a UPUW site is able to deliver water as intended due almost entirely to UPUW Activity programming. In a "remediated" case, a site is able to deliver water as intended only due to WASCO remediation of failed UPUW infrastructure, as described in the previous section. In a "failure" case, WASCO has been so far unable to remediate failed UPUW infrastructure, and the site is unable to deliver as intended.

Our conclusion about the final implementation at each site from this perspective is included below in Table 3, along with our expectations about the benefits reaching intended household beneficiaries. The variation between sites in terms of ultimate implementation success means that we expect the result of an evaluation of the UPUW Activity as a whole will subsume substantial between-site variation, although through discussions with MCC we understand that site-specific estimates are not a necessary output of this evaluation. Further, given the degree to which WASCO remediation has been necessary to bring each site up to improved operation, the impact estimated for the UPUW Activity *as a whole* will be challenging to fully attribute to MCC, given the significant resources WASCO has invested to remediate design and construction issues.

Table 3. Implications of UPUW Activity implementation for evaluation

Site	Result	Expected situation at site
Package 1 - Maseru	Success	Compact interventions appear to have met objectives and have not required substantial remediation by WASCO.
Package 2 - Semonkong	Success	Compact interventions have met objectives, despite ongoing snags that have not yet affected delivery of water. Marginal WASCO remediation has occurred. As community had no piped water before, the works represent a major change from pre-Compact status.
Package 5 - Mapoteng	Success	Compact interventions generally working as planned, and have increased supply though anecdotally falling slightly short of providing year-round access
Package 3 - Qacha's Nek	Success	Compact interventions generally meeting objectives, despite some WASCO remediation with regard to water quality equipment. Increased supply capacity and reservoirs likely providing improved supply to community.
Package 3 - Quthing	Remediated	Compact interventions generally not meeting objectives. WASCO remediations have allowed the plant to meet community demand, and report that Compact works mainly prolonged the period of time for which the WTW could meet community demand (which was already the case pre-Compact). New reservoir appears to have improved reliability.
Package 4 - Mokhotlong	Remediated	Compact interventions appear to have led to some increased supply, but intake was a failure, requiring unsustainable temporary remediation by WASCO. The lower sections of the town are being supplied via the break-pressure tank while the new high-level sections do not require the pressure reduction provided by a break pressure tank. The Compact works are being used, but demand growth has moved to a different area. Compact-funded chlorine dosing appears to be working properly, likely increasing water quality.
Package 4 - Leribe	Remediated	Compact interventions have increased supply but substantially less than design capacity, and at risk of supply interruption during high river flow periods; thus, the works are not meeting objectives, even though providing a small benefit to the community. Reservoirs have anecdotally increased reliability. Even current operation due to a substantial amount of WASCO remediation to Compact works. New reticulation has reportedly led to about 400 new connections.
Package 3 - Mafeteng	Remediated	Compact interventions have met some of the objectives but have also required a significant level of WASCO remediation. Supply and quality may have improved, but modestly relative to expectations.
Package 4 - Butha-Buthe	Failure	Compact interventions have not met objectives and are generally not in use. While supply has increased slightly relative to the past (in rainy reason), it is not meeting demand, and falls short of targets.
Package 3 - Mohale's Hoek	Failure	Compact interventions not meeting their objectives, nor community demand. Supply may be subject to rationing given issues at the WTW. Site has required significant WASCO remediation. Water quality likely not impacted as WTW using old equipment. Improvements appear limited to water storage in new reservoirs.

5.3 Previous data collection

Previously, MCC had contracted NORC to undertake an evaluation of the MCC water sector projects, including rural and urban. NORC's evaluation ultimately proceeded only with the rural water sector interventions, though an evaluation design report and baseline report were also produced for urban water. Baseline (2010) and midline (2012) data collection was conducted by the Lesotho Bureau of Statistics (BoS). The survey instrument used for these data collection waves is referred to as the impact evaluation multi-purpose survey (IEMS). Since the SI evaluation pursues a different design approach than the previous NORC evaluation, we assessed the IEMS datasets specifically for usability for this new design and found that the previous data collection cannot be leveraged for this evaluation.

The main limitations that do not allow SI to leverage this data for the current design include (1) a lack of information to track households from previous rounds, and (2) insufficient coverage of UPUW Activity sites for the current design. In regard to the first limitation, the data does not include identifiers that could be used for tracking households, and household lists are not available to SI. GPS points from the two IEMS datasets are in different formats, do not correspond between the two datasets, and many points appear in unexpected locations, including outside of the national borders of Lesotho. In the absence of viable GPS points, enumeration area codes could theoretically be used to place IEMS households in their survey clusters. However, EA codes are truncated in the IEMS datasets provided to SI and thus could not be matched with the EA sampling frame obtained from the Lesotho BOS. It is likely that the only party who would have had access to the full EA codes is the Lesotho BOS. SI also attempted to matching the IEMS data with the EA sampling frame at a higher level (community or constituency) but this was not successful either given differences in names between the two files. With the six- to eight-year lag, even if this information was available, re-contact rates could be low.

With respect to the second limitation, the IEMS 2012 (midline) was focused on Maseru and Teyateyaneng only, given delays in implementation at other UPUW sites. The baseline data contains a sample size for many of the UPUW sites that is smaller than the required sizes calculated for SI's design (n=455 for all UPUW sites together), and only 7 of the 10 UPUW sites were included.

For these reasons, SI cannot use the data in order to follow up with previous households, but SI will attempt to use the IEMS data as a descriptive benchmark.

6 EVALUATION DESIGN

6.1 Overview

SI will implement a mixed-methods design to measure the impacts of the UPUW Activity (Table 4), and answer evaluation questions covered by the summative evaluation (Table 5). SI will conduct a quasi-experimental design (QED) with matching to estimate the effect of the activity on households, supplemented by qualitative data collection with communities in each site, as well as analysis of administrative data. Qualitative data collection will precede household surveying in order to provide insight into which sites are most appropriate to include in the quasi-experimental design, versus those better-suited to an ex post customer survey. As described in greater detail in the following sections, the final design as it pertains to measuring impact on households will be based on the following decision process:

- Qualitative data collection will be conducted first, to assist in making a final determination about which sites are appropriate to include in the quasi-experimental design (i.e. those where there is a reasonable counterfactual, attributable to MCC). All sites will include qualitative data collection, as well as analysis of administrative data pertaining to that site.
- Sites included in the quasi-experimental design will then proceed according to the methodology described in this report, for designs A and B.
- Sites not included in the quasi-experimental design will include a customer survey, to collect information about household experiences of the interventions, and estimate the current level of outcomes of interest in those sites.

This approach prioritizes the most suitable areas for an impact evaluation design, while ensuring that other sites are included in the evaluation for accountability and learning purposes.

To evaluate the effect of the MP on Lesotho's garment and textile industry, we will analyze administrative data and assess trends in employment over time as well as conduct key informant interviews and site visits with a purposively selected sample of firms to gain further insight into changes in water usage by water-intensive industries in Lesotho. To evaluate the effect of the MP and UPUW Activity on SMEs in Maseru, we will analyze administrative data and conduct key informant interviews.

Given the ex-post nature of this evaluation, in each of the next three sections detailing our approach for households, industry, and SMEs, we also discuss trade-offs in technical rigor, feasibility, and cost. Along with our methodology, we outline risks and proposed mitigation strategies where relevant.

Table 4. Overview of Summative Evaluation Components and Data Collection Approaches

Group	Component	Approach	Timing	Units	Sample size
Household	**Qualitative**	Focus groups	6/18-7/18	Communities	48
Household	**Qualitative**	KIIs	6/18-7/18	LHLDC, chiefs	12
Household	**Administrative**	Administrative	4/18-8/18	Households	n/a
Household	**Quasi-Experiment**	Matching	9/18-10/18	Households	3760
Household	**Customer Survey**	Survey	9/18-10/18	Households	1400
Industry	**Time series**	Administrative	3/18-5/18	Firms	n/a
Industry	**Case Studies**	Case studies	6/18	Firms	5
Industry	**Case Studies**	KIIs	6/18	LNDC, MTI	2-4
SMEs	**Time series**	Administrative	3/18-5/18	Enterprises	n/a
SMEs	**Qualitative**	KIIs	6/18	MTI, other	4-6

Table 5. Summative Evaluation Questions and Design Components

Evaluation Question	Design Component
7a. To what extent has access to quality water increased? What activities, if any, has WASCO conducted to encourage households to connect to the network?	Household: focus groups, administrative data (access), survey WASCO: qualitative interviews (process evaluation)
7b. To what extent are community members (including businesses such as manufacturing firms) using water from the urban water network and how has this changed since the Lesotho Compact started?	Household: focus groups, administrative data (consumption), survey Industry & SMEs: administrative data (consumption), qualitative interviews
7c. To what extent are community members experiencing cost and time savings, or reductions in water-related illness?	Household: focus groups, administrative data (waterborne illness), survey

6.2 Households

To identify the impact of the project, we wish to compare the outcomes of individuals who have received benefits of the interventions against the counterfactual – the outcomes for these same individuals, if they had not been exposed to the interventions. As this cannot be directly observed, the counterfactual must be estimated with a valid comparison group. The gold standard in impact evaluation to estimate the counterfactual involves random assignment. Randomization removes selection bias, so that both groups are comparable on observable and unobservable traits prior to the intervention, such that later differences between groups can be attributed to the intervention. However, randomization is rarely possible in cases of large, system-level infrastructure programs. These interventions are geographically bound to certain areas, often providing varying levels of benefit to a large population, and beneficiaries cannot always be pinpointed with precision *a priori*. In these cases, a quasi-experimental approach must be pursued to evaluate impact.

Further, the design must address issues that could threaten the validity of the estimated impact, especially given that several years have elapsed between the completion of the interventions and the summative evaluation. One of these issues is the potential for selective out-migration from urban or peri-urban neighborhoods in which interventions were implemented. Out-migration could be stimulated by a number of factors, including the relationship between access to infrastructure and rising housing or rental prices. If present, such selective out-migration could result in unobservable differences between matched treatment households and comparison households, which could potentially bias estimates of impact. In order to address this, SI will conduct qualitative data collection prior to conducting household surveys for the quasi-experimental design. The purpose of this qualitative data collection with respect to addressing potential bias is two-fold: first, interviews with knowledgeable local stakeholders and focus groups with community members will provide insight on the extent of in- and out-migration over the last few years in specific communities, as well as perceived reasons for these movements. While some information at the regional level on internal migration is available in Lesotho, there is no such information readily available for smaller administrative areas or within-city migration between townships. As such, qualitative data collection will help to provide area-specific insights for the sites of interest to this study.

SI will limit the quasi-experimental evaluation to sites where there is a plausible case for a valid counterfactual, meaning that sources of potential bias can be identified and mitigated such that impact estimates are internally valid and attributable to MCC. Qualitative data collection will take place first, the results of which will thus be

used to determine which of the UPUW sites should be included in the quasi-experimental design. In sites that are not included in the quasi-experimental design, SI will conduct a customer survey to collect household experiences of the interventions in those sites, especially considering the variation in how interventions in different places were implemented, and to estimate the current level of outcomes of interest in those sites. Thus, for those sites not included in the QED, there will not be a comparison group nor a direct before and after comparison. Data collected from earlier stages of the evaluation could be used in several sites as a benchmark for comparison, administrative data from WASCO can provide insight into changing consumption patterns over time for interviewed customers, and the survey can ask some recall questions about their experiences before and after the intervention. Nonetheless, the customer surveys will be one time period cross-sectional surveys without a comparison group.

Relatedly, focus groups with community members will allow SI to explore household decision-making and preferences around water infrastructure, which will provide SI with a better understanding of the variables likely to be the most relevant for self-selection into the treatment group for design A, such that they can be used to strengthen the matching approach outlined below.

6.2.1 Methodology

To measure the impact of the UPUW Activity on households in certain sites, to be determined ultimately following qualitative data collection, we will carry out a quasi-experimental design, through which a counterfactual group will be constructed using propensity score matching (PSM). This approach will ensure that comparison households are comparable to treatment households on a set of observable traits based on their ability to predict selection into the treatment group. The impact of the urban water interventions was hypothesized to occur through two potential pathways – by improving access (through new connections), and through improved supply (for existing connections). Thus, for this evaluation, we can conceptualize two types of treatment households: (A) households newly connected to the network, and (B) those with existing connections prior to the interventions. Groups A and B would be expected to benefit in different ways and to different degrees. As such, each requires its own comparison group. The first group corresponds directly to estimated impacts in the UPUW ERR, while there are no such estimates for existing customers. We discuss our proposed design for each beneficiary group separately below; ultimately, considering a variety of factors related to the evaluation's ex post nature and other challenges, not all sites will be subject to evaluation using both designs. A summary of the design is provided in Table 6, with further detail in following sections.

Table 6. Evaluation designs for UPUW Activity

Design A – improved access	Design B – improved service
Treatment (T): **Newly connected households** Comparison (C): **Unconnected households**	Treatment (T): **Connected to Metolong supply** Comparison (C): **Not connected to Metolong supply**
A selection of UPUW sites, TBD	**Maseru only**
Matching	Matching

Design A: Impact of improved access

First, for households with a new connection, our evaluation would measure the impact of improved access. In this scenario, the treatment is conceptualized as access to a new tap on premises. Comparison households can be defined as households that were unconnected prior to the intervention and are still unconnected now. This comparison group can be constructed by drawing a sample of unconnected households from the same areas as the newly connected households, all within a certain distance from the main. In this way there are treatment and comparison households in the same areas, rather than treatment and comparison areas. The

study will be powered to detect changes for all of the included sites as a whole, rather than to develop site-specific estimates of impact.

As we do not have valid baseline data to use, we will collect information to conduct the matching through a household survey. We must match on variables unaffected by the interventions to obtain an unbiased estimate of the propensity that each household self-selected into treatment. We expect these will include various measures of demographic and physical household characteristics, e.g. physical building materials, household composition, distance from the network, and household head characteristics. Given the time between pre-intervention and now, such variables (typically seen as time invariant or at least stable over the short-term) may have changed, but we will attempt to reconstruct pre-intervention status through our survey. Many of those factors will also serve as proxies of socioeconomic status, but other variables can also be included such as how households purchased their plot or dwelling. To control for changes in ownership versus renting status over time, we can also include an indicator for a household's pre-intervention status, as well as a continuous variable for the number of years since becoming the owner of the dwelling, as applicable.

The PSM approach is chosen to efficiently match observations on multiple variables (Rosenbaum and Rubin 1983). Given the lack of baseline data and households' self-selection into treatment, other design choices are limited. The matching procedure is carried out as follows. First, household selection into treatment is modeled using a logit or probit regression. The model assigns each household a propensity score, based on the variables used to model selection into treatment. After the propensity scores are generated, we test balance to ensure that households with similar scores are indeed similar. We also assess the overlap between the propensity scores assigned to the treatment group, and those assigned to the comparison group – this is called the region of common support. In general, we would expect the treatment group to have higher propensity scores, but we need to have a sufficiently large region of common support as households outside this region for which we cannot identify a match must be dropped from the analysis. After testing for balance and common support, we use the scores to match treatment and comparison households. There are several algorithms that can be used to implement the matching procedure. We will utilize nearest-neighbor matching with replacement and caliper width – i.e. using comparison households with similar propensity scores, allowing comparison households to serve as such for multiple treatment households, and setting a maximum threshold of the score within which neighbors can be identified. We will check the robustness of our matching approach by running other matching algorithms and comparing results.

By matching on propensity scores, we construct treatment and comparison groups that are balanced along the observed characteristics (Heckman 1997). However, as with any quasi-experimental approach, one of the main limitations with PSM is that we are unable to account for unobservable factors that might influence a household's self-selection into treatment, making them systematically different than comparison households. The sequencing of data collection in this evaluation with qualitative first is partially designed to obtain more information that will assist SI in estimating the selection equation to reduce potential bias from unobservable factors. After the matching is completed, we can estimate the impact on the outcomes of interest by using a simple difference between treatment and comparison groups. The advantage of this compared to regression approaches is that it is non-parametric, i.e. it does not assume a certain functional form. We will make use of Stata's packages for propensity score matching, including *pscore*, *psmatch2*, and *pstest* (Becker 2002; Leuven and SIanesi 2003).

We discuss challenges, limitations, and risks in a later section, but it is important to note here that it may be difficult to identify which household connections are specifically due to the interventions, and thus identifying post-intervention connections based on the installation date may serve as a reasonable proxy, while also noting the number of household provisions provided by the Compact versus the total number of new

connections that are found to be in this post-intervention date range. Also, this design presumes a sufficient number of newly connected and unconnected households in each site. The ERR assumed that across UPUW sites, 40 percent of households had been unconnected prior to the intervention. The number of new connections in each site can be examined by SI by examining the WASCO customer database in more detail.

Design B: Impact of improved service

The second group of household beneficiaries are those who had a tap prior to the interventions and are still connected.[5] The comparison group for these households is simple to conceptualize but challenging to identify practically. Conceptually, the appropriate comparison group is comprised of households who have also been connected since before the interventions, but *who were not exposed to* the intervention. We know that most interventions in the UPUW sites were system-level and had the potential to impact households across the network, beyond households in their immediate vicinity. For this reason, we do not believe that we can separate areas of the majority of UPUW sites into "exposed" and "not exposed". The only exception to this is Maseru, where we learned during the process evaluation that there are some areas of the city that have not been connected to the new Metolong supply. We return to Maseru after the discussion below.

For UPUW sites other than Maseru, we considered several options, described below, and have concluded that none are feasible given the limitations of each (a summary is provided in Table 7).

We considered a before-after design, comparing measurements from a sample of households with existing connections, pre- and post-intervention. Pre-intervention status would be reconstructed either through an existing dataset or recall questions in a survey. Neither option appears feasible. The IEMS data is not usable for our evaluation in its current form and expecting households to recall quantities of the key outcome variables after many years is unreasonable. Other sources of baseline data are not available. Neither the Lesotho DHS 2009 nor the Lesotho Bureau of Statistics 2011 demographic survey measures all the indicators needed for this evaluation.

In the absence of a with-without scenario at each site, along with the absence of a viable baseline, we are led to a third potential option, which is to look for entire sites that may serve as a valid comparison. Such sites would have to be comparable to UPUW sites at baseline and unaffected by the MCC interventions. WASCO service centers that did not receive UPUW interventions include Thaba-Tseka, Maputsoe, and Peka. However, we assume that these sites would not represent valid comparisons because there are likely systematic reasons why these areas were not included in the Compact. For example, our understanding is that Thaba-Tseka was a quasi-dormitory town for the Katse Dam construction and had a substantial injection in terms of infrastructure from the Lesotho Highlands Development Authority (LHDA). As such, it is not a typical town and the present population could be substantially less than at peak of Katse construction, with a surplus capacity in some of the urban services.

Table 7. Challenges in identifying a comparison for connected households

Potential comparison	Limitation
Baseline survey	IEMS surveys cannot be used as a valid baseline for these areas (see 5.3).
Pre/Post recall	Recall period too long to expect valid results. It has been four years since the end of the intervention and approximately 10 since pre-intervention.

[5] For the purposes of the evaluation, it would be ideal to define this as households who have been consistently connected since before the intervention and exclude households who may have been intermittently disconnected, to exclude the possibility of on-and-off status as a confounder. However, it is not clear at this point whether sufficient information will be available to exclude such households from our sampling frame; the feasibility of applying this criterion can be assessed after receiving customer data from WASCO.

Within-site	UPUW interventions system-level with all connected households having the potential to benefit. The only exception is Maseru, where some areas were not connected to the new Metolong supply.
Between-site	The UPUW was essentially a nation-wide program, and the minority of urban areas without the program are expected to be different in ways that are likely to affect their suitability as valid comparisons.

For Maseru, we have learned that there are areas of the city which have not yet been connected to the new Metolong supply. Most of these areas are fed from small treatment plants with low-lying reservoirs and have yet to be connected to the high terminal reservoirs. These would only be connected through pressure-reducing valve systems which were not included in the Compact. We have recently confirmed with WASCO which specific townships remain unconnected and have verified that there will be an adequate number of connected households therein to form a comparison group.

Given the lack of baseline information, the key assumptions that must be made in order to use this group as a comparison in the impact analysis with only post-intervention data are (A) that non-Metolong exposed households experience a level of service comparable to the level they experienced at baseline, and (B) that Metolong-exposed households and non-Metolong exposed households experienced similar levels of service prior to the intervention. The primary risk of Assumption A is that over the last ten years, service could have degraded because of further infrastructure aging, or alternatively improved because of other actions taken by WASCO in the interim; further, the net balance of either of these two possibilities may vary even within that area. Moreover, these households could potentially benefit indirectly from the Metolong areas, by accessing improved water from other areas. We need to be able to claim that the households in that area are not systematically different from those exposed to the Metolong supply in ways correlated with the treatment. On the latter point, we know from process study KIIs that WASCO did not prioritize one set of townships over another for connection to Metolong, originally requesting from MCA-Lesotho that all of the city be connected. According to respondents, MCA-Lesotho only denied this request under the justification that interconnecting the unconnected areas was outside the scope of the existing project, directing WASCO to the Government of Lesotho who has not been able to allocate the funds for interconnection to date. We believe this likely removes this potential source of selection bias, although we will verify that nothing appears to be systematically different about these areas during FGDs with local households prior to quantitative data collection.

In order to further test these assumptions and inform the decision of whether or not to pursue Design B in Maseru, we will plot longitudinal consumption data from the WASCO EDAMS database in townships connected to the Metolong Dam and in those unconnected to the Metolong Dam from 2007 to 2018, separated out by consumer type. Although consumption is not a perfect representation of service quality, this analysis will at least verify if any differences in consumption existed at baseline and how trends compared pre-intervention, as well as since the Metolong Dam has come online. Second, we will seek to explain consumption trends and explore any other differences in service provision through focus groups in each area sequenced in advance of quantitative data collection. Based on a combined analysis of these two data sources, we will determine whether to proceed with Design B.

Attribution

As described above, the counterfactual in Design B was assessed to be infeasible in all sites other than Maseru. Design A represents the attempt to estimate the project's impact on a portion of the intended beneficiaries. Therefore, in most areas where Design A would be conducted, the estimated impact will not capture the full MCC project impact. It is, however, a valid counterfactual for a portion of the MCC project,

since in those sites, provisions for new connections were included as part of the intervention. In other words, this is a valid counterfactual, but only for a portion of the intended beneficiaries such that any impact estimates it generates could be considered a lower bound of the MCC impact in these sites. The exception to this is Semonkong, which was previously served by the Department for Rural Water Service (DRWS), in which an entirely new water network and treatment plant were installed, such that Design A would capture the full impact of the MCC project in that site.

Attribution will not be possible in the same way in sites subject only to a customer survey, given the lack of a comparison group. Nonetheless, the survey – in combination with qualitative and administrative data – will attempt to identify the causes for any observed changes in the data. Even in the best-case scenario, in some sites observed changes would only be partially be attributable to MCC, because of WASCO remediations that were done to address shortcomings of some of the installations; in these cases, SI understands that MCC is still interested in pursuing the approach, as the evaluation will estimate the impacts of the program theory of change, even where remediation was needed.

Customer Survey

In sites that are not included in the QED, SI will conduct a customer survey. The customer survey will sample respondents directly from the WASCO database. The sample will not be limited to those connected pre-intervention, as there would be an equal interest in collecting information on new and existing customers, aligning with the hypothesized impact pathways for the Compact interventions. The customer survey will be finalized at the same time as the household survey and will be informed by results of the qualitative data collection. The objectives of the customer survey will be to estimate the current level of key outcomes among current WASCO customers, including new and existing customers, relative to the interventions. The customer survey will be an abbreviated and slightly modified version of the household survey, given the narrower scope of this activity. As described above, there will be no direct comparison households for these customers. However, using past data collected SI may be able to develop comparative benchmarks – both in terms of the previous levels of the outcomes of interest, as well as the composition of connected households in terms of socioeconomic status – that can be used to compare against that of the current customers sampled from the WASCO database. Lastly, if analysis of administrative data combined with findings from the qualitative data collection reveals that new connection targets fall below those expected through the Compact or by WASCO, or that unconnected households consistently cite monthly costs as a barrier to obtaining a connection, SI may consider implementing a willingness to pay (WTP) module within the survey. This will be decided in collaboration with MCC following analysis of the qualitative data and administrative WASCO data.

Outcomes

Inclusive of all options pursued, our evaluation will be focused on measuring the impact of the program on two primary outcomes, to align our strategy with the MCC ERR for the UPUW Activity: time savings (as a result of reduced water collection times), and diarrheal illness (as a result of increased consumption of quality water). We also include two other outcomes in our evaluation as secondary outcomes, given their close relationship to the primary outcomes. The first is per capita daily water consumption, which we include because in MCC's project logic, increased consumption precedes improved health outcomes. The second is monthly household water expenditures. While there is no ERR assumption for this indicator, the project logic assumes households will experience cost savings because of improved water access and service. These indicators will allow us to better explain any impacts observed on the primary outcomes and validate assumptions in the project logic. SI has experience collecting data for these outcomes on multiple other urban water sector evaluations. The outcome definitions and measurement strategy are described below.

Table 8. Outcomes of interest for household-level impacts

Outcome	Definition	Measurement
Time	**Average time (minutes) spent collecting water per household per day**	**For each source collected outside the home, round-trip time to the source (including queuing) is multiplied by the frequency of trips and standardized to minutes per day. Questions are asked in reference to 'average' behavior for current season.**
Health	**Prevalence of diarrheal illness among children under 5, last two weeks**	**Using a household roster administered to capture demographic information about all household members, follow-up questions will be asked for each child<5 in every household. For each child, we will ask whether they experienced any episodes of diarrheal in the last two weeks (14 days).**
Consumption	Liters per capita per day	Amount of water consumed on the last water bill (as applicable) will be recorded and converted from cubic meters per month to liters per day. For sources collected outside the home, the number of liters collected per source per trip will be multiplied by the frequency of trips and standardized to liters per day. The amounts will be added and divided by the total number of household members.
Expenditures	Expenditures on water, from all sources, per household per month	Amount of water consumed on the last monthly water bill (as applicable) will be recorded in Maloti. For sources collected outside the home, the amount paid per container will be multiplied by the number of containers per trip and the number of trips, standardized to Maloti per month. The amounts will be added. For households who report re-selling water to neighbors, we will calculate a second version of this indicator which subtracts the monthly amount the household receives from neighbors to whom they re-sell water from their tap, as applicable.

Note: Primary outcomes in bold; these align with MCC UPUW ERR and are the basis of the sample size calculations.

6.2.2 Timeframe of exposure

The endline survey will take place in 2018. This timing implies that households have been exposed to the UPUW Activity interventions for at least four years, in most cases, as the contracts for UPUW Activity construction ended in 2014.[6] In some cases, some components of the works will have been completed prior to that date. However, in several cases, as described in the process evaluation findings, WASCO has been remediating issues and challenges often due to ineffective design or construction of the UPUW works, and thus households have been exposed for varying amounts of time to better or worse situations depending on the site and situation remedied. There is essentially no precedent in the literature in terms of what to expect over such a lengthy period of time for this type of intervention. In general, in the absence of other interventions, environmental conditions, other large-scale shocks, or major changes in household practices or preferences, we would expect an accumulation of benefits for the two primary outcomes over time. We use the assumptions from MCC's ERR to guide our expectations about the potential impact of the UPUW Activity, and to guide our sample size calculations. Note that it may be unreasonable to assume that no other interventions could confound our results over the period of a decade since the beginning of the Compact. However, this can only

[6] Post-Compact ITT says Package 2 and Package 3 works were completed in the first half of 2014.

be validated through additional discussions with WASCO, data collection with community leaders or household beneficiaries, and research into water and sanitation interventions in UPUW areas.

6.2.3 Study Sample

For design A, our study sample is comprised of households connected after the completion of the interventions (treatment), and households that have remained unconnected since prior to the interventions (matched comparisons). To help ensure comparability between the treatment and comparison groups, eligibility for either group will include residence in the study area since prior to the interventions. Further, eligibility would be limited to households within the eligible distance from network according to WASCO policy for connecting households. These criteria reduce the potential for systematic differences between treatment and comparison households, especially once matched, for reasons other than the interventions. As such, there are no treatment and comparison *areas*, but rather treatment and comparison *households* within the same areas.

For design B, our sample is drawn from those with connections from before the interventions were implemented – treatment households are drawn from areas exposed to the new Metolong supply, while our comparison households are drawn from areas that are not connected to the new supply.

Beneficiary Groups

As described earlier, qualitative data collection will be conducted in order to determine the final set of sites included in the quasi-experimental evaluation. The results from the qualitative data collection will provide the information needed to make the final decision. However, at this stage we anticipate that the areas that will be most suitable for an impact evaluation include Semonkong and the peri-urban areas of Maseru. In addition, we anticipate that Design B will remain viable in Maseru urban. For new customers in Maseru urban, as well as all Package 3 through 5 UPUW sites, we expect to conduct a customer survey.

Thus, we anticipate that household beneficiaries will be covered by different parts of the evaluation, as follows:

Table 9. Beneficiaries covered by quantitative data collection components, by site

Site	Quantitative data collection component
Maseru urban	**QED (Design B) – Household surveys** Beneficiaries: Existing customers
Maseru peri-urban (Mazenod, Roma, Morija) and Teyateyaneng	**QED (Design A) – Household surveys** Beneficiaries: New customers
Semonkong	**QED (Design A) – Household surveys** Beneficiaries: New customers
Mafeteng	**Customer survey only**
Mohale's Hoek	**Customer survey only**
Quthing	**Customer survey only**
Qacha's Nek	**Customer survey only**
Leribe	**Customer survey only**
Butha-Buthe	**Customer survey only**
Mokhotlong	**Customer survey only**
Mapoteng	**Customer survey only**

Sample size and power

Our sample size calculations are presented below. We calculated sample size requirements for detecting estimated reductions on time savings and diarrheal illness using Stata's *-power-* command.

The estimates for Design A and B below are based on standard inputs for power calculations, including a test size of 0.05 (α) corresponding to 95% significance, and power of 80%, corresponding to a probability of correctly concluding the program has an effect when one exists. Note that we do not factor in clustering, based on the assumption that we will randomly select households from the sites of interest. To inform the calculations, we used available values from the IEMS data or Lesotho DHS 2009 data and applied ERR targets to estimate the expected change in outcomes.[7] In general, lesser impacts would be expected on households with existing connections; thus, we scaled the expected impacts by half for design B.[8] After calculating the sample size requirements, we inflated by 25% to account for potential loss outside of the region of common support during the matching procedure. Sample size estimates for the customer survey are based on equations to estimate a population proportion and population mean, with 95% significance and levels of precision (margin of error) specified.

Our estimate of the total sample size required is 5,160. This includes Design A and B for Packages 1 and 2 and a customer survey for Packages 3-5. For Design A, we use the larger sample size requirement for diarrheal illness for our planning, meaning that we could detect even smaller true effects on time savings. The final sample size for Semonkong is lower than the calculated sample size based on our expectations of the maximum number of new connections available there and baseline expenditures are given as zero since they received water free of charge previously from the department of rural water supplies (DRWS). For Design B, the sample size for detecting the estimated impact on diarrheal illness is prohibitively large. Given the sample sizes used for planning, MDES is also shown below.

For the customer survey, the sample size allows for calculation of *levels* of the outcomes shown, and is not designed to test changes over time, as it is a one-time ex-post survey. Further, packages 3 through 5 are treated as a pooled sample, given our understanding that MCC is not interested in separately powering data collection for each site, which is aligned with the fact that the theory of change for the Compact interventions was not substantially different for each of those sites (the same outcomes and same targets were built into the M&E plan and ERR across all UPUW sites), even though interventions were not the same across sites. The level of precision used in this sample size calculation thus applies to the sample as a whole; precision for each site will be lower (i.e. wider margin of error).

It is critical to note that the estimated sample size overall rests on the assumptions made about the ultimate result of the qualitative data collection and will be finalized after the qualitative component is completed.

[7] The ERR target for diarrheal illness actually applies to diarrheal illness *mortality*, but we apply it here to prevalence of diarrheal illness. While the ERR also hypothesized a benefit stream of time savings due to decreased time spent on care-seeking for diarrheal illness, but there was no analogous target set and the model incorporated only the expected reductions in time spent collecting water.

[8] Given the specific outcomes included in the ERR (diarrheal illness and time savings), we expect based on the existing literature that a switch to a piped connection could drive large changes in these outcomes, that would be less pronounced for households that had a tap already, unless, for example, the tap was not providing any water to the household or providing highly contaminated water to the household. We scaled the impacts by half, given that we expect outcomes to be attenuated among households who had an existing connection and no corresponding target is provided in the ERR or MCC M&E documents. If larger impacts occur, our design will be able to pick those up given our assumptions. A six-minute time savings is already small in terms of how that time could be re-allocated for any given household toward productive activities, and a 15% reduction in diarrheal illness already leads to a prohibitively large sample size.

Table 10. Sample size calculation inputs

Design & outcome[9]	μ1	Target	μ2	σ_p	m.e.	Sample size	+25%	T	C
Design A									
Time Savings	47	-50%	24	67	-	270	**338**	169	169
Diarrheal Illness	.129	-30%	.09		-	2,012	**2,516**	1258	1258
Design B									
Time Savings	23	-25%	17	26	-	592	**740**	370	370
Diarrheal Illness	.106	-15%	.09	-	-	10,840	**13,550**	6775	6775
Customer Survey[10]									
Time collecting water	24	-	-	26	5	**352**	-	-	-
Diarrheal Illness	.09	-	-	-	.015	**1,398**	-	-	-

Table 11. Minimum detectable effect sizes for Designs A and B

Design & outcome[11]	n	μ1	MDES (δ)	μ2	n	μ1	MDES (δ)	μ2
Design A					**Semonkong**			
Time savings	2,012	47	(-)8.37	38.6	500	47	(-)16.82	30.18
Diarrheal illness	2,012	.129	-0.039	.09	500	.129	(-).0953	.0337
Water consumption	2,012	55	(-)6.87	62	500	55	13.81	69
Water expenditures	2,012	30	(-)3.75	26	500	0	7.53	7.53
Design B								
Time savings	592	23	(-)6	17		-	-	-
Diarrheal illness	592	.106	(-)0.0813	0.0247		-	-	-
Water consumption	592	97	22.37	119		-	-	-
Water expenditures	592	82	(-)18.91	63		-	-	-

Table 12. Final estimated sample size requirements

Site	Design A (Beneficiaries: New customers)	Design B (Beneficiaries: Existing customers)	Customer Survey (Beneficiaries: All Current Customers)
Maseru urban	-	**740**	-
Maseru peri-urban (four towns)	**2,520** (630 per site)	-	-
Semonkong	**500**	N/A	-
Packages 3-5 (pooled)	-	-	**1,400** (175 per site)
TOTAL = 5,160	**3,020**	**740**	**1,400**

*Note: sample sizes are rounded up to accommodate multiple sites, e.g. 2516/5 sites=503.2; 504*5=2520*

[9] Mean and SD values for time savings are from IEMS 2012 dataset. Mean collection time for non-piped households 47 with SD 71, and for piped 23 with SD 26 and pooled SD of 67. Pooled SD is used for design B calculation. Diarrheal illness values are from Lesotho DHS dataset: 10.6% for households with improved source, 12.9% for households with unimproved source.

[10] Equation for estimating a population proportion: N = (p*q)/(me/z)^2, where q=1-p, z=1.96 corresponding with desired 95% significance level, and me=desired margin of error. Equation for estimating population mean: N= (z*s^2)/me^2, where z=1.96 for desired 95% confidence level, s^2 is equal to the population standard deviation (estimated using IEMS midline data), and me=margin of error.

[11] Time savings from IEMS 2012 dataset; water consumption and expenditures from NORC baseline report. Consumption for all households with a tap 97 lpcd; water bill 82 Maloti per month. Rural & peri-urban are combined in the NORC report, given as 14.3 lpcd consumption, inclusive of all drinking water sources; expenditures reported categorically but average is approximately 30. SDs for consumption and expenditures set equal to mean.

Sampling Frame & Sampling Strategy

For this evaluation, we propose a combination of approaches for our sampling strategy that are meant to balance logistical feasibility and resources, such that we attempt to maximize the utility of existing data, only resorting to more costly and time-consuming listing activities where necessary. Specifically, our sampling strategy relies on two main approaches including (1) sampling connected households using the WASCO customer database, and in the ideal case (2) using a geospatial sampling approach to sample comparison households. Based on information available in WASCO's customer database this ideal method of developing a control sample frame is only feasible at present in Maseru and Mazenod.[12] In areas that SI elects to follow through with an IE methodology, but which lack GIS information for connected households, we will need to resort to listing exercises to develop a sample frame if WASCO does not collect and provide to SI GPS information for customers before the time that sampling for household surveys would need to occur (anticipated to be late August or early September 2018). Sampling will be conducted in advance of quantitative data collection. Stratification is not built into this design, though it could be considered later based on a more comprehensive assessment of the quality of GPS coordinates in the customer database.[13]

Sampling from WASCO customer database

To sample connected households for both designs A and B, we plan to draw a random sample from WASCO's customer database. This database includes unique combinations of variables such as account number, meter number, connection installation date, customer contact information, status of connection (active/inactive), township/service center, and walk route on which the customer's meter is read. For customers in Maseru, Mazenod, and occasionally Maputsoe there are also GPS coordinates. To construct our sampling frame, we will need to specifically determine what constitutes a post-intervention connection; this will likely vary by site depending on completion dates. Relatedly, we will have to address the possibility of anticipation effects, by analyzing trend changes in customer connection rates using the WASCO data.

To sample unconnected households, we need a sample frame of all unconnected households. Traditional methods would dictate costly listing exercises in targeted geographical locations and would normally necessitate a two-stage cluster sampling methodology which usually results in a loss of power, requiring a larger sample size. Although such an exercise will likely still be required in sites that lack GPS coordinates in the customer database, our approach in urban Maseru can make use of Open Street Map (OSM) layers that are publicly available for Lesotho. In areas for which GPS coordinates are available we will define the geographic areas of interest (e.g. clusters or EA codes) for this evaluation, based on the placement of the network for the corresponding service center, and conduct a geospatial sampling approach in each site. The approach is further described below.

Geospatial sampling approach

The country of Lesotho benefits from the existence of OSM base layers in shapefile format. In addition to the typical features included in OSM downloads such as roads, waterways, and points of interest, the Lesotho download contains building footprints. This is very unusual, and exists presumably due to Lesotho's small size, although it is still a major effort to create almost 1,000,000 building footprints.

[12] At this time, about 90% of households with active connections in urban Maseru and Mazenod have corresponding GPS points. Almost no households in the remaining study areas have corresponding GPS points in the customer database.

[13] We may consider stratifying our sample in each site by distance to the transmission main which serves as a proxy of customers' willingness to pay as connection fees are based on distance. Stratification would not be feasible if some strata did not contain enough connected or non-connected households. This can only be assessed after the WASCO database is in hand.

Using this data, we propose to employ a geospatial sampling approach to conduct listing and sampling of comparison households for both designs A and B in at least urban Maseru, where GPS coordinates in the EDAMS database are most complete. Our proposed approach is meant to circumvent the time-consuming and expensive process of creating a household listing from which to sample comparison households. In sum, we would use OSM base layers to identify residential buildings, overlay the GPS points of WASCO customers, assess and reconcile overlap between those two layers, and extract households to form the sampling frame for each comparison group.[14] We would overlay GPS coordinates of WASCO customer data on OSM maps indicating residential buildings, and extract residential buildings without WASCO connections to serve as our sampling frame for the comparison group for design A. For design B, the approach would be analogous, but would instead extract connected households in areas we define as comparison (not connected to the new Metolong supply). We would then export the sampling frame and using Stata, draw a random sample from it, using replicable and transparent code.

To create a residential building sampling frame, it is necessary to include only those footprints classified as residential. Only about 36% of the building footprints are classified by type, which includes classifications such as “residential”, “hut”, “commercial”, and “industrial”. There are three other layers included in the Lesotho OSM download that can be used to classify those building footprints whose type is blank: 1) Land use polygons, 2) Points of Interest points and polygons, and 3) Places of Worship points and polygons. Each of these layers can be used to spatially exclude building footprints that for example are churches, schools, hotels, or government buildings. The remaining footprints will be considered as potentially residential based on three criteria: 1) they have been actively classified as a type of residential building, 2) they are unclassified, but exist within a land use that is classified as a type of residential land use and are not a point of interest or a place of worship, or 3) they are unclassified, and exist outside of any land use polygon are not a point of interest or a place of worship. The spatial layer can then be sorted, and a sample of buildings can be selected.

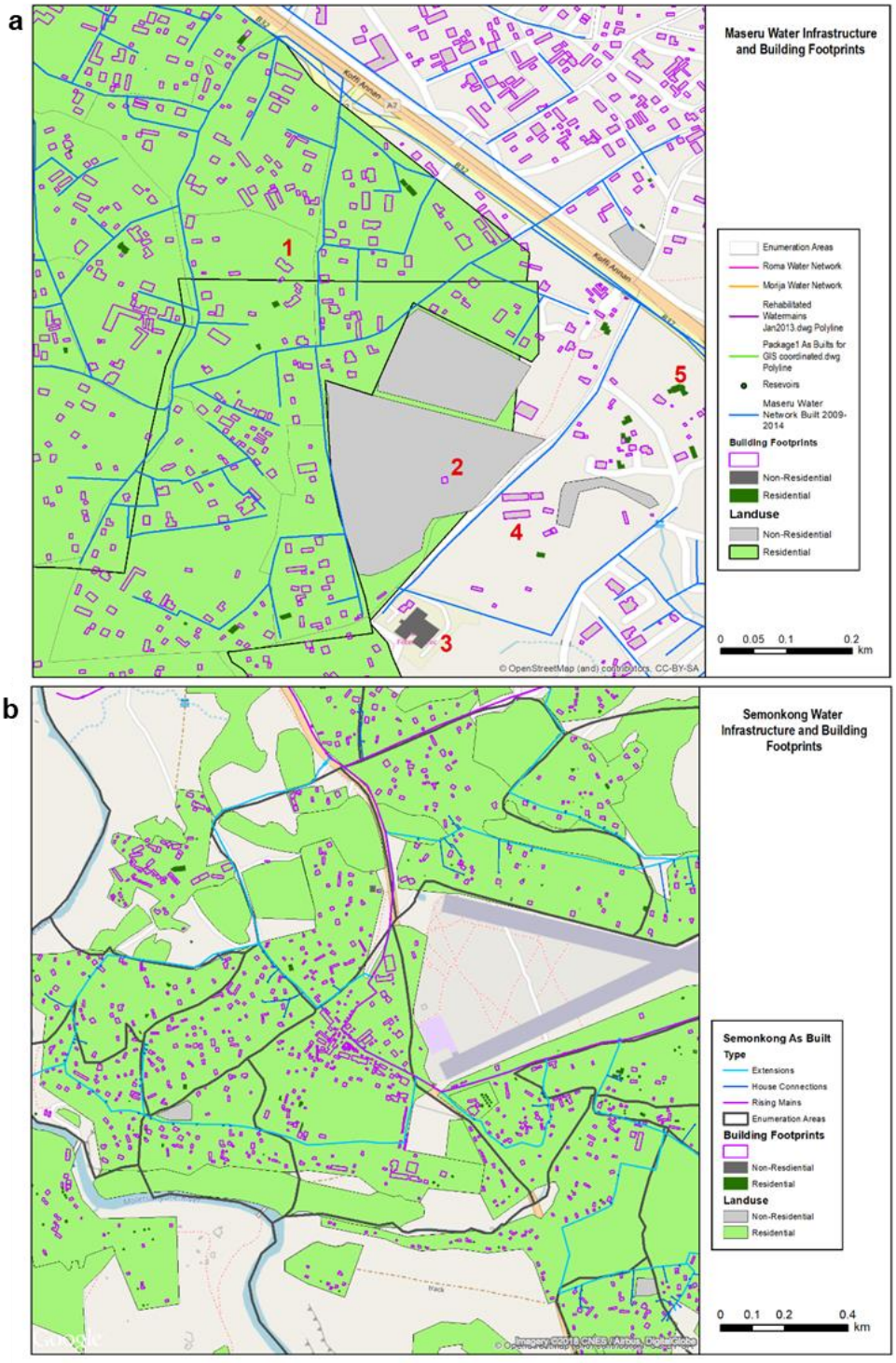
The building classification strategy is illustrated by maps showing Maseru and Semonkong in Figure 3 (though we note that GPS coordinates are not available at this time for Semonkong and this example is illustrative of the procedure that can be carried out for areas where GPS coordinates are available). For Maseru:

- Building 1 is classified as residential. It is unclassified in OSM but falls within residential land use.
- Building 2 is classified as non-residential. It is unclassified in OSM but within non-residential land use.
- Building 3 is classified as non-residential. It is it classified as non-residential in OSM.
- Building 4 is classified as residential. It is unclassified in OSM and outside of any land use polygon.
- Building 5 is classified as residential. It is classified as residential in OSM.

In Semonkong, it appears that the land use polygons were drawn using the buildings as a guide. Virtually no buildings are outside the residential land use polygon. Most have no classification, but almost all would be included in the frame, since they are inside the residential land use polygon. Once the selected building footprints have been selected, it is possible to create a digital map showing each one. The digital map can be georeferenced and used offline on a mobile device using a free or low-cost mobile application, allowing field staff to navigate to the correct household, and conduct the interview.

[14] The procedure of overlaying these layers may also help us in identifying relevant geographical landmarks for the treatment households sampled from the WASCO database.

Figure 3. (a) Maseru building footprints, (b) Semonkong building footprints



6.2.4 Qualitative Data Collection

SI will conduct qualitative data as part of the household component of this evaluation, including key informant interviews and focus groups. As described earlier, the purposes of qualitative data collection in this evaluation will be to (i) identify potential sources of bias in estimating impacts, particularly related to internal migration in the interim between intervention completion and this evaluation; (ii) identify factors that can be measured and included in the matching selection equation to strengthen comparability of treatment and comparison groups; and (iii) to supplement household surveying with additional learning regarding household decision-making in response to interventions and obtain richer, narrative descriptions regarding household experiences of impacts (or unintended consequences). The timeframe of exposure related to the focus groups is identical to the household survey.

The desired information on internal migration and local housing markets cannot be ascertained from existing sources, necessitating additional qualitative data collection. The information that is readily available is reported at the regional level and does not include information for the time following the intervention completion. For example, a Lesotho Bureau of Statistics (BOS) Demographic Survey from 2011 shows that internal migration generally stagnated between 2006 and 2011, apart from slightly higher migration from highlands and rural areas to industrial areas in the lowlands, consistent with historical trends, and largely driven by the search for employment opportunities. The primary destinations of internal migrants appear to be Maseru and Maputsoe/Leribe. Information for the time period through 2016 is not yet available as the 2016 Lesotho census results have not yet been publicly released. Further, information about housing and rental prices is not readily available in Lesotho. Preliminary research shows property transfers are rare and much of the market operates informally. Formal transactions are typically limited to higher income residents, informative mostly if coupled with selective out-migration in same areas, for which data is not readily available. It is also not clear to what extent access to infrastructure has influenced any internal migration. Lastly, no data is readily available on composition of internal movers and/or whether movements *within* urban areas has changed, rather than between regions in Lesotho. Qualitative data collection will thus provide some useful information about households' perceptions of, and decision-making in response to, changes in the housing market.

Key informant interviews: The first two purposes will be served by key informant interviews with knowledgeable local stakeholders and focus groups with community members in each of the UPUW sites. SI will recruit key informants from the Lesotho Housing and Land Development Corporation (LHLDC), as well as local chiefs in the study areas. The former is responsible for the creation of affordable housing in Lesotho, and many low-income residents purchase their land plots from local chiefs[15], who are likely to be knowledgeable about movements in and out of the area over time. Key informant interviews will be carried out by SI local staff together with by qualified personnel hired by local data collection subcontractors, trained by the successful bidder (to be selected through a competitive bidding process) and SI.

Focus group discussions: Focus groups will provide valuable information toward all three of the objectives listed above. First, focus groups will provide useful information with respect to households' perceptions of internal migration over the last few years, including whether it has happened, from/to where individuals have moved, and the reasons motivating these movements. Second, our PSM design rests on our ability to construct a model that can predict selection into treatment reasonably well. Focus group discussions can give us greater context about why some households opted to connect and others did not, strengthening the quantitative methodology. Thirdly, focus groups can provide useful context for our interpretation of the household survey results, especially given the various limitations around the lack of baseline and potential

[15] Lesotho Housing Profile (UN-Habitat, 2015)

confounders. Focus group discussions may allow us to understand with greater nuance household beneficiaries' experiences of the Compact-funded infrastructure over a long period of time, i.e. whether through immediate, tangible benefits at some sites or service disruptions at others. This is especially important considering the process evaluation findings regarding the potentially negative consequences of some of the installations. Given the complex dynamics of water use in urban areas, focus groups are an important source of contextualization to understand *why* and *how* households make decisions about the main outcomes of interest, e.g. water consumption and expenditures. This includes a nuanced understanding of why households use multiple sources for various domestic activities, the extent to which spillover (indirect benefits to non-exposed households) may be occurring, how factors such as reliability, quality, convenience, and price interact in this specific context to influence household decisions about water use, and whether and how households experience downstream benefits (e.g. how any potential time savings were allocated). Focus groups will be carried out by qualified personnel hired by local data collection subcontractors, trained by the firm and SI collaboratively.

Besides these uses for focus groups, which are directly related to the project logic and validity of our identification strategy, focus groups could additionally be used to inexpensively explore unanticipated consequences of the MP and UPUW Activity. An unanticipated consequence frequently hypothesized by process study respondents was vandalism of the Metolong downstream conveyance system (DCS) by rural communities during a drought period, resulting in rumored increased consumption of free water in these communities for domestic and economic purposes to this day along with drastic increases in NRW for WASCO. If deemed worthwhile by MCC, we propose to conduct two focus groups with these communities to understand how they may be unintended beneficiaries of the program and provide further context to any post-Compact changes in NRW experienced by WASCO.

Sequencing: Qualitative data collection will occur prior to any household surveying. Limitations of this sequencing include that focus group participants may not be those ultimately sampled for the household survey which means, logistically, that recruitment for focus groups requires an additional and separate effort, and also that focus groups will not be able to follow up on and explore particular findings from the quantitative surveys, nor that perspectives from the focus groups would align directly with information provided in the household surveys. An additional limitation of this sequencing is that it pushes back the timeline of household surveying to September-October 2018. Ultimately, this may affect the timeline of the evaluation as a whole, it also accommodates the winter season in July-August, during which household data collection would likely face logistical challenges in any case. The benefits of conducting the qualitative data collection first, as described in this section above, outweigh these limitations.

Sample size: Key informant interviews with local chiefs and community focus groups will be conducted in all UPUW sites. At this stage we are unaware how many local chiefs we will need to interview to adequately understand housing and migration dynamics in each area, but we plan to solicit suggestions on this question from bidders to ensure adequate coverage. Focus groups will be conducted with households in both the treatment and comparison groups for each of Designs A and B, to gather the breadth of information needed to fulfill the three objectives of the qualitative data collection. Based on published literature, we expect that thematic saturation from focus groups can be mostly reached within two to three focus groups. Thus, we would target this many groups per treatment and control group per site, plus two or three along the DCS if requested by MCC. The successful bidder would be required to demonstrate their ability to carry out the intended number of focus groups; if the need arises to balance time and cost constraints, we can reduce the number of focus groups in each site.

Focus group participants: Each focus group would have between 8-12 participants. Intended participants will be members of households resident in the study areas since pre-intervention; it is likely that this definition will be further narrowed to female individuals based on SI's prior experience elsewhere about the individuals most knowledgeable about water affairs in the household and effective focus group dynamics. A final determination will be made with the input of the data collection partner, to be selected through a competitive bidding process.

Use of qualitative results: With qualitative data collection conducted first, the findings from the qualitative component will feed directly into the final design of the quasi-experimental component. For example, it will only be suitable to conduct the quasi-experimental evaluation in areas where a credible counterfactual can be estimated, such that the impact can also be credibly attributed to MCC. Therefore, if the qualitative results suggest that in some areas, internal migration (or other factors) are likely to bias the estimation of results, or if SI identifies the presence of any major confounding factors, such areas will be considered for exclusion from the quasi-experimental component. If some sites are, ultimately, excluded from the quasi-experimental design, a customer survey can be implemented instead in that site. Illustrative topics for the focus group discussions is presented below (Table 15), and a draft instrument in attached as Annex 4-2.

Table 13. Key informant interviews

Key informant	Estimated sample size
Lesotho Housing and Land Development Corporation (LHLDC)	1-2 Key informants
Chiefs	10 chiefs (estimate based on 10 urban sites) Subject to input from data collection partner, TBD

Table 14. Estimated number of focus group discussions

UPUW site	Design A	Design B (Maseru urban only)
Maseru urban + Maseru peri-urban	**10 focus groups** Newly connected: 1 Unconnected: 1 In each of the 5 sites: Maseru urban, Roma, Morija, Mazenod, Teyateyaneng	**2 focus groups** Metolong-supplied: 1 Not connected to Metolong: 1
Semonkong (1 site)	**4 focus groups** C: 2, U: 2	n/a
Other UPUW	**32 focus groups** Newly connected: 2 Unconnected: 2 In each of 8 sites: Qacha's Nek, Quthing, Mohale's Hoek, Mafeteng, Leribe, Butha-Buthe, Mokhotlong, Mapoteng	n/a
Maseru area, rural residents of villages near downstream conveyance from Metolong	n/a	**2 focus groups**
Total: 50 focus groups	**46 focus groups**	**4 focus groups**

Table 15. Illustrative focus group topics, by participant type

Participant type	Connected households (Design A Treatment; Design B, all)	Unconnected households (Design A Comparison)
Selection into treatment; Experience of interventions; Potential spillover effects	- Reasons for requesting a connection - Previous water source situation - Perceived changes in water service - Notable 'shocks' (e.g. major interruptions) - Expectation of benefits versus reality - Perceived water quality; changes over time - Water consumption; seasonality - Alternative source use (current and past); reliability, quality, convenience, and price - Perceived benefits over time	- Demand for connection, willingness to pay - Current water source use; seasonality - Factors influencing source choice; reliability, quality, convenience, and price - Seasonality of water use practices - Barriers to connection - Indirect benefits/spillovers - Perceived changes over time in any of the above
Other WASH practices	- Household water storage - Household water treatment - Sanitation & hygiene in household - Seasonality of these practices - Cost of coping behaviors - Perceived changes over time	- Household water storage - Household water treatment - Sanitation & hygiene in household - Seasonality of these practices - Cost of coping behaviors - Perceived changes over time
Bias & confounding factors	- Internal migration over the last decade - Housing market and rental prices - Awareness of MCC interventions - Other interventions / WASH programs	- Internal migration over the last decade - Housing market and rental prices - Awareness of MCC interventions - Other interventions / WASH programs
Outcomes	- Time & cost savings, diarrheal illness - Re-allocation of time or money - Distribution of benefits - Unanticipated effects	- Time & cost savings, diarrheal illness - Re-allocation of time or money - Distribution of benefits - Unanticipated effects

Qualitative Data Quality & Data Processing

SI has robust procedures for ensuring data quality for qualitative research. Instruments are translated by our local data collection partner staff or translators, and then back-translated by a third party to ensure that the intended meaning and concepts are communicated accurately. We ensure field presence during training of qualitative staff and participate in leading trainings on the focus group guide. SI collaborates closely with data collection partners on a qualitative fieldwork manual that serves as a reference for focus group facilitators. The manuals contain comprehensive information about the tool and study objectives, along with professional and ethical conduct, best practices in facilitating and note-taking, and expected standards for transcription.

SI's presence in the field also includes oversight of piloting activities for qualitative data collection, after which a full transcription and translation is completed, along with a debrief of the results of the pilot activities, a review of revisions that need to be made to the guide(s), and recommendations to interviewers on addressing any issues that may arise during fieldwork. SI's local coordinator will also attend a sample of the FGDs to monitor the work of the data collection firm's focus group facilitators and provide SI with insight into the conduct of the work, the professionalism of the facilitators, and the general progress of the fieldwork.

Likewise, SI schedules an early check-in after the first one to three focus groups to discuss with the data collection firm any issues that need troubleshooting and discuss substantive results from the early focus groups to provide suggestions of any areas to probe or avoid, clarifications on the research objectives and scope of the discussion, and to provide responses based on any emerging themes. SI will conduct other periodic, structured check-ins at other points during the qualitative fieldwork similar in scope to the first, with

the number and timing depending on the scale of the fieldwork. SI often asks local data collection partners to submit a summary of each focus group to accompany each transcript, so as to orient our evaluation team with the facilitators' impressions of the discussion, as they relate to the study objectives. This is used in combination with the transcripts to analyze and interpret the qualitative findings.

Once transcripts and FGD summaries are submitted to SI, along with audio recordings of the focus groups and participant forms listing the demographics and consent for each participant, we review thoroughly and return a structured form back to the data collection firm with any clarifications requested. Following these quality control measures, SI analyzes the transcripts by applying codes and then conducting thematic analyses on the content of the transcripts. Our analysis plans are described in more depth in the next section.

6.2.5 Quantitative Data Collection

Our design will be conducted on a sample of households surveyed in 2018. While previous data collection was conducted in 2010 and 2012, our assessment is that those data are of limited usability for our evaluation, as described above in Section 5.3. To conduct household surveys, we propose to use the IEMS tools as a starting point, while revising and adding to them to ensure that all relevant topics are addressed based on our own best practices from similar evaluations, specifically in terms of measuring the outcomes of interest and key covariates for our impact analysis. The main areas where we have included additional questions to measure key concepts in a more sophisticated fashion are regarding drinking water storage, drinking water treatment, coping and aversion behaviors, and diarrheal illness.

To ensure the best value possible for high-quality data collection for our evaluation, we will conduct a competitive procurement including Basotho and South African firms for a data collection subcontract. We will also ensure that the firm ultimately selected is capable of handling the logistics of the effort, given its size and the likelihood that it may stretch into winter months in Lesotho where travel can become difficult. Preliminary market analysis during our scoping trip, process evaluation, and evaluation design process has revealed a variety of firms who have been employed to conduct similar work in this space in the past. Each of the firms and any additional firms which we can identify will be invited to submit proposals to collect our primary data.

Quantitative Data Quality

For every impact evaluation (IE) in SI's portfolio, we utilize a systematic quality assurance framework for data collection. This system is comprised of a comprehensive set of guidance documents, templates, examples, and check-points. Our system aggregates learning from nearly a decade of experience conducting IEs worldwide, and incorporates social science standards and best practices. It includes comprehensive guidance documents for planning data collection and monitoring data quality, to ensure these activities are carried out with the highest level of integrity. Our team uses check-lists that accompany guidance on high-quality data collection to ensure compliance with all minimum requirements and best practices. We also use tools and procedures outlined in our guidance in working with our local data collection partners, which sets clear expectations for field teams, allows SI to monitor performance throughout data collection, and contributes to building local capacity for high quality data collection. Below we outline a few of the major data quality activities that we employ as standard practice in our survey work that we will include in our efforts in Lesotho.

SI's data quality monitoring is not limited to troubleshooting problems but can also be used to provide positive feedback to field staff, increasing morale and enhancing performance and retention, especially important in large-scale data collection activities where teams may be traveling long distances and working long hours, and where retention of quality enumerators and supervisors over time helps to continually build institutional knowledge and sustains data quality over time. SI's data quality approach is three-fold, and focuses on (a)

prevention of errors, (b) quality control in the field, and (c) independent data-quality monitoring. This approach is summarized in Table 16, with some of the main components further described below.

Table 16. SI data quality control approach

(a) Error Prevention

QC Measure	*Responsible*
Data collection team structure	*SI requires in procurement; local data collection partner implements*
Field team training, pre-testing, and piloting	*SI and partners collaborate*
Electronic data collection and build-in quality controls	*SI codes the forms; partners tests and provides feedback*
Field manuals and protocols	*SI sets required content; partners develop; SI reviews and contributes*

(b) Field-based Quality Control

QC Measure	*Responsible*
Supervisor observation	*Partners implement; SI monitors observation forms*
Co-enumeration	*Partners implement; SI analyses forms and IRR*
Supervisor form checks	*Partners implement, report major issues to SI; SI reconciles major issues*
Team Debriefs	*Partners implement, reporting major issues to SI; SI reconciles major issues*

(c) Independent Monitoring

QC Measure	*Responsible*
High frequency checks	*SI conducts, reports issues to partners; partners reconcile*
Monitoring Electronic Data Collection Controls	*SI monitors and analyses; reports issues to partners; partners reconcile*
Back-Checks	*SI monitors and analyses; reports issues to partners; partners reconcile*

SI requires that supervisors observe a minimum of 10% of the surveys conducted. Supervisors shadow enumerators to ensure that they are asking questions appropriately, accurately recording direct observations (e.g., functionality), and acting in accordance with expected professional standards. Supervisors are required to check on a nightly basis the forms submitted by each of the enumerators on their team. Data collection teams are required to hold periodic debriefs (e.g., every few days) to review any logistical or technical challenges faced in the interim period since the last debrief, to ensure relevant knowledge, tips, and clarifications are shared team-wide, and additionally as an opportunity for supervisors and coordinators to offer positive feedback and reinforcement to field teams as a further measure to help ensure quality and maintain morale. Independent monitoring measures are undertaken by SI as well during and after data collection. SI conducts high-frequency checks on a daily basis for the first 1-2 weeks of data collection and every few days thereafter. These checks are focused on the progress of data collection, adherence to sampling protocols, and early detection of outliers or anomalies, as well as to detect early habits forming among the enumerator team that may affect data quality. SI also conducts and monitors back-checks on about 10% of the sample to verifying that household was indeed visited and interviewed, and to ensure the reliability of information captured by enumerators.

6.2.6 Secondary Data

Our methodology for characterizing household-level outcomes of interest will utilize two main secondary data sources. First, the WASCO customer database is required both for sampling, and for proper analyses on its own. Both will be used to analyze trends over the relevant period of time for this evaluation, and to contextualize and triangulate results from the household survey. Both of our proposed designs require WASCO customer information to construct a sample frame for connected and unconnected households (the latter by combining with our geospatial sampling approach). As much of the evaluation design relies on this dataset, it will be vital to maintain our positive working relationship with WASCO during the evaluation.

Second, we will use data from Lesotho's current electronic health management information system, DHIS2. SI recently obtained data from the DHIS2 database including water borne illness inpatient visits, outpatient

visits, and deaths disaggregated by age group (below or above five years old) and sex. This data is mostly available from 2010-present, although SI is verifying if it is possible to obtain older data. Data on typhoid is sparse compared to the data on diarrheal illness, but it nonetheless represents the entirety of data in MoH's possession. Although a known limitation of this data is that most cases of diarrheal illness are treated without seeking formal care, this data may still allow us to track trends in severe cases over the life of the Compact.

6.2.7 Analysis Plan

Our evaluation involves estimating the impact of improved access and service on households in UPUW sites through analysis of household survey data, contextualized by administrative and qualitative data. The household survey data will allow us to estimate the presence and quantity of any impact on household-level beneficiaries, while the administrative and qualitative data help us answer 'why' and 'how' questions, and also assist us to understand the role of MCC relative to other factors influencing or potentially confounding any of the observed changes in our outcomes of interest.

Quantitative data analysis

We will first present descriptive statistics for our sample, including outcomes and covariates, disaggregated by UPUW site, for each of the two designs. These descriptive statistics will be shown for both the raw and matched data, to demonstrate the relative balance of treatment and comparison groups before and after matching. We will also present the results of the matching procedure, including the final matching model, the region of common support, and sensitivity tests using alternative matching algorithms.

Our main analysis will involve a comparison of the mean outcome of each treatment and comparison group, following the matching procedure. This analysis estimates the average treatment effect on the treated (ATT). We will present the average treatment effect for the two primary outcomes (time savings and diarrheal illness), as well as for the two secondary outcomes (water consumption and expenditures). We will also conduct other analyses to assess the robustness of our results to other methods. Options for robustness checks include weighted regressions, using the inverse of the propensity scores, or a Heckman selection model. In addition, we expect that comparison households near treatment households may be likely to experience some indirect benefits of the intervention, i.e. spillover effects. Ignoring the potential for spillover could lead to an underestimation of the program's true impact or, in the extreme case, a conclusion of no impact. As mentioned earlier, stratifying our sample in each site and allowing for one strata outside the distance where connections can be made (beyond 500m from the main) may be feasible depending on the density of connections in each potential stratum, which can be assessed after WASCO data is obtained. Further, we can evaluate the influence of distance to the transmission main as a factor influencing self-selection into treatment (by adding it to our selection model in the matching procedure), and its influence on outcomes of interest by including it as a covariate in our regressions.

Secondary & qualitative data analysis

We will analyze secondary and qualitative data specifically in the context of the quantitative findings. We will analyze longitudinal data from WASCO on customer connections, water consumption, and water expenditures for domestic customers, for years at least 2008 to present. We will chart overall trends and disaggregate trends by UPUW site. We will provide descriptive analysis and, if possible depending on the quality of data provided, could conduct an interrupted time series analysis (ITSA) testing for significant changes in these trends before and after the completion of the UPUW interventions. ITSA is more challenging when gradual roll-out of programs occurs and/or when large anticipatory effects are present in the data, and thus the feasibility of this approach will depend on the extent to which we see anticipatory effects in customer

connections before or during the Compact period. For qualitative data, we will develop a codebook prior to reading any transcripts highlighting key themes, and we will code excerpts with each relevant code. All excerpts with the same code can then be analyzed to extract themes that explain key issues including: (1) factors leading to self-selection into treatment; (2) experiences of the interventions over time; (3) other potential reasons for changes in the outcomes; (4) exposure to other interventions over time; (5) uncover any unanticipated effects. Findings will be used to contextualize and explain quantitative results.

6.2.8 Limitations and Risks

Selection bias: Quasi-experimental designs (QEDs) vary in approach and carry their own set of risks and limitations. In general, they are more susceptible to selection bias relative to experimental designs, as the starting point is a group of households we know have been exposed to a program, and comparison households must be identified through statistical procedures that make use of observable traits. WASCO's connection fees increase with distance from the main, which can be used as a proxy for willingness to pay. Nonetheless, as with all QEDs, we can only account for observable traits and remain vulnerable to the influence of any unobservable traits that are correlated with a household's self-selection into treatment or outcomes of interest.

Lack of baseline: The most significant limitation of this impact evaluation is the lack of baseline data for treatment and comparison households. Without baseline data for any of these groups, our analysis can only be based on a comparison *at present*, which implicitly assumes that the conditions of the comparison group would have been expected to represent their same situation a decade ago, prior to the implementation of any UPUW interventions, save for any information we can gather through recall during the household survey. Given the dynamic changes happening in urban areas in Lesotho, this assumption must be carefully considered when interpreting results. Our ability to benchmark current comparison group values against *a priori* expectations may be limited to the use of DHS and NORC data collected prior to the intervention. Both datasets can provide some indication of the level of some key outcomes and important covariates among those unconnected prior to the interventions. We cannot fully mitigate the lack of a baseline, although we do not believe that this invalidates the ex-post comparison of treatment and control areas.

Time lag: SI was contracted to conduct this evaluation well after the intervention had finished. On our projected mobilization schedule for data collection, it will have been over four years since improvements in outcomes of interest may have been expected to begin to take shape. On an ideal timeframe of exposure, we may have conducted endline data collection more proximally to the commissioning of the new and upgraded infrastructure. Our current timeframe of exposure makes respondent recall problematic, which is particularly damaging in the absence of adequate baseline measurements. Our delayed timeframe does present some advantages—long-term behavioral changes in water storage and consumption as well as in WASCO management of the new infrastructure have had adequate time to manifest, and although attribution to MCC will be limited in areas where significant WASCO remediation was required, this remediation will now at least allow for verification of MCC's intended theory of change that would not have been possible at an earlier time. MCC's own ERR assumptions posit benefit streams for at least twenty years, so we have a rare opportunity to test the continued prevalence of benefits as the new infrastructure has had the chance to wear and degrade, though this is subject to the other limitations described above.

WASCO data availability: Although we have succeeded in obtaining nearly all of the data we requested from WASCO, the completeness of variables necessary for locating customers for sampling is lacking in all areas except for urban Maseru and Mazenod. As a result, with currently available information, we would only be able to consider a sample frame using geospatial methods for Design A and Design B in Maseru, leaving any remaining Design A sites that are suitable for an IE subject to expensive listing costs in order to develop a

sample frame. This limitation is mitigated by the fact that many of these sites may, in fact, turn out to be more suitable for a performance evaluation. The sites suitable for an IE have the potential to be relatively small peri-urban areas for whom a listing exercise may be achievable at a moderate cost and could furthermore include marginal questions targeted at verifying that assumptions about mobility and the local housing market required for the validity of our identification strategy to hold.

Confounding: As described earlier, the UPUW sites vary substantially in terms of the overall success of implementation, and the degree to which WASCO had to intervene to remedy issues with design and construction of the works. It follows that in different sites, MCC's direct role in contributing to any observed impacts will vary. With an evaluation approach that aims to measure impacts of the UPUW Activity as a whole, we will not necessarily be able to parse the extent to which any detected effects are attributable to MCC as compared to WASCO-led improvement or remediation efforts. Nevertheless, detected effects would still provide validation of MCC's theory of change and proof of what similar approaches, if implemented successfully in similar contexts, may be able to achieve.

The extent to which other donors, NGOs, or public or private organizations in Lesotho have conducted relevant interventions in any of the UPUW sites over the last few years could also threaten our ability to attribute any observed changes to the MCC interventions. A list of current WASCO projects is listed on their website.[16] In addition, the extent to which other organizations have implemented WASH interventions systematically among either our treatment or comparison households will jeopardize our ability to attribute impacts to MCC. We suggest the need to remain open about potential changes to our design if we discover the presence or influence of such interventions. Otherwise, assuming we proceed with the evaluation as designed, we will gauge households' exposure to other programs through our surveys and focus groups.

Site-specific sample size: We have made the assumption that there are a sufficient number of households in each UPUW site to satisfy the sample size needs of the evaluation in each treatment and control group. In the absence of site-specific administrative data from WASCO, we only have occasional qualitative information from area managers' participation in our process evaluation to validate this assumption. While there appears to be no risk of inadequate connections in large cities with successful UPUW implementation, like Maseru, smaller cities or cities where UPUW implementation went poorly, like Semonkong and Butha-Buthe, are at risk of failing to satisfy required sample sizes. Still, we don't view this limitation as posing a large risk to our evaluation. In the case of Semonkong, all connections are post-intervention connections given the fact that the UPUW Activity funded an entirely new plant. If these new connections are short of the required sample size, we would plan to include the entire treatment population in the evaluation. Sites other than Maseru and Semonkong are combined in our recommended design, such that shortages in available sample size at one site could potentially be ameliorated by oversampling in another site.

Spillover: Given that even unconnected households in this area may indirectly benefit (i.e. by using the taps of neighbors with connections), there is the potential for spillover in these areas. With qualitative data collection sequenced before quantitative data collection, we will be able to assess from focus groups with unconnected households whether they may be experiencing any indirect benefits from the interventions. Considering the limitation of including comparison households from outside the eligible areas for connection – mainly that this is likely to introduce additional bias – the preferable option is to maintain the design as proposed, while attempting to measure and quantify any potential spillover. Mechanisms for spillover (e.g. through neighbors or other means) will be explored in the qualitative data collection, and to the extent possible, measured in the quantitative surveys.

[16] A list of current WASCO projects is described on WASCO's website: http://www.wasco.co.ls/current-projects/

Water quality: For both system-level and household-level water quality, there are no baseline values, targets, or monitoring data from the Compact against which to benchmark current water quality. In general, the assumption that a household's connecting to the WASCO network represents a shift to better quality water is likely to hold but is complicated especially in urban settings where households may substitute water sources of varying quality for different uses, use multiple sources concurrently, and may re-contaminate treated water within the household through storage and retrieval. Water quality testing would represent an extra component of our measurement strategy that will be costed as part of procurement of a local data collection firm. Apart from this, we will request WASCO data on water quality at the system level at each UPUW site, and measure households' perceived quality. However, we note that system-level quality does not indicate point-of-use quality, and that household survey-based perceptions of quality have, in other studies, been shown to be poor predictors of microbiological contamination. As such, these measures would serve to triangulate, rather than as primary indicators of water quality.

6.3 Industry

To characterize industry level impacts we plan to combine non-experimental quantitative and qualitative methods to respond, in part, to evaluation question 7b. Benefits to firms will be evaluated through case studies and interviews with firm owners or managers. Benefits to employees will be measured only through administrative data for employment.

6.3.1 Methodology

Quantitatively, we will analyze pre- and post-Compact trends in a time series of industrial water consumption and employment data from at least 2008 through the present. Qualitatively, we will develop case studies of five purposively sampled garment and textile firms to contextualize the quantitative analysis and hear directly from firm leadership (owners and managers) about how they have perceived and reacted to the results of the Metolong Program. Case studies will involve key informant interviews (KIIs) and site visits to factories. Quantitative data analysis will include all industrial firms in Maseru's industrial estates, and case studies will integrate disaggregated quantitative data for the selected firms. While there is no counterfactual for this component, we will analyze (1) actual data trends observed versus MCC's ERR assumptions, and (2) firms' perception of the counterfactual situation in the absence of the new supply provided by the Metolong Program. Our approach will yield quantitative insights that are relevant to the evaluation question and contextualized by qualitative information from key industrial stakeholders.

6.3.2 Timeframe of exposure

Given the timing of this *ex post* evaluation, our data collection and analysis will occur about 2.5 years following commissioning of the Dam and commencement of increased bulk water supply. Our analysis of industry impacts thus focuses on changes that have occurred since December of 2015, when the Metolong Dam was commissioned. This timing is beneficial for the evaluation objectives, since the most significant expected outcomes (investments, expansion, increased employment) are expected in the intermediate- and long-term. The shorter-term outcome of preserved employment can also be assessed using a combination of the administrative data and qualitative interviews with firm owners and managers.[17]

[17] We know from project documents and scoping trip interviews that there has been no fabric mill opened in Lesotho, as was the main assumption behind the preserved employment benefit from the Metolong Program ERR. Nonetheless, indications are that employment was preserved nonetheless as a result of the most recent AGOA extension. SI will still assess whether employment was preserved and will note whether or not the Metolong Program contributed to this beyond that of AGOA, based on interviews with firm leadership.

In order to be able to assess meaningful changes post-intervention, we will assess data from at least 2008 through present. This will allow us to look at nearly a decade's worth of consumption and employment data leading up to the Metolong Dam commissioning, to provide sufficient context about pre-Compact trends. Administrative data will allow us to look at trends in consumption and employment over the entire period of time following the completion of the MP. With regard to qualitative data collection, we will only be able to capture immediate shifts in firms' perceptions about changes in the bulk water supply through recall, though their perspective now given the longer elapsed period of time is arguably more valuable, given that there has been sufficient time for relevant management and investment decisions to have been made and implemented.

6.3.3 Quantitative Data Collection

We have requested and received monthly administrative data, for 2008 to present, for water consumption, expenditures, employment, wages, and factory space for industrial firms in Lesotho, described below. Where possible and as described below, we have received data that separates out these variables by relevant location and sector.

Water Consumption and Expenditures: We will analyze data on industrial water consumption (cubic meters) and industrial water expenditures (Maloti) from WASCO. The billing and consumption data recently received from WASCO will allow for us to parse out trends for industrial consumers alone (excluding other non-domestic consumers such as commercial and government), focusing on industrial consumers in Maseru.

Employment, Wages, Factory Space: We have also received, and will analyze, industrial employment data (total and sex-disaggregated by sector, where available), factory space (square meters), and wages (Maloti) for industrial firms supported by the Lesotho National Development Corporation (LNDC). SI was informed during our scoping trip by an industry stakeholder that LNDC is the most reliable source for this data, since it is the main parastatal organization in charge of promoting industrial development in Lesotho. Only employment and minimum wage data is available historically, and even among these it is not always possible to distinguish Maseru-based trends from national trends. Still, however, the data will permit us to characterize national industrial employment and minimum wage trends by sector and verify whether or not the amount of factory space currently in use reflects what was projected in the MCC ERR. Factory space was not provided longitudinally, so only one figure is available for each industrial firm in the administrative dataset.

6.3.4 Qualitative Data Collection

We plan conduct KIIs and site visits with five firms from the garment and textile industry. These primary data will be supplemented by secondary document review of project and publicly available industry sources (e.g. LNDC annual reports) to fill gaps in or contextualize information from interviews and site visits.

Sample Frame: The sample frame from which we will purposively select enterprises will include all garment and textile firms operating out of either the Thetsane Industrial Area or the Tikoe Industrial Estate. We plan to use the data obtained from LNDC to form the sampling frame from which to select these firms. In the event that the data available from LNDC is only available in aggregate form, our backup method for constructing the sample frame will be to request the full register of companies in Lesotho maintained by the Lesotho Ministry of Trade and Industry (MTI)[18]. We would filter this registry based on physical addresses that indicate company residence at Thetsane or Tikoe and sector codes that indicate a company is involved in the garment and textile

[18] See: www.companies.org.ls

industry. As a last resort, we will use a non-comprehensive list of garment and textile firms located at the two industrial areas obtained by SI from a representative of the Lesotho MTI during our scoping trip.[19]

Sampling: We will sample five firms purposively for this component of the evaluation. Where information is available to do so either from scoping trip interviews or WASCO data showing the largest consumers of water, we will attempt to prioritize firms with the greatest water needs or water consumption pre-Compact. Secondarily, we will attempt to obtain a diverse cross-section among key characteristics such as sub-sector within the textile and garment industry (knit, woven, denim, etc.), total industrial area, and primary market served (e.g. South Africa, U.S., etc.), depending on the availability of this data from LNDC records. The objective of this secondary goal is to collect perspectives from firms using water for a variety of productive purposes including steaming, washing, bleaching, and dyeing. A summary of the KII topics for firm owners or top managers is presented in Table 17. We also intend to conduct a small number of key informant interviews with representatives from LNDC and MTI, to further contextualize – and obtain a macro view of – changes in the operating environment of industry over the last few years and the extent to which water was or still is a constraint. These individuals would likely be able to comment on other key factors influencing industry growth, employment, and expansion.

[19] This list is included as part of a synopsis publication shared by Mark Bennett at the South African Trade Law Centre (tralac) 2017 Annual Conference. See: https://www.tralac.org/news/article/11501-lesotho-s-textiles-apparel-and-footwear-manufacturing-industry.html

Table 17. Illustrative key informant interview topics for Industry Case Studies

Theme	Illustrative topics
Pre-Compact	Historical challenges, perceptions of supply adequacy pre-Metolong Program
Supply chain	Description of supply chain, from raw inputs to final product Input processing & value; general water needs; wastewater treatment
Current water usage	Current water consumption Productive uses of water/water as an input Water use for other purposes (e.g. employee consumption) Perception of current supply reliability, quality, adequacy Current challenges with water supply, mitigation, coping costs
Perceived impacts of MP	Investment, expansion, employment Changes in water consumption, water expenditures Discussion of perceived counterfactual Discussion of observed trends in quantitative data Factors mediating impact (facilitated, inhibited) Unintended consequences Lessons learned
Future expectations	Constraints to further investment or expansion Water supply versus other constraints

6.3.5 Analysis Plan

We will analyze data for all firms within the Thetsane and Tikoe industrial states, both in aggregate and disaggregated by firm, where possible. The latter will allow us to pinpoint whether specific firms may be driving aggregate trends. Prior to analysis, SI will review all administrative datasets for anomalies such as extreme outliers prior to using them. If any anomalies are found, SI will follow up to clarify or seek explanations, from WASCO and LNDC. If these anomalies are left unexplained, SI will conduct analyses with and without these points to test the sensitivity of the conclusions to extreme outliers. For relevant indicators (e.g. expenditures, wages), SI will conduct any necessary standardizations be able to make comparisons over time. The objective in the time series analysis of our quantitative data will be to separate random noise or cyclical/seasonal trends over time from discernable patterns over time, related to the Metolong Program. We will describe trends in industrial water consumption and employment over time, with specific reference to changes after December 2015. We will highlight any shifts in trends that appear to occur at other points in time that may affect our ability to conclude any impact of the Metolong Program, such as other concurrent political or economic events. Using employment data, we will assess whether trends correspond to projections in MCC's ERR calculations for the Metolong Program, and if not, provide suggested inputs for updating the ERR. To the extent possible depending on the nature of the trends in the data, SI will analyze changes in these key indicators using time series analysis methods (e.g. auto-regressive, moving average, or Box-Jenkins models). The objective of applying time series analysis methodologies is to go beyond description of trends, to analyze whether there are statistically significant changes in the trends before and after a time point of interest.

For the KIIs conducted as part of our qualitative case study methodology, SI will take detailed notes during interviews and will transcribe them in full following the interviews. Site visits around factories selected for these case studies, provided access is permitted by factory owners/management, will allow SI to directly observe how increased water supply is being used as part of manufacturing processes. We anticipate that site visits will be guided and facilitated by the key informants, which will also allow for further follow-up for topics discussed during the KII. We intend to conduct preliminary analyses of the administrative quantitative data

prior to conducting KIIs and site visits at the factory premises, so that we are able to display our analysis to each key informant and ask relevant follow up questions to contextualize and observed trend changes. SI will conduct a thematic analysis of these KII notes wherein we will develop a codebook to capture key themes related to the project logic hypotheses and underlying assumptions.

The quantitative and qualitative data from the industry component of this evaluation will allow SI to validate the key program hypothesis for industry – that water supply is a necessary condition for expanded employment and firm creation in the textile and garment industry in Lesotho – or otherwise identify factors that explain why this assumption did not hold. We will integrate findings from the quantitative and qualitative analyses, so that key quantitative trends are displayed and contextualized with findings from the case studies. We will analyze and report on the information shared by key informants as relevant for each hypothesis, interpreting responses in the context of information available through project documentation, project monitoring information, and direct observation during site visits.

6.3.6 Limitations and Risks

Completeness of administrative data: Administrative data has been obtained through WASCO and LNDC. As described above, not all of the data requested is available in the form anticipated. Factory space is not provided historically, with only a single figure provided for each firm in the LNDC database. Wages specific to firms are not available and were obtained instead through the minimum wage gazette (in hard copy).

No counterfactual: As mentioned earlier, from an internal validity point of view our methodology does not identify an explicit counterfactual—that is, the state of affairs for industrial firms in the Maseru area in the absence of the bulk water supply provided by the MP to which their current state of affairs could be compared. Our methodology will be able to identify trends before and after the supply became available, but it will not fully be able to parse out to what extent these trends are owed to the supply of water and to what extent they are owed to unobserved macro or micro-level variables – except through discussions with key informants. Insights from our qualitative case study approach may yield intuitions that can help explain the trends and put them in context, but there is no experimental or mathematical construct that can isolate the effect of the program. With that being said, a design option with such a construct is not available to SI. The intervention affected practically the whole of the textile and garment industry in the Maseru area, so there is no comparison group to which the affected firms can be compared. As described earlier, our analysis will compare observed trends to estimated “without-program” trends from MCC’s ERR. We will discuss observed and projected trends with firm leadership during key informant interviews, to obtain their informed perspective about how water consumption, expenditures, employment, and investment might have differed in the absence of the program.

Firms’ willingness or ability to participate: As for our qualitative case study methodology, we anticipate some potential risks in obtaining the needed information from selected firms. First, firms are busy with normal operations or alternatively may experience survey fatigue if they have been approached by numerous others for various data collection purposes, and thus selected firms may not agree to participate. We will select more than five firms for the case studies and pursue back-ups if any of the prioritized five are not available. Further, it may be difficult to schedule time with the right individual(s) at each firm who can provide insight into the relevant topics. The intended respondent will either be the owner or top decision-maker at each firm (depending on who is based in Lesotho), though we anticipate in some cases this individual may refer us to others who have specific knowledge about water consumption or usage at each firm. It may be a logistical challenge to schedule interviews, or even to complete them in a single visit. Our team will persist with up to three attempts to make initial contact with the appropriate respondents at each firm, and then will conduct up to three visits in order to complete each interview.

In addition, respondents may not be able to recall specifics about the situation from several years ago, and similarly turnover of staff at the firm may limit institutional knowledge among current firm staff. While we anticipate that our interviews will not address issues considered sensitive to these firms, there is a chance that respondents are hesitant to answer questions fully or in an unbiased manner if they perceive any potential reputational risk to the firm (e.g. discussions of current wastewater practices in the absence of a treatment facility), or otherwise do not feel authorized to provide responses or facilitate site visits without explicit permission from a supervisor. Targeting the owner or top-most decision maker to some extent averts the latter risk. We will attempt to mitigate the former by making clear the objectives of the evaluation, the evaluation team's affiliations, and the specific purview of the questions we are interested in answering.

Purposive sampling and external validity: Finally, our decision to purposively select firms for the case study will mean that our findings will pertain specifically to the firms interviewed or similar large, exporting firms and may not be representative of the industry as a whole (though our quantitative analysis will include analysis of all firms' data in aggregate). As our study is naturally interested in the heaviest water users of the textile and garment industry in Lesotho, it is not a risk to the study that our qualitative findings will not be representative to all firms, including those that are smaller and non-exporting.

6.4 Small & Medium Enterprises

Our approach to assessing changes in water use among SMEs in Maseru (evaluation question 7b), is to analyze administrative data and conduct qualitative interviews with knowledgeable sector stakeholders.

6.4.1 Methodology

We have opted not to pursue an impact evaluation design or a more intensive data collection effort for this component, given our assessment of the trade-offs in technical rigor, attribution, feasibility, and cost. Overall, we considered four options for evaluating effects on SMEs: (1) a module within the household survey; (2) a separate SME survey; (3) using surveys conducted by others to obtain the needed information; and (4) administrative data analysis and KIIs only. We detail our recommended approach (option #4) below and describe our assessment of each of the other three options relative to our selected methodology.

Our approach entails analyzing commercial water connections (changes in rates of connection over time), commercial water use (consumption and expenditures), and trends in formal registration of commercial entities in the Maseru area over time. Due to constraints in the administrative data available to SI (data obtained from the Ministry of Trade and Industry (MTI) does not currently include information on firm size, such that SMEs could be identified, either on the basis of employment or annual turnover), we will not be able to restrict this quantitative analysis to SMEs only. We will supplement this with qualitative key informant interviews specific to stakeholders knowledgeable about SMEs, including key officials from the Ministry of Trade and Industry (MTI), and other stakeholders in this sector in Lesotho, such as donor-funded business development support projects and relevant small business associations in the Maseru area.

6.4.2 Timeframe of exposure

For SMEs, the timeframe of exposure to the interventions is equivalent to the households – approximately four years since the completion of the works, and nearly ten since before the Compact. While this may present some challenges with informants' recall of the pre-intervention situation, we expect that benefits flowing to SMEs could take longer, in general, to manifest than at the household level, contingent on dependent processes like hiring employees or attracting more investment. The amount of time elapsed, to have allowed such effects to manifest, may be beneficial to the evaluation, especially as long-term trends can be assessed

using the WASCO data and key informants can reflect, ideally, on the relative contribution of the MCC interventions versus other business environment factors that could influence outcomes of interest.

6.4.3 Study Sample

All data for relevant enterprises from the WASCO database will be assessed. In order to take advantage of an unexpected research opportunity we conducted the SME-level interviews prior to obtaining administrative data. Key informants were purposively sampled, as specific individuals are best placed to respond to questions regarding changes in enterprise water consumption and use, water as a constraint to enterprise growth (in relation to other constraints), and other business environment factors influencing employment and investment. We will analyze the qualitative data from the interviews in the context of the administrative data, conducting follow-up interviews as necessary to fill any perceived gaps in understanding SMEs reaction to changes brought about by the urban water programming.

6.4.4 Secondary Data

WASCO's customer database will allow us to assess trends in water use among commercial entities in Lesotho over time. Although we had planned to use information from the MTI company registry[20] to limit our analysis of the WASCO commercial data to SMEs alone, excluding larger businesses, we have determined since obtaining raw data from this registry that we will ultimately be unable to exclude these entities from the larger set of commercial WASCO customers. The MTI business registry database will allow us to assess trends in new business registration over time, although we will again be unable to restrict our analysis to SMEs based on the information available to us in the registry. A 2008 MTI report claimed water was a greater constraint for SMEs in peri-urban areas of Maseru, which we can attempt to verify by looking at consumption trends in Maseru urban versus peri-urban using the WASCO customer database.

The advantage of using administrative data is that it allows for an objective picture of network water usage over a long time horizon; on the other hand there are distinct drawbacks as it cannot provide insight into how businesses manage or allocate their water usage depending on changes in supply, nor can it speak to coping or aversion expenditures on the part of SMEs, who – like households in urban areas – may be substituting to meet their needs with other, lesser quality and/or more expensive sources, or engaging in other coping behaviors. In this specific case, the administrative data can also only characterize commercial water usage in broad strokes, without the ability to differentiate between trends among smaller and larger companies. Assessing use of network water thus provides an important, but not full, picture of commercial water use.

6.4.5 Qualitative Data Collection

Given the limitations we've identified in directly obtaining the perspective of enterprises, to fill this information gap, we have opted to supplement this quantitative analysis with key informant interviews (KIIs) with individuals that have a broad perspective on enterprise growth, constraints to growth, trends over time, and importantly, other factors that could explain any observed trends, e.g. changes in the political climate, economic regulations, or other national-level shocks. We used a recent survey conducted by Finscope of micro, small, and medium enterprises (MSME) in Lesotho as a point of departure to select small business associations, civil society organizations, and non-governmental organizations providing support to small businesses (Finscope 2016). We then requested references from each of these organizations to any other organizations in Lesotho that they believed would be knowledgeable about growth and water usage of SMEs. Below we list some of the main organizations working in this space in Lesotho; which we attempted to interview

[20] MTI Business registry: http://www.companies.org.ls/

during a trip to Lesotho to gather administrative data. We have thus far conducted four KIIs spread among three of these organizations, which we may supplement with additional interviews if needed and as feasible during future data collection trips.

Table 18. Key informants for SME component

Entity	Description/rationale
Ministry of Small Business Development, Co-operatives, and Marketing*	Nascent Ministry which absorbed staff and mandate from the Ministry of Trade and Industry to provide government support to SMEs. High-level perspective on trends in business registration, retention, and growth, especially in relation to other factors that may influence enterprise behaviors.
Basotho Enterprise Development Corporation (BEDCO)[21]*	Provides business development support services for small and medium-sized enterprises in Lesotho.
Lesotho Chamber of Commerce and Industry (LCCI)[22]	Promotes SMME growth and entrepreneurship; investment in Lesotho; participation of Basotho-owned businesses and export-oriented industrial development. Collaborates with the SMME Network (see below).
Small, Micro and Medium Enterprises Network (SMME Network-Lesotho)[23]*	Consortium of eleven business development service (BDS) providers, mainly focusing on the empowerment of SMME initiatives and activities.
Federation of Lesotho Women Entrepreneurs (FLWE)[24]	The FLWE has had a mandate since 2013 to enable women to contribute to poverty reduction and economic growth initiatives through supporting women-led businesses. Recent news suggests that the primary mechanism through which they attempt to fulfill this mandate is microfinance, suggesting a focus on SMEs.
Lesotho Hotels and Hospitality Association (LHHA)[25]	Most of the LHHA's members are small and medium local tourism operations, which may rely on adequate water supply and quality for routine operations and customer satisfaction. While these SMEs are not in the manufacturing sector, this organization was a referral from another key informant; time allowing during future data collection trips, we will invite the LHHA for an interview.
Lesotho Enterprise Assistance Program (LEAP)[26]	Program funded through the World Bank, started in 2014, which provides business development support to micro, small, and medium enterprises (MSMEs)

* Already interviewed by SI

6.4.6 Quantitative Data Collection

Primary quantitative data collection will not be conducted as part of this evaluation component. We considered implementing a survey module as part of the household survey, as well as a stand-alone SME survey. However, there are several limitations to these approaches. First, there is a lack of baseline data against which to compare any survey data we would collect, through either approach. MCC and the World Bank both implemented enterprise surveys before the interventions were completed, but each has its own limitations. The MCC enterprise survey contains a prohibitively small number of manufacturing SMEs in urban areas, and many of the useful indicators are encoded in ways that cannot be used for analytical purposes. The World Bank surveys contain a larger number of urban manufacturing firms, but do not measure water-related

[21] Website is not functional. See: https://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapId=81852393
[22] LCCI website: http://www.lcci.org.ls/entrepreneurship-and-innovation/
[23] SMME Support network website: http://smmesupportnetwork.co.ls/
[24] See: http://genderlinks.org.za/programme-web-menu/lesotho-economic-empowerment-initiatives-changing-womens-lives-2013-05-13/
[25] LHHA website: http://www.lesothohotelsandhospitality.com/
[26] World Bank 2017: http://www.worldbank.org/en/news/feature/2017/03/07/better-business-lesotho-enhances-competitiveness-in-new-growth-sectors; http://www.worldbank.org/en/news/press-release/2017/03/24/lesothos-efforts-to-promote-private-sector-led-growth-and-jobs-lifted

indicators as comprehensively as the MCC enterprise survey. We also assessed the results of an MSME survey conducted by Finscope post-Compact in 2015-2016 across Lesotho, with a sample of largely informal, micro-enterprises, identified via an approach typical of household surveys (enumeration area listing, then household surveys) (Finscope 2016). These data are not public, though we have been in contact with the Finscope team and may be able to request their data. The report is available online and shows that water is ranked low on a list of constraints to growth among MSMEs in Lesotho. However, they do not present disaggregation of these figures for SMEs and there is no pre-Compact wave of this data.

Survey recall is an option, but suffers the same limitations as described in depth with respect to the household survey; the long time-lag between pre-intervention and endline is likely to result in noisy, lumpy, or potentially unreliable information regarding prior water use from other sources and prior coping behaviors. It is tempting to layer an SME module on the household survey effort already planned, but we assume MCC's interest is primarily in relation to formal SMEs, while the household survey would result in a sample largely comprised of informal, micro-enterprises.[27] Further, the intended respondent for the household survey will not necessarily be the business owner in the household and may not have answers to enterprise-related questions tacked on to the household survey, potentially necessitating a follow-up visit to speak to the business owner in the household. A separate SME survey would be technically feasible, using a sample frame that could be generated using data from the Ministry of Trade and Industry (MTI) enterprise registry, or using WASCO's customer database, but is subject to the same limitations described above.[28]

Further considerations of the project logic and ERR have led us away from survey approaches to this evaluation component. The problem diagnostic and theory of change for SMEs is generally unclear in Compact decisional documents. The discussion in those documents mostly centers around industry beneficiaries, with only cursory mention of expected impacts on businesses, and no detailed due diligence or beneficiary analysis specific to SMEs. The hypothesis is that improved water service would retain and attract more enterprises to Maseru, specifically those using water as an input (though we note that the evaluation question narrowly pertains to water use among businesses). There is limited discussion regarding the status of SME water use at the beginning of the Compact or detailed water demand projections over time. The MCC investment memo noted that: "the value of water as an input for commercial development in urban areas is difficult to assess." The project logic for Package 1 includes hypothesized benefits of reduced manufacturing costs and increased manufacturing opportunities, but these do not distinguish between industry and SMEs. Based on available information, we assume MCC's hypothesized impacts pertain to manufacturing SMEs in Maseru. Indicators related to enterprises were eventually dropped from the M&E plan, though the hypothesized benefit was retained in the economic model. Further, although the MCC ERR's fourth hypothesized benefit stream relates to SMEs, it appears unlikely that any information gathered about SME water use or experience of improved supply would feed into a revised ERR unless the model was adjusted substantially.

We note our willingness to reconsider survey approaches to this component, depending on MCC's interests and contingent on clarification of some of the gaps in the theory of change for SMEs noted above. However, we encourage a critical assessment of the required resources against the planned uses of such data, as well as trade-offs in technical rigor, as described above.

[27] The recent Finscope survey reports that through a listing of 336 enumeration areas (including 47 thousand households), they identified 4 thousand owners of micro, small, and medium business owners. After applying sampling weights, they estimated that among all MSMEs in Lesotho, only 3.3 percent were owners of small or medium enterprises, with the vast majority micro-enterprises (<5 employees).

[28] We consulted with the team that oversaw enterprise surveys in Lesotho and Swaziland, and gathered useful information and lessons learned from their team. This includes the fact that multiple respondents are often required to obtain responses to all of the survey questions, though we expect that in our case, a survey that is more streamlined toward manufacturing firms asking primarily questions related to water use may be more easily targeted to a single respondent.

6.4.7 Analysis Plan

We will analyze data for all formally registered businesses with data in WASCO's customer database, as well as the MTI business registry. Our analysis will focus on changes in trends with regard to specific outcomes related to SMEs, based on the information available to SI and the focus of the evaluation questions, including: water use (consumption and expenditures) of enterprise customers of WASCO in the Maseru area, and employment and new business registration among businesses in the Maseru area. With the administrative data available to SI from WASCO and MTI, it will only be possible to discern aggregate commercial trends in consumption and registration quantitatively, although we will rely on qualitative responses to contextualize these trends with perceived outcomes among SMEs according to respondents. As with the industry data, prior to analysis, SI will review all administrative datasets for anomalies such as extreme outliers prior to using them. If any anomalies are found, SI will follow up to clarify or seek explanations. We will describe trends over time, with specific reference to changes after December 2015, given the completion of the MP and UPUW works.

Findings from the key informant interviews will supplement the quantitative analysis, by providing useful insight into observed trends, perspectives from the enterprise perspective to the extent possible, and contextual information regarding other factors in the business environment that could be driving any observed trends in the quantitative data. Likewise, we expect that findings from the key informant interviews will allow us to draw out additional lessons learned with regard to water demand and use among SMEs in similar urban contexts. SI will analyze detailed notes taken during KIIs, transcribed in full following the interviews. SI will conduct a thematic analysis of these KII notes wherein we will develop a codebook to capture key themes related to the project logic hypotheses and underlying assumptions.

6.4.8 Limitations and Risks

Completeness of administrative data: As discussed in previous sections, the administrative data currently in SI's possession for the quantitative component of the SME methodology does not include all of the variables SI would parse SME trends in employment, water consumption, and water expenditure out from broader, aggregate commercial trends. The data available will still provide objective historical data on commercial water consumption and formal registration against which qualitative findings can be compared and contextualized. Even if these trends are not particularly informative, qualitative data alone could provide useful information regarding the SME-level project logic and ERR assumptions.

Availability of key informants: Our qualitative methodology relies on the perspective of 4-6 key individuals who are well-informed about SME commerce in Lesotho in a broad sense and about any effect that water supply has on businesses. The interviews conducted so far did not pose any major limitations; the FLWE was not available to meet with SI during a recent trip to Lesotho and future attempts will be made to engage.

Firm-level perspective will not be obtained directly: Our methodology aims to capture broad trends in urban manufacturing SMEs without actually surveying these firms directly. To the extent that their perspectives are not represented by key commercial stakeholders who purport to know about sector-wide trends, our methodology will not capture an accurate representation of the role that the MP and UPUW Activity have played in firms' operations and decision-making. As described earlier, we do not believe that it would be cost-effective for our study to capture a representative sample of these firms in light of the limitations described above. Adding a module to our household survey would likely identify informal and micro-enterprises that were not of primary interest to the evaluation, and a full listing exercise that would capture such a sample would be expensive in the context of how important of a beneficiary SMEs were to the projects. Focusing on firm-level

perspectives with a qualitative exercise would inevitably be unrepresentative and likely biased toward the perspective of easily contacted firms.

No counterfactual: By selecting a non-experimental methodology for SME-level beneficiaries, we are introducing error into our characterization of program outcomes by not presenting them in the context of the counterfactual – the state of affairs for SMEs in the absence of the MP and UPUW Activity. This means that we will be unable to fully isolate the role that the MCC projects played, from historical effects and confounding factors, aside from presenting what key stakeholder perceive the state of affairs may have been. However, a counterfactual approach for SME-level beneficiaries is impractical. Without the ability to randomize enterprises into groups that would benefit or not from the MP and UPUW Activity, any counterfactual-based approach would need to rely on baseline data collection. The baseline data collection on enterprises funded by MCC does not have a sufficient sample for our purposes and includes variables encoded in a way that does not render them useful (these variables have been shared with MCC separately). So, it is not possible for us to utilize such an approach in our own methodology. In both our quantitative and qualitative methodologies, however, we will aim to minimize the risk this poses to the validity of our evaluation by comparing trends in the pre-intervention period using administrative data and probing in qualitative interviews for respondents' perceptions of other factors that could explain any observed changes over time in SME water use. These may include macroeconomic shocks, changes in the legal framework governing SME operation, or any local policy or international donor efforts that were concurrent with MCC's programming.

6.5 Economic Analysis

Based on the methodology outlined above, SI anticipates that we can provide updated inputs for the ERR with regard to household benefit streams (in the UPUW Activity ERR), and industry employment trends (in the Metolong Program ERR). Our data collection approach for the household-level analysis will allow us to provide inputs for diarrheal illness reduction and time savings. We note that the ERR included assumptions for diarrheal illness mortality rather than morbidity. While we will attempt to triangulate our survey findings with administrative data from the Ministry of Health, diarrheal illness estimates are more reliable and could be used as a proxy to update this component of the ERR. SI will also be able to comment broadly on the assumptions made within the ERR, using our updated literature review as well as, eventually, findings from the household survey and focus groups, as well as data analysis and case studies with industry beneficiaries. The intention of this broad commentary would be to help MCC inform future ERRs for urban water sector interventions.

With regard to cost, we can use the information available in MCC and LMDA budget documents to update actual project costs. However, as described in depth above, we note that not all observed impacts may be attributable to MCC, given the substantial investment made by WASCO in remediating infrastructure at many sites. Further, the specific amount invested by WASCO to remediate sites is not available to SI at present. We may be able to use planned data collection plans outlined in this report to attempt to gather this information; however, based on the process evaluation fieldwork we are not optimistic we will be able to obtain precise estimates of this amount.

7 REPORTING & DISSEMINATION

As part of this summative evaluation, SI will prepare an evaluation report (inclusive of key findings from the process evaluation), a presentation, and anonymized datasets. We will also prepare a Findings Brief to summarize the evaluation results, and an online StoryMap, both in order to convey results to a wider array of stakeholders in non-technical language. Draft reports will be shared with MCC and key stakeholders for review and comment according to MCC's Evaluation Clearance Process. Reports and datasets will be made available for public access through MCC's Evaluation Catalog. Following the submission of the draft evaluation report, DC- and Lesotho-based dissemination efforts will also be coordinated by SI with MCC's approval and guidance. The final public report will incorporate feedback from MCC and local stakeholders, including any statement of difference/support.

7.1 Deliverables

The contract deliverables in SI's scope of work for this evaluation include those outlined below.

Table 19. Evaluation deliverables & anticipated schedule

Task	Deliverable	Schedule
Obtain Evaluation Data	Draft data collection firm TORs	June 1, 2018
Obtain Evaluation Data	English questionnaires and training manuals (draft & final)	Drafts complete; final to be completed August 2018
Obtain Evaluation Data	Travel SOW and Trip Report for each visit to Lesotho	April 2018 trip completed Additional trips anticipated: June 2018 (qual); September 2018 (quant); 2019 TBD (dissemination)
Obtain Evaluation Data	IRB package	June 2018
Obtain Evaluation Data	Summary of pilot test	July 2018 (qual); September 2018 (quant)
Obtain Evaluation Data	Written minutes of meetings with data collection firm(s)	June-October 2018
Obtain Evaluation Data	Any deliverables from survey firm	June-October 2018
Obtain Evaluation Data	Written summary of quality control checks	November 2018 (preliminary) Final with final report
Develop Final Evaluation Report, Update Economic Analysis	Draft and Final Evaluation Report, including documentation of local stakeholder and MCC feedback and response and any Public Statement of Difference/Support	Draft: February 2019 Final: May 2019
Prepare Data Files	Final raw and analysis files, anonymized following MCC guidelines	June 2019
Conduct Dissemination	Final .ppt files for presentations	2019, TBD

7.2 Dissemination Plan

SI's dissemination plan includes two main components. The first includes in-person presentations to key stakeholders in Lesotho and Washington D.C. to present the findings of the evaluation, including all components (process and impact). We will use the in-person presentations to socialize and explain evaluation findings in a way that is understandable and useful to relevant stakeholders. At MCC's discretion, these presentations may include MCC- or LMDA-facilitated sessions to allow these stakeholders to discuss the

evaluation findings, conclusions, and any lessons learned among themselves with key members of the evaluation team present and available to clarify and expand upon important, policy-relevant points. Stakeholders external to the project will be invited to both presentations, as desired and permitted by MCC. Key project stakeholders we expect to attend dissemination events include WASCO, LMDA, the Commissioner of Water (COW), the Ministry of Water, and potentially other donors involved in funding the Metolong Program. Other non-project stakeholders who may be interested to attend dissemination events or receive findings include: LEWA, DRWS, USAID, DFID, UN Lesotho, and the World Bank.

In addition to these presentations, SI will produce and distribute (a) a web-based ArcGIS StoryMap, displaying the evaluation's findings in an interactive, innovative manner, and (b) a Findings Brief no more than five pages in length which summarizes the evaluation's findings for a lay audience, and which can be translated to Sesotho if requested by MCC. These two products will be intended to stand on their own as reporting products with the full, written report providing more detailed findings. With MCC's permission, both products will be made available via links in the MCC Evaluation Catalogue and on a relevant, project-specific page on SI's website.

8 ADMINSTRATIVE

8.1 IRB requirements and clearances

SI has an in-house Institutional Review Board (IRB) that reviews applications for research with human subjects. Prior to conducting data collection, study protocols are submitted to the IRB with draft instruments and informed consent statements. Depending on the nature of the study and the study population, the application may be designated for full-board or expedited review. It is our expectation that this study will be designated for expedited review. Any changes to the protocols, draft instruments, and consent statements made either in response to MCC review or following enumerator training and piloting will be re-submitted to the IRB for an updated, final approval. SI will also ensure that, as applicable, any required local permissions are obtained before data collection.

8.2 Data access, privacy, and documentation

SI's process for respecting privacy of respondents during data collection, transfer, storage, analysis, disposal and dissemination is governed by SI's data security guidelines, which are aligned with MCC's microdata guidelines. SI will adhere to MCC's open data policy in preparing data for publication. All primary quantitative ex-post data collected as part of the evaluation will be submitted to MCC according to the most updated version of MCC's Disclosure Review Board (DRB) process. SI will submit qualitative data with direct identifiers removed to the extent feasible without posing risks to participants, and in keeping with the informed consent, and as permitted by the SI IRB.

8.3 Evaluation team roles and responsibilities

The SI evaluation team has several core team members that will work together to design and implement the summative evaluation, analyze the data, and produce final reports. Team composition is detailed in Table 20 below. Besides these core team members, other SI staff members or consultants may contribute to the project to provide surge, short-term research or administrative support.

Table 20. Evaluation team roles and responsibilities

Name	Role	Responsibility
Danae Roumis	Program Manager	Lead engagements with MCC and local counterparts. Lead design for evaluation and data collection. Oversee all data collection and analysis, coordinate team members inputs, supervise completion of all deliverables. Lead author of written deliverables.
Basab Dasgupta	Sr. M&E Advisor	Review IE design, data collection instruments, and final report; provide high-level technical advisory support and quality control.
Robin Clanahan	Sr. Analyst/Water	Consult with SI regarding implications of infrastructure functioning for IE design and analysis. Comment on water-related aspects of data collection instruments.
Pheello Pompong	Research Assistant (*Local coordinator*)	Liaise with in-country stakeholders regarding data needs. Assist in the oversight of preparation for and execution of data collection. Contribute to the development of data reports.
Miguel Albornoz	Jr. Analyst	Lead project coordination. Assist Program Manager with implementation of evaluation plans; work closely with WASCO to obtain and use customer data; primary point of contact for local data collection firm; contribute substantively to data analysis and report-writing.
Andrea Hur	Admin. Assistant	Assist Jr. Analyst with project coordination, data collection, and coordination with MCC during data collection and data analysis.

8.4 Evaluation Timeline

Table 21 presents the anticipated evaluation timeline, based on updates to the schedule given the sequencing of data collection with qualitative conducted and completed prior to quantitative data collection. Qualitative data collection will be completed during July 2018. Using the results of the qualitative data collection, final decisions will be made about which sites to include in the impact evaluation, and quantitative data collection will begin in September 2018. Following the anticipated completion of household surveying by early November 2018, data cleaning and final quality control checks will be conducted prior to analysis, which will extend until draft final report submission in mid-February 2019. Following integration of comments from MCC and local stakeholders, the final draft will be submitted at the end of March 2019. At MCC's discretion we will conduct dissemination events in Washington D.C. and Lesotho either of preliminary findings, following MCC evaluation management committee (EMC) revisions to the draft evaluation report, or of final findings, following EMC approval of a final evaluation report.

Table 21. Evaluation timeline

Round	Data Collection	Data Analysis	Draft Report	Final Report
Endline	Qual: start end of 7/18 Quant: 9/18-11/18	Qual: 8/18-9/18 Quant: 11/18-1/19	2/15/2019	3/29/2019

9 REFERENCES

Adane, Metadel, Bezatu Mengistie, Girmay Medhin, Helmut Kloos, and Worku Mulat. 2017. “Piped Water Supply Interruptions and Acute Diarrhea Among Under-Five Children in Addis Ababa Slums, Ethiopia: A Matched Case-Control Study”. PLoS ONE 12 (7). PLoS ONE, e0181516. https://doi.org/10.1371/journal.pone.0181516.

Alarcon Falconi, Tania M., Alexandra V. Kulinkina, Venkata Raghava Mohan, Mark R. Francis, Deepthi Kattula, Rajiv Sarkar, Honorine Ward, Gagandeep Kang, Gagandeep Kang, Vinohar Balraj, and Elena N. Naumova. 2017. “Quantifying Tap-to-Household Water Quality Deterioration in Urban Communities in Vellore, India: The Impact of Spatial Assumptions”. International Journal of Hygiene and Environmental Health 220 (1). International Journal of Hygiene and Environmental Health, 29-36. https://doi.org/10.1016/j.ijheh.2016.09.019.

Alexandra V. Kulinkina, Karen C. Kosinski, Alexander Liss, Michael N. Adjei, Gilbert A. Ayamgah, Patrick Webb, David M. Gute, Jeanine D. Plummer, Elena N. Naumova, 2016, 'Piped water consumption in Ghana: A case study of temporal and spatial patterns of clean water demand relative to alternative water sources in rural small towns', Science of The Total Environment, vol. 559, pp. 291-301 DOI: 10.1016/j.scitotenv.2016.03.148

Ashraf, Nava, James Berry, and Jesse Shapiro. 2010. “Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia”. https://doi.org/10.3386/w13247.

Bain, Robert, Ryan Cronk, Jim Wright, Hong Yang, Tom Slaymaker, and Jamie Bartram. 2014. “Fecal Contamination of Drinking-Water in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis”. PLoS Med 11 (5), e1001644. https://doi.org/10.1371/journal.pmed.1001644.

Banda, B. M., S. Farolfi, and R. M. Hassan. 2007. “Estimating Water Demand for Domestic Use in Rural South Africa in the Absence of Price Information”. Water Policy 9 (5). Water Policy, 513. https://doi.org/10.2166/wp.2007.023.

Bartram, Jamie, and Sandy Cairncross. 2010. “Hygiene, Sanitation, and Water: Forgotten Foundations of Health”. PLoS Med 7 (11), e1000367. https://doi.org/10.1371/journal.pmed.1000367.

Becker, Sascha O., and Andrea Ichino. “Estimation of average treatment effects based on propensity scores,” The Stata Journal, 2, no. 4 (2002): 358–377.

Bello, H.M, M.R Malefane, and S. Babatope-Obasa. 2010. “An Estimation of Water Demand Function in Some Peri-Urban Settlements of Lesotho: HOUSEHOLD WATER SUPPLY AND USE IN SOME PERI-URBAN SETTLEMENTS OF LESOTHO” 10 (1). Household Water Supply and use in some Peri-urban S. https://doi.org/www.researchgate.net/publication/266893215.

Bennett, Mark. 2006. "Lesotho’s Export Textiles & Garment Industry." In *Textile and Clothing Industry in Sub-Saharan Africa*, Herbert Jauch / Rudolf Traub-Merz (Eds.) Available at: http://library.fes.de/pdf-files/iez/03796/12lesotho.pdf

Bennett, Mark. 2017. "Africa’s Inspired Apparel Sourcing Hotspot: Lesotho’s Textile, Apparel & Footwear Manufacturing Industry." Synopsis Publication – 2017 tralac Annual Conference https://www.tralac.org/news/article/11501-lesotho-s-textiles-apparel-and-footwear-manufacturing-industry.html

Bisung, Elijah, and Susan J. Elliott. 2018. “Improvement in Access to Safe Water, Household Water Insecurity, and Time Savings: A Cross-Sectional Retrospective Study in Kenya”. Social Science & Medicine 200. Social Science & Medicine, 1-8. https://doi.org/10.1016/j.socscimed.2018.01.001.

Bivins, Aaron W., Trent Sumner, Emily Kumpel, Guy Howard, Oliver Cumming, Ian Ross, Kara Nelson, Joe Brown, and Joe Brown. 2017. “Estimating Infection Risks and the Global Burden of Diarrheal Disease Attributable to Intermittent Water Supply Using QMRA”. Environ. Sci. Technol. 51 (13). Environ. Sci. Technol., 7542-51. https://doi.org/10.1021/acs.est.7b01014.

Blum, Annalise, Clair Null, and Vivian Hoffmann. 2014. “Marketing Household Water Treatment: Willingness to Pay Results from an Experiment in Rural Kenya”. Water 6 (12). Water, 1873-86. https://doi.org/10.3390/w6071873.

Boakye-Ansah, A. S., G. Ferrero, M. Rusca, and P. van der Zaag. 2016. “Inequalities in Microbial Contamination of Drinking Water Supplies in Urban Areas: The Case of Lilongwe, Malawi”. Journal of Water and Health 14 (5). Journal of Water and Health, 851-63. https://doi.org/10.2166/wh.2016.258.

Bogetic, Zeljko. 2006. “International Benchmarking of Lesotho’s Infrastructure Performance”. https://doi.org/10.1596/1813-9450-3825.

Cairncross, S., and P. J. Kolsky. 1997. “RE: WATER, WASTE, AND WELL-BEING: A MULTICOUNTRY STUDY”. American Journal of Epidemiology 146 (4). American Journal of Epidemiology, 359-60. https://doi.org/10.1093/oxfordjournals.aje.a009279.

Davis, Jennifer, Alice Kang, Jeffrey Vincent, and Dale Whittington. 2001. "How Important is Improved Water Infrastructure to Microenterprises? Evidence from Uganda." World Development 29, no. 10.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, Vincent Pons, and William Pariente. 2011. “Happiness on Tap: Piped Water Adoption in Urban Morocco”. SSRN Journal. https://doi.org/10.2139/ssrn.1803576.

Ercumen, Ayse, Benjamin F. Arnold, Emily Kumpel, Zachary Burt, Isha Ray, Kara Nelson, and John M. Colford. 2015. “Upgrading a Piped Water Supply from Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched Cohort Study in Urban India”. PLoS Med 12 (10), e1001892. https://doi.org/10.1371/journal.pmed.1001892.

Fan, Victoria Yue-May, and Ajay Mahal. 2011. “What Prevents Child Diarrhoea? The Impacts of Water Supply, Toilets, and Hand-Washing in Rural India”. Journal of Development Effectiveness 3 (3). Journal of Development Effectiveness, 340-70. https://doi.org/10.1080/19439342.2011.596941.

Farole, Thomas and Deborah Winkler, eds. 2014. Making Foreign Direct Investment Work for Sub-Saharan Africa: Local Spillovers and Competitiveness in Global Value Chains. Directions in Development. Washington, DC: World Bank. doi: 10.1596/978-1-4648-0126-6. License: Creative Commons Attribution CC BY 3.0

Finscope 2016. Micro, small and medium enterprises (MSME) survey. Lesotho, 2016. Accessible at: http://www.finmark.org.za/wp-content/uploads/2016/11/finscope-lesotho-pocket-guide_en.pdf

Galiani S, Gonzalez-Rozada M, and Schargrodsky E, 2009. "Water Expansions in Shantytowns: Health and Savings," Economica, London School of Economics and Political Science, vol. 76(304), pages 607-622, October.

Gamper-Rabindran, Shanti, Shakeeb Khan, and Christopher Timmins. 2008. “The Impact of Piped Water Provision on Infant Mortality in Brazil: A Quantile Panel Data Approach”. SSRN Journal. https://doi.org/10.2139/ssrn.1129044.

Gibbon, Peter. 2003. "AGOA, Lesotho ‘Clothing Miracle’ & the Politics of Sweatshops." Review of African Political Economy 30, no. 96.

Gulyani, Sumila; Talukdar, Debabrata; Kariuki, Mukami. 2005. Water for the urban poor: water markets, household demand, and service preferences in Kenya (English). Water Supply and Sanitation Sector Board discussion paper series; no. 5. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/169601468044990414/Water-for-the-urban-poor-water-markets-household-demand-and-service-preferences-in-Kenya

Gwimbi, Patrick; Thomas, Timothy S.; Hachigonta, Sepo and Sibanda, Lindiwe M. 2013. Lesotho. In Southern African Agriculture and Climate Change: A Comprehensive Analysis. Chapter 4 Pp. 71-109. Washington, D.C.: International Food Policy Research Institute (IFPRI). http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/127788

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd, “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,” The Review of Economic Studies 64, no, 4 (1997): 605-654.; and Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score,” Econometrica 71, no. 4(2003): 1161-1189.

Hunter, Paul R. 2009. “Household Water Treatment in Developing Countries: Comparing Different Intervention Types Using Meta-Regression”. Environ. Sci. Technol. 43 (23). Environ. Sci. Technol., 8991-97. https://doi.org/10.1021/es9028217.

Jeandron, Aurélie, Jaime Mufitini Saidi, Alois Kapama, Manu Burhole, Freddy Birembano, Thierry Vandevelde, Antonio Gasparrini, Ben Armstrong, Ben Armstrong, Sandy Cairncross, and Jeroen H. J. Ensink. 2015. “Water Supply

Interruptions and Suspected Cholera Incidence: A Time-Series Regression in the Democratic Republic of the Congo". PLoS Med 12 (10), e1001893. https://doi.org/10.1371/journal.pmed.1001893.

International Monetary Fund (IMF). 2004. Lesotho: Selected Issues and Statistical Appendix. Country Report No. 04/23. Available at: https://www.imf.org/en/Publications/CR/Issues/2016/12/30/Lesotho-Selected-Issues-and-Statistical-Appendix-17150

International Monetary Fund (IMF). 2006. Kingdom of Lesotho: Poverty Reduction Strategy Paper, Prioritization, and Cost Matrix. IMF Country Report No. 06/143, April 2006.

International Monetary Fund (IMF). 2012. Kingdom of Lesotho: Poverty Reduction Strategy Paper— National Strategic Development Plan. IMF Country Report No. 12/102 March 2012.

J-PAL. 2012. "J-PAL Urban Services Review Paper." Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.

Kamlana, Kelebogile Theodora. 2014. "Water security with specific reference to the Kingdom of Lesotho." MSS diss., University of Pretoria.

Kimuyu, Peter, and Steve Kayizzi-Mugerwa. 1998. "Enterprise Response to Deficient Infrastructure in Kenya." Discussion Paper. Institute of Policy Analysis and Research. Nairobi

Kravitz JD, Nyaphisi M, Mandel R, Petersen E. Quantitative bacterial examination of domestic water supplies in the Lesotho Highlands: water quality, sanitation, and village health. Bulletin of the World Health Organization. 1999;77(10):829-836. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2557741

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2011. "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions" 126 (1), 145-205. https://doi.org/10.1093/qje/qjq010.

Kumpel, Emily, and Kara L. Nelson. 2013. "Comparing Microbial Water Quality in an Intermittent and Continuous Piped Water Supply". Water Research 47 (14). Water Research, 5176-88. https://doi.org/10.1016/j.watres.2013.05.058.

Lall, Sanjaya. "FDI, AGOA and Manufactured Exports by a Landlocked, Least Developed African Economy: Lesotho." Journal of Development Studies 41, no. 6 (2005): 998-1022.

Lesotho Bureau of Statistics (BOS), 2009. 2006 Lesotho Population and Housing Census Analytical Report, Volume IIIA Population Dynamics.

Lesotho Bureau of Statistics (BOS), 2011. 2011 Lesotho Demographic Survey Analytical Report.

Leuven, Edwin, and Barbara Sianesi. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing," Boston College Department of Economics, 2003, revised 19 Jul 2012, available online: http://ideas.repec.org/c/boc/bocode/s432001.html.

Martínez-Santos, Pedro (2017): "Does 91% of the world's population really have "sustainable access to safe drinking water"?, International Journal of Water Resources Development, DOI: 10.1080/07900627.2017.1298517 http://dx.doi.org/10.1080/07900627.2017.1298517

Masupha, Tholoana Marcellina. 2007. "Water management at a textile industry: a case study in Lesotho." MSc diss., University of Pretoria.

McDonald, Robert I., Katherine Weber, Julie Padowski, Martina Flörke, Christof Schneider, Pamela A. Green, Thomas Gleeson, Stephanie Eckman, Stephanie Eckman, Bernhard Lehner, and Deborah Balk. 2014. "Water on an Urban Planet: Urbanization and the Reach of Urban Water Infrastructure". Global Environmental Change 27. Global Environmental Change, 96-105. https://doi.org/10.1016/j.gloenvcha.2014.04.022.

Ministry of Finance. Kingdom of Lesotho Poverty Reduction Strategy, 2004/2005 – 2006/2007. Maseru

Ministry of Finance. Kingdom of Lesotho Poverty Reduction Strategy, 2010/2011 – 2011/2012. Maseru

Ministry of Health and Social Welfare - MOHSW/Lesotho and ICF Macro. 2010. Lesotho Demographic and Health Survey 2009. Maseru, Lesotho: MOHSW and ICF Macro. https://dhsprogram.com/publications/publication-fr241-dhs-final-reports.cfm

Ministry of Health/Lesotho and ICF International. 2014. Lesotho Demographic and Health Survey 2014. Maseru, Lesotho: Ministry of Health/Lesotho and ICF International, https://dhsprogram.com/publications/publication-fr309-dhs-final-reports.cfm

Ministry of Trade and Industry (MTI), 2008. The State of Small Enterprise in Lesotho, A study undertaken by SBP for the Ministry of Trade and Industry, Cooperatives and Marketing, available at: http://www.sbp.org.za/uploads/media/FINAL_Lesotho_State_of_Small_Enterprise_Survey_23_10_08.pdf

Morris, M., C. Staritz, and J. Barnes. 2011. “Value Chain Dynamics, Local Embeddedness, and Upgrading in the Clothing Sectors of Lesotho and Swaziland.” International Journal of Technological Learning, Innovation and Development 4 (1–3): 96–119.

Nelson, Kara and John Erickson. Intermittent supply in the context of efforts to improve piped drinking water supply in Latin America and the Caribbean: Lessons from a case study in Arraiján, Panama. IDB Technical note, IDB-TN-1137, January 2017.

Null, Clair, Jorge Garcia Hombrados, Michael Kremer, Robyn Meeks, Edward Miguel, and Alix Peterson Zwane. 2012. “Willingness to Pay for Cleaner Water in Less Developed Countries: Systematic Review of Experimental Evidence”. https://doi.org/10.23846/sr1003.

Onda, Kyle, Joe LoBuglio, and Jamie Bartram. 2012. “Global Access to Safe Water: Accounting for Water Quality and the Resulting Impact on MDG Progress”. IJERPH 9 (12). IJERPH, 880-94. https://doi.org/10.3390/ijerph9030880.

Onjala, Joseph & Ndiritu, Simon Wagura & Stage, Jesper. 2014. Risk perception, choice of drinking water and water treatment: Evidence from Kenyan towns. Journal of Water, Sanitation and Hygiene for Development. 4. 268. 10.2166/washdev.2014.131.

Onjala, Joseph, Simon Wagura Ndiritu, and Jesper Stage. 2014. “Risk Perception, Choice of Drinking Water and Water Treatment: Evidence from Kenyan Towns”. Journal of Water, Sanitation and Hygiene for Development 4 (2). Journal of Water, Sanitation and Hygiene for Development, 268. https://doi.org/10.2166/washdev.2014.131.

Overseas Development Institute (ODI). N.d. Aid for Trade in Lesotho: ComMark’s Lesotho Textile and Apparel Sector Programme. Available at: https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/5948.pdf

Pickering, Amy J. and Jennifer Davis (2012). Freshwater Availability and Water Fetching Distance Affect Child Health in Sub-Saharan Africa. Environmental Science & Technology 46 (4), 2391-2397, DOI: 10.1021/es203177v

Pullanikkatil, Deepa, and Kevin C. Urama. 2011. "The effects of industrialisation on water quality and livelihoods in Lesotho." International Journal of Environmental Engineering 3, no. 2.

Rosenbaum, PR, and DB Rubin. “The central role of the propensity score in observational studies for causal effects.” Biometrika 70, no. 1 (1983): 41-55.; and Rubin, Donald B. “Matching to Remove Bias in Observational Studies”, Biometrics, 29, no. 1 (1973): 159-183.

Spencer, James H. 2008. “Household Strategies for Securing Clean Water”. Journal of Planning Education and Research 28 (2). Journal of Planning Education and Research, 213-24. https://doi.org/10.1177/0739456x08321793.

Staritz, C., and M. Morris. 2012. “Local Embeddedness, Upgrading and Skill Development: Global Value Chains and Foreign Direct Investment in Lesotho’s Apparel Industry.” Working Paper 32, ÖFSE, Vienna

Stoler, Justin, Raymond A. Tutu, and Kiana Winslow. 2015. “Piped Water Flows but Sachet Consumption Grows: The Paradoxical Drinking Water Landscape of an Urban Slum in Ashaiman, Ghana”. Habitat International 47. Habitat International, 52-60. https://doi.org/10.1016/j.habitatint.2015.01.009

Thompson, John, Ina T Porras, James K Tumwine, Mark R Mujwahuzi, Munguti Katui-Katua, Nick Johnstone and Libby Wood (2001). Drawers of Water II: 30 years of change in domestic water use & environmental health in east Africa. IIED.

Toerien, DF, and MT Seaman. 2011. "Ecology, water and enterprise development in selected rural South African towns." Water SA 37, No. 1

United Nations Human Settlements Program (UN-HABITAT). 2015. Lesotho Housing Profile. Nairobi, October 2015.

Waddington, Hugh, and Birte Snilstveit. 2009. “Effectiveness and Sustainability of Water, Sanitation, and Hygiene Interventions in Combating Diarrhoea”. Journal of Development Effectiveness 1 (3). Journal of Development Effectiveness, 295-335. https://doi.org/10.1080/19439340903141175.

Wason, D., Hall, D. (2004) Poverty in Lesotho 1993 to 2002.An overview of Household Economic Status and Government Policy, CPRC Working Paper No. 40. 48 pp. ISBN 1-904049-39-7.

Whittington, Dale, Xinming Mu, and Robert Roche. 1990. “Calculating the Value of Time Spent Collecting Water: Some Estimates for Ukunda, Kenya”. World Development 18 (2). World Development, 269-80. https://doi.org/10.1016/0305-750x(90)90052-y.

World Bank. 2015. Lesotho - Systematic country diagnostic (English). Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/371421468188678379/Lesotho-Systematic-country-diagnostic

Wright, Jim, Stephen Gundry, and Ronan Conroy. 2004. “Household Drinking Water in Developing Countries: A Systematic Review of Microbiological Contamination Between Source and Point-of-Use”. Trop Med Int Health 9 (1). Trop Med Int Health, 106-17. https://doi.org/10.1046/j.1365-3156.2003.01160.x.

10 ANNEXES

A1 UPUW SITE NETWORK MAPS

[Redacted]

A2 PENDING DATA REQUESTS

[Redacted]

A3 EVALUATION BUDGET

[Redacted]

A4 DATA COLLECTION INSTRUMENTS

See attachments:

A4-1. Household Survey, draft tool

A4-2. Household Focus Group Guide, draft tool

A4-3. Industry Key Informant Interview Guide, draft tool

A4-4. Small & Medium Enterprise Key Informant Interview Guide, draft tool

A4-5. Local chiefs Key Informant Interview Guide, draft tool

A4-5. LHLDC Key Informant Interview Guide, draft tool

A5 STAKEHOLDER COMMENTS AND EVALUATOR RESPONSES

All comments below from MCC M&E:

Comment text	SI
Note that MCC's definition of beneficiary only applies to people (who have an increase in their standard of living as a result of an MCC intervention).	Utility referenced as a potential beneficiary, and language in EDR clarified to separate utility from other beneficiaries as traditionally defined by MCC.
The EMC discussed whether we need to explore wastewater issues as a specific potential unintended consequence of the Metolong investment. Questions were raised about whether the issue was identified during due diligence and whether we needed to look more broadly at constraints to production (if we find it hasn't increased).	Following further discussions on this topic with MCC, SI reached agreement to address topic within industry interviews.
Site-specific estimates [of impact] are not necessary [for UPUW Activity].	Understood. UPUW sample size calculations assume that impact estimates are not expected to be site-specific.
Re: assumption that "non-Metolong exposed households experience a level of service comparable to the level they experienced at baseline", has SI gathered any information to confirm this is the case?	Language in EDR clarified to explain how SI will verify this before collecting data.
Do we know why [areas of Maseru not receiving supply from new Metolong WTW] weren't connected [to the new supply] ultimately?	Language in EDR clarified in discussion about Design B and relevant assumptions underlying the design.
Column labeled "n" in Table 11: What does this represent?	It is the expected number of households that make it into the final sample (see Table 10 under "required sample size" – the number of households required has been inflated for matching, as not all households will "find" matches).
The EMC is interested in [water] quality as an important variable for understanding water-related illness. We would want to measure the full range from source, to point-of-use, to stored containers. We don't have targets for improvements, but can we use international standards to calculate needed sample sizes and cost out this addition?	Yes, we can consider WQ measurements to at least compare to WHO recommendations and use in our analysis of diarrheal illness. In contrast to the more expensive laboratory methods used on the Tanzania evaluation, we propose considering a lower-cost alternative. The main drawback (aside from not being gold-standard methodology) is that it won't allow for the calculation of geometric mean of E. coli samples, for instance. There are two main methods we can consider using: an mWater kit, and a Compartment Bag Test (CBT). The mWater kit is aligned to tell you if the sample meets WHO requirements for drinking, plus a high range plate count that will tell you if it's 100x or 1000x the limit (surface water could be in that limit, though in this case we may not expect something that high even from stored water). The CBT gives you a numerical Most Probable Number (MPN) estimate of E. coli colony forming units, but it can be deceiving because if you look at the 95% confidence intervals often the possible result spans more than one WHO risk category. It does have the qualitative benefit of reporting along a scale of 'how bad it is'. It does not go as high as the mWater kit at the high end. Tentative per-test estimates of cost have been obtained; we will consult the companies that produce these kits and develop sample size estimates to incorporate into the RFP for local data collection partners.
What about some of the KIIs you planned to do that didn't pan out because those people didn't feel informed?	Individuals from MTI and LNDC indicated that they did not have any relevant knowledge for our study, and that the Ministry of Small Business Development, Cooperatives, and Marketing was a more relevant informant. This Ministry was previously housed within MTI, but when it became its own Ministry it took individuals who were well informed about SMEs within MTI along with it. The other group who felt uninformed was the Lesotho Chamber of Commerce and Industry, and we take them at their word that they do not have relevant insight for the study. Informants generally feel there is very little to say about the effect of the interventions on SMEs.

Comment text	SI
We're interested in time savings from collecting water but also time saved from being sick and caring for the sick (if the EMC agrees to focus on <5 children's health, then maybe we'd focus on time savings from caring for them when they're sick).	The survey instrument includes questions on sick time and caregiving, both for children under five and other members of the household. Note, however, that the sample size assumptions are based on ERR targets, which relate specifically to time savings from collecting water. This allows us to pick up effects of a certain size, given our assumptions. If the effect is actually bigger (because of additional time savings from other activities), we can also detect this.
Would baseline admin data on water consumption and expenditure help strengthen the match for Design B?	Yes, absolutely. Baseline data for these outcomes would strengthen the match for Design B. Some of this information can be extracted from the WASCO EDAMS database, if WASCO agrees to provide us with customer-level monthly consumption and expenditure data. Baseline data would have strengthened the match for Design A as well, but the baseline conditions for those households is non-piped, so that information cannot be obtained from the WASCO database.
1Can you provide some insight into your proposed sample sizes for the qualitative data? The discussion topics look great for the focus groups but long—would you attempt to cover all in all groups, or split them up somehow?	Generally, focus groups would be allocated in order to achieve thematic saturation, but we balance the breadth of the exercise in our proposed sample sizes against considerations of time, feasibility, and cost. In terms of the discussion topics, the intention would be to cover all in the groups. In our recent experience, these topics can be covered in focus groups that last between 90-120 minutes. Our intention will be to hold focus groups that last no longer than 2 hours. Focus groups will have detailed guides with many probes to guide the facilitators in the discussion, but as with all qualitative interviews the discussion is expected to progress more naturally than a structured survey, and various topics may be addressed in a different order between different focus groups. The main instruction to facilitators is normally that all topics should be addressed, but they don't have to follow the guide in order or ask every single question in the guide, as long as the discussion addresses all topics within the allotted time frame.
Regarding the possibility of using interrupted time series, was there a single point when Metolong came online in an appreciable way?	Based on information gathered during the process evaluation, we are using December 2015 as the date that Metolong came online in an appreciable way.
You mention that the way the enterprise survey was coded prevents you from using it. Are there particular issues or variables that are problematic?	The main issue with the dataset is that most continuous variables are coded into bins with unintelligible labels. Presumably this was a form of de-identification. The value of tracking down the raw data may be limited, given that the sample includes many enterprises that would not be relevant to our evaluation—only 7.7% of the sample (52 households) consists of manufacturing enterprises in urban areas, and some of these may need to be eliminated when the number of employees they have becomes clearer. [Additional information on specific variables with problematic coding were provided to MCC].
Since your power calculations considered the entire Urban Water Activity, but package 1 is also connected to Metolong, I wondered how you were thinking of framing the results—i.e., will you describe everything in terms of Urban and Metolong together?	In general, for package 1 the intervention is conceptualized as Metolong and UPUW activities in combination. However, there are slightly different emphases in Design A and B. Design B is mainly concerned with where the new Metolong supply was delivered – the comparison includes households that had existing connections, to which the Metolong supply was not delivered. In this case, attribution would mainly go to the Metolong program. Even for Design B, however, it will be important to take into consideration where UPUW infrastructure was completed, which we should be able to overlay using the GIS maps once we identify which areas are not being supplied by Metolong. We should be able to discern whether, for example, all UPUW infrastructure was completed in the Design B treatment areas, vice versa, or a mix of the two. The challenge in disentangling the effect of the UPUW infrastructure in Maseru for Design B, is the fact that interventions to repair pipes or install reservoirs are likely to have effects for households beyond their immediate vicinity, as it improves water supply in the system broadly.
Ex post matching on time-invariant observables could be acceptable if home ownership and status are relatively stable in	MCC raises a concern about the validity threat posed by the potential for selective out-migration from certain urban neighborhoods. SI understands this concern to pertain mainly to the relationship between increases in

Comment text	SI
the area. However, if property values have increased in the intervention areas with poorer people moving out and wealthier people moving in, then comparing people just inside and just outside range of the network might raise concerns (a similar concern, which we exchanged about before, is whether people living in the treatment communities who have chosen not to connect to the network after all this time are comparable to those who chose to connect). We wondered whether we could use administrative data to explore changes in property transfers and rentals but a colleague who worked on our land intervention in Lesotho doesn't think the admin records would support this. Since it's unlikely we could use admin data to assess the stability of the housing market, another option might be a listing exercise where we use a knowledgeable community member as a key informant to learn about tenure of ownership, building of additions and which households have young children. If we find that there has been a lot of migration or expansions of homes, then we could still consider doing a household survey but focus it more on willingness to pay issues than trying to get precise measurements of health impacts.	access to infrastructure and housing or rental prices that could selectively drive some households out of study areas. If present, such selective out-migration could result in unobservable differences between treatment households and comparison households, which could bias estimates of impact. A Lesotho Bureau of Statistics (BOS) Demographic Survey from 2011 shows that internal migration generally stagnated between 2006 and 2011, apart from slightly higher migration from highlands and rural areas to industrial areas in the lowlands, consistent with historical trends, and largely driven by the search for employment opportunities. The primary destinations of internal migrants appear to be Maseru and Maputsoe/Leribe. A Lesotho Housing Profile (UN-Habitat, 2015) shows some important differences in dynamics of housing market between low- and high-income residents in Lesotho, and between some other regions in Lesotho. Information gaps include: • Recent migration data not yet available: Data or reports from the 2016 census are not readily available to show trends since 2011, which aligns with the post-intervention period. Even where data is available, it is not disaggregated specifically to study areas of interest. • Data on movers & influence of infrastructure: Not clear to what extent access to infrastructure has influenced these dynamics. No data readily available on composition of internal movers and/or whether movements within urban areas has changed. • Housing market data not readily available: Data on the housing market and rental prices in Lesotho is not readily available. Preliminary research shows property transfers are rare and much of the market operates informally. Formal transactions limited to higher income residents, informative mostly if coupled with selective out-migration in same areas, for which data is not readily available. ***EDR revisions:*** *Broadly, the result of these comments and the ensuing discussion with MCC was the revision of the EDR in the following ways:* • *Qualitative data collection sequenced before quantitative, to better understand dynamics of internal migration and changes in the housing market as related to infrastructure. This includes focus groups with households for all components of the design and key informant interviews with local chiefs and agencies with specific knowledge of relevant land and housing issues.* • *More selectiveness in which sites are included in Design A, now limited to Maseru peri-urban areas and Semonkong only, with all other UPUW sites for Packages 3 through 5 receiving only a customer survey. Design A originally was proposed for Maseru urban and other UPUW sites but removed to account for anticipated results of the qualitative; the final designation of which sites will be included in which components of the design will follow the results of the qualitative data collection.* • *Customer survey added to final EDR, and was not originally proposed, based on related discussion with MCC regarding interest in estimating the current status of outcomes of interest among all customers, while acknowledging limitations in terms of attribution of impact.* *Completing qualitative data collection prior to the start of any quantitative data collection will allow us to further investigate household decisions to take up treatment (or not), and to explore whether migration has been occurring selectively over the past few years from the perspective of long-time residents of study areas and additional key informants. The findings from the qualitative component will be instrumental in determining whether there is a credible counterfactual in each site. If it is found that in some*

Comment text	SI
	sites, an IE is not feasible, we will have made progress toward a PE in each of those sites. *One advantage of this approach is the ability to understand better what kinds of observable and unobservable factors influence exposure to the treatment, which can help determine the feasibility of an IE, and in cases where one is feasible, facilitate more effective matching. As industrial areas, smaller towns, and peri-urban areas may respond differently, these factors might differ by site, further allowing us to strengthen the approach for the context. These may include where individuals found out about their plot (or rental property), and from whom they purchased land (for owners), and how individuals purchased their plot (for owners, e.g. cash or bank loan). Initial research indicates that these answers can differ substantially for lower and higher income households, and this variable could be a useful proxy for pre-intervention socioeconomic status, especially combined with other variables. In the meantime, we will also attempt to obtain information pertaining to internal migration from the 2016 census as soon as it is available.*
If the matching proves viable, we want to make sure we understand exactly what we will be measuring with the methodology proposed, i.e. is this simply going to give us a comparison of connected and non-connected HHs? Is that the right measure for evaluating the impact of our project?/Can we attribute the differences estimated with the proposed methodology to MCC?	This pertains mostly to the counterfactual as defined for Design A, in which the proposed methodology will match connected and unconnected households to identify the impact of the program on households who have connected post-interventions. The project's impact was hypothesized through two potential pathways – either through new connections, or through improved supply for existing connections. For a number of reasons, the counterfactual in the latter group was assessed to be infeasible in all sites other than Maseru. Design A represents the attempt to at least estimate the project's impact on a portion of the intended beneficiaries. Therefore, in areas where Design A would be conducted, the estimated impact will not capture the full MCC project impact. It is, however, a valid counterfactual for a portion of the MCC project, since in those sites, provisions for new connections were included as part of the intervention. In other words, this is a valid counterfactual but only for a portion of the intended beneficiaries, and would be considered a lower bound, of sorts, of the MCC impact in these sites. The exception to this is Semonkong, which was previously served by the Department for Rural Water Service (DRWS), in which an entirely new water network and treatment plant were installed, such that Design A would capture the full impact of the MCC project in that site. In some areas the effect will only partially be attributable to MCC, because of WASCO remediations that were done to address shortcomings of some of the installations; in these cases, SI understands that MCC is still interested in pursuing the approach, understanding that the evaluation will estimate the impacts of the program theory of change, even where remediation was needed.
There isn't much on willingness to pay in the questionnaire now. Do you have a set of questions you've used before that we could easily adapt for use in Lesotho?	WTP questions could be borrowed from similar existing literature if that became part of the scope – however, this would depend on what specific questions would MCC or WASCO have about WTP. For example, potentially this could target areas where take-up has been lower than expected. Or it could be contextualized to expected tariff increases. Discussions between MCC and SI agreed to table this question until the qualitative data collection is completed, and potentially could include WTP questions in the customer survey if relevant.
We want to make sure that the indicators in the household survey will allow us to report on some of our key common indicators for	Revised in household survey.

Comment text	SI
WASH. We're in good shape on most of them but I noticed that our access to improved water indicator includes "protected dug well" as an option and this doesn't appear in the questionnaire. This is also one of the categories of improved water under the JMP definition (along with "unprotected dug well" as an unimproved source) so I propose to include them as separate categories for the relevant questions.	
In terms of water consumption, can we use WASCO data to look at trends in Industrial/Commercial water consumption in terms of cubic meters per month?	Yes, that information has now been obtained and can be analyzed using WASCO administrative data as described in the EDR.
It looks like non-revenue water isn't covered under the evaluation. It isn't one of the indicators specifically referenced under question 7 about short-term and intermediate outcomes but it's related to Q7d and since it is part of the program logic, one of the indicators in the M&E Plan, and specifically referenced in the Compact, we should at least examine the trend. Given SI's data requests to WASCO, I thought it would be covered under the evaluation but want to confirm and document the plan.	Revised language in EDR to explicitly include NRW, and relevant information has since been obtained from WASCO.

SOCIAL IMPACT, INC.
2300 CLARENDON BLVD. SUITE 1000
ARLINGTON, VA 22201

SOCIAL IMPACT

ADVANCING DEVELOPMENT EFFECTIVENESS