IMPACT EVALUATION DESIGN REPORT

GHANA POWER COMPACT - GRIDWATCH

Elaborated by:
Susanna Berkouwer (UC Berkeley)
Pierre Biscaye (UC Berkeley)
Prabal Dutta (UC Berkeley)
Noah Klugman (UC Berkeley)
Matt Podolsky (UC Berkeley)
Steve Puller (Texas A&M)
Jay Taneja (UMass – Amherst)
Catherine Wolfram (UC Berkeley)

Submitted to the
Millennium Challenge Corporation
Original Submission: August 20, 2018

List of Acronyms

DiD	Difference-in-Differences
ECG	Electricity Company of Ghana
EE	Energy Efficiency
EFOT	ECG Financial and Operational Turnaround
ERR	Economic Rate of Return
GDP	Gross Domestic Product
GIS	Geographic Information System
GLSS	Ghana Living Standards Survey
GOG	Government of Ghana
GSS	Ghana Statistical Service
IPP	Independent Power Producer
KWh	Kilowatt Hour
LV	Low Voltage
M&E	Monitoring and Evaluation
MCC	Millennium Challenge Corporation
MiDA	Millennium Development Authority
NEDCo	Northern Electricity Distribution Company
NFOT	NEDCo Financial and Operational Turnaround
PMC	Project Management Consultant
PSP	Private Sector Participation
RD	Regression Discontinuity
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SAM	Social Accounting Matrix
SMEC	SMEC Holdings Limited
TBD	To be determined

Tables and Figures

Table of Contents

1. INTRODUCTION & BACKGROUND

In urban areas of Ghana, more than 90% of households are already connected to the electric grid. In these settings, a primary issue is the reliability of the grid rather than the lack of access to electricity. Between 2012 and 2015, persistent power failures in Ghana negatively affected its economy and gave rise to the term “Dumsor,” meaning “lights off-on” in the local Akan language. According to the most recent World Bank Enterprise Surveys (2013), **61.2% of firms in Ghana saw electricity reliability as a major constraint**, with firms reporting an average of over 700 hours of outages annually, compared to 1.5 hours for firms in the U.S.

Anecdotally, we know that frequent outages constrain the economic well-being of households and small businesses by reducing the benefits from, and discouraging investments in, welfare-improving appliances (like fans, refrigerators, or income-generating assets like sewing machines). To mitigate the impacts of these outages, customers make large investments in substitutes for high quality grid electricity. Investments in backup generators and stabilizers potentially crowd out productive investments. Yet the scope of the grid unreliability problem is not well understood. Governments and utilities are forced to make important operational and investment decisions without accurate information on the frequency, duration, and geographic extent of power outages, the extent to which blackouts and brownouts affect households and businesses (and, indirectly, the local economy), or the most cost-effective ways to improve grid reliability.

MCC’s $498 million Ghana Power Compact (“the Compact”), signed in August 2014, supports the Government of Ghana’s goal of increasing electricity access and reliability while attracting private-sector investment. The Compact is scheduled to be completed by September 2021. The Compact currently includes four projects:

- ECG Financial and Operational Turnaround (EFOT) Project,
- Regulatory Strengthening and Capacity Building Project,
- Access Project, and
- Energy Efficiency and Demand Side Management Project.

The GridWatch team, comprising both engineers and economists, is engaged in two high-level activities:

(1) **Designing and deploying a suite of technologies** to measure outage duration, outage frequency, and voltage quality and to determine outage location on the low-voltage distribution level of the electrical grid at households and firms.

(2) **Designing and implementing an impact evaluation** to measure the socioeconomic benefits of a reduction in outage duration and frequency and an increase in voltage quality as experienced by end-users.[1] The team will focus on evaluating the EFOT Project.

The technologies included in the GridWatch suite and our approach to piloting these technologies is described in our Phase I general work plan, included in Annex A for reference. This report focuses on our approach to the impact evaluation and discusses how the GridWatch technologies will support that evaluation.

At a high level, the team aims to measure primary Compact outcomes, including the frequency (SAIFI) and duration (SAIDI) of outages[2], and voltage level irregularities and then evaluate the socioeconomic impacts of improvements in those outcomes due to the Compact. The research team and MCC developed a set of evaluation questions and outcomes of interest, and the research team, in close collaboration with MiDA, MCC, ECG and other Compact stakeholders, has refined these questions and developed a proposed approach to answer them. This report outlines the initial questions and the evolution to the proposed evaluation design. In short, we are proposing two impact evaluation approaches, one which relies on known differences in reliability across Accra (the "Dumsor Priority Feeder" approach) and another which will measure the socioeconomic impacts of the Compact more directly (the "Compact Quasi-Experimental" approach) but which is necessarily less concrete given that we do not yet know what the impacts of the Compact investments will be on power reliability and/or power quality. The GridWatch devices will be deployed to measure these impacts with a higher degree of granularity and precision than is currently available, which will allow us to adapt the evaluation approach to up-to-date information on Compact impacts. This highlights one of the benefits of deploying GridWatch technologies – the ability to collect rich, highly geographically and temporally precise data on outages and power quality.

---

[1] MCC reports delineate evaluations as follows: "**impact evaluation** is a study that measures the changes in income and/or other aspects of well-being that are attributable to a defined intervention. Impact evaluations require a credible and rigorously defined counterfactual, which estimates what would have happened to the beneficiaries absent the project. **Performance evaluation** is a study that seeks to answer descriptive questions, such as: what were the objectives of a particular project or program, what the project or program has achieved; how it has been implemented; how it is perceived and valued; whether expected results are occurring and are sustainable; and other questions that are pertinent to program design, management and operational decision making," (Albert, et al., 2014).

[2] The GridWatch suite is designed to detect all types of outages, including those originating in the low-voltage, medium-voltage, and high-voltage systems. Using spatial and temporal analytical tools, the technology is able to detect when a large number of devices in a particular area experience an outage at the same time, for example if these are all connected to the same infrastructure. The technology can thereby distinguish between an LV outage and an MV or HV outage with a high degree of confidence.

2. OVERVIEW OF THE COMPACT AND THE INTERVENTIONS EVALUATED

In this chapter, we provide context for the evaluation of the ECG Financial and Operational Turnaround Project by describing the Ghana Power Compact, the project and its activities, and the mechanisms through which the activities are expected to affect outcomes, as described in the MiDA Ghana Monitoring and Evaluation Plan ("M&E Plan"). We also summarize the ex-ante Economic Rate of Return (ERR) that MCC calculated at the Compact level to describe the expected benefits and costs of all Compact activities.

A. Overview of the Ghana Power Compact

On August 5, 2015, MCC signed a $535.6 million compact with the Government of Ghana (GOG). The Governments of the United States and Ghana contributed US$498.2 million and US$37.4 million, respectively. The Compact, the second with the GOG, aims to develop a more reliable and efficient electricity grid to support economic growth and poverty reduction. The Compact went into force on September 2016, and originally contained six projects: (1) the ECG Financial and Operational Turnaround (EFOT) Project to improve the quality and reliability of electricity through reduced outages, thereby improving utility financial health, credit-worthiness and cost effective service delivery; (2) the NEDCo Financial and Operational Turnaround (NFOT) Project to similarly develop NEDCo into a utility that can drive economic growth in Northern Ghana, (3) the Regulatory Strengthening and Capacity Building Project to promote greater transparency and accountability for results in the sector, and enhance evidence-based decision making among sector institutions; (4) the Access Project, which is designed to increase the number of new legal connections for small businesses in selected market centers, (5) the Energy Efficiency and Demand Side Management Project, which aims to reduce strain on the country's generation capacity by implementing four demand-side energy efficiency activities, and (6) the Power Generation Sector Improvement Project, which will support measures aimed at opening up Ghana's power sector to make it attractive to private investors for additional generation capacity. The Compact is scheduled to close on September 2021. Following the Mid-Term Review of the Program in February 2019, the NFOT and the PGSI Project were de-scoped, and the funds reallocated to the other projects, particularly the EFOT Project.

According to the MiDA Ghana M&E Plan, "[t]he Program Objectives are to: (i) increase private sector investment and the productivity and profitability of micro, small, medium and large scale businesses; (ii) increase employment opportunities for men and women; (iii) raise earning potential from self-employment; and (iv) improve social outcomes for men and women. Prior to the achievement of these high level objectives it is envisaged that a set of hierarchically lower level but interrelated objectives (outcomes) of power generation, distribution and access will be achieved. These outcomes are expected to trigger (cause) the aforementioned program objectives (effects) including (a) increased availability, reliability and expansion of cost-effective generation, and (b) reduced "hidden costs" and increased (re)investment and expansion in the power sector."

B. Overview of the EFOT Project and Current Implementation Plan[3]

The ECG Financial and Operational Turnaround Project's objective is to improve the quality and reliability of electricity through reduced outages and cost-effective service delivery by ECG, reduce aggregate technical, commercial and collections losses, and to ensure ECG can serve as a creditworthy and credible off-taker under power purchase agreements. The project objectives will be achieved by reducing implicit subsidies (created by losses, underpricing, and under-billing) and ensuring cost-recovery and reinvestment in the distribution sub-sector through introduction of Private Sector Participation (PSP) in the governance and management of ECG, and through infrastructure and foundational investments designed to reduce losses and improve service quality.

There are five Activities under the EFOT Project.

(1) Private Sector Participation (PSP) Activity: This Activity will provide support for the design and execution of an acceptable ECG PSP transaction. Funding for this Activity is intended to cover the following interventions:
   - Transaction Advisory services to, among others, advise the Government on the design and implementation of an international tender to select an acceptable ECG PSP Provider, supporting the Government until financial close;
   - Assistance with targeted communications strategy, outreach, and consultation to gain the support of stakeholders;
   - Consultation with management and employees of ECG to gain support for PSP; and
   - Consultancy to design the institutional set-up for the acceptable ECG PSP transaction.

(2) Modernizing Utility Operations Activity: This Activity is designed to introduce modern tools to ECG, build the capacity of ECG's staff to use the tools, and provide a robust communication network for ECG. Specific interventions include:

[3] This section draws heavily from the MiDA Ghana M&E Plan, with additions to reflect current activities.

- Installation of a Geographic Information System (GIS) based distribution management system, grid digitization, and customer census to record and store basic data;
- Installation of an Enterprise Resource Planning (ERP) system and integration with existing enterprise applications for the purpose of facilitating the flow of information between business functions within ECG and managing connections to outside stakeholders;
- Provision of technical assistance to strengthen Project implementation through the hiring of qualified advisors;
- Upgrade of data center and communications network to assist ECG in creating a data center compatible with current industry standards and to better manage the network;
- Institutionalizing gender responsiveness to support gender auditing, development of a gender policy at ECG and support activities for strengthening institutional capacity of ECG to implement a gender policy and enhance the capacity of female employee associations through knowledge sharing, networking, and the development of internships and mentoring to university students in science and technology, particularly women; and

(3) Reduction in Commercial Losses and Improvement of Revenue Collection Rates Activity: The Activity addresses commercial and collection losses. Specific interventions to address commercial losses include:

- Strengthening the loss control program by providing the loss control units at ECG with the means (training, tools, and equipment) to more effectively reduce commercial losses.
- Installation of automated meter readers at special load tariff service locations and on selected non-special load tariff service locations in the ECG Target Regions as well as installation of metering at critical nodes of the distribution system in the ECG Target Regions to provide ECG the ability to identify and monitor where technical and commercial losses are occurring.
- Dependent on the availability and release of contingency funds:
    - Creation of service connection standards and normalization of existing services to update existing standards with a new design; train ECG personnel to enforce the new standards; and repair and upgrade non-conforming services .
    - Replacement of legacy credit meters with pre-payment meters in the ECG Target Regions to improve collection efficiency and timely closing of monthly financial statements.

(4) Technical Loss Reduction Activity: The interventions under this Activity will result in lowering thermal losses for the primary and secondary distribution systems in the ECG Target Regions. Specific interventions include:

- Updated distribution design and construction standards based upon currently accepted best practices to ensure compliance with international best practice for low loss and economical designs.
- Low voltage (LV) bifurcation and network improvements to reduce the length of the low voltage circuits to ensure they do not exceed a length that affects the quality of service and a technical loss threshold.
- Installation of two (2) bulk supply points with feeders to existing primary substations to ease overloading based on the current demand forecast and to avoid rolling brownouts.
- Installation of seven (7) primary substations with interconnecting sub-transmission links and medium voltage offloading circuits to help reduce technical losses and avoid extended outages caused by failures or maximum capacity reached at geographically adjacent substation.
- Dependent on the Distribution Master Plan and Capacitor Placement Study:
    - Introduction of reactive power compensation for primary substations to optimize power levels at 33/11 kV substations.

(5) Outage Reduction Activity: The Outage Reduction Activity will improve service and increase sales. The interventions under this Activity include:

- Installation of outage management system to identify outage locations and causes and serve to reduce outage frequencies and durations.
- Provision of specialized vehicles, tools, and equipment required for fault clearance and restoration of outages in the ECG Target Regions.
- Dependent on the Distribution Master Plan and Capacitor Placement Study:
    - Sectionalizing study of ECG target regions and automation of medium voltage networks and system control and data acquisition expansion to locate sectionalizing devices in the 11 kV network to reduce the geographic area affected by outages.

Unreliable power can be a major constraint to growth of businesses in Ghana. To effectively support the growth requirements of the economy, the Compact must address key challenges in the distribution of power. ECG, the leading distribution company in Ghana, incurs high technical and commercial losses, which deter private investment and may lead to low quality of service and higher cost of electricity service provision.

The EFOT Project is designed to achieve short, medium, and long-term goals. In the short term, the Project will improve the financial sustainability of ECG by reducing commercial losses, improving billing and collection, ensuring cost reflective tariffs, implementing regular automatic tariff adjustments, and improving financial management generally within ECG. Ghanaians in the ECG Operational Area should experience improvements in power reliability and quality under EFOT through the reduction of unplanned outages and distribution technical losses, which will result in an improved voltage profile. Similarly, improvements in outage response time, reductions in the cost per kWh of electricity generation, and reinvestment and maintenance in capital expenditures will transform the operations management of ECG. Consequently, ECG will regain the ability to adhere to commercially agreed time limits for payments of bills to IPPs. However, vital institutional capacity building activities must complement the PSP option to achieve planned improvements in short term outcomes. In addition, sufficient infrastructure investments, including private sector participant contributions, are needed to achieve loss reduction targets within ECG.

The medium term outcomes include improving the financial viability of ECG, positioning ECG as a credible off-taker, improving satisfaction among ECG customers, and enhancing investment capacity within ECG to achieve sustainable service delivery. Tariff reductions and improvements in power reliability and quality will reduce customers' reliance on diesel and petrol generators, resulting in an increase in electricity consumption from ECG.

In the long term, more reliable and higher quality electricity service is likely to lead Ghanaian businesses in the ECG Operational Area to increase investments, expand their sales, and hire more employees. Achievement of this outcome under the EFOT Project, however, depends on coordinated actions from all other Compact program components, including sector expansion to meet demand. Deterioration in conditions under which businesses operate in Ghana (e.g., inflation, interest rates and credit availability, exchange rates) and other factors outside of the Compact scope may impact the likelihood of investment and deter private sector investment in the power sector.

MCC engaged the GridWatch team to pilot a novel low-cost technology designed to measure outages, and power reliability and quality more generally. Many of the EFOT Project activities are designed to improve the reliability and quality of electricity service in the ECG Operational Area, and particularly in the ECG Accra East and Accra West Regions. For example, the Outage Reduction Activity is designed to both limit the frequency of outages (SAIFI) and minimize the duration of outages (SAIDI) when they occur. Based on our conversations with SMEC, the PMC, we also understand that some of the components of the Technical Loss Reduction Activity, such as the LV bifurcation and network upgrades, are likely to lead to more consistent voltage quality and fewer outages both across and within customers. The team aims to provide highly geographically and temporally precise measurements of outages, restorations, and voltage quality and use these as the basis for a rigorous impact evaluation rather than ECG SAIDI and SAIFI data which have been found to suffer from significant data quality issues.

The following table summarizes the EFOT project goals, as well as general Compact goals and the corresponding approaches that we envision using to measure outcomes.

Table 1: Concordance between Program Goals and Evaluation Measurement Approach[4]

Program Goal	Measurement Approach
Intermediate Outcomes (technical)	
Distribution system losses, technical losses, percentage of pre-payment customers	ECG data on technical losses*
Sum of durations, in customer-hours, of all customer interruptions in a quarter / Total number of customers connected to network in the same quarter in Accra East and Accra West. (SAIDI)	GridWatch measurements
Sum of customer-interruptions in a quarter/ Total number of customers connected to network in the same quarter in Accra East and Accra West. (SAIFI)	GridWatch measurements
Commercial losses	ECG data on bill repayment rates
Objective Level Indicators	
Percent of firms citing electricity as a major obstacle to doing business	Enterprise Survey and existing GSS data on numbers and types of firms by area**
Average value of sales losses due to electricity outages as percentage of revenue	Enterprise Survey
Average annual kWh of diesel generation consumed by registered firms as a percentage of total kWh electricity consumed	Enterprise Survey, including questions about generator ownership and estimates of operating times
(i) increase private sector investment and the productivity and profitability of micro, small, medium and large scale businesses; (ii) increase employment opportunities for men and women; (iii) raise earning potential from self-employment; and (iv) improve social outcomes for men and women	Household and Enterprise Surveys and existing GSS data

* While the GridWatch team will try to coordinate with ECG, MiDA and other evaluators to understand how these outcomes are being measured, and look for correlations between these outcomes and the ones on which we are focusing, they are beyond the scope of the impact evaluation.

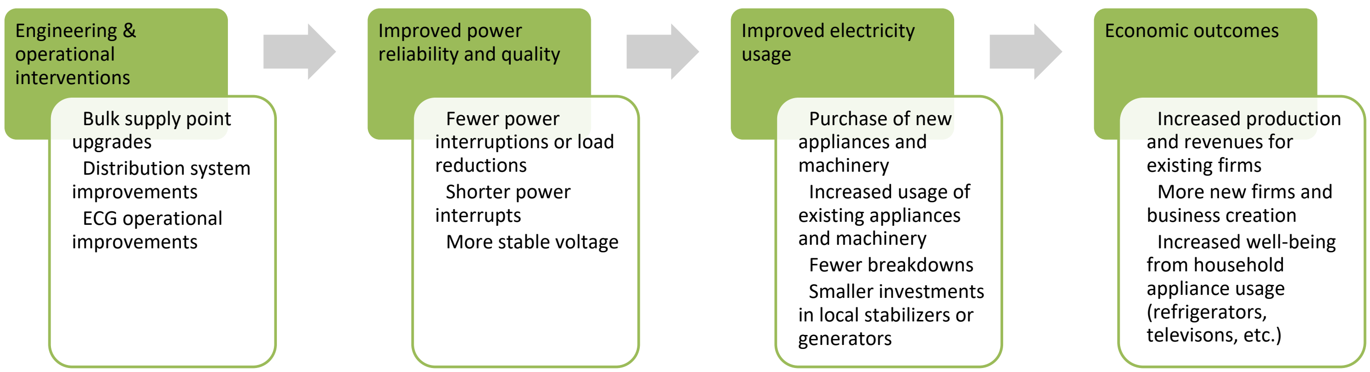
** Some measurement approaches appeal to revealed preference logic. For example, a good way to understand whether firms see electricity and outages specifically as an obstacle to doing business would be to observe differences in the number of new firms that, all else equal, emerge in areas with low outages as compared to areas with high outages.

[4] Program goals are as described in Annex 1 of the Monitoring and Evaluation Plan.

C. Theory of Change

To better understand the mechanisms by which the MCC investment program is supposed to foster change and deliver positive outcomes to beneficiaries in Ghana, the GridWatch team spent time discussing the relationships between specific investments and outcomes with MCC and MiDA. We also reviewed the ERR. To develop the causal chains that can ultimately be tested in the impact evaluation, the team created the following figure to illustrate the causal relationships that could be traced between Compact activities and measurable impacts.

Figure 1: Theory of Change



The ERR also identified a virtuous circle, where institutional changes at ECG together with Compact investments lead to improved operations by reducing commercial and technical losses, generating higher net earnings, and higher levels of retained earnings, which are then applied to the operational sustainability of the system and lead to improved service quality. This describes a potentially important dynamic impact of Compact investments, although one that will be particularly hard to measure quantitatively. We intend to analyze the first element of the chain – a relationship between improved service quality and bill payment rates.

D. Cost Benefit Analysis & Beneficiary Analysis

MCC's ERR analysis applies several methodologies to estimate likely Compact impacts, which are expected to be roughly 3-4% of Ghana's GDP. The analysis focused on the economic costs of power outages by quantifying their impacts on productive activities. For example, MCC analysts developed a fixed-price multiplier from an existing Social Accounting Matrix (SAM) and then applied the multiplier to estimates of the increases in electricity likely to be generated by the Compact. Analysts also developed estimates of the cost of outages in terms of dollars per kilowatt-hour for firms by estimating reductions in value added due to the costs of operating a backup generator and losses of value added due to the temporary cessation of economic activity. They then multiplied these estimates by the expected increase in kilowatt-hour production generated by Compact investments.

The ERR notes that, “the primary beneficiaries of the EFOT Project are consumers of electricity engaged in productive activity in the ECG Target Regions. These regions generate over 22 percent of the gross domestic product (GDP) of Ghana and represent more than 23 percent of ECG’s total customers. The proposed interventions are expected to reduce losses in added value in terms of lost income to the owners of businesses (or owner-operators as the case may be for informal activities) and wages because of service disruptions.”

The impact evaluation will provide estimates consistent with the ERR, particularly the second analytical approach, as we will measure the effect of reduced outages on firm operations. We will estimate direct impacts of outages on sales and profitability, and attempt to understand the mechanisms through which reduced outages lead to these changes by measuring the impacts on firms’ decisions to invest in electricity-using assets or to reduce investments in substitutes for grid electricity, such as generators.

We note several points of departure from the analysis in the ERR. First, while the ERR focused on outages, we intend to measure the impacts of voltage fluctuations as well. Low or fluctuating voltage can lead to economic losses by destroying electrical appliances, and the fear of voltage issues can prevent firms or households from investing in profit or welfare-enhancing equipment. To consider a simple example, a seamstress may opt to use a manual-powered sewing machine if he or she fears that voltage issues will destroy an electric machine. With an electric machine, however, the seamstress might be able to produce more products per hour. Because the GridWatch technologies can measure the voltage quality and frequency experienced by the customer, we are able to design a study to measure the impacts of unstable voltage as well as outages.

As a second point of departure, the ERR focuses on short-run losses at existing firms. We will measure the impact of outages on the number of new businesses created. For example, if all new firms in a particular industry are required to purchase a backup generator to survive, this essentially serves as a tax to being in business, which, on the margin, will reduce new business creation.

Third, we will also measure lost consumer surplus at households. Anecdotally, we have heard of households spending significant amounts of money on electrical equipment to protect household appliances against surges and voltage fluctuations. Households may also experience significant welfare reductions from power outages, for example through the expiration of foods or medications that require refrigeration.

E. Literature Review

i. *Summary of the existing evidence*

Several existing working papers study the effects of the Dumsor crisis on socioeconomic outcomes in Ghana. Hardy and McCasland (2019) find that one additional blackout day among small firms is associated with an 11% decrease in weekly profits, even though firm owners respond to blackouts by shifting production to non-blackout days. Abeberese, Ackah, and Asuming (2017), in a policy brief commissioned by the International Growth Center (IGC), found that power outages during the 2012-2015 Dumsor crisis had a significant negative impact on productivity among small Ghanaian firms, and that one extra day of outages is associated with a 1% reduction in labor productivity and total factor productivity (TFP). Their analysis, however, relies on the assumption that all firms within a given city were equally likely to experience a power outage at any particular time. We are not aware of any additional rigorous quantitative evidence on the effect of the Dumsor crisis on socioeconomic outcomes in Ghana. Beyond Ghana, there is a small literature on the socioeconomic effects of poor reliability. Allcott, Collard-Wexler, and O'Connell (2016) and Fisher-Vanden, Mansur, and Wang (2015) study firm responses to scheduled rolling blackouts in India and China, respectively, and find that the negative effect of outages on firm productivity is mitigated to some extent by a firm's ability to store inputs over time and reallocate production to non-outage hours.

ii. *Gaps in literature*

The papers in the Ghanaian context, primarily Hardy and McCasland (2019) and Abeberese, et al. (2017) relied on the assumption that power outages were distributed randomly across firms and households. Any claim around causality relies on the assumption that firm blackouts are randomly assigned conditional on a time control, which is unlikely in this setting given anecdotal evidence that blackouts are often concentrated in areas with limited political or economic influence. It is up for debate whether this assumption would hold in the Ghanaian context, where power outages are largely at the discretion of ECG, either directly (through rolling blackouts) or indirectly (through differential investment in low voltage infrastructure). There is still no rigorous, causally identified estimate of the effect of improved reliability in Ghana on socioeconomic outcomes. Allcott, et al. (2016) and Fisher-Vanden, et al. (2015) both focus on short-run outcomes and are not able to address the possibility that, for example, persistent outages discourage new firms from entering a market in the first place. In addition, neither paper provides evidence on the effect of more frequent and longer duration outages, nor do they study unannounced electricity outages caused by infrastructure failures, which could have a higher impact than the effect they identified due to firms' inability to prepare for unannounced outages by storing inputs or shifting working hours. Due to these gaps in the literature, we believe that a rigorous, quantitative study on the effects of both unannounced low-voltage outages and high-frequency rolling blackouts in the Ghanaian context would be a meaningful contribution to the literature.

iii. *Policy relevance of the evaluation*

Over the next several decades, almost all of the increase in energy demand worldwide is expected to come from developing countries. In response to rapidly growing electricity demand in Sub-Saharan Africa, particularly in its cities, electricity companies in the region will be evaluating approaches to improve the reliability of their urban electricity networks. In addition, utilities, as well as international donor agencies, will want to know how to do so most effectively. However, there is limited information on both the effect of network investments on reliability and the effect of improved reliability on socioeconomic outcomes. This impact evaluation intends to address both issues. It will provide evidence on the effectiveness of transformer upgrades on reliability, and it will then provide both short-run and medium-run estimates of the effect of these reliability improvements on household well-being and firm performance. With better estimates of these effects, public actors can make more informed decisions about which infrastructure investments may generate the largest economic returns.

3. EVALUATION DESIGN

A. Evaluation Questions

This section outlines the high-level questions that the impact evaluation seeks to address. Subsequent sections describe the approach and specific methodology that we will employ to develop answers.

The Phase 1 Task Order Statement of Work, developed collaboratively by MCC and the research team, identified seven questions as a starting point for the impact evaluation. We list them here with details on how we intend to answer the questions through the evaluation or explanations for why we have decided not to include them. Our comments are informed by subsequent discussions with a number of stakeholders, including MiDA, MCC, ECG, government officials and Mathematica, although we do not mean to implicate any other organization in our conclusions.

1. What are the economic and socio-economic benefits of access to reliable power on customers, including households and enterprises? How are these benefits distributed?

Observations: These questions will form the core of the evaluation. Customers primarily consume electricity through electrical appliances, such as refrigerators or televisions, and machinery. For customers to benefit from improved power reliability, we would first expect to see an improvement in their electricity usage through these channels, for example by operating their appliances for more hours, by purchasing better quality equipment, or by spending less money on replacements for reliable grid power, such as backup generators or voltage stabilizers. We will analyze whether and how reliability improvements affect electricity use of firms and households through all of these channels to understand the mechanisms through which reliability may be affecting household incomes and firm profits.

For households, we are primarily interested in changes in income. For example, improvements in electricity reliability could result in increased productivity of home-based self-employment, increased labor force participation, improved health outcomes, which could in turn lead to improved personal productivity, expenditure switching that induce productivity or income effects. In addition, we will also analyze impacts on a variety of socioeconomic variables (some of which, such as health and education, may be considered as outcomes in their own right), as reliability may lead to improved incomes through these channels. For example, changes in educational outcomes, changes in household saving (investment behavior) may subsequently affect household income. We plan to estimate the relative importance of these channels by also evaluating the impacts of electricity reliability changes on these various intermediate socioeconomic outcomes. This will allow us to explore which channels may be most important for driving any observed changes in income we observe, and to better understand the mechanisms through which electricity reliability may affect incomes and economic growth in the long run.

For firms, we are primarily interested in analyzing impacts on economic outcomes such as incomes and profits. In addition, our goal is to understand the channels through which reliability may affect business performance, such as capital investments, hours of operations, use of labor, and productivity. We intend to study primarily small firms as these are likely to have a direct economic impact on low-income households.

Based on discussions with ECG, MiDA and the PMC, we have also decided to evaluate the economic and socio-economic benefits of access to higher quality power (for example, power with fewer voltage surges). We have learned that some of the investments are targeted at power quality more than reliability. Also, we have learned that customers are negatively impacted by poor power quality.

2. What happens within households and businesses when the power goes out? When it comes back on?

Observations: These questions address the very short-run decisions that customers make in response to outages. Answering them could reveal, for example, how long it takes for consumption to return, after an outage, to the customer-level norm for a given day of the week and time of day. This approach would require access to high frequency electricity consumption data to understand these very granular, short-run decisions. We do not think those data exist for a representative set of customers in ECG's service territory. Instead, to address these questions, we can observe how short-term decisions on individual outages accumulate to longer run outcomes in electricity consumption and, ultimately, in the economic and socio-economic impacts identified in question 1.

3. How long does it take households and businesses to make lumpy investments in power consuming technology when the reliability of the grid improves?

Observations: We intend to include retrospective survey questions that ask customers to describe when they invested in new appliances or machinery. By correlating their answers with major changes in power outages and quality, we will be able to answer this question.

4. What is the Program's overall impact on the profitability and productivity of enterprises? What are the mechanisms or channels through which these impacts occur?

Observations: The methodology described below aims to develop unbiased estimates of power reliability elasticities and power quality elasticities – the percentage change in key Compact outcomes for a given percentage change in power reliability and a given percentage change in power quality. We will then be able to use the elasticity estimates E_i, together with GridWatch measurements of changes in power reliability and quality, to calculate the socioeconomic gains from the Compact. Specifically, we will aim to run the following calculations:

$$\%\Delta SEO_i = E_i \times \%\Delta PQR \qquad (1)$$

Where SEO_I is socioeconomic outcome i, E_i is the elasticity estimated using the methodologies described below ($E_i = \frac{\%\Delta SEO_i}{\%\Delta PQR}$) and PQR is the measured power quality or power reliability. Note that we will use GridWatch technologies to measure the effects of the Compact investments on primary outcomes, such as power quality. For firms, the main socioeconomic outcome of interest is profits, but we will also measure impacts on costs, revenues, and aspects of business operations as possible channels through which reliability may affect profits.

We have decided that it is also important to answer the household analog of this question: **What is the Program's overall impact on household incomes and well-being?** Among households that improve their usage of electrical appliances, we are interested in quantifying the extent to which this usage affects income as the primary outcome variable, but also in evaluating impacts on other socioeconomic outcomes such as health, labor force participation, home-based self-employment, and education which may be important channels through which improved reliability affects incomes.

5. To what extent do small, medium, and large firms respond to more reliable, accessible, and/or higher quality power by:
    a. Expanding or intensifying production
    b. Expanding employment
    c. Investing in expanded plant or other fixed assets and/or different production technologies reliant on electricity

Observations: Among firms that improve their usage of electrical appliances and machinery, we will quantify the extent to which this usage affects economic outcomes such as production, employment, and complementary investments. We will evaluate these outcomes directly by including survey questions for firms. We may not be able to obtain separate estimates for large firms in particular given our design methodology and the limited number of large firms likely to exist in the districts targeted for improvements under the EFOT Project. Note that although the question asks about "accessible" power, we do not intend to address access issues, as we understand that this will be covered by a companion evaluation that is planned for the Access Project.

6. How do small, medium, and large firms respond to higher electricity tariffs? How does the change in tariff affect the profitability and productivity of businesses?

Observations: We understand that MiDA is embarking on an analysis along these lines based on recent tariff changes. We will coordinate with MiDA to understand the results of their analysis, as it could be interesting to identify how large a change in power reliability must be to generate comparable changes in electricity consumption or profitability, relative to the impacts of the recent tariff change. The main way that current tariffs might affect our ERR analysis could be on the cost side, as improved reliability may result in increased electricity consumption, and to the extent that tariffs do not reflect costs this would increases losses to the electricity distributor/provider. We will account for any such increase in costs in the cost benefit analysis.

7. Are customers notified ahead of schedule of their outages? What is the differential impact on customers between known and unknown outages? What is the impact of known versus unknown outages on customer relations?

Observations: One of the advantages of the GridWatch App is that it displays a real-time map of outages occurring across Accra.

Figure 2: Live Outages Displayed within the GridWatch (DumsorWatch) App

Note: The outages and the alert displayed in this image are simulated to preserve privacy from the ongoing pilot study.

With wide deployment of the App, customers will know, for example, that there is an outage at their home before they leave their place of employment. We can also evaluate the share of outages recorded by GridWatch that were pre-announced through the various ECG channels. We are wary, however, of trying to disentangle the impacts of known and unknown outages as we do not yet know what share fall into each category. If the vast majority of outages are unknown, for example (as is the case for unplanned outages, which make up the majority of outages), then we will not be able to measure the impact of the known outages. To evaluate the impact of outage reduction on customer relations, we would like to add the following evaluation question: **How do reduced power outages and improved power quality affect customer bill payments?** Residential and commercial customers may increase both the speed with which they pay their bills and the share of billed amounts they pay, if outages fall and power quality improves. This will lead to lower collection losses for ECG. Previous work by one of us (Dzansi, Puller, Street and Yebuah-Dwamena, 2019) documented a negative relationship between bill payment and outage exposure using billing data from ECG. We would like to extend this analysis with updated data.

We describe the methodology we will use to answer these questions in the next subsection. In the process of answering them, we will be able to document other important results listed below.

M1. How does being assigned on a high-priority feeder line affect power reliability?

Observations: ECG has assigned "exempt" status to feeder lines that contain, for example, hospitals, government buildings, and military bases, where the cost of low reliability is relatively high. Exempt status means that these feeder lines are generally not subject to load shedding. In addition, ECG district engineers report that other feeder lines are considered to have "sensitive" status based on specific infrastructure or customers served by those feeders. Sensitive feeders are subject to some load shedding but are prioritized over ordinary feeders when load shedding allocation decisions are made, and are also prioritized in responses to outage issues. Exempt and sensitive feeders can together be considered "high-priority" feeders, though the designation of "sensitive" feeders is less formal and the degree to which these feeders are prioritized over ordinary feeder lines is unclear. The GridWatch technology will allow us to quantify the degree to which these high-priority designations improve reliability for the customers connected to these feeder lines. This will address the question about the distribution of outages posed above in 1.

M2. How does the installation of network infrastructure affect power reliability and quality?

Observations: The line bifurcation project will result in both the injection of new transformers and the replacement of aging and failing existing infrastructure. In Accra, the root causes of low-voltage power quality issues are not broadly known. We are interested in measuring any improvements in power quality in the immediate short-term for services in the grid areas addressed by the line bifurcation project. A large improvement in power quality for these services would indicate that low-voltage power quality issues stem from network configuration problems rather than issues with actual service connections themselves. The team will use GridWatch technology to monitor voltage quality, grid frequency stability, service interruption frequency, and service interruption duration for services located below the planned grid improvements. The hypothesis is that services under the network infrastructure improvements will both have more stable voltage and frequency and may experience fewer and shorter service interruptions.

B. <u>Evaluation Design Overview</u>

We seek to estimate the effect of decreases in SAIFI and SAIDI and increases in voltage quality on firm outcomes and household well-being. To do this, we first need to establish a source of quasi-random variation in SAIFI, SAIDI, and voltage quality across different firms and households. There are several reasons to believe that power quality and reliability may be endogenously related to socioeconomic outcomes. For example, ECG may consider the economic importance of local businesses or the wealth of local households in determining which feeders to supply during load shedding events. Moreover, if ECG has installed and maintained the electric infrastructure more comprehensively in areas with higher levels of economic activity, more remote customers may have both lower voltage and lower socioeconomic outcomes. If this is true, simply analyzing cross-sectional variation in socioeconomic outcomes could lead to a biased estimate of the impacts of power reliability and quality. To avoid this spurious correlation, we plan to implement two approaches to identify plausibly exogenous variation.

The **Dumsor priority feeder approach** takes advantage of the fact that ordinary households and small businesses may happen to be connected to a feeder that serves a priority customer while a statistically equivalent (in terms of observed characteristics) group of neighboring households and small businesses happen to be connected to feeders that experienced many more hours of outage during the Dumsor period (which is reported to have begun as early as 2012, and ended in 2016). Based on conversations with ECG, we have identified 31 priority feeders in Accra East & West that are designated as “Exempted Essential Feeders” and experience significantly less load shedding than other feeders; they are intended to be exempt from all but the most severe load shedding. This designation would have been most relevant during the Dumsor crisis period. Exempt feeders generally serve critical government or public infrastructure, such as hospitals, police headquarters, and ministry buildings.

The key to our approach to study the impact of power reliability is that households and firms that are statistically similar in terms of observed characteristics received significantly different load shedding during Dumsor, and may continue to experience different electricity reliability. We have analyzed billing data for a subset of post-pay customers in Accra East to determine that there are approximately 1,300 residential customers and 500 small business customers connected to the priority feeders. These customers will be matched to statistically comparable customers who are not served by priority feeders. Our preliminary proof-of-concept investigation focused on Accra East suggests that average monthly electricity consumption prior to the Dumsor period is similar among residential customers on priority and non-priority feeders.

We aim to establish “treatment” (i.e., priority feeder) and “control” (i.e., non-priority feeder) groups that are comparable along a number of socioeconomic dimensions in addition to monthly electricity consumption. To do this, we will identify sites around the boundaries of areas served by priority and ordinary feeders. Because the groups of households or businesses across these sites are neighbors, they are expected to be statistically equivalent in terms of observed characteristics. We will test for balance across treatment and control sites using socioeconomic data from the Ghana Statistical Service (details on the site selection process are below in the description of the study sample). If groups are not perfectly comparable, we will consider using matching approaches to ensure pre-treatment balance. While we can construct two groups of households and small businesses that are similar in many respects, the feeder status impacted power reliability during the Dumsor period differently for these two groups. In general, customers on non-priority feeders experienced an average of 120 hours of load shedding per month in 2015 while customers on priority feeders experienced an average of only 19 hours of monthly load shedding.

An advantage of the Dumsor priority-feeder approach is that it relies on known differences in outages across customers. However, it does not directly measure the impact of Compact investments. The second **Compact quasi-experimental approach** will. In particular, we will use the precise and granular measures of SAIDI, SAIFI, and voltage quality collected by the GridWatch technologies to identify a source of quasi-experimental, Compact-period variation. For example, we may learn that customers on priority feeders continue to experience fewer interruptions, perhaps because the equipment on their feeders did not experience as much stress from continued cycling during the Dumsor period.

Another possibility is that the placement of new infrastructure will allow us to implement a Difference-in-Difference (DD) design. The PMC, in its role as engineering designer and construction supervisor under the EFOT project, has agreed to conduct line bifurcation by adding a transformer in settings where 1) the distance between a transformer and a grid endpoint exceeds 500 meters, or 2) the load factor on any transformer exceeds 70%. We propose a Difference-in-Difference (DD) design, based on comparing outcomes in grid service areas around injection sites with outcomes in grid services in comparable control sites. We randomly select "control" service areas in the zone 300-600m from any transformer, in parallel to the first criterion for line bifurcation site selection based on distance to nearest transformer, which is likely highly correlated with power quality. We will estimate the impact of line bifurcation transformer injections on quality and reliability of power by using GridWatch sensors to measure resulting changes in SAIDI, SAIFI and voltage quality in areas that receive the injections and in the randomly selected similar control sites.[5]

The econometric identification assumption here does not require strict adherence to the decision rule. Instead, there are two testable requirements. First, the control and treatment sites must satisfy the parallel trends assumption, that is, control and treatment sites had similar levels of reliability over time prior to the line bifurcation. We have been collecting GridWatch outage data in the relevant sites for more than a year now, and using these data we have tested and confirmed that indeed there are no detectable differences in power quality across the treatment and control sites. Second, the sites must not have been chosen on any factors correlated with the outcomes of interest, which in our case are economic indicators. This would be a concern if, for example, the PMC incorporated economic indicators such as income across all the different transformers when making their decision about where to inject new transformers. Through extensive conversations with the PMC, we feel confident that this is the case. Our understanding is that the PMC did not even have access to economic information when they were making their injection decisions. Further, we will be able to assess whether project completion is correlated with economic outcomes by comparing results where we use an indicator variable, based on the original sites targeted for line-bifurcation, to instrument for project completion (essentially estimating the local average treatment effect coefficients) to results where we do not instrument with the original site indicator variable. We therefore feel

[5] For the Compact quasi-experimental approach to be most effective, the infrastructure placement should comply with these discontinuous cutoffs. In order to evaluate the extent to which contractors supervised by the PMC are complying with designs, we will work closely with the PMC to understand their compliance and verification processes. We may also decide to send our own field staff to observe randomly selected construction projects.

confident that the DD approach will be rigorous. We feel confident that any differences in power quality or economic outcomes between control and treatment sites will be caused by the transformer injection activities.

Our preference is to use the DD design as the Compact quasi-experimental approach, so much of the remainder of this section describes its implementation details. If that is not possible, or if the data from the GridWatch devices suggest that there are larger plausibly exogenous differences in outages, we will pursue a different quasi-experimental approach. At the very least, we can compare customers in areas with high outages to customers in areas with low outages while controlling for as many observable differences between customers as possible.

We will use these two sources of variation – priority feeders during the Dumsor period and the as-yet-to-be-determined quasi-experimental variation during the Compact – to estimate the causal effect of improvements in power reliability and quality on socioeconomic outcomes. First, we will estimate the extent to which power reliability and quality affect firm and household electricity consumption directly through its impacts on the purchasing and usage patterns of electrical appliances or machinery. We will then estimate whether this affects economic outcomes such as firm productivity, profits, or revenues, or household income.

For several reasons, we believe it would be ideal to pursue both the Dumsor priority feeder and the Compact quasi-experimental approach as they have unique advantages and drawbacks. The Dumsor priority feeder analysis takes advantage of large, observed cross-customer differences in outages. There is no uncertainty in whether or not we can conduct this analysis to measure the impacts on households and firms. On the other hand, the analysis will involve surveying customers over 4 years after the major Dumsor period outages (which began in 2012, peaked around 2015, and largely ended by end of 2016) and may lack precision as we are not able to collect baseline data on individual customers. To the extent that outages and electricity reliability remain a significant issue in Accra and that the level of these issues varies between priority and non-priority feeders, however, our priority feeder analysis approach will also allow us to analyze impacts of continued differences in reliability on households and firms since the Dumsor crisis. The Compact quasi-experimental approach may allow us to collect baseline data (depending on which quasi-experimental approach we use) and would measure direct impacts of Compact investments. However, there is the risk that the main impacts of the Compact investments are to increase reliability for all customers. For example, if the main reliability improvements are based on the bulk supply investments, it is possible that nearly all customers in Accra experience fewer outages. While this would be a highly favorable outcome for Ghanaian customers, it would leave us with very little variation to use to evaluate the direct impacts of the Compact investments. In this case, following the logic outlined in equation (1) above, we can use the elasticities estimated using the Dumsor priority feeder approach together with the Compact-based improvements in reliability measured with GridWatch to calculate socioeconomic gains from the Compact. If we are able to implement both the Dumsor priority feeder approach and the Compact quasi-experimental approach, we will have two measures of the same elasticities, which will enhance the credibility of our estimates.

C. Quantitative Approach

i. Methodology

We next outline how we will implement the Dumsor priority feeder and the Compact quasi-experimental approaches, focusing on the LV bifurcation difference-in-difference as the specific Compact quasi-experimental approach.

For the Dumsor priority feeder approach, we will use a Difference-in Differences (DD) design. First, we will validate the empirical strategy by testing if customers on priority and non-priority feeders exhibit parallel trends in the outcome variables prior to Dumsor. Second, we will estimate treatment effects of being exposed to fewer outages during Dumsor.

Using a panel of household and firm-level data both prior to and during Dumsor, we will estimate the following regression equation:

$$\boldsymbol{Y_{it} = \gamma PRIORITY_i + \delta DUMSOR_t + \beta PRIORITY * DUMSOR_{it} + \sum \theta X_i + \varepsilon_{it}} \quad (2)$$

where Y_{it} measures electricity consumption, appliance ownership, income, profits, or socioeconomic outcomes for firm or household i. The treatment variable $PRIORITY_i$ is a measure of whether the customer is served by a priority feeder. $DUMSOR_t$ is an indicator of whether the data are from the Dumsor period. X_i is a set of controls such as industry (for firms) or household size (for households) that will improve the precision of our estimates. ε_{it} captures unobserved factors that affect electricity and socioeconomic outcomes for households over time. β is the parameter of interest because it measures the impact on the outcome variable of receiving priority power reliably during the Dumsor period. In general, we will test the hypothesis that being on a priority feeder during the Dumsor period had a positive impact (defined based on the outcome, for example an increase in household electricity consumption or firm profits or a decrease in spending on electricity protection) on the outcome measure under consideration against the null hypothesis of no impact.

This baseline model can be expanded in a variety of ways to analyze specific channels through which reliability impacts outcomes. For example, we can estimate "dynamic treatment effects" to estimate how the impact of reliability varies over time. We can also measure reliability as a continuous variable (such as hours of load shedding per month) so that we can estimate the elasticity of the outcome variable with respect to reliability. Our ability to estimate these elasticities will depend on the completeness of our load shedding data.

For some variables, such as those collected by our own surveys and not through archival data, we will not have panel data, and so will instead estimate the following cross-sectional specification:

$$Y_i = \beta PRIORITY_i + \sum \theta X_i + \varepsilon_i \quad (3)$$

where Y_i measures the outcome variable, the variable $PRIORITY_i$ is a measure of whether the customer is served by a priority feeder and X_i is a set of controls. β is again the parameter of interest and measures the impact of receiving priority power reliability on the outcome variable. The validity of the cross-sectional specification will be supported by the parallel trends tests on the outcome variables for which we have panel data.

We generally hypothesize a positive impact of receiving priority power reliably on household and firm outcomes, and will test this using a wide variety of outcome variables. Outcomes considered will include electricity consumption, appliance ownership and use, spending on electricity protection and reliability coping mechanisms (including generators), household income and labor participation, and firm operations, costs, revenues, and profits. The hypothesized impact of being on a priority feeder during the Dumsor crisis will vary depending on the outcome being considered - an "improvement" in the outcome might be either an increase or a decrease in the variable. The key outcomes of interest are economic outcomes variables – household incomes and firm profits. However, we also include various other socioeconomic outcomes which may be expected to have an impact on incomes and profits.

The table below presents our primary outcome variables succinctly:

Table 2: Summary of Outcome Variables

Outcome Group	Socioeconomic Variable
Energy use	Electricity consumption
Energy use	Appliance ownership and use
Energy use	Spending on protection against outages
Household	Income
Household	Labor participation
Firm	Number of Employees

Firm	Revenues
Firm	Profits

To increase precision of our estimates, control variables may include a variety of household and firm characteristics, any relevant differences in electricity access not captured by the priority feeder dummy variable, and the timing of the survey. Key household control variables are likely to include measures of household size and education-level, and key firm control variables are likely to include industry. Our identification approach assumes that exogenous variables other than the electricity network should not differ systematically across treatment and control sites. However, we may also include EA-level controls to further increase precision of our estimated impacts. Our empirical approach will not be biased as long as all other factors that determine the outcome (e.g. tariffs, income, weather) are affecting that outcome in both our treatment and control groups.

For the LV bifurcation quasi-experimental approach, we will use a Difference-in-Difference (DD) design. Intuitively we will compare customers who are within a very similar distance from an existing transformer (e.g., around 500 meters), and thus on average will have similar characteristics, *except* that some of these households will have their line bifurcated with a transformer injection and thus will experience higher power reliability and quality. This strategy exploits the fact that although LV lines become discretely eligible for bifurcation if they extend more than 500 meters from an existing transformer (or the load factor exceeds 70%), not all such LV lines have been selected for injections. In this sense, both the treatment and control groups consists of customers eligible for bifurcation, and are otherwise similar except that treatment sites benefit from a line bifurcation with a transformer injection. We implement the DD strategy by estimating the following regression equation:

$$Y_i = \beta_0 + \beta_1 * BIFURCATION_i + \beta_2 * TIME_i + \beta_3 TREAT_i + \varepsilon_i \qquad (4)$$

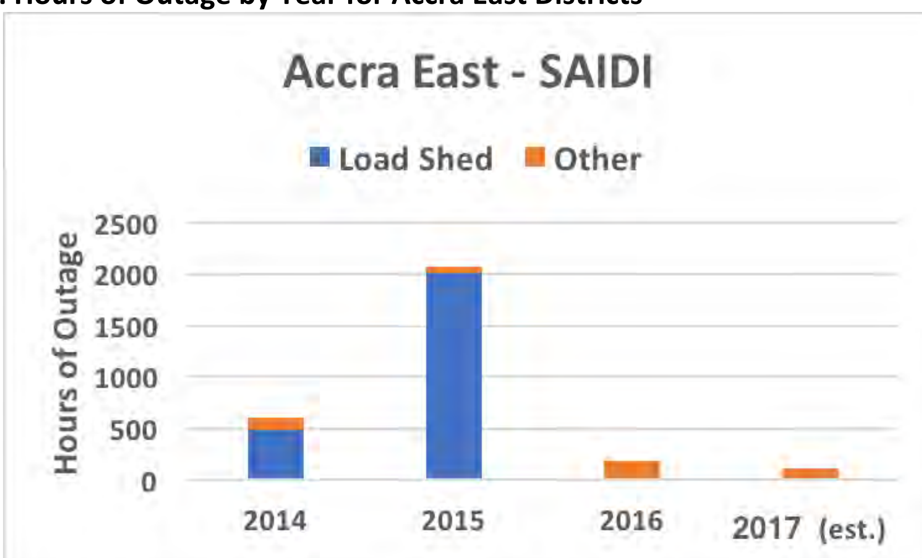
where Y_i measures the same outcomes that we discuss above. $BIFURCATION_i$ is an indicator of whether the customer is in a site selected for bifurcation and $TIME_i$ is an indicator for whether the customer is being observed after the line bifurcation work has been completed. The coefficient of interest is β_3, which measures the effect of treatment – being in a site selected for bifurcation after the bifurcation work has been completed. This coefficient is the estimated causal impact of bifurcation eligibility on the outcome. By combining this estimate with metrics of the first-stage impact of bifurcation on power reliability and quality, we will estimate the causal effect of power reliability or power quality on the outcome. If we conduct a baseline survey for the LV bifurcation approach, we will be able to estimate a panel version of the above equation.

ii. Timeframe of exposure

We are proposing three rounds of surveys, as outlined in Table 5 below. The first survey will be conducted in Q1-Q2 2020 targeted at households on priority feeders as well as the identified set of control households. The second survey will be conducted over the same timeframe targeted at households in treatment LV bifurcation areas and nearby control areas a similar distance from existing transformers. This will serve as a baseline for the Compact quasi-experimental approach. The third will be in Q1 2022 as an endline survey for the Compact quasi-experimental approach.

Figure 3 below describes the recent trend in outages in Accra East based on data collected from ECG. This highlights that the level of outages has decreased substantially since 2015, although the average Accra resident still experiences 100 hours per year without power.

Figure 3: Hours of Outage by Year for Accra East Districts



Source: ECG's Incident Management System call volume data.

For the Dumsor priority feeder approach, we will survey customers approximately 4 to 5 years after the major disruptions occurred. Since we are primarily asking about capital investments, such as appliance ownership, this lag should not be a problem. We intend to ask customers whether they own an appliance or machine and the year of purchase. If customers in the unprotected areas have increased purchases in 2016 and 2017, we will be able to reconstruct that behavior. Our survey of households and firms will include recall questions about experiences during the Dumsor crisis period, as well as questions about current electricity experiences and perceptions of whether Dumsor is back or is coming back. Identification of the impacts of differences in reliability during the Dumsor crisis will be based on respondents who lived in the same location during Dumsor. Other respondents in the priority feeder sample will provide evidence of differences in electricity reliability across priority and non-priority feeders more generally, and not just during the Dumsor period. We will also use archival data, described more fully below. Some of the surveys we will access are from 2015-16, so we will be able to measure the impacts of the Dumsor period directly.

Similarly, assuming that the LV bifurcation work takes place in late 2019 to early 2020 and that we conduct the endline survey in 2022, we will survey customers approximately 2 to 3 years after their reliability has improved. We feel that 2 years is long enough for customers to have adapted to the new levels of higher reliability, but not so long that the impacts of the Compact have faded.

iii. Study Sample

The unit of analysis for our evaluation will be either a household or firm, depending on the research question. They will be clustered within small geographic zones according to their low-voltage network connections.

We determine the appropriate sample size for each analytical approach using statistical power calculations. Our research design will have adequate "power" if we can be reasonably sure that the observed differences in outcomes across the treated (e.g., high reliability) and control (e.g., low reliability) groups were caused by the quasi-random differences in power reliability or quality. The more powerful our research design, the smaller the effect we can detect with conventionally accepted levels of confidence.

1. Dumsor Priority Feeder Approach

We will implement the Dumsor Priority Feeder Survey with limited information about reasonable detectable effect sizes. We explain the variation in outages above. For outcome variables that will be measured as proportions, we can use existing information on the share of households or firms that currently own a refrigerator, estimates of the effects of reliability and straightforward calculations to determine the necessary sample size. For example, to evaluate the null hypothesis that reliability had no impact on refrigerator ownership, i.e. $P_1 = P_2$ where P_1 is the share of households on priority feeders that own a refrigerator and P_2 is the same for households that are not on priority feeders, the test statistic is the z-score:

$$z = \frac{(P_1 - P_2)}{Std.Err}$$

where

$$Std.Err. = \sqrt{P \times (1-P) \times \left[\left(\frac{1}{N_1}\right) + \left(\frac{1}{N_2}\right)\right]} \; for \; P = \frac{P_1 \times N_1 + P_2 \times N_2}{N_1 + N_2}$$

For continuous variables such as electricity consumption, the power calculation formulas we use are based on the sampling properties of the treatment effect estimate described by equation (3) above. Some additional notation is useful when describing the variance estimator for the β coefficient. Let σ^2 measure the variance of the residual in the estimating equation. Let the number of customers in our study be denoted N and let P denote the proportion of the sample receiving the treatment.

We denote the Type I error rate (α) to be the probability of rejecting the null hypothesis when it is true. The conventionally accepted Type 1 error rate α = 0.05. The desired level of statistical power (κ) measures the probability that a difference of a given magnitude will be correctly detected. Power greater than 0.80 is generally accepted as adequate. To achieve this level of power, given our chosen (α), it must be that the true population average effect is greater than or equal to some minimum detectable effect (MDE):

$$MDE \geq (t_\alpha + t_{1-\kappa}) \sqrt{\frac{1}{P(1-P)} \frac{\sigma^2}{N}}$$

Table 3 summarizes the minimum detectable effect sizes for our target sample size based on applying these formulae.

Table 3: Sample Sizes for Select Variables – Dumsor Priority Feeder Analysis

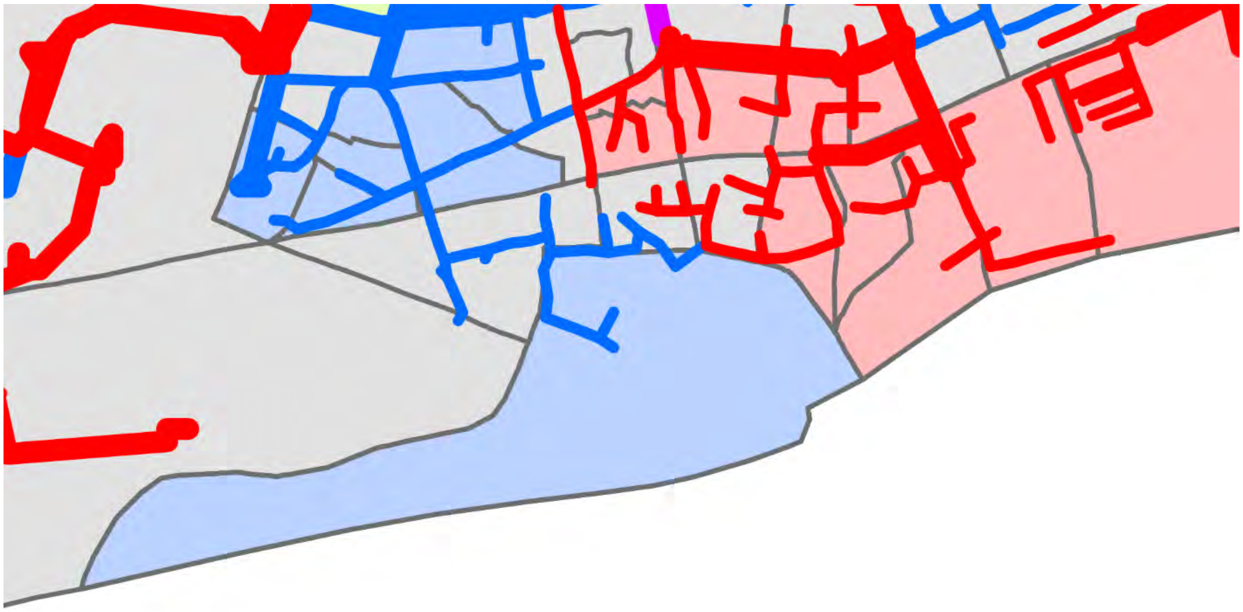
	Baseline Estimate	Source	Sample Size		Minimum Detectable Effect
			High Power Quality	Low Power Quality	
Refrigerator Ownership	0.76	2010 Ghana Citizens Report	750	750	0.043
Computer ownership	0.22	2010 Ghana Citizens Report	750	750	0.042
Bill repayment	0.18	ECG Data	1300	3000	0.025

(share fully paid)					
Electricity consumption (kWh)	TBD	ECG Data	TBD	TBD	TBD

Sample frame and sampling strategy – The sample frame and sampling strategy will vary depending on whether we are using archival data or survey data. For archival data from the GSS, we will use data at the enumeration area (EA) level. We have merged data on the distribution of GSS EAs with data on feeder designations and grid network GIS data from ECG and the PMC and are working to produce maps of the feeder and low-voltage electricity grid across districts in Accra so that we can understand the spatial distribution of feeders with different designations. The majority of exempt feeders are in Makola district, but Roman Ridge, Legon, Kaneshie, and Teshie also have at least one exempt feeder. Based on our mapping of exempt and ordinary feeders and associated LV lines we have been able to identify EAs that are served by different types of feeders – exempt only, ordinary only, or a combination of both. This process allows us to match physical points on the electricity grid with firm- and household-level outcomes. ECG billing data are already mapped to feeders. For analyses using ECG or GSS data, we will use the entire set of available customers since there is no cost to collecting additional data. For this reason, there is no sampling strategy.

For the survey, we will identify a set of clusters of customers on the priority feeders and a set of comparable clusters of customers on non-priority feeders using baseline (pre-Dumsor period) data from GSS and from ECG. To reduce selection issues, for our PF evaluation site selection we are primarily interested in boundary areas, as households and firms in these areas should be expected to be more similar than if they were drawn from non-adjacent areas. We have so far identified 24 boundary exempt-only EAs and 66 boundary ordinary-only EAs, where “boundary” means that an EA borders another EAs served by the opposite type of feeder. There are also 44 EAs served by both types of feeders. For households and firms surveyed in the latter type of EA, we will be able to use GPS location information to determine ex post whether they are served by an exempt or an ordinary feeder.

The screenshot below shows an example of the designation of EAs in Makola district. The thick lines represent feeder lines and the thin lines represent LV lines, with red indicating exempt designation and blue indicating ordinary designation. The gray lines show the boundaries of EAs. Red shading indicates the EA is exempt-only, blue shading indicates the EA is ordinary-only, and gray shading indicates the EA is served by both exempt and ordinary feeders. The white area at the bottom is the ocean.



Using EAs as survey sites for the PF evaluation has several advantages. First, we are able to use GSS data to conduct baseline balance tests across EAs, in order to confirm that neighboring households or firms were indeed statistically similar prior to the Dumsor period. We can also potentially use these data to supplement our evaluation and the survey data we collect. Second, EAs are recognized and well-understood sampling areas and our survey implementation partners at the University of Ghana are confident in being able to conduct household and firm sampling within EA boundaries in Accra.

We have collected baseline (pre-Dumsor crisis) data from the 2010 census to conduct balance tests on these EAs for a selection of household socioeconomic characteristics reported in the census. Initial results suggests that the different categories of EAs are generally statistically indistinguishable. We do not currently have baseline data on firms with which to conduct balance tests, as the GSS' Integrated Business Establishment Survey (IBES) firm census and follow-up surveys conducted in 2013 and 2014 did not use EAs for sampling. The GSS do not have clear documentation available for matching the sampling areas for the firm data to EAs, but we are working with them to try and accomplish this. If it appears that any of the priority and non-priority areas were significantly different prior to the Dumsor period, we will explore using matching techniques to select a sample that is balanced at baseline.

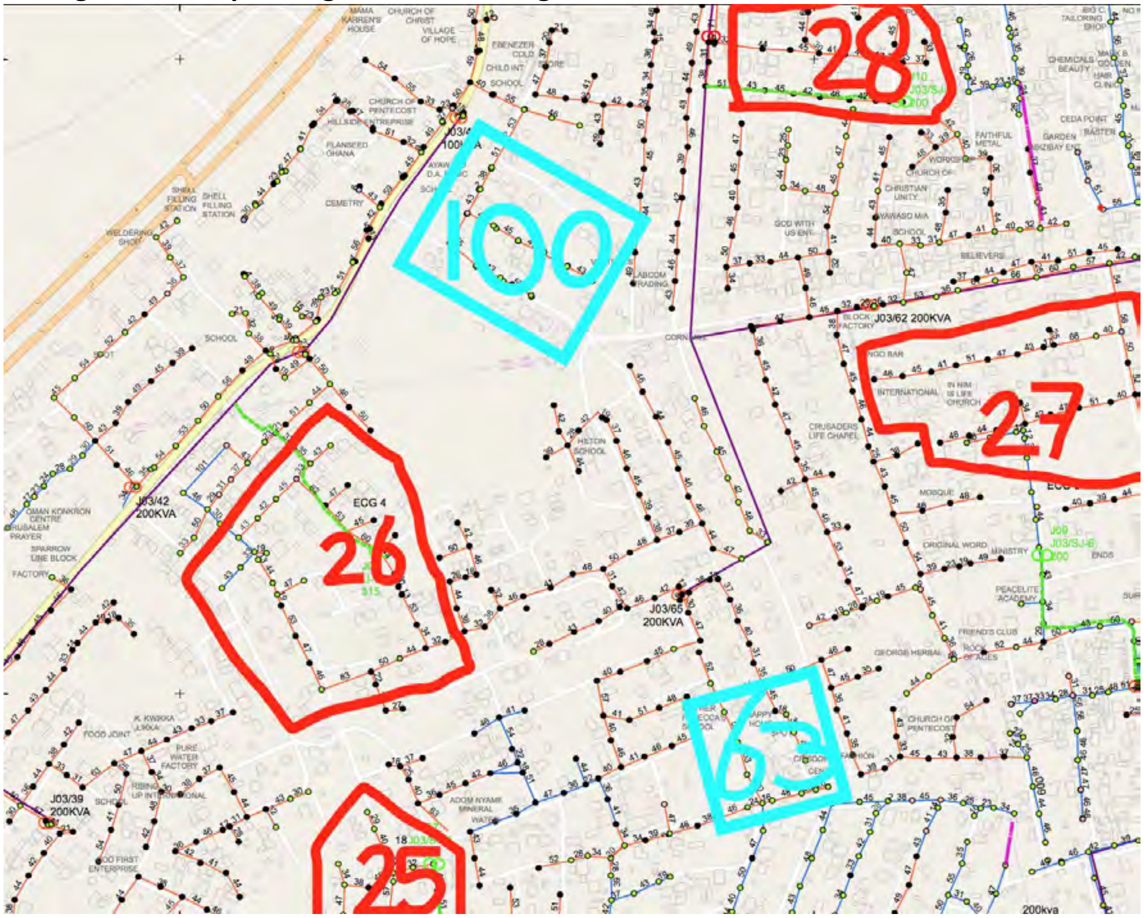
After testing for balance across these "boundary" EAs, we will draw a random sample of these EAs in which to interview households and firms. Within EAs, sampling of households and firms will be done quasi-randomly by random walk through the relevant neighborhoods. This method has been used in a large number of development economics studies and is accepted to generate a representative sample of the population. If time and budget allows, we will aim to conduct a listing of firms within the selected EAs to analyze differences in the composition of businesses across survey sites. Budget and timing constraints prevent us from pursuing a listing of both households and firms and the random sampling of respondents from those listings, but a well-implemented random walk should ensure the selection of respondents in each site is essentially random and is representative. If we are able to conduct a firm listing by site, we may stratify our random walk approach to select firms to ensure certain subsets of firms are included in the sample.

2. Compact Quasi-Experimental Approach

The sample size, sample frame, and sampling strategy for the LV bifurcation DD approach are much less concrete as they are conditional on several outcomes. We describe the approach to making decisions on these factors in this section. At a high level, we need to understand the cross-customer variation in outages experienced as a result of the LV bifurcation work, assess reasonable effect sizes that would result from this variation, and then calculate the sample size required to measure these effect sizes.

The GridWatch technologies will be crucial to understanding the cross-customer variation in outages experienced as a result of the LV bifurcation work. We have created maps based on data from the PMC on target areas (i.e., areas that are receiving a new transformer because either the local one is overloaded or customers are too far from the transformer) and comparable non-target areas (see Figure 4), as described above. We will continue to collect data from GridWatch and assess how power reliability and quality varies between target and non-target areas. We have also collected survey data from randomly selected firms and households within LV bifurcation treatment and control sites (during the deployment of the GridWatch sensors) and analyzed these data to verify that site selection is not correlated with socioeconomic outcomes.

Figure 4: Sample Target and Non-Target Areas



We will use information from the priority feeder analysis to determine what the expected effect size is from the LV bifurcation. For example, we know that during the Dumsor period, customers on protected feeders experienced roughly 100 more hours of outages per month. Imagine that we estimate a 10% difference in refrigerator ownership between customers on protected and unprotected feeders based on the priority feeder survey. Imagine too that the GridWatch data reveal a 25 hour per month difference in outages between target and non-target areas as a result of the LV bifurcation. A simple linear interpolation suggests that we would expect a roughly 2.5% difference in refrigerator ownership between target and non-target areas. We would then perform power calculations, as discussed above, to determine the necessary sample size to measure an MDE of 2.5%.

Note that there is no theoretical reason that we would expect the effect to scale linearly. It's possible, for instance, that for smaller differences in outages, customers are less likely to assume that they reflect systematic differences and so less likely to make investment decisions on the basis of them. On the other hand, if the LV bifurcation improves voltage as well as outages, we may expect a larger effect size. To be conservative (as power calculations are inherently noisy), we might target a slightly lower MDE than suggested by the linear interpolation. The issue of linear effects of reliability on outcomes is less relevant to our evaluations as we are primarily going to be evaluating the effects of dummy treatment variables. To the extent that we might also conduct some analyses using continuous treatment variables (e.g., number or length of outages per period), we might explore different functional forms such as a polynomial in the treatment variable.

We will work with the PMC to verify that construction adheres to the current list of planned line bifurcation injection sites, and that no injections take place in the sites we have selected as controls. For example, we will ask for construction logs. We are still assessing whether it will be cost-effective to verify independently that the PMC and the contractors conducting the LV bifurcation work adhere to the design drawings and installs new equipment as planned. We could do this, for example, by sending field officers to a randomly selected subset of the construction sites to verify the existence of a new transformer and protected conductors. We will collect the data from GridWatch on resulting power reliability and quality differences, so if the PMC or the construction contractors do not adhere to the drawings, these data will suggest few if any differences between target and non-target areas. On the other hand, it may be independently interesting to understand whether the GridWatch data reveal no differences because the equipment wasn't installed or because the equipment was installed correctly but it did not lead to the expected improvements in power reliability and quality.

iv. Data

We will rely on a series of primary and secondary data sources for both baseline characteristics and endline outcomes. Data on baseline characteristics will be used to assess balance between areas on protected feeders and areas on unprotected feeders. The following table outlines the types of information we will collect and the expected sources.

Table 4: Data Sources

Variable	Level of variation	Source	Required Steps
		Primary Sources	
Outage indicator	Transformer by hour	GridWatch technologies	Deploy technologies to sample areas
Voltage quality (units)	Transformer by hour	GridWatch technologies	Deploy technologies to sample areas
Electricity consumption (spending)	Customer by month	Household and firm survey	Perform survey
Electrical appliance ownership and use	Customer	Household and firm survey	Perform survey

Generator ownership and use	Customer	Household and firm survey	Perform survey
Electricity protection device ownership (e.g., stabilizer) and spending	Customer	Household and firm survey	Perform survey
Income	Customer (Household only)	Household and firm survey	Perform survey
Educational attainment	Customer (Household only)	Household and firm survey	Perform survey
Labor Participation	Customer (Household only)	Household and firm survey	Perform survey
Revenues	Customer (Firm only)	Household and firm survey	Perform survey
Profits	Customer (Firm only)	Household and firm survey	Perform survey
Hours of operation	Customer (Firm only)	Household and firm survey	Perform survey
Employees	Customer (Firm only)	Household and firm survey	Perform survey
Secondary Sources – Outcomes			
Electricity consumption (kWh)	Customer by month	ECG	Work with MiDA to get customer-level billing data for relevant sample
Bill payment and account balance	Customer by month	ECG	Work with MiDA to get customer-level billing data for relevant sample
Census of outage events	Feeder, start and end time	ECG	Work with MiDA to get outage data for relevant sample
Line bifurcation injections	Survey site	PMC	Work with PMC to monitor construction timing and completion
Socioeconomic characteristics – households and firms	Enumeration area, possibly household and firm	2017 Ghana Living Standards Survey 7	Available
Business characteristics	Neighborhoods to match to enumeration areas	2019 Integrated Business Establishment Survey Phase II (IBES-II)	Available conditional on manual name match between neighborhoods and enumeration areas
Secondary Sources – Baseline Characteristics			
Socioeconomic characteristics - households	Enumeration area (1,500 in Accra)	2010 Census	Available
Socioeconomic characteristics – households and firms	Enumeration area, possibly household and firm	2012 Ghana Living Standards Survey 6, Ghana Statistical Service	Available
Business characteristics	Neighborhoods to match to enumeration areas	2013-2014 Integrated Business Establishment Survey Phase I (IBES-I)	Available conditional on manual name match between neighborhoods and enumeration areas

Notes: Outage indicator and frequency variables measured by GridWatch will be available starting in roughly August 2018, scaling up to cover larger areas. Note that we will ask customers about their electricity consumption by month and also analyze data from ECG. Data from ECG will allow us to cover a more complete sample, while asking survey respondents will allow us to assess that important outcome variable in concert with all the other information collected through surveys.

1. Primary Data Collection

We will perform three types of primary data collection: technical data, household-level surveys, and firm-level surveys.

Technical data on power reliability and quality will be collected using the GridWatch suite of technologies, which in Accra will consist of several thousand downloads of the DumsorWatch App and several hundred deployed PowerWatch devices, linked and analyzed through cloud computing software. The technical details are outlined in Annex A. Currently, there are several hundred PowerWatch devices deployed in treatment and control areas for the line bifurcation work. There are no specific plans to deploy PowerWatch devices near our priority feeder survey areas, but since there are fewer feeders Accra-wide than the number of PowerWatch devices to be deployed, we are cautiously optimistic that we will be able to use devices deployed for other purposes to learn about outages on priority and non-priority feeders.

The household surveys will collect information on socioeconomic characteristics of the household, including household size, income, and education of the respondent. The firm surveys will collect information of business activities, revenues, and profits, as well as information on constraints firms face and measures taken in response to electricity reliability issues. For both firms and households, we will also collect detailed data on their electricity usage, including electricity spending, appliance ownership and usage, ownership and usage of electricity protection devices, including generators, and impacts of reliability issues.

Survey instrument – The household and firm survey instruments will be developed based on well-tested existing instruments previously used by members of the GridWatch team in studies in other countries, as well as through pilots. (Annex B contains the existing survey used in pilot deployments of the GridWatch and PowerWatch devices.) Once we have developed the full Dumsor Priority Feeder Survey, we will conduct an extensive pilot in non-sample areas of Accra prior to launch of the survey. The household survey instrument is expected to cover between 5-10 modules, such as household demographic characteristics, economics, electricity usage, and perceptions of power reliability and quality experienced, and will take around 1 hour to complete. The instrument for the firm survey will be similar in length and content, but will focus more strongly on productive assets, productivity, and revenues. Respondents in the priority feeder survey samples who have been in the same location since the height of the Dumsor crisis will also be asked a series of recall questions about their household or firm during the Dumsor crisis. These questions will mirror the set of current questions respondents are asked about electricity use, appliance ownership, electricity protection and responses to outages, and impacts of reliability issues. The surveys will be performed by a team of field officers using SurveyCTO on electronic tablets.

Rounds and timing – We envision conducting data collection for the Dumsor Priority Feeder Survey between January - May 2020. We will also collect baseline data for the LV bifurcation DD approach during this time period. If power calculations and post-construction GridWatch measurements suggest that the DD approach or another quasi-experimental approach will yield valuable estimates, an endline survey will happen after the Compact closes, roughly in early 2022, once households and firms have had access to improved power reliability and quality and have had time to adjust their electricity consumption patterns.

Respondent(s) within the sample unit – The Dumsor priority feeder DD design will be conducted at the neighborhood (Enumeration Area) level. While the exact sampling procedures will depend on cross-consumer variation and the number of neighborhoods deemed eligible for the study using secondary data, we expect to conduct between 5-50 surveys in each neighborhood. The LV bifurcation DD design will use a transformer as the unit of observation, which is a relatively smaller unit of observation than a neighborhood. We will therefore seek to conduct between 2-10 surveys at each transformer for the line bifurcation study. Note that one neighborhood could contain, for example, 5 transformers, in which case the density of surveys relative to population would be similar across the two designs. In both cases, we will implement a sampling strategy that aims to ensure that the sample is representative of the broader sample unit population.

Staff – Enumerators will need to have experience conducting household or firm surveys, an understanding of the Accra context and ability to engage respondents respectfully and clearly in this context, and a familiarity with SurveyCTO software for data collection purposes.

Data processing – All data will be collected electronically using SurveyCTO. Data will be stored on a secure server and analyzed at UC Berkeley by graduate students with experience analyzing survey data.

Data quality – Survey data will be downloaded on a daily basis and responses will be analyzed for consistency across questions and field officers. This way, any discrepancy or unexpected responses can be investigated and adjustments can be made to the survey protocol as they come up. In addition, a random 5% sample will be selected for in-depth backcheck re-surveys of stable questions (questions for which the answers are not expected to change substantially within the span of several weeks) to ensure that responses are consistent.

Table 5: Summary Table - Surveys

Data collection	Timing	Sample Unit/ Respondent	Sample Size	Relevant instruments	Exposure Period
Priority Feeder Survey	1/2020 – 5/2020	Households and firms	~ 1,500 firms ~ 1,500 households	See sample in Annex B	48-53 months

Baseline Line Bifurcation Survey	1/2020 – 5/2020	Households and firms	~ 1,000 firms ~ 1,000 households	See sample in Annex B	0 months
Endline Line Bifurcation Survey	1/2022 – 6/2022	Households and firms	TBD	See sample in Annex B, likely to be updated based on analysis of Priority Feeder Survey data	24 months (assuming LV bifurcation work commences ~Q1 '20)

2. Secondary Data Sources

Both approaches require significant secondary data. For the priority feeder approach, we will exploit the list of high-priority feeders that ECG maintains. We have obtained information on exempt and "sensitive" feeders across Accra East and Accra West. We have also obtained data on outages by feeder for the period 2012-2015, and have requested more recent data. As detailed in Table 4, we will also access several GSS surveys. Data collected before 2015 will help us assess balance on observable characteristics between areas that include customers who were on priority feeders compared to areas that include customers who were not on priority feeders.

We will use GSS data from 2016 or later to measures various outcomes, including number of firms (perhaps differentiating between electricity intensive sectors and less electricity intensive sectors), employment and various household characteristics. Note that one benefit of the priority feeder approach is that we can exploit additional archival data sets as we uncover them, as long as these data are georeferenced.

For the Compact quasi-experimental approach analyzing impacts of line bifurcation, we have obtained the most current and complete data on transformers and their attributes from the PMC, including load, location, and distance to furthest endpoint, for our study region. We will use these data to determine which transformers (and their nearby customers) are likely to receive new infrastructure and which are less likely to receive new infrastructure. To determine whether the cut-off is adhered to, we will obtain additional secondary data from the PMC once construction has been completed, listing where and when construction occurred, and what exactly was completed at each site. If this is deemed cost-effective, we may also choose to verify the accuracy of the administrative construction data by complementing these with on-site field surveys.

v. Analysis Plan

1. Data Assembly

The project includes several categories of data. Here we describe how we will collect and manage the data, evaluate data quality, and use the data to test the validity of our empirical strategy.

Primary household data – We will collect data via survey on household ownership of appliances, electricity consumption, and purchase and use of generators and electricity protection devices. We will also collect basic household sociodemographic information, including age, gender, education, and income for all household members. We will assess the data for coding inaccuracies and compare our survey data to secondary data from the Ghana Living Standards Survey that is collected by the Ghana Statistical Service.

Primary firm data – We will collect data via survey on firm characteristics such as date of establishment, number of employees, hours of operation, ownership of appliances, electricity consumption, and purchase and use of generators and electricity production devices. We will assess the data for coding inaccuracies and compare our survey data to data from the Non-Farm Enterprise Questionnaire of the Ghana Living Standards Survey and to data from the Integrated Business Establishment Survey.

Secondary firm and household data – The Ghana Living Standards Survey include micro data with rich geographic granularity that can be used to investigate the impact of the Dumsor period on a rich set of household and firm outcomes. GLSS round 6 occurred prior to Dumsor and GLSS round 7 occurred after the Dumsor crisis had subsided. In addition, the 2010 Census includes some basic sociodemographic information at the EA level. We have obtained the original microdata for all 3 surveys and are working to access data on geographic identifiers for GLSS 7 from the Ghana Statistical Service. The Integrated Business Establishment Survey I (IBES I) includes a census of all business in Ghana in 2013, along with limited information about the firms. The IBES II includes greater detail about a random sample of firms. We have obtained these data from the GSS along with available geographic identifiers, which must be matched manually to EAs and specific locations.

Primary grid data – The GridWatch technologies will collect sub-hourly data at the transformer level for areas of Accra West where we deploy the devices. These data include whether the power is currently on, location, time, and voltage quality at a plug in a household.

Secondary grid data – We will assemble data on load shedding by feeder from situational reports by ECG. These data provide us with a secondary source for outage data and allow us to evaluate the effects of historical outages on household and firm outcomes.

2. *Implementation of Research Design*

In order to assess the validity of the research design for the Dumsor priority feeder approach, our difference-in-differences approach requires that trends in the outcome variables follow parallel trends in the period prior to Dumsor. For outcomes where we have panel data, such as electricity consumption, we will test the assumption using standard techniques from the economics literature. For outcomes where we have snapshots of pre- and post-Dumsor data, we will construct balance tables to test for similar distributions across priority and non-priority feeders. We will use both archival data (from the Ghana Living Standards Survey) and primary survey data that we collect to test for balance across a rich set of household and firm characteristics.

Assessing the validity of the research design for the LV bifurcation approach will similarly involve comparing trends in the outcome variables in the period prior to the compact investments. We will use a similar approach to the one just described using variables through mid-2019 (prior to the compact investments).

D. Challenges

i. *Limitations of interpretations of the results*

Both approaches – Dumsor priority feeder and Compact quasi-experimental – rely on comparing economic outcomes across areas within Accra. It's possible that outages influence economic decision-making systemically. For example, an entrepreneur who lives in an area that experiences frequent outages may not be willing to invest in electrical equipment at his or her firm, even if the firm is in an area that experiences fewer outages. In general, if people form expectations about the likelihood of outages based on more than their own local experience, our approaches are likely to underestimate the benefits of Compact investments that lead to improved power reliability and quality. It is very difficult to measure systemic changes, such as changes in people's perceptions, as it is hard to find an appropriate control group.

As discussed above, we are unlikely to be able to capture the dynamic feedback impacts of improved reliability. There may be a virtuous circle, whereby fewer outages lead to higher bill repayment rates, and higher levels of retained earnings are then applied to continue improving the performance of ECG's system, which in turn leads to improved service quality. This describes a potentially important dynamic impact of the Compact's investments, although one that will be particularly hard to measure quantitatively. We note that we intend to analyze the first element of the chain – a relationship between improved service quality and bill payment rates

Another limitation of our study is the timeframe of the endline. It is possible that the infrastructure that the consultant is proposing to construct could lead to long-term improvements in power reliability and quality in our treatment areas. It may be worthwhile to return to these study sites in the long-term (starting 4-5 years after treatment) and conduct long-term endline surveys in order to estimate the long-term effect of reliability on socioeconomic outcomes.

ii. Risks to the study design

We see few risks to the Dumsor priority feeder approach, which is one of the main reasons we are pursuing it. It relies on known cross-customer variation in outages and will take advantage of archival data to the extent possible. One challenge is that this approach will ask survey respondents to recall information (e.g., about when they purchased a backup generator) and recall-based surveys can be imprecise.

Another risk is that the ECG data on outages during the Dumsor period are not in fact descriptive of customers' experiences. We will match GridWatch data collected beginning in 2019 with ECG data to assess potential discrepancies in the data. For example, if we see that ECG is underreporting outages by a factor of two, we will need to adjust the Dumsor period estimates to reflect this discrepancy.

The largest risk for the line-bifurcation DD approach is that construction will not follow the current selection of construction sites. For example, the consultant may complete network infrastructure improvements outside of the selected line bifurcation sites, or fail to complete installations where they were in fact required according to the rules. We plan to continue our efforts to communicate the importance of adhering to the eligibility cutoffs through in-person meetings, and request data at every stage to document this process. We hope that this will improve adherence to the rule. In addition, even with imperfect compliance, the initial selection of line bifurcation injection sites still can be used as instruments in an Instrumental Variables (IV) approach. Thus, even if adherence is not perfect, as long as the location of construction still aligns with these sites in the majority of cases, we will be able to obtain an accurate estimate of the impact of line bifurcation on power quality

Another potential challenge to the DD approach is low statistical power. Especially relative to the Dumsor period, the changes in outages brought about by line bifurcation may be imperceptible to most customers, especially in the short run.

An additional potential challenge to the Compact quasi-experimental approach is the information study we have been conducting in Achimota. Providing information collected by GridWatch to ECG could influence operation and maintenance decisions that affect system reliability over the life of the Compact. This may raise concerns about whether we will be able to conduct an independent impact evaluation of Compact investments. We feel confident that we can use standard econometric methods to conduct an impact evaluation to learn by how much Compact investments have influenced reliability and, in turn, how they have affected socio-economic outcomes in Accra, while also providing ECG with GridWatch data. We seek to evaluate two separate interventions. First, we would like to understand how Compact investments affect reliability and voltage. Second, we would like to understand whether providing GridWatch data to ECG improves their maintenance operations and, as a result, reliability and voltage in that same region. In accordance with standard econometric methods, we will cross-randomize these interventions. This will allow us to estimate the causal effect of each intervention on reliability separately, which is statistically identical to the causal effect that would be estimated if we only implemented one of these two interventions. In addition, with the cross-randomization approach our ability to identify the causal effect of the line bifurcation investments will not depend in any way on how ECG responds to the information we provide. This is why we believe that sharing the information will not jeopardize our ability to conduct an independent impact evaluation of compact investments. In addition, the information study will not last for the length of the Compact, and we have in fact already completed our initial phase of this study. There should therefore be even less risk that providing information to ECG during this study should impact the results of our evaluation of the line bifurcation investments, as that work has not yet begun.

Finally, a risk to the Compact period analysis, whether it is the DD approach or another quasi-experimental approach, is GridWatch equipment failure. We have built redundancy into our approaches to measuring local power reliability and quality. Each of the GridWatch technologies will alert the research team in near real-time if they experience performance issues. These technologies can nearly always be debugged from Berkeley and patches to the software can be sent from Berkeley to Ghana, reducing system downtime. Further, we adhere to best practices for software and hardware development by performing design reviews, code reviews, and laboratory tests before a technology or an update to a technology enters the field. Finally, we visualize incoming data streams, which allow high-level deployment management (i.e., which sensors are running and location, current outages and location, number of app users) to be performed by members of the research team without formal engineering expertise.

E. Summary of Methodology to Estimate Post-Compact Economic Rate of Return (ERR)

As described above, the MCC's ERR analysis focused on measuring the impact of outages on firms through (a) their need to run backup generators, and the expenses associated with that, and (b) lost sales. The impact evaluation will provide estimates consistent with the ERR as we will measure the effect of reduced outages on firm-level outcomes, including sales and the cost of operating a backup generator.

To use our estimates to measure economy-wide impacts of the Compact, one would need to extrapolate from the firms in our sample to the population of firms impacted by the Compact. If we succeed in implementing a sampling strategy that ensures our sample is representative of the population, then this is a reasonable extrapolation to make. If not, this may involve some re-weighting or extrapolations. Next, one would then need to measure the change in outages experienced by the population of firms. GridWatch measurements can provide more insight on this than any previous analysis. As our estimates will provide elasticity estimates, one could then multiply these (appropriately weighted to reflect the population of firms) by the change in outages measured using GridWatch to obtain estimates that are directly comparable to the ERR.

4. ADMINISTRATIVE

A. Summary of IRB Requirements and Clearances

All survey instruments, consent forms, and data collection protocols for this evaluation will be submitted for review by the Committee for the Protection of Human Subjects (CPHS) which is the Institutional Review Board (IRB) for the University of California, Berkeley (UCB). We note that after a detailed and iterative review process, we have already obtained UCB IRB approval for our Phase 1 activities. This approval authorizes us to deploy the DumsorWatch App and the PowerWatch devices in Accra, Ghana and collect data, as well as enroll up to 8,000 individuals at households or firms in the study by conducting an initial Enrollment Survey. We will request amendments to this protocol once we have developed the baseline survey for the Dumsor priority feeder approach, which should be a more streamlined process with the Phase 1 IRB already approved.

B. Data Protection

At a high level, all data collected will be encrypted both in flight (*i.e.,* during any network transmission) and at rest (*i.e.,* persistent storage). All data handling procedures will be in compliance with all appropriate Federal and UCB IRB regulations. Specific data protection methods are outlined below.

Surveyors in Ghana will collect data using encrypted tablets on a secure Open Data Kit-based platform, such as SurveyCTO. The tablets will be encrypted with passwords to prevent a third party from accessing the data. Each Surveyor will send the data to the secure SurveyCTO server at the end of each day.

To maintain confidentiality, surveyors will be prohibited from using these tablets for purposes outside the survey. Data uploaded to the SurveyCTO server will be encrypted automatically. Data downloaded from the SurveyCTO server will be stored in password-protected locations on the researchers' computers. In the case that any hard copies are used in the surveying activities, these will be stored in locked cabinets. For security purposes, all data will be kept in encrypted files.

Any sensitive data that are collected and need to be shared among the researchers will be securely stored within a standalone encryption container, such as a Cryptomator vault, *before* being uploaded to any secure cloud-based storage service such as Box. These fully encrypted containers will encrypt both filenames and their contents. Any fully encrypted container synced to cloud folders will be shared only amongst immediate members of the evaluation team. In addition to their cloud service credentials, these team members will only be able to decrypt the container via a secure randomly generated passphrase shared via a password management tool such as LastPass. Access to the underlying data will thus require both a protocol user's Box and LastPass account credential to access the encrypted file share and its unlocking password, respectively. Local working copies of the sensitive data may then be synced to researchers' computers, where they will be secured by user passwords and disk-level encryption.

If raw audio files are collected via the PowerWatch sensors (for purposes of detecting AC mains "hum"), these will be stored on the SD card of PowerWatch sensors over the course of the deployment. When PowerWatch is collected, the files will be removed from the SD cards and placed in an encrypted database. The SD cards will then be erased. Any raw audio files captured by the GridWatch app (for the same power presence detection reasons) will be stored on the phone of the GridWatch participants. As with all data above, these files will be encrypted both in flight and at rest. These files are removed when the app is uninstalled at the end of the study.

C. Preparing Data Files for Access, Privacy and Documentation

The privacy of every participant in the data collection process will be treated with the utmost respect and care throughout the evaluation. Datasets provided will follow both IRB and MCC Data Documentation and Anonymization Requirements, including those specified in the MCC Evaluation Microdata Documentation and De-Identification Guidelines. In keeping with the spirit of MCC's emphasis on transparency, findings and data that meet the aforementioned privacy requirements will be published and shared with the broader development research and donor communities, as outlined in the dissemination plan below. Properly anonymized and privacy-protected datasets will be released within 6 months of their respective final reports, after iterative rounds of consultation with the MCC Disclosure Review Board. Household- and firm-level data will require low levels of effort to prepare for publication and release. Identifiers will be removed from the private identifiable information. After such removal, the information can be published and used for future research studies or distributed to other investigators. Data collected through the DumsorWatch App and PowerWatch devices are particularly useful to researchers when mapped to geographic areas, and anonymized versions of the data (without any locational information) will be less productive to other researchers. The research team will explore ways of editing or anonymizing the data that is in line with IRB confidentiality and anonymity requirements but that will still prove useful to outside researchers. However, once a protocol is agreed upon, this task will likely be easy to streamline due to the clean and highly formatted nature of the data collected. This task will therefore likely require medium-level levels of effort.

Identifiers will be removed from the data upon download and stored separately. Responses will be assigned a code number, and the list connecting personally identifiable information with this number will be kept secure and will be destroyed five years after the completion of the study. Identifiable and coded information will be secured via both password protection and encryption on the surveyors' tablets, in the SurveyCTO survey server and on researchers' computers. All identifiers will be coded using a master list. This list will be stored in a password protected and encrypted database on a server and will be made available only to members of the study team. It will be kept for a period of five years after study completion, after which it will be destroyed.

D. Dissemination Plan

In November 2017, we met with senior members of the Ministry of Energy, including the Acting Chief Director and the Director for Generation and Transmission. Both expressed frustration with the poor quality and credibility of the electricity reliability data that they currently use to make key investment and project decisions. They emphasized a need for higher quality data, especially when paired with rigorous socioeconomic research, to better understand the economic impacts of ECG's investments in Ghana's electricity infrastructure. We plan to continue our engagement with senior management at the Ministry to ensure dissemination of results with key decision makers, so that our findings can lead directly to improvements in ECG investments and operations.

Our close and extensive collaboration with Kenya Power over the past 2 years has revealed a similar urgency to improve the quality of data that policymakers use to make key operations and infrastructure investments decisions in Kenya. This is reflected in the fact that Kenya Power and GridWatch are currently working to integrate GridWatch sensing into Kenya Power's mobile app, with the eventual goal to use the resultant data stream as an input into their operational systems.

We also intend to engage citizens and disseminate our research to members of civil society. On our research trip in November 2017, we engaged with thought leaders in the non-governmental sector (including the Executive Directors of KITE and ACEP), as well as senior academics at GIMPA and the University of Ghana. These stakeholders all agreed on the importance of enhancing independent data collection as a way to hold government players more accountable and improve their understanding of the socioeconomic and distributional effects of poor reliability in the Ghanaian context. We also met with Kobina Aidoo, primary author of the Dumsor Report, a citizen initiative that distributed information about outages to citizens during the Dumsor crisis via social media outlets (Briggs and Aidoo, 2018). As the Dumsor crisis grew in severity, public interest in the Dumsor Report grew rapidly, and Kobina's network of social media and civil society contacts enabled the report to reach widespread prominence. Since the GridWatch app displays a real-time map of outages occurring across Accra, widespread downloads of the app will enable direct citizen engagement with and understanding of the provision of public goods by the public sector. We feel confident that we will be able to use these established channels to disseminate our data and findings to the general public.

We also note that we intend to develop additional dissemination activities once we have developed and agreed upon a Statement of Work for Phase II evaluation activities.

E. Evaluation Team Roles and Responsibilities

The first principal investigator (PI) for this project is Professor Catherine Wolfram, the Cora Jane Flood Professor of Business Administration at the Haas School of Business, University of California, Berkeley. Professor Wolfram is a Research Affiliate at the Energy Institute at Haas and the Program Director of the National Bureau of Economic Research's Environment and Energy Economics. Catherine's role focuses on conducting socioeconomic research and impact evaluation of a number of electricity interventions happening in Accra.

The second PI is Prabal Dutta, an Associate Professor in Electrical Engineering and Computer Sciences at University of California, Berkeley. He has been recognized with an NSF CAREER award, an Alfred P. Sloan Research Fellowship, an Intel Early Career Award, and a Popular Science Magazine Brilliant Ten Award. Prabal's role focuses on contributing to the high-level architecture and design of the technology used in the research project.

Professor Steve Puller is the PERC Professor of Free Enterprise in the Department of Economics at Texas A&M University. Professor Puller is a Research Associate with the National Bureau of Economic Research and a Research Affiliate with the International Growth Centre. Steve's role focuses on the implementation of the priority feeder and the LV bifurcation regression, discontinuity designs, the assembly and analysis of the household-level and firm-level surveys, and estimating the impact of reliability on socioeconomic outcomes.

Professor Jay Taneja is an Assistant Professor in Electrical and Computer Engineering at the University of Massachusetts Amherst. Jay will oversee the technical components of the project, and will be a primary contact in collaborations with employees of the utility in Ghana.

Matt Podolsky, Managing Director of the UCB Technology and Infrastructure for Emerging Regions group and Associate Director for Data Analytics at DIL, is serving as the data collection expert and co-fulfilling senior engineering duties. Matt will be responsible for the management of the technology development and deployment (fixed sensors, GridWatch app, and data collection system).

Karen Notsund, Associate Director at the Energy Institute at Haas, will (1) assist the PIs and team with reports and dissemination of results, (2) manage deliverables and HR for Energy Institute personnel, and (3) coordinate efforts for this project and its intersection with other funding sources for different phases of this project.

Noah Klugman, a Ph.D. candidate in Computer Science and Engineering at UCB, is a Graduate Student Researcher on the project. Noah is the first author on the original GridWatch paper and is the lead developer of the technologies deployed in this project. Under the supervision of Professor Dutta, Noah will be responsible for implementing the GridWatch app, the PowerWatch fixed sensors, and the automatic incentive payment schemes.

Susanna Berkouwer, a Ph.D. candidate in the Department of Agricultural and Resource Economics at UCB, is a Graduate Student Researcher at the Energy Institute at Haas. Under the supervision of Professor Wolfram, Susanna's role focuses on conducting the socioeconomic research and impact evaluation, and assisting with the implementation of field activities and surveys.

Pierre Biscaye, a Ph.D. student in the Department of Agricultural and Resource Economics at UCB, is a Graduate Student Researcher for the GridWatch project. Similar to Susanna, Pierre's role focuses on conducting the socioeconomic research and impact evaluation and assisting with the implementation of field activities and surveys.

F. Table: Evaluation Timeline & Reporting Schedule

The following table documents the expected timeline for data analysis and reporting. We will elaborate on it once we have developed and agreed on a Statement of Work for Phase II evaluation activities, including detailed tasks and deliverables.

Table 6: Reporting Timelines

Name of Round	Data Collection	Data Cleaning and Analysis	First Draft Report Expected	Final Draft Report Expected	Dissemination of Microdata
GridWatch Data Collection	7/2018 – ongoing	Ongoing	1/2020	4/2020	10/2020
Priority Feeder and Line Bifurcation Survey	1/2020 – 5/2020	3/2020 – 3/2021	3/2021	6/2021	9/2021
Endline Line Bifurcation Survey	1/2022 – 6/2022	2/2022 – 2/2023	2/2023	5/2023	8/2023

5. REFERENCES

- Abeberese, A. B., C. G. Ackah, P. O. Asuming, 2017. "How did the 2012-2015 power crisis affect small and medium manufacturing firms in Ghana?" *International Growth Centre*. Policy Brief #33305.
- Albert, J., S. Alikhan, L. Etter, M. Jeuland, and A. Wyatt, 2014. "Jordan compact – Water network, wastewater network, and As-Samra expansion projects." *Impact Evaluation Design Report.*
- Allcott, H. and A. Collard-Wexler, S. D. O'Connell, 2016. "How Do Electricity Shortages Affect Industry? Evidence from India." *American Economic Review.* 106 (3): 587-624.
- Briggs, R. and K. Aidoo, 2018. "Underpowered: Rolling blackouts in Africa disproportionately hurt the poor." *African Studies Review*, forthcoming.
- Dzansi, J., S. L. Puller, B. Street, B. Yebuah-Dwamena, 2018. "The Vicious Circle of Blackouts and Revenue Collection: Evidence from Ghana." *International Growth Centre.* Working paper E-89457-GHA-1.
- Fisher-Vanden, K., E. Mansur, and Q. Wang, 2015. "Electricity Shortages and Firm Productivity: Evidence from China's Industrial Firms." *Journal of Development Economics*. 114: 172-188.
- Hardy, M. and J. McCasland, 2019. "Lights Off, Lights On: The Effects of Electricity Shortages in Small Firms in Ghana." *Working Paper.*
- Lee, D. S. and T. Lemieux, 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.
- World Bank Group, 2013. "Ghana Enterprise Survey."

6. ANNEXES

A. Work Plan
B. Pilot Survey
C. Stakeholder Comments and Evaluator Responses
D. Evaluation Budget