STIMULATING ELECTRICITY DEMAND TO ENHANCE SUSTAINABLE HUMAN DEVELOPMENT IN SUB-SAHARAN AFRICA

June M. Lukuyu
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Power and Energy Commons

Recommended Citation

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
STIMULATING ELECTRICITY DEMAND TO ENHANCE SUSTAINABLE HUMAN DEVELOPMENT IN SUB-SAHARAN AFRICA

A Dissertation Presented

by

JUNE M. LUKUYU

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2023

Department of Electrical and Computer Engineering
STIMULATING ELECTRICITY DEMAND TO ENHANCE SUSTAINABLE HUMAN DEVELOPMENT IN SUB-SAHARAN AFRICA

A Dissertation Presented

by

JUNE M. LUKUYU

Approved as to style and content by:

Jay Taneja, Chair

David Irwin, Member

Prashant Shenoy, Member

Nathan Williams, Member

Christopher V. Hollot, Department Head
Electrical and Computer Engineering
To my late grandmother, Charlotte Wambugu, who taught me that the strength of a woman is not measured by the impact that the hardships in her life have had on her; but by the extent of her refusal to allow those hardships to dictate what and who she becomes. Your memory lives on in our hearts.
I would like to express my deepest gratitude to my Ph.D. advisor, Dr. Jay Taneja, for his invaluable support, guidance, and feedback throughout this journey and for challenging me to grow as an independent researcher and also personally. He truly embodies what it means to be a great researcher, educator, mentor, and human being. I am deeply indebted to my committee members, Dr. David Irwin, Dr. Prashant Shenoy, and Dr. Nathan Williams, for their insightful comments, suggestions, and unique viewpoints that went into improving this thesis. Additionally, this endeavor would not have been possible without the generous support from the Rockefeller Foundation and UK Aid, who funded my research.

Working at STIMA Lab has been an overwhelming joy, and I would like to thank my labmates, Santiago, Aggrey, Zeal, Bob, Simone, Kiki, and Kyatha. Their support, kindness, and friendship made these past four years much more enjoyable and kept me sane throughout the process. I would also like to thank my colleagues from the e-GUIDE initiative, from whom I have learned a lot during this journey. I have also been able to supervise some brilliant undergraduate students during my Ph.D., whose enthusiasm and diligent work helped my projects progress. I particularly thank Anvita and Xavier, whose work is all included in this thesis.

I would also like to extend my sincere thanks to everyone I have closely collaborated with externally: Dr. Rob Fetter and Dr. P.P Krishnapriya (Duke University), Dr. Gunther Bensch (RWI), Dr. Ryan Kennedy and Myriam Shiran (University of Houston), Dr.
Rebekah Shirley (WRI), Dr. Vijay Modi and Terence Conlon (Columbia University), and
the teams from Equatorial Power and The Minigrid Innovation Lab at CrossBoundary.
Working with them has been an amazing experience! A special thanks to everyone
at the Energy for Growth Hub, especially Todd Moss, Rose Mutiso, Katie Auth, Jake
Kincer, and Grace Tamble, for providing invaluable support, access to an incredible
research network, for teaching me how to translate my research into practical ideas and
insights, and nurturing my writing skills. This network has undoubtedly made me a
better researcher.
I have also been fortunate to have the incredible support of so many friends, particularly
Priscah, Anne and Paul, Damaris, Purity, Will and Kelly, Sam and Lynn, Micah and
Elizabeth, and Bosco. I have treasured their encouragement and friendship throughout
this journey.
Lastly, words cannot express my gratitude to my family: my partner, Sego; my dad
and mum, Ben and Margaret; my siblings, Angela and Andrew; and my nieces, Neema,
Nadya, and Charlotte. Their unwavering support, belief in me, and constant love got
me through the late nights, the hard times, and every milestone these past four years.
This endeavor would not have been possible without them. I am also grateful for all the
support I received from the rest of my extended family.
ABSTRACT

STIMULATING ELECTRICITY DEMAND TO ENHANCE SUSTAINABLE HUMAN DEVELOPMENT IN SUB-SAHARAN AFRICA

FEBRUARY 2023

JUNE M. LUKUYU

B.S., SMITH COLLEGE

M.S., LOUGHBOROUGH UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Jay Taneja

Sub-Saharan Africa continues to aggressively pursue universal electricity access to drive economic development, improved health, literacy, food security, and gender equality. However, increasing electricity supply and household connections is only one important step towards achieving economic transformation. Countries in the region must also strive for increased use of energy services, not only by households but also for income-generating productive uses and community facilities. Unfortunately, unlike the other regions in the world, increasing electricity access rates in sub-Saharan Africa have not been met with complementary growth in electricity consumption density. Therefore, contrary to the energy efficiency zeitgeist in industrialized regions, countries
in sub-Saharan Africa need to increase their electricity consumption density for income-generating purposes to grow their economies and, as such, for consumers to realize the full benefits of electrification.

The body of my thesis work evaluates strategies that are aimed at advancing inclusive economic growth and human development by stimulating electricity use from clean energy technologies while subsequently improving the sustainability of electrification programs in sub-Saharan Africa. Specifically, I evaluate strategies in three key domains in sub-Saharan Africa: electricity supply, transportation, and agriculture. These strategies seek to either facilitate organic consumption growth or target the conversion of existing fossil fuel-based technologies into electric for income-generating purposes and improved livelihoods. The results of this body of work could be useful in informing policy frameworks in the region aimed at making progress toward achieving multiple United Nations Sustainable Development Goals.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Thesis contributions</td>
<td>3</td>
</tr>
<tr>
<td>1.1.1 Examining Customer Profiles and Patterns for Decentralized ...</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2 Stimulating Organic Electricity Consumption Among Mini-grid Customers in East Africa</td>
<td>4</td>
</tr>
<tr>
<td>1.1.3 Converting Fishing Boats for Electric Mobility to Serve as mini-grid Anchor Loads</td>
<td>6</td>
</tr>
<tr>
<td>1.1.4 Planning for future transportation electrification in Nairobi considering local mobility and grid infrastructure constraints</td>
<td>7</td>
</tr>
<tr>
<td>1.1.5 Using Satellite Data to Detect Diesel-Powered Irrigation for Guiding Electrification in Ethiopia</td>
<td>8</td>
</tr>
<tr>
<td>1.2 Thesis Outline</td>
<td>9</td>
</tr>
<tr>
<td>2. BACKGROUND</td>
<td>11</td>
</tr>
<tr>
<td>2.1 Why Demand Stimulation?</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Lessons from History</td>
<td>12</td>
</tr>
<tr>
<td>2.3 What Makes Demand Stimulation a Unique Problem in Sub-Saharan Africa?</td>
<td>14</td>
</tr>
<tr>
<td>3. RELATED WORK</td>
<td>17</td>
</tr>
<tr>
<td>3.1 Facilitating Organic Consumption Growth</td>
<td>17</td>
</tr>
<tr>
<td>3.1.1 Appliance financing programs</td>
<td>18</td>
</tr>
<tr>
<td>3.1.2 Electricity tariff subsidy programs</td>
<td>18</td>
</tr>
<tr>
<td>3.2 Productive Uses of Electricity</td>
<td>19</td>
</tr>
<tr>
<td>3.2.1 Electricity for Transportation</td>
<td>20</td>
</tr>
<tr>
<td>3.2.2 Electricity for Agriculture</td>
<td>22</td>
</tr>
</tbody>
</table>
4. CUSTOMER PROFILES AND PATTERNS FOR DECENTRALIZED CUSTOMERS IN EAST AFRICA ........................................... 24
   4.1 Background and Motivation ........................................... 24
   4.2 Data .................................................................. 27
      4.2.1 Demographic and socioeconomic data ....................... 28
      4.2.2 Electricity consumption and payment data .................. 29
      4.2.3 Study areas .................................................... 30
   4.3 Methodology .......................................................... 36
      4.3.1 Linear models .................................................... 36
      4.3.2 Daily load profile segmentation ............................... 38
   4.4 Consumption patterns and electricity spending behavior .......... 39
   4.5 Conclusions and policy implications ............................... 49
      4.5.1 Implications for business model development ............... 51
      4.5.2 Implications for demand side management .................. 53
5. STIMULATING ORGANIC ELECTRICITY CONSUMPTION GROWTH AMONG MINI GRID CUSTOMERS IN EAST AFRICA .......... 55
   5.1 Background and Motivation ........................................... 55
   5.2 Methods and Data .................................................... 58
      5.2.1 Study Area ....................................................... 58
      5.2.2 Study Design ..................................................... 58
      5.2.3 Data ............................................................... 60
      5.2.4 Methods .......................................................... 63
   5.3 Analysis .................................................................. 69
      5.3.1 Appliance financing program: Effects on consumption ...... 69
      5.3.2 Appliance financing program: Time-of-Use Effects ........... 76
      5.3.3 Appliance financing program: Loan repayment ............... 78
      5.3.4 Appliance financing program: Economic analysis ............ 82
      5.3.5 Appliance financing program: Robustness analysis .......... 84
      5.3.6 Appliance financing program: Customer perspective ......... 86
      5.3.7 Tariff subsidy program: Effects on consumption ............. 86
   5.4 Discussion ............................................................... 89
   5.5 Conclusion .............................................................. 91
6. CONVERTING FISHING BOATS FOR ELECTRIC MOBILITY TO SERVE AS MINI-GRID ANCHOR LOADS ............................ 93
   6.1 Background and Motivation ........................................... 93
   6.2 Data and Methodology ................................................ 95
      6.2.1 Data collection and description .................................. 95
      6.2.2 Sizing of an electric outboard motor and battery system ... 101
      6.2.3 Residential and commercial demand estimation ............. 102
      6.2.4 Modeling ice factory power demand ........................... 103
      6.2.5 mini-grid operation model considering stochastic electric boat charging load ............................................ 104
6.3 Analysis ........................................... 107
   6.3.1 Residential and small commercial demand profile .... 107
   6.3.2 Electric outboard motor and battery sizing ............ 107
   6.3.3 Impact on mini-grid operation .......................... 109
   6.3.4 Impact on economics of mini-grid project .............. 111
   6.3.5 Economic impact for boat owners ....................... 113
   6.3.6 Demand response .................................... 115

6.4 Discussion and Future Work ............................. 116
6.5 Conclusion ......................................... 118

7. PLANNING FOR FUTURE TRANSPORTATION ELECTRIFICATION IN NAIROBI CONSIDERING LOCAL MOBILITY AND GRID INFRASTRUCTURE CONSTRAINTS ........................................ 119
   7.1 Introduction ......................................... 119
   7.2 Literature review ...................................... 121
   7.3 Transport electrification pathway scenarios .............. 125
   7.4 Methods .............................................. 126
      7.4.1 Model overview .................................... 126
      7.4.2 Data sets and sources .............................. 138
   7.5 Analysis .............................................. 141
      7.5.1 Estimating the mobility of different EV fleets ....... 141
      7.5.2 Potential implications of future vehicle electrification on Nairobi’s grid under different scenarios ............ 145
      7.5.3 Potential implications of future vehicle electrification on the city’s load profile and national generation under different scenarios ........................................ 151
   7.6 Discussion ............................................ 152
      7.6.1 Caveats of the model and limitations of our work ..... 153
      7.6.2 Future work ........................................ 154

8. USING SATELLITE DATA TO DETECT DIESEL POWERED IRRIGATION FOR GUIDING ELECTRIFICATION IN ETHIOPIA ........................................ 155
   8.1 Background and Motivation ............................. 155
   8.2 Methodology ......................................... 159
      8.2.1 Ground Truth Data .................................. 160
      8.2.2 Defining Features ................................... 163
      8.2.3 Time series feature extraction and selection ......... 168
      8.2.4 Data pre-processing .................................. 169
      8.2.5 Classifiers .......................................... 169
      8.2.6 Diesel-powered irrigation detection ................. 171
      8.2.7 Performance Metrics ................................ 174
   8.3 Analysis .............................................. 175
      8.3.1 Single-stage ternary classification .................... 175
      8.3.2 Two-stage binary classification ....................... 177
      8.3.3 Feature Contribution ................................ 179
8.3.4 Spatial Analysis ........................................ 180
8.3.5 Validation of Results ................................. 182

8.4 Discussion and Future work ............................. 184
8.5 Conclusion ............................................... 186

9. THESIS SUMMARY .......................................... 188

APPENDICES

A. SOCIOECONOMIC STATUS COMPARISON OF OFF-GRID AREAS USING NIGHTTIME LIGHTS LUMINOSITY .................. 189
B. EFFECTS OF APPLIANCE FINANCING PROGRAM ................. 192
C. MATCHING QUALITY INDICATORS AND ALTERNATIVE MATCHING METHODS ........................................ 197
D. ELECTRIC VEHICLE TECHNICAL OPTIONS ..................... 206

BIBLIOGRAPHY .................................................. 209
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Summary of data used in the study</td>
<td>28</td>
</tr>
<tr>
<td>4.2 The weighted relative wealth of populations in the administrative units where the SHS and mini-grid customers are located.</td>
<td>33</td>
</tr>
<tr>
<td>4.3 Demographic and socioeconomic characteristics of decentralized customers.</td>
<td>33</td>
</tr>
<tr>
<td>4.4 Appliances owned by decentralized customers.</td>
<td>36</td>
</tr>
<tr>
<td>5.1 Appliance purchases through the appliance financing program</td>
<td>62</td>
</tr>
<tr>
<td>5.2 Demographic and socioeconomic characteristics of mini-grid customers</td>
<td>63</td>
</tr>
<tr>
<td>5.3 Differences between treated, control, and matched control customers</td>
<td>67</td>
</tr>
<tr>
<td>5.4 Repayment rates for other consumer goods</td>
<td>80</td>
</tr>
<tr>
<td>5.5 Customer characteristics predicting on-time appliance loan repayment</td>
<td>81</td>
</tr>
<tr>
<td>7.1 Description of the scenarios considered for this study</td>
<td>126</td>
</tr>
<tr>
<td>7.2 Data sources, type, region and publication year.</td>
<td>140</td>
</tr>
<tr>
<td>7.3 Change in Nairobi’s load factor under different scenarios.</td>
<td>152</td>
</tr>
<tr>
<td>8.1 Number and total area of irrigated plots by method of irrigation</td>
<td>162</td>
</tr>
<tr>
<td>8.2 Number of 250 × 250 m² pixels per class</td>
<td>163</td>
</tr>
<tr>
<td>8.3 Comparison of model performance and classification approaches</td>
<td>176</td>
</tr>
<tr>
<td>8.4 Top ten features contributing to model performance</td>
<td>180</td>
</tr>
<tr>
<td>8.5 Performance of k-NN model considering the spatial proximity of class pixels</td>
<td>182</td>
</tr>
<tr>
<td>8.6 Bootstrap Matthews correlation coefficient estimates and confidence intervals for classifiers in the two-stage classification approach</td>
<td>183</td>
</tr>
<tr>
<td>B.1 Month level dynamic treatment effects of appliance financing program: Alternative regression specifications</td>
<td>192</td>
</tr>
<tr>
<td>C.2 Month level dynamic treatment effects of appliance financing program: Alternative matching methods</td>
<td>200</td>
</tr>
<tr>
<td>C.1 Rosenbaum sensitivity analysis for average treatment effect of appliance financing program</td>
<td>203</td>
</tr>
<tr>
<td>D.1 Electric vehicle options technical information</td>
<td>207</td>
</tr>
<tr>
<td>D.2 Electric vehicle options technical information (continued)</td>
<td>208</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.1</td>
<td>Annual average electricity consumption versus electricity access rate. Sources: IEA World Energy Balances [8] and World Bank Data Bank [113]</td>
</tr>
<tr>
<td>4.1</td>
<td>Proportion of population connected to decentralized systems. Notes: The circles within the boxplots represent the mean proportion of the population connected, while the lines represent the median.</td>
</tr>
<tr>
<td>4.2</td>
<td>Temperature and precipitation patterns of areas where mini-grid and SHS customers. Notes: Temperature and precipitation data for SHS customer areas is from Jan - Dec 2018, and temperature and rainfall data for mini-grid customer areas is from Jan - Dec 2017</td>
</tr>
<tr>
<td>4.3</td>
<td>Distribution of the monthly expenditures on electricity consumption for mini-grid customers and solar home system customers</td>
</tr>
<tr>
<td>4.4</td>
<td>Distribution of payment per top-up for mini-grid customers and solar home system customers</td>
</tr>
<tr>
<td>4.5</td>
<td>Distribution of the monthly consumption for mini-grid customers and solar home system customers</td>
</tr>
<tr>
<td>4.6</td>
<td>Daily load profile patterns of mini-grid and solar home system customers</td>
</tr>
<tr>
<td>4.7</td>
<td>Proportion of average daily consumption consumed during evening peak hours (5pm - 10pm). Bottom panels: Percentage of sample that is in each bin. Note: *p &lt; 0.1; **p &lt; 0.05; ***p &lt; 0.01</td>
</tr>
<tr>
<td>4.8</td>
<td>Number of hours in a day when hourly consumption was above 5% of the total daily consumption. Bottom panels: Percentage of sample that is in each bin. Note: *p &lt; 0.1; **p &lt; 0.05; ***p &lt; 0.01</td>
</tr>
<tr>
<td>5.1</td>
<td>Average treatment effect of appliance financing program. Notes: Robust standard errors are clustered by village; error bars represent a 95% confidence interval. Includes household fixed effects, calendar-month fixed effects, and relative week fixed effects. *p &lt; 0.1; **p &lt; 0.05; ***p &lt; 0.01</td>
</tr>
<tr>
<td>5.2</td>
<td>Treatment effect of individual appliance uptake. Notes: Error bars represent a 95% confidence interval; robust standard errors are clustered by village. Includes household fixed effects, calendar-month fixed effects and relative week fixed effects. *p &lt; 0.1; **p &lt; 0.05; ***p &lt; 0.01</td>
</tr>
<tr>
<td>5.3</td>
<td>Effect of appliance uptake on daily load profiles. Notes: Top panels are normalized daily load profiles; bottom panels show the average effect of individual appliances, distinguished by hour of the day. Error bars represent a 95% confidence interval, calculated from robust standard errors clustered by village. The regressions that generated the coefficients illustrated here include household fixed effects, calendar-month fixed effects, and relative hour fixed effects. *p &lt; 0.1; **p &lt; 0.05; ***p &lt; 0.01</td>
</tr>
<tr>
<td>5.4</td>
<td>Cumulative distribution functions of loan balances for customers who bought appliances</td>
</tr>
</tbody>
</table>
5.5 Average daily consumption prior to and after full loan repayment. Includes those customers who paid loans in full, regardless of whether the payment was within the loan term \( n = 131 \) (35\%).

5.6 Financial analysis for appliance financing program.

5.7 Effects of tariff reduction program.

5.8 Price elasticity of electricity demand for tariff subsidy customers.

5.9 Normalized daily load profile of residential and small business customers prior to and after a 75% tariff reduction.

6.1 Flowchart of the study methodology.

6.2 Example fishing boat on Lolwe Island. Inset: boat tracking device attached to the boat before deployment.

6.3 Probability distributions of the arrival and departure times of all 77 recorded fishing trips.

6.4 Select daily traces of fishing boats. Three departure villages are each denoted with an X.

6.5 Average daily load profile of clusters among (a) small business, (b) residential, and (c) home business customers.

6.6 Daily load profile of residential and small business customers.

6.7 Cumulative distribution of boat battery usage from two 9.1 kWh BMW i8 lithium ion battery packs connected in parallel.

6.8 mini-grid supply and demand curve on an average day of the year (a) without ice factory and electric boat charging load and (b) with ice production of 13,000 kg/day of and 102 boats charged at 15 charging stations.

6.9 Change in maximum daily electric boat charging demand as a function of number of charging stations and daily ice production.

6.10 Net Present Value, annual diesel consumption and annual excess PV supply as a function of number of charging stations and daily ice production. Note that the number of boats are maximized in each scenario.

6.11 Payback period on purchase of 40 HP Torqeedo Deep Blue electric outboard motor and 18.2 kWh lithium-ion battery system.

6.12 Impact of shifting electric boat charging load.

6.13 Impact of charging electric boats to 100% vs. 80%.

6.14 Distribution of signal strength as a function of distance from the island. High signal strength = 4; No signal strength = 0.

7.1 Flowchart showing the model and analysis framework.

7.2 Comparison of distribution of curb weights of the registered private vehicles in Kenya and the electric vehicle options included in our study.

7.3 Energy consumption curves with speed.

7.4 Low range anxiety probability distributions.

7.5 Low range anxiety probability distributions.

7.6 Probability distributions of departure times.

7.7 Simulated origin-destination private vehicle movements.

7.8 Simulated commercial vehicle movements.

7.9 Simulated paratransit vehicle movements.

7.10 Connections in the low range anxiety scenario compared to the high range anxiety scenario.

7.11 Zone level baseline aggregate transformer load and loading.

7.12 Number of zones with transformer overloading.

7.13 Zones with transformer overloading in high range anxiety scenario.

7.14 Changes to Nairobi demand profile with added EV charging load for different scenarios.

8.1 Overview of study methodology.
8.2 Location of cultivated plots in (a) East of Lake Tana in Amhara region and (b) East of Lake Ziway in central Oromia region, Ethiopia. ............... 161
8.3 Patterns of average tropospheric NO\textsubscript{2} measurements for diesel irrigated and non-diesel irrigated pixels .................................................. 165
8.4 Comparing tropospheric NO\textsubscript{2} column measurements and ground NO\textsubscript{2} concentrations in Accra, Ghana .................................................. 166
8.5 Empirical cumulative distribution functions of time constant features for pixels irrigated with and without diesel. ................................. 168
8.6 Single-stage ternary classification approach ........................................... 173
8.7 Two-stage binary classification approach .............................................. 173
8.8 Confusion matrices of models in the single-stage ternary classification approach .................................................................................... 177
8.9 ROC curve for Random Forest model in single-stage ternary classification approach ............................................................................ 178
8.10 Confusion matrices of models in the two-stage binary classification approach ..................................................................................... 179
8.11 10-fold cross validation ROC curve of diesel irrigated class for \(k\)-Nearest Neighbor model in two-stage binary classification approach. Note: value of \(k\) is 5. ................................................................. 180
8.12 Locations of test set pixels by classification ......................................... 181
8.13 Distribution of average weighted \(F_1\) scores from permutation test of \(k\)-NN model in two-stage classification approach ........................................ 184
8.14 Predictions of diesel-powered irrigation activity in the Amhara region of Ethiopia. .......................................................... 185
A.1 Average effect of installing decentralized systems on average nighttime lights radiance band ................................................................. 191
B.1 Distribution of frequency of electricity top-ups per month among treatment and control customers before and after appliance delivery. Note: Bottom panel shows the average treatment effect of the program. Error bars represent a 95% confidence interval; *\(p < 0.1\); **\(p < 0.05\); ***\(p < 0.01\) 194
B.2 Weeks from appliance delivery that coincides with the lean season. Notes: Top panel shows the months of the year that the weeks from appliance delivery in the bottom panel fall on for majority of the sample during each week. ................................................................. 195
B.3 Average treatment effect of program among customers whose reported primary source of income is from agriculture and those whose reported primary source of income is from commerce, services or salaried work. Notes: Error bars represent a 95% confidence interval; robust standard errors are clustered by village. Includes household fixed effects, calendar-month fixed effects and relative week fixed effects. Note: *\(p < 0.1\); **\(p < 0.05\); ***\(p < 0.01\) 195
B.4 Average weekly consumption among residential and business customers prior to and after appliance delivery. Note: Includes only a subset of the appliance financing customers for whom we have information on their connection type. ................................................................. 196
B.5 Treatment effect of taking up individual appliances and treatment effect of taking up multiple types of appliances. Notes: Error bars represent a 95% confidence interval; robust standard errors are clustered by village. Includes household fixed effects, calendar-month fixed effects and relative week fixed effects. ................................................................. 196
C.1 Average treatment effect of appliance financing program from kernel matching technique. Standard errors are bootstrapped; error bars represent a 95% confidence interval. Includes household fixed effects, calendar-month fixed effects, and relative week fixed effects. Note: *p < 0.1; **p < 0.05; ***p < 0.01.

C.2 Average treatment effect of appliance financing program from radius matching technique. Standard errors are bootstrapped; error bars represent a 95% confidence interval. Includes household fixed effects, calendar-month fixed effects, and relative week fixed effects. Note: *p < 0.1; **p < 0.05; ***p < 0.01.

C.3 Propensity scores for households in treatment (appliance uptake) and matched control (no uptake) groups indicating domain of common support.
Access to modern, reliable, affordable electricity is the backbone of any thriving economy. In the 1920s, the world’s largest economies, especially Europe, the United States, and the Soviet Union, aggressively pursued electrification as a driver for economic development [147, 180]. To date, there are no high-income countries with low electricity access rates. In addition to economic growth, improved energy access correlates with other development indicators, including improved health, literacy, food security, and gender equality [138, 136]. As such, following Sustainable Development Goal 7 (SDG 7), there have been sustained, concerted efforts globally towards ensuring access to “affordable, reliable, sustainable, and modern energy for all, particularly in sub-Saharan Africa (SSA), where 75% of the people without access to electricity reside [124]. Though it is imperative to increase electricity connections, there is a growing consensus that having an electricity connection is an incomplete measure of electricity access. Increased electricity supply and connections need to be accompanied by affordable and adequate electricity consumption and financially sustainable power systems to realize the intended benefits of electrification [112]. Unfortunately, electricity consumption remains low even with increasing electricity access rates in most SSA countries [88, 82].

Africa accounts for over 3% of global electricity demand [266], with an average annual per capita electricity consumption of 483 kWh. To put this into context, no
high-income countries today have annual per capita electricity consumption below 3,000 kWh, the median being 6,720 kWh [18]. As shown in Figure 1.1, unlike other regions with a striking exponential trend in average electricity consumption per capita with electricity access, the progression of average electricity consumption per capita in SSA with increasing access rates have plateaued. More so, countries in SSA at access rates over 70% are not exhibiting the growth in electricity consumption that we see in other regions at the same access rates. This trend raises the vital question of why electricity consumption growth in SSA is slower than in other regions and how electricity demand can be actively stimulated, given the landscape of available tools and technologies.

**Figure 1.1:** Annual average electricity consumption versus electricity access rate. Sources: IEA World Energy Balances [8] and World Bank Data Bank [113]

For SSA countries to grow their economies and reduce inequality, governments and development partners must develop policies and strategies to stimulate electricity demand, particularly in marginalized communities [195, 200]. Demand stimulation
is also essential for electricity providers - including utilities and off-grid companies - the majority of whom are struggling with financial sustainability [261, 207]. Growing electricity consumption enables a virtuous cycle – electricity providers generate more revenue to cover the costs of their investments and the ongoing costs of maintaining reliable power systems. This increase in revenue, in turn, lowers the unit cost of electricity, subsequently enabling further growth in electricity consumption and improved livelihoods.

1.1 Thesis contributions

In this thesis, we evaluate strategies to advance inclusive economic growth and human development through sustainable electricity use from clean energy technologies while improving the sustainability of electrification programs in sub-Saharan Africa. We consider five main contributions to this work:

1.1.1 Examining Customer Profiles and Patterns for Decentralized Electricity Systems in East Africa

The emergence of decentralized generation and storage, coupled with smart metering and payment digitization, has reinforced the rapid adoption of decentralized (off-grid) systems as electrification pathways in rural areas of SSA to supplement grid infrastructure, which is expensive to construct, slow to expand, and suffers frequent outages [100]. Whereas the electricity grid in SSA has long been dominated by the public sector—in the form of heavily subsidized, often state-owned utilities—private start-ups, have taken up the task of reaching those beyond the grid, with the dominant modern technologies being solar home systems (SHS) and mini-grids (MG). With the accelerated adoption of mini-grids and solar home system solutions, it is becoming increasingly important to
understand how their customer bases consume and spend electricity.

In this contribution, we use a novel data set of hourly electricity consumption data and financial expenditures for electricity, as well as socio-economic and demographic data of about 13,000 solar home system customers and 2,000 mini-grid customers; we present an initial comparison of the financial expenditures for electricity and electricity consumption patterns among mini-grid and solar home system customers as well as a discussion of the policy implications of our results. Our analysis shows that, on average, mini-grid customers spend substantially less on electricity than solar home system customers and make smaller, more frequent monthly payments, making cost recovery impossible without subsidization or implementing strategies to increase the revenue-generating potential of these mini-grids. We also find that higher monthly spending among mini-grid customers correlates with off-peak consumption and a higher load factor in the daily demand profile, contrary to the behavior of solar home system customers. These results have policy implications for business model development and demand-side management strategies of these off-grid platforms.

1.1.2 Stimulating Organic Electricity Consumption Among Mini-grid Customers in East Africa

In many settings, mini-grids offer a combination of affordability, reliability, and capacity for productive use of power, more so than most solar home systems and some central grids. Yet the economic sustainability of mini-grids relies on achieving target usage levels, and consumption data to date (evidenced by the results of our analysis from the first contribution) suggest that they may be commercially unsustainable due to consistently low demand for power once installed—and that newly-connected recipients cannot take full advantage of access.

In this contribution, we evaluate two strategies to address this shortfall that threatens
sustainable electricity access in these settings: Using a uniquely fine-grained data set spanning 29 villages in East Africa, we test whether credit constraints and the cost of electricity hinder demand growth among mini-grid-connected households. This study revealed a substantial demand for appliance financing among mini-grid customers and that customers are also very price sensitive. We find that households that purchased appliances under a financing program increased consumption by up to 66 percent compared to matched controls. However, a sensitivity analysis suggests this estimate is somewhat sensitive to bias from unobservable characteristics, and the increase is not sustained. While most customers in the program do not repay loans in full, we find that, on average, customers repay about 78 percent of the loan amount.

When we analyze developers’ return on investment, we find that the profitability of appliance financing programs at a market cost of capital, similar to those evaluated in this study, depends substantially on the types of appliances on offer. We also find that even at a subsidized cost, only appliances that were bought for business purposes, such as refrigerators and grain mills, improved the mini-grid system load factor by 10% as they were used primarily during off-peak hours, suggesting that the efficacy of appliance financing programs in enhancing the techno-economic performance of mini-grids hinges on offering appliances and equipment that also improve economic outcomes for end-users. For the tariff subsidy program, we find that lowering the cost of electricity by up to 75 percent substantially increased consumption, albeit with mixed signals for whether overall revenue could be maintained at a lower tariff. Overall, neither program by itself facilitated the sustained use of mini-grid power and adequate revenue generation, highlighting the need for further research into strategies that enable electricity consumption for livelihood development while driving a profitable business model for mini-grids serving low-income communities.
1.1.3 Converting Fishing Boats for Electric Mobility to Serve as mini-grid Anchor Loads

To evaluate a novel strategy to stimulate electricity use for human development in addition to bolstering mini grid business models in rural communities, this contribution evaluates the potential for converting diesel-based fishing boats in Lake Victoria to electric motor and battery-based systems that can provide a crucial anchor load for a nascent 650 kWp hybrid solar-battery-diesel mini-grid. We conduct a survey among fishing boat operators \((n = 69)\) to characterize the target population and deploy a custom tracking system to measure fishing boat movement patterns. Using these primary data and secondary data on customer consumption, we select a candidate electric mobility system, create synthetic loads of residential and business customers, and construct technical and financial models of the complete mini-grid system. We then use these models to evaluate the excess capacity on the mini-grid for electric boats, the trade-offs between electric mobility and manufacturing on the mini-grid, and the impacts of demand response capabilities for charging the boats. We find that electric boat charging contributes to at least 17% more daily consumption, resulting in substantial technical and financial value to the mini-grid system, though perhaps at the cost of additional use of the system’s backup diesel generator. We find that adding shifting capabilities to electric boat charging can save up to 6% of diesel expenditures at little to no impact on the system’s Net Present Value. We also find that fishing boat owners could save up to $20,000 in 5 years from avoided fuel costs. We combine these mini-grid-scale evaluations with design considerations for a future boat tracking system, guiding mini-grid designers and operators to incorporate the potentially attractive load class of electric mobility systems.
1.1.4 Planning for future transportation electrification in Nairobi considering local mobility and grid infrastructure constraints

The growing climate crisis is making a complete transformation of the world’s energy systems increasingly urgent. Transportation is one of the most significant emissions drivers globally. Technology advancements and rapid reductions in technology costs are driving the deployment and adoption of electric vehicles (EVs) globally, as a viable alternative to gasoline-powered vehicles, in response to environmental and public health concerns. In Sub-Saharan Africa (SSA), EVs are still very nascent, and their adoption pathway is far less clear. Countries in the region present a set of unique challenges, including the dominance of used car imports, rapid growth in motorization, significant power grid constraints, poor urban air quality, a unique transportation culture, and low vehicle affordability, that require a thoughtful approach to EV adoption to balance the trade-offs between economic growth, sustainable development and meeting climate and public health goals. In this contribution, we utilize a data-driven approach to evaluating vehicle electrification pathways for an urban city in sub-Saharan Africa, considering the unique local transportation culture and mobility patterns, as well as the constraints of the distribution grid. Our study is carried out in Nairobi, the most populous city in Kenya. We explore the impacts of different electrified transportation options on the electricity distribution grid, taking into account several non-binding vehicle electrification targets outlined locally and globally. Specifically, we investigate how various types of EVs (private, commercial, motorcycles, and paratransit vehicles) and different consumer behaviors (low and high range anxiety) could affect the loading on distribution transformers under different scenarios. We leverage various independent data sources and stochastic models to fill in the data gap of transportation data, which is fundamental for this work. Hourly simulations of electrified transportation are carried
out for Nairobi, incorporating granular models for driving patterns, charging decision models, and comprehensive modeling of the load on distribution transformers. Our results suggest that electrification of transportation in Nairobi could be beneficial for the technical performance of the power system, as evidenced by improvement in the load factor under specific scenarios upon adding load from EV charging. However, our results support the argument that the distribution network will be a significant bottleneck to the broad diffusion of EVs in Nairobi – we find that under all scenarios, at least 1% and up to 8% of the zones will experience transformer overloaded with added EV charging load.

1.1.5 Using Satellite Data to Detect Diesel-Powered Irrigation for Guiding Electrification in Ethiopia

To ensure that electrification stimulates economic growth, electricity demand constraints must be addressed at every stage of electrification, especially the planning stage. Demand stimulation has increasingly become an integral part of electrification planning in sub-Saharan African countries. Agriculture is the dominant source of economic growth in Ethiopia. As such, one of the goals of Ethiopia’s National Electrification Plan is to maximize the development impact of electrification by prioritizing grid access to areas with the highest potential for irrigation and agricultural processing. In this contribution, we develop a novel approach to mapping areas with existing diesel-powered irrigation in Ethiopia by combining ground survey data with satellite-measured pollution data and supervised binary classification techniques. These areas can then be targeted for electrification and serve as the grid’s first productive use load. Combustion of fossil fuels from the motorized pumps releases pollutants (mainly nitrogen oxides ($NO_x$), carbon monoxide (CO), and particulate matter (PM)) into the atmosphere, which can be measured remotely.
We conduct on-the-ground surveys to collect a first-of-its-kind comprehensive data set on the locations and measurements of cultivated plots in Ethiopia, as well as their crop cultivation and irrigation practices, including the method of obtaining water for irrigation in the case of irrigated plots. Then, we present two classification approaches for classifying cultivated plots into three classes: not irrigated, irrigated using diesel pumps, and irrigated with other methods (non-diesel). The first approach directly classifies all observations into three classes. The second approach is a dual-stage binary classification, first classifying irrigated and non-irrigated observations and then classifying the true irrigated observations into two classes: irrigated using diesel pumps and irrigated with other methods (non-diesel). We then evaluate the precision and recall of our models and find that the best performing model in the second classification approach achieves a precision of 100% and recall of 94%, meaning that it correctly identifies 94% of the diesel irrigated observations, and all the observations that it identifies as diesel irrigated are in fact diesel irrigated.

### 1.2 Thesis Outline

The remainder of the thesis is structured as follows: Chapter 2 provides a background on demand stimulation, lessons learned from the history of demand stimulation, and discusses what makes demand stimulation a unique problem in SSA. Chapter 3 provides background on existing strategies and policies in SSA aimed at facilitating electricity use for income generation and subsequent socio-economic development. Chapter 4 evaluates the electricity consumption patterns of decentralized energy customers in East Africa. Chapter 5 evaluates strategies that stimulate organic electricity consumption among mini-grid customers in East Africa. Chapter 6 presents a novel technique for converting fishing boats for electric mobility to bolster mini-grid business models.
presents a data-driven approach to evaluating potential vehicle electrification pathways in Nairobi using a behaviorally realistic model. Chapter 8 presents a novel technique for mapping areas with existing diesel-powered irrigation in Ethiopia. Finally, chapter 9 presents the thesis summary.
2.1 Why Demand Stimulation?

The main goal of electrification is to enable activities that use power. However, electricity service must be reliable and affordable to achieve this goal. In addition, customers should be able to access and afford domestic and commercial appliances that use the electricity provided. Electricity access programs in sub-Saharan Africa have made great strides in increasing electricity generation and customer connections. However, many newly-connected customers consume limited amounts of electricity, with limited growth over time [89], resulting in cumulative demand far less than supply. New customers cannot grow their consumption because they often have limited access to or cannot afford appliances that consume electricity and have limited income to support electricity purchases. To alleviate this chicken-and-egg problem, it is imperative to implement demand stimulation programs that either facilitate new customers to organically grow their consumption over time or develop ancillary businesses that consume electricity directly from the grid/mini-grid, such as the strategies considered in this study. Adding consumption to the system creates a virtuous cycle whereby developers can recoup system costs from increased revenue and therefore afford to lower the unit cost of power, enabling customers to grow their consumption further and
ultimately realize the intended benefits of electricity access.

2.2 Lessons from History

Demand stimulation has a rich history. The rural electrification journeys of several developed and developing countries in the 20th century were riddled with the same electricity demand challenges that are being experienced in SSA today, such as unaffordable costs of both electricity and the electrical appliances to connect in low-income communities, as well as lack of knowledge of how to leverage electricity services to their benefit. It is valuable to examine the historical approaches taken to tackle these challenges, as they can guide electricity consumption growth in SSA.

Appliance financing programs

By 1932, only 10% of the predominantly low-income rural farming communities in the United States had access to electricity compared to over 90% of the urban population[37]. Private-owned utilities were reluctant to invest in the rural market where consumption was low and costs were high as it would be difficult to recoup the costs of their investment [206]. The US government, therefore, took on the responsibility of providing electricity services to the remaining population by forming both the Rural Electrification Agency (REA) and the Electric Home and Farm Authority (EHFA) in 1935, which worked in partnership to grow both supply and demand concurrently in rural areas. The EHFA provided low-cost, long-term financing services to electricity customers to purchase approved appliances and approved tariffs set by partnering utilities who connected these customers. In addition, they organized roadshows to educate customers on the value, safe operation, and maintenance of electrical appliances [59]. By 1953, 90% of US farms had access to and were consuming electricity [194].
**Tariff subsidy programs**

National rural electrification programs in developing countries such as Nepal, Bangladesh, Costa Rica, and the Philippines, aided by the US, adopted the rural electric cooperatives model to distribute electricity to rural customers [202, 90]. In addition to offering microfinance loans to their members to purchase appliances, the cooperatives, which were responsible for producing, selling, consuming, or distributing energy, set up a tariff structure that cross-subsidized low-income domestic and agricultural customers by charging them tariffs below the cost of service, with industrial and commercial customers who were charged tariffs above the cost of service. The cooperatives then received grants from the government to meet revenue shortages [205, 237]. This strategy ensured that low-income electricity customers could afford the cost of electricity.

**Promoting productive uses of electricity**

Another acclaimed rural electrification success story is Vietnam. In just over 30 years, Vietnam successfully grew its electricity access rate from 2.5% to about 96% between 1975 and 2009. One important feature that contributed to the success of Vietnam’s rural electrification program is the prioritization of economic growth from the very beginning. Policymakers prioritized the electrification of areas with high potential for growth in productive uses of electricity, which increased government revenue and household incomes, leading to higher consumption, thus promoting the overall financial viability of rural electrification [97]. On the contrary, the key drivers of electricity access attached to SDG 7 focus primarily on the percentage of people with electricity access at home and not at all on the productive uses of electricity.
2.3 What Makes Demand Stimulation a Unique Problem in Sub-Saharan Africa?

Although country-specific factors and conditions played a role in the success of the demand stimulation programs discussed in section 2.2, these approaches provide valuable lessons that could be applied in different geographical contexts. However, certain current landscapes in SSA, trivial in the historical cases discussed, make demand stimulation a unique problem in SSA.

1. **Development of new electrification technologies:** The US and European countries relied on public utilities and subsequently public funds to extend the national grid to rural areas and provide universal electricity access by the mid of the 20th century [206]. Owing to technological advancements in photovoltaic systems and improved battery systems, there has been a rapid growth in low-cost off-grid alternatives to grid electrification in the past decade, most notably solar lighting devices, solar home systems, and mini-grids [103]. The International Energy Agency estimates that currently, 33 million people have access to electricity with decentralized systems and that they are the least-cost pathway to deliver universal electricity access by 2030 in SSA for 60 percent of households that do not currently have access. That would entail about 200,000 mini-grids and millions of solar home systems [7].

2. **Relatively lower income levels that call for a hybrid public/private approach to electrification:** Countries in SSA are going through electrification at far lower income levels today compared to developed countries when they were going through electrification – the average income in the US was about US $6,000 per capita per year at a 20% electrification rate, whereas by 2018, the
average income in SSA was about US $1,600 per capita per year [216]. Delivery of electricity services to meet the universal electricity access target of 2030 would require an investment of about US $24 - 49 billion per year [162]. This means that there will likely be a need for significant private sector investment in off-grid rural electrification, given the limited funds from public sources and, at the same time, a greater need for subsidies to ensure affordability for low-income people [277]. The high financial risk associated with investing in low-income areas and the increased participation of the private sector, therefore, means there is a much higher emphasis on returns on investment for stakeholders investing in rural electrification in SSA.

3. **Different digital and financial infrastructure**: The innovation, development, and growth in digital and financial technology solutions, particularly devices and software that collect, store, process, and transfer digital data, has been a real game changer in the off-grid energy sector in SSA [170]. In particular, the creation of mobile-money services and digital smart meters has enabled the innovation of Pay-As-You-Go (PAYG) systems, which provide electrification pathways to low-income customers by offering a reprieve from bulk upfront or monthly electricity payments by allowing for smaller regular payments over time through mobile banking. It has become a crucial component of solar home system business models [2, 17].

4. **The concurrent decarbonization push and confounding factor of energy efficiency**: The world is currently involved in a massive shift from fossil fuels towards renewable energy to try to curb climate change. Therefore, efforts to increase electricity access globally are coupled with climate change mitigation strategies. This was far from a concern when most industrialized nations were
undergoing electrification in the early 20th century – coal was the dominant fuel of the industrial revolution [196]. Demand stimulation is counter-intuitive to the energy efficiency zeitgeist that calls for reducing electricity use to reduce fossil fuel use. While it is possible that demand stimulation strategies could contribute to greenhouse gas emissions, there have been recent developments in efficient and super-efficient appliances which can enable the implementation of demand stimulation strategies in an innovative, efficient way [63, 209].
CHAPTER 3

RELATED WORK

This chapter provides a background on existing strategies and policies in sub-Saharan Africa aimed at facilitating electricity use for income generation and subsequent socio-economic development.

3.1 Facilitating Organic Consumption Growth

Despite households in SSA willing to spend more than 10% of their monthly budget to consume electricity[239], both urban and rural electricity customers are exhibiting slowing electricity demand growth [88]. While it is possible that some customers could actively choose to decrease their consumption, others may be precluded from enjoying the benefits of using electricity to improve their livelihoods due to a myriad of issues such as unpredictability and low levels of income, poor reliability of power, lack of access to financing to purchase appliances, unavailability of appliances, the prohibitive cost of electricity or a combination of these factors [28, 27, 172]. Current strategies and policies in SSA aimed at facilitating organic consumption growth of grid and off-grid electricity customers are geared towards addressing these barriers.
3.1.1 Appliance financing programs

The usage of electricity fundamentally depends on consumers obtaining the necessary appliances and equipment to connect. The costs of equipment and appliances are often too high for low-income customers to afford, with many currently living in extreme poverty, which is below $2 a day. In addition, customers living in remote locations may lack access to efficient and compatible appliances and equipment, particularly mini-grid customers. Many do not have access to formal credit and, as such, rely on informal credit such as borrowing from family and friends [15]. Consequently, the lack of credit history and collateral makes it challenging to offer financing to low-income electricity customers to purchase appliances. Electricity providers are uniquely placed to provide their customers with financing for appliances since they possess rich demographic and socio-economic data on their customers from pre-electrification surveys, historical electricity payment data, and established payment platforms, as well as collateral stakeholders who must understand those who do default on their loans. Evidence from a recent study showed that the potential profitability of appliances financing programs hinges on financing appliances that can be used for income generation, suggesting that stakeholders must understand the drivers for appliance uptake and how customers use and desire to make use of electricity services [185].

3.1.2 Electricity tariff subsidy programs

Even if appliances were given to electricity customers for free, likely, many low-income electricity customers might not afford to use some, if not most of them, as they may find electricity payments too high. With the prevalence of low connection rates in SSA [27, 23], many utilities and off-grid companies have been forced to subsidize connection charges in low-income areas or fully finance them and allow customers to pay
back in installments. This means that they rely on subsidies, which are often limited, and revenue from electricity payments for cost-recovery, which could take over 40 years [88]. In the case of utilities, many African governments set regulated and uniform tariffs, often below cost-recovery levels, to ensure fairness and affordability [222]. However, mostly privately owned mini-grids emphasize cost recovery, as they are ultimately accountable to their investors. As such, they are forced to charge significantly higher tariffs, which only about 10 – 15% of their customers can afford [211]. The fundamental question is if mini-grid operators lowered their tariffs, would their customers grow demand enough to ensure the financial sustainability of the mini-grid? Ultimately, electricity is of no benefit to customers if they cannot use it. In a limited study, the Mini-grid Innovation Lab found that lowering tariffs by 50 – 75% negatively impacts the mini-grids financial viability by reducing net present value (NPV) by 13% despite significant growth in consumption [184], which means that for tariff reduction programs to be successful, they must be accompanied by hybridized approaches such as external subsidies, cross-subsidization or tapping into alternative revenue streams such as integrating productive electricity uses [222].

3.2 Productive Uses of Electricity

Electricity use in rural settings in SSA is dominated by electricity for lighting, mobile phone recharging, and other media and information technologies such as radio and television. However, additional existing and unexplored energy uses in agriculture, transportation, and cooking could be electrified to generate local economic development of newly-connected rural communities and, simultaneously, ensure the financial sustainability of rural electrification projects.
3.2.1 Electricity for Transportation

Affordable and reliable rural transportation and mobility are key drivers of local development. It enables economic livelihood and other benefits such as healthcare, education, markets, and social interactions. Most of the rural population in SSA walk, cycle, or use public transportation mainly in the form of two and three-wheeler vehicles, as they are cheaper and adapted to the poor road infrastructure [249].

Today, nearly all transport activities in SSA are non-electric. The rapidly growing global market for electric vehicles owing to the development of cheaper, more efficient batteries, coupled with the global climate goals of lowering emissions and the growing burden of fuel dependency, have propelled several initiatives focusing on electric mobility, particularly of two-wheelers in rural areas of SSA [26]. However, these initiatives have been limited to small pilots that involve recharging electric two and three-wheelers at privately owned solar charging stations. Adopters of electric motorcycles and electric outboard motors in pilot studies in East Africa have reported daily profit margins of up to $20 [249, 4]. There are also undoubted fuel cost savings for adopters of electric vehicles in rural and urban settings of SSA [225, 251].

As the EV (Electric Vehicle) market in SSA continues to grow, in addition to their environmental and economic benefits to consumers, electric mobility as either a dispatchable load, distributed storage to the grid, or both, presents a latent opportunity to shore up utility and mini-grid business models in both rural and urban settings, by improving capacity utilization and subsequently enabling a decreased LCOE and added revenue generation for the electricity providers [96].
Electric Mobility as Flexible Demand

The use of electric mobility loads as flexible demand is a well-studied topic in the literature. Previous work has used a variety of techniques for improving the use of this flexible resource, including better predictions of arrivals using fluid dynamic models [14], optimal charging schedules when faced with unknown future demand [254], a Markov Decision Process (MDP) framework for making charging decisions [288], heuristic algorithms for an NP-hard construction of the EV charging problem [290], genetic algorithms that consider grid parameters [10], model-free reinforcement learning techniques to coordinate multiple charging stations [148], and a technique based on Distributed Resource Allocation for smoothing grid operations using EVs [181]. While some or all of these techniques may apply to our scenario, they have been primarily evaluated on centralized electricity grids, which exhibit substantially different constraints than decentralized mini-grids, and for electric cars, which have different usage patterns and requirements as compared to boats.

In mini-grids, demand response (DR) may be used for optimizing grid operations through load scheduling and load control to achieve higher efficiencies, saving fuel for backup needs, decreasing operation expenses, providing grid resilience, and delaying the need for further investments [191]. Various DR strategies using EV charging have been explored, such as influencing the behavior of the users by reducing the EVs’ trip distance or trip time shifting [243], day-ahead planning of EV schedules due to fluctuations in daily renewable generation [40], time-varying pricing intended to shift load from peak to off-peak periods [193], incentive-based DR programs that support vehicle-to-grid for load shifting and congestion management [218] and controlled EV charging DR programs [242].
### 3.2.2 Electricity for Agriculture

Most people living in rural areas in SSA, primarily poor, depend on small-scale agriculture for all or part of their livelihoods. Therefore, these communities’ agricultural value chain investments will likely contribute to individual and community development. In rural areas of SSA, the critical potential sources for electricity demand are in replacing inefficient fossil-fueled and manual irrigation and post-harvest and primary processing activities such as milling, cassava processing (chipping), drying, chilling and cold storage, meat processing, and oil extraction [204, 16]. Current research efforts have focused on exploring the economic viability of integrating some of these productive agricultural uses in the context of both grid extension and mini-grids.

It is estimated that electricity demand from agriculture will double, reaching about 9 GW by 2030, with irrigation being the single largest potential source of rural agricultural electricity demand in comparison to agro-processing demand [16]. The financial viability of targeted grid extension to power irrigation activities is highly contingent on favorable financing terms, such as having access to concessional loans or grants for capital investment and the ability to lower costs and sell sufficient power. Since grid tariffs are regulated, subsidies may be required to realize grid extension projects to power agricultural activities. That said, investments in rural electrification and agriculture integration are justified by a positive net present value (NPV) when the overall socio-economic benefits such as job creation, better outcomes for connected households, improvement in agricultural yields, and market access are considered. Farmers also gain substantial profits from increased yields [16]. The seasonality of irrigation loads can significantly impact the cost recovery of electricity supply investments. As such, it is essential to mitigate this risk by allowing for multi-cropping and complementing irrigation loads with post-harvest processing loads.
In off-grid communities, post-harvest processing loads, such as maize milling, rice milling, and cassava grating, could decrease mini-grid Levelized cost of electricity (LCOE) by up to 42% from improved capacity utilization, subsequently enabling tariff reductions by 6 – 19% while still earning a 15% internal rate of return (IRR) for investors. At the same time, crop processors could make a substantial profit over the investment lifetime [16, 146, 233]. Besides post-harvest processing loads, electric chilling and cold storage technologies for preserving perishable goods, such as milk and fish, could contribute to food security and increased incomes by reducing food wastage and, at the same time, improve the technical performance of the mini-grid and generate added revenue for the mini-grid operators [57, 33]. However, just as in the case of the utilities, the financial viability of integrating agricultural loads in mini-grids is sensitive to capital investment costs [16], suggesting that replacing existing non-electric agro-processing equipment in already financially-sustainable businesses may be preferable and more sustainable than creating and growing new businesses.
CHAPTER 4

CUSTOMER PROFILES AND PATTERNS FOR

DECENTRALIZED CUSTOMERS IN EAST AFRICA

4.1 Background and Motivation

Off-grid electricity systems, that is, Solar Home Systems (SHS) and mini-grids, are emerging as an essential driver of rural energy access, complementing grid extension. The World Bank estimates that tens of millions of people now have access to electricity through SHS, and 47 million are connected to mini-grids globally [75, 76]. Further to the global universal electrification goal of the United Nations Sustainable Development Goal 7, the IEA projects that by 2030, 470 million people in sub-Saharan Africa will need to receive electricity from off-grid solutions (roughly 60% from mini-grids and the balance from SHS) [125]. However, in contrast with traditional grid extension, planning for off-grid electrification in sub-Saharan Africa has primarily been an informal, bottom-up, market-driven, and decentralized phenomenon. It is constituted ad hoc with different stakeholders, such as off-grid system manufacturers, Rural Electrification Agencies, donors, and financial institutions, who could benefit from both the off-grid market and customer intelligence [220]. Yet, there is still limited information about the behavior, preferences, and consumption patterns of off-grid electricity customers in the region. People adopt these systems based on a combination of their financial capabilities,
their volition for these electricity services relative to their needs (i.e., their demand), and the availability of the services where they live, which presents a fast-changing picture that is difficult to piece together because it is a composite of different private operations. Nonetheless, some robust studies \cite{140, 25, 11} have explored how customers’ demographic, socioeconomic and geographic characteristics influence the adoption of SHS and mini-grids.

The economic viability of off-grid systems is contingent on the premise that off-grid customers can afford to pay for and use the electricity services or systems provided. As such, mini-grid companies, in particular, have a vested interest in applying the knowledge of the consumption patterns of their customer base in consumption prediction for mini-grid sizing; over-sizing systems via overprediction reduces the financial viability of the systems while under-sizing systems via underprediction lowers the reliability and technical sustainability of the systems. The lack of publicly-available smart meter data has resulted in a limited body of work that examines the electricity spending and consumption behavior of off-grid customers in these regions. Regardless, a limited body of work \cite{276, 275, 69, 29} has analyzed the load profile patterns of mini-grid customers to inform the optimization of mini-grid designs. On the other hand, some studies look at the load profile patterns of SHS customers to explore the potential for load sharing with neighboring households \cite{60} and the implications of average daily load estimation error on system sizing \cite{161}. However, comparing and contrasting these studies and analyzing the characteristics of decentralized customers as a whole is challenging. In addition, to the best of our knowledge, no study examines how the ability to pay for electricity of both groups of off-grid customers, characterized by their financial spending patterns, relates to their electricity consumption behavior.

This chapter provides a first-of-its-kind examination of the financial expenditures, electricity consumption patterns, and behaviors of both mini-grid and SHS customers.
We situate our work in East Africa, a region with fast growth in both technologies. Specifically, we examined the socioeconomic and demographic characteristics, electricity consumption, and spending behavior of over 2,000 mini-grid customers in Tanzania in 2017 and over 13,000 SHS customers in Kenya in 2018. We perform our analysis with the caveat that we cannot prove that our samples of SHS and mini-grid customers represent the entire markets of SHS and mini-grid customers. The samples are biased towards the customers available to businesses that sell and operate these systems and reflect the many unquantifiable factors that went into the distribution of these modes of electrification. Although the data used in this study is for SHS and mini-grid customers residing in different countries over different time periods, due to the novelty and rapid growth of these technologies and the unique nature of the data sets, which include similar variables across these different customer types, we believe a relative analysis of the two customer groups and their electricity spending patterns and consumption behaviors can be valuable for understanding broad similarities and differences among off-grid customers, projecting growth, and informing policy, regulation, and investment in these sectors.

We describe the data sets and methodology used to characterize decentralized customers’ electricity consumption patterns and behavior in Section 4.2 and 4.3, respectively. Subsequently, we present and discuss the results of our analysis in Section 4.4. We conclude by discussing the policy implication of this study for different stakeholders in the off-grid sector, such as off-grid companies, policymakers, electricity planners, investors, and development partners in Section 4.5.
4.2 Data

The East African Community – especially Kenya, Tanzania, Uganda, and Rwanda – has emerged in recent years as a hub for the innovation and adoption of off-grid systems within sub-Saharan Africa, accounting for over half of the total global investments in off-grid solar in 2017 [126]. One reason for the relatively higher growth of the off-grid solar market in East Africa is the parallel growth in mobile money in countries like Kenya and Tanzania and relatively inexpensive smart meters, which have given rise to innovative business models such as flexible Pay-As-You-Go (PAYG) solutions that make off-grid electricity affordable and accessible to people with low and uncertain incomes.

This study leverages data from BBOXX, an off-grid solar company operating in 11 African and Asian countries [19]. The product in this study is a 50 $W_p$ smart SHS, with a 17 $Ah$ energy storage unity and a remote monitoring system. Typical SHS products in the region range from 8 $W_p$ to 200 $W_p$, with systems 50 $W_p$ and above categorized as large SHS. Households that purchased larger systems are more likely to live in peri-urban areas than those that purchase smaller systems [107]. The SHS customers in our analysis were in on a PAYG fee-for-service plan since the beginning of their contracts, allowing them to only pay for the energy they use [127].

Further, we examine mini-grid customers from 22 mini-grid sites in Tanzania, ranging in capacity between 5 $kW_p$ and 50 $kW_p$ and between 20 and 220 connections per site. The mini-grid billing systems also employ a PAYG model, whereby customers purchase units of energy, much like a traditional utility model, consequently linking mini-grid revenues to the electricity consumption of customers. The only difference between the SHS and mini-grid PAYG system is that SHS customers can only purchase and consume up to the energy limit \(^1\) of the SHS package in their contract. At the same time, there

---

\(^1\)Energy limit is the amount of energy that can be drawn from the battery each day and is determined by the estimated energy demand of the appliances used by the customers.
is no ceiling on the amount of electricity that mini-grid customers can purchase and consume.

This study uses two categories of data for both SHS and mini-grid customers: a demographic and socioeconomic data set and a behavioral data set, which is summarized in Table 4.1. The demographic and socioeconomic data are collected before a customer purchases a solar home system or is connected to a mini-grid. The behavioral data are collected once customers are connected to the off-grid system and consist of electricity consumption and payment data.

**Table 4.1: Summary of data used in the study**

<table>
<thead>
<tr>
<th>Customer type</th>
<th>Country</th>
<th>Year</th>
<th>No. of customers</th>
<th>No. of observations</th>
<th>No. of customers</th>
<th>No. of observations</th>
<th>Cons. data observations</th>
<th>Payment data observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHS customers</td>
<td>Kenya</td>
<td>2018</td>
<td>86,121</td>
<td>86,121</td>
<td>13,562</td>
<td>91,066,351</td>
<td>125,925</td>
<td></td>
</tr>
<tr>
<td>MG customers</td>
<td>Tanzania</td>
<td>2017</td>
<td>3,186</td>
<td>3,186</td>
<td>2,347</td>
<td>12,489,774</td>
<td>112,600</td>
<td></td>
</tr>
</tbody>
</table>

### 4.2.1 Demographic and socioeconomic data

Demographic and limited socioeconomic information was available for 86,121 SHS customers, and a more comprehensive data set was available for 3,186 mini-grid customers. This information enables us to compare the demographic characteristics of SHS and mini-grid customers based only on the intersection of variables in both data sets, which include: the age and gender of the household head, the primary source of income for the household, and the previous source of lighting.

Given this limited socioeconomic data available to characterize these off-grid customers, we leverage census microdata for the study areas available from the Integrated
Public Use Microdata Series- International [178], to explore the relative socioeconomic status of the populations in the areas where the SHS and mini-grid customers are located. We examine information on asset ownership, housing characteristics, and access to utilities (i.e., water, electricity, and sewer services). This asset-based approach to determining socioeconomic status has been widely used in previous studies [227, 174, 68, 228] as an appropriate measure of household wealth. We use the 2009 Kenya IPUMS census microdata with 895,230 households and the 2012 Tanzania IPUMS census microdata with 950,776 households. While we could not obtain the exact locations of the SHS customers, the demographic data set also included the name of the nearest school to each SHS customer. We, therefore, utilize a data set of school names and locations in Kenya to obtain the location of the closest school to each SHS customer.

On the other hand, the demographic data set of mini-grid customers included GPS coordinate information for each customer. We calculated a weighted average of each indicator based on the proportion of SHS in each of the 87 constituencies in Kenya in which they are located. We then compared it to a weighted average of the corresponding indicators based on the proportion of mini-grid customers in each of the 20 districts in Tanzania where they reside. We assume that the relative differences in wealth among these areas do not change in the period leading up to the study.

4.2.2 Electricity consumption and payment data

The SHS in this study have an integrated central unit that, in addition to its function as a charge controller and energy storage unit, is also a measurement and data communication unit. It records electricity consumption in 15-minute intervals and directly uploads readings to a central server. The electricity payments made by the customers are linked to their consumption data via the SHS product International Mobile Equipment Identity (IMEI) number. Mini-grid developers install electronic smart meters on their
customers’ premises, which record electricity consumption in 15-minute intervals and upload the readings to a central server. The electricity payments made by the customers are linked to their consumption data via the smart meter number. While we obtained the solar home system consumption readings at a 15-min resolution, we obtained the mini-grid consumption readings at an hourly resolution. For this reason, our unit of analysis of the consumption data is hourly.

For this study, we analyze a year’s worth of data for both sets of decentralized customers. For the SHS customers, we use data from Jan - Dec 2018, after the company transitioned to a pay-as-you-go system in late 2017. On the other hand, we used data for mini-grid customers in Tanzania from Jan - Dec 2017 because mini-grid developers launched interventions that selectively influenced the electricity consumption behavior of some customers in the mini-grid sites used in this analysis at the start of 2018. In total, we obtained both consumption and payment data for 13,562 SHS customers and 2,347 mini-grid customers. As previously mentioned, publicly available smart meter data for off-grid customers are sparse; previous studies on mini-grid customers [29, 276, 275] have considered sample sizes of up to 900 customers.

### 4.2.3 Study areas

The SHS customers in our study are distributed across 87 constituencies in Kenya, while the mini-grid customers are distributed across 20 wards in Tanzania. First, we consider the representation of the decentralized systems in these areas. SHS and mini-grid companies carefully select customers to develop economically viable business models. SHS companies target customers who can afford their products in the regions where they operate. On the other hand, mini-grid companies first choose a village, and then in a particular village, they connect customers that can afford the connection fees. It is, therefore, likely that mini-grids serve a larger proportion of the population in a
given area and may connect a more socioeconomically diverse population. We test this hypothesis by calculating the population ratio within a 2 km radius of the central point of mini-grid sites connected to the mini-grid and the proportion of the population within a 2 km radius of the schools near SHS customers. We leverage the WorldPop dataset, a gridded geospatial population count data set [280]. We examine the distributions of these proportions in Figure 4.1. We observe that, on average, the mini-grids connect 45% of the population in the regions where they operate. At the same time, the SHS customers make up about 35% of the population in the areas where they reside.

We also examine the relative socioeconomic status of populations in the study areas by calculating the International Wealth Index (IWI), an asset-based index (between 0 and 100) of a household’s economic status that is comparable across time and countries in the developing world [241]. It measures wealth by looking at a household’s possession of durables, access to basic services, and characteristics of the house in which it is living. The methodology of constructing the IWI is described in detail in [241]. Using Equation 4.1, we calculated each household’s IWI score in the administrative units where each customer group is located. Then, we calculated the average IWI for SHS and mini-grid administrative unit areas by weighting the scores by the square root of the total population size of the administrative units.
where $\beta_n$ are the re-scaled asset weights for each asset $x_n$

The areas where the SHS and mini-grid customers in our study are located are socioeconomically similar, with average IWI values of 28.0 and 28.6, respectively. For further context, [104] calculated Kenya’s 2017 average IWI as 43.7 and Tanzania’s as 34.8. Therefore, the regions where both off-grid customer groups in our study live are below the average country-level economic statuses. We also consider the use of nighttime luminosity as a proxy for socioeconomic indicators. It is a supplementary approach or substitute for existing measures of socioeconomic activity that is increasingly used by social scientists and development practitioners in regions where socioeconomic data are sparse or unavailable [48, 176, 273]. However, we find that nighttime lights by themselves would likely not be an effective method or proxy for measuring economic activity in off-grid regions as has successfully been done in urban areas or electrified regions with higher population densities (See Appendix A).

When we also examine the agricultural assets in the study areas, we observe that areas with mini-grid customers are more agriculturally intensive, evidenced by the ownership of larger herds of livestock (Table 4.2). This is also the case when we only examine the socioeconomic and demographic characteristics of the SHS and mini-grid customers in our study, summarized in Table 4.3. We observe that a plurality of households with SHSs derives their primary source of income from salary or wage work. In contrast, a plurality of households connected to a mini-grid derives their primary source of income from agricultural sources. This would suggest that mini-grid households may be more subjected to income fluctuations due to seasonality, particularly during the lean season. At the same time, SHS customers are likely to have a steadier income stream. This could
significantly impact when mini-grid households can pay for electricity. We, therefore, control for seasonality in our subsequent analyses.

**Table 4.2:** The weighted relative wealth of populations in the administrative units where the SHS and mini-grid customers are located.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Admin units with mini-grid customers</th>
<th>Admin units with SHS customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>5,937,589</td>
<td>16,889,744</td>
</tr>
</tbody>
</table>

**Asset-based wealth index**

Average Weighted International Wealth Index

<table>
<thead>
<tr>
<th>Variable</th>
<th>Admin units with mini-grid customers</th>
<th>Admin units with SHS customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number of cows</td>
<td>8.3</td>
<td>1.8</td>
</tr>
<tr>
<td>Average number of goats</td>
<td>8.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Average number of sheep</td>
<td>5.3</td>
<td>0.7</td>
</tr>
<tr>
<td>Average number of poultry</td>
<td>6.9</td>
<td>5.4</td>
</tr>
</tbody>
</table>

**Agricultural assets**

**Table 4.3:** Demographic and socioeconomic characteristics of decentralized customers.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SHS customers (SD in parentheses)</th>
<th>mini-grid customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of head of household</td>
<td>37.7 (11.9)</td>
<td>41.4 (11.8)</td>
</tr>
<tr>
<td>Female head of household</td>
<td>18%</td>
<td>15%</td>
</tr>
</tbody>
</table>

**Primary source of income for household**

<table>
<thead>
<tr>
<th>Source</th>
<th>SHS customers</th>
<th>mini-grid customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural</td>
<td>24%</td>
<td>47%</td>
</tr>
<tr>
<td>Business/enterprise</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>Salary or wage work</td>
<td>48%</td>
<td>12%</td>
</tr>
<tr>
<td>Trading/commerce</td>
<td>1%</td>
<td>21%</td>
</tr>
<tr>
<td>Other</td>
<td>7%</td>
<td></td>
</tr>
</tbody>
</table>

**Previous source of lighting**

<table>
<thead>
<tr>
<th>Source</th>
<th>SHS customers</th>
<th>mini-grid customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batteries</td>
<td>9%</td>
<td>1%</td>
</tr>
<tr>
<td>Lantern</td>
<td>55%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Candles</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>SHS</td>
<td>-</td>
<td>23%</td>
</tr>
<tr>
<td>Other</td>
<td>16%</td>
<td>20%</td>
</tr>
</tbody>
</table>

Further, we consider the weather patterns of the study areas during the analysis period. We examine the daily average temperature and precipitation in the areas where SHS users are located in 2018 and in the areas where mini-grid customers are located in
2017. As shown in Figure 4.2, we find that the higher/lower temperature seasons and rainy/dry seasons coincide in the areas where SHS and mini-grid customers are located across the different years. However, the areas where SHS customers reside have higher rainfall values on average during this period. The similarity in weather patterns further enables a broad comparison between the two customer groups. Nonetheless, we control for variations in weather in our subsequent analyses.

![Figure 4.2: Temperature and precipitation patterns of areas where mini-grid and SHS customers. Notes: Temperature and precipitation data for SHS customer areas is from Jan - Dec 2018, and temperature and rainfall data for mini-grid customer areas is from Jan - Dec 2017](image)

SHS customers are mainly residential customers, with only six indicating to be a company, while about 18% of the mini-grid customers are small commercial (enterprise) customers. We also note that most households with SHSs previously used lanterns for lighting, while about a quarter of the households connected to a mini-grid previously
used a SHS for lighting. A common hypothesis is that economic growth shifts energy users up the "energy ladder," suggesting that mini-grid customers have a higher ability to pay for electricity. We test this hypothesis in our subsequent analyses.

Concerning appliances, the SHS customers can choose an appliance package, which would include a set from LED light bulbs, a rechargeable torch, a radio, a television, and a multiple phone charger with five connectors. The power rating of each appliance is summarized in Table 4.4. While we do not have information on the specific appliance ownership of the SHS customers, we have data on their energy limits. The energy limit, which is the amount of energy that can be drawn from the battery each day, is determined by the estimated energy demand of the appliances used by the customers, measured in watt-hours (Wh). We calculated the mean energy limit of the SHS customers to be 64.46 Wh/day with a standard deviation of 32 Wh/day and a range of 50 - 204 Wh/day. On the other hand, mini-grid customers own a range of appliances. Their appliance ownership and use are only limited by the availability of appliances and their financial ability to purchase them. Nonetheless, as shown in Table 4.4, most appliances owned by mini-grid customers are similar to the SHS appliances i.e., light bulbs, phone chargers, radios, and televisions; 55% of the mini-grid customers reported owning only these appliances.

While both customer groups are on a pay-as-you-go electricity payment model, their tariff structures differ. SHS customers use a subscription tariff, which means that their electricity payments are a function of how much energy they consume, that is, their energy limits (kWh) and the length of their contracts with the service provider. On the other hand, some mini-grid customers are on a block tariff, whereby lower tariffs incentivize increased electricity use. The others are on a time of use tariff, where lower tariffs incentivize off-peak electricity use. Since our study analyzes customers’ time-of-use behavior, we explicitly control the effect of the time-of-use tariff among mini-grid customers. Given the differences between the subscription and block tariff
Table 4.4: Appliances owned by decentralized customers.

<table>
<thead>
<tr>
<th>Appliance</th>
<th>Power/Energy rating (W)</th>
<th>Appliance ownership (%) of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHS customers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LED light bulb</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Rechargeable torch</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Radio</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Television</td>
<td>13 - 18</td>
<td></td>
</tr>
<tr>
<td>Phone charger</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Mini-grid customers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light bulb</td>
<td>7 - 30</td>
<td>75%</td>
</tr>
<tr>
<td>Phone charger</td>
<td>4</td>
<td>68%</td>
</tr>
<tr>
<td>Radio</td>
<td>18 - 50</td>
<td>23%</td>
</tr>
<tr>
<td>Television</td>
<td>36 - 120</td>
<td>29%</td>
</tr>
<tr>
<td>Sound system</td>
<td>25 - 100</td>
<td>13%</td>
</tr>
<tr>
<td>Satellite TV service</td>
<td>18 - 20</td>
<td>13%</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>470 - 760 Wh/day</td>
<td>5%</td>
</tr>
<tr>
<td>PC/printer/photocopier</td>
<td>60 - 225</td>
<td>2%</td>
</tr>
<tr>
<td>Shaver/hair clipper</td>
<td>11</td>
<td>1%</td>
</tr>
<tr>
<td>Other</td>
<td>up to 1 kW</td>
<td>10%</td>
</tr>
</tbody>
</table>

Structures, the nature of the mini-grid tariff structure could drive higher consumption among residential mini-grid customers compared to SHS customers, beyond customers’ ability to pay. This suggests that the results of our analysis could underestimate how much electricity residential mini-grid customers consume relative to SHS customers for similar spending patterns. We caveat this in the discussion of our results.

4.3 Methodology

4.3.1 Linear models

We characterize and compare the electricity consumption behavior of decentralized customers by considering two electricity parameters: the proportion of total daily electricity consumed during the evening hours of 5 pm - 10 pm, which alludes to the timing of consumption, and the total hours of high consumption, indicated by the number of hours in a day when consumption is over 5% of the total daily consumption. This parameter alludes to the sharpness or flatness of the daily load profile. We investigate
these parameters as a function of electricity spending. We utilize ordinary-least squares (OLS), linear regression models given in Eq. 4.2 and Eq. 4.3 to estimate how the variation in these electricity parameters changes with the monthly electricity spend and how this relationship compares between SHS and mini-grid customers:

$$Y_{it} = \alpha_i + \beta_{1i}X_{it} + \beta_{2i}D_{1i} + \epsilon_{it} \quad (4.2)$$

where $Y_{it}$ is the mean electricity parameter for SHS customer $i$ in month $t$, $X_{it}$ is the electricity spending by SHS customer $i$ in month $t$. We control for weather patterns, represented by rainfall $D_{1i}$ as well as seasonal effects by including a month fixed effects, $\alpha_i$.

$$Y_{it} = \alpha_i + \beta_{1i}X_{it} + \beta_{2i}D_{1i} + \beta_{3i}D_{2i} + \beta_{4i}D_{3i} + \epsilon_{it} \quad (4.3)$$

where $Y_{it}$ is the mean electricity parameter for mini-grid customer $i$ in month $t$, $X_{it}$ is the electricity spending by mini-grid customer $i$ in month $t$. Given the range of appliance ownership and the differences in tariff structures among mini-grid customers, we control for appliance ownership, $D_{2i}$, which we code as 1 for customers with the basic appliances, i.e., appliances offered to SHS customers and 0 for households with additional appliances to those offered to SHS customers, as well as tariff structure, $D_{3i}$, which we code as 1 for a block tariff, where customers are charged a different price depending on how much electricity they consume and 0 for a time-of-use tariff, where customers are charged different rates during certain hours of the day. We also control for weather patterns, represented by rainfall $D_{1i}$, as well as seasonal effects by including a month fixed effects, $\alpha_i$. 
4.3.2 Daily load profile segmentation

We also look at the mean monthly load profile patterns of SHS and mini-grid customers and how they compare. Several studies such as [215, 87, 283] have used data mining techniques to analyze daily load profile patterns of electricity consumers within a heterogeneous population. The consensus among these studies is that k-means clustering is the best-known and most frequently applied partitioning clustering technique. We, therefore, use a k-means clustering approach to the normalized mean monthly load profiles of decentralized customers over an entire year and examine how the load profile patterns compare between SHS and mini-grid customers. We construct the normalized mean monthly load profiles by taking the mean electricity consumption for each hour of the day over each month for which the customer has a complete set of data (24 hourly values). To emphasize the load shape patterns rather than the amplitude absolute value, we normalize the profiles by dividing each hourly mean by the mean daily electricity consumption so that the integral of each profile is one. The objective of this clustering technique is to split our data set, \( x \) comprising \( n \) patterns, into \( k \) clusters, \( C_1, \ldots, C_k \) such that similar load profile patterns are placed in the same cluster \( x_i, x_j \in C_k \) and dissimilar load profile patterns are grouped into different clusters, minimizing the within-cluster sum of squares given by:

\[
\sum_{i=1}^{k} \sum_{x_i \in C_k} \|x_i - x_k\|^2
\]  

(4.4)

where \( \|x_i - x_k\|^2 \) is the euclidean distance between the pattern \( x_i \) and its closest cluster centroid \( x_k \). Determining the number of clusters, \( k \), is challenging and often unclear from the data set itself. Studies on electricity consumption data segmentation using k-means clustering have used indices such as the elbow method [275, 145] and gap analysis [140] to determine the optimal number of clusters in a data set. We use an approach presented
in [46], which evaluates 30 such indices and offers the most appropriate number of clusters for the data set based on the majority of indices.

4.4 Consumption patterns and electricity spending behavior

In East Africa, off-grid business models struggle to create a cost-efficient market that strikes a balance between financial sustainability and the ability of customers to pay for power. To investigate how off-grid customers’ ability to pay for power compares, we consider the distribution of monthly expenditures over a year for SHS customers, residential mini-grid customers, and small commercial mini-grid customers. We also examine their spending behavior by considering the frequency and amount of each payment. We consider each month of spending and consumption for each customer independently as opposed to the average monthly values per user because of the high variability in both monthly spending and monthly consumption values of each customer over the year in both customer groups. As shown in Figure 4.3, our results underpin the notion that mini-grid customers are likely more financially constrained than SHS customers and subsequently have a lower electricity purchasing power. We observe that, on average, SHS customers spend substantially more on electricity than both residential and small-commercial mini-grid customers at the lower end of the electricity spending distribution; 26% of the monthly electricity expenditures of SHS customers are under $5, compared to about 68% and 52% of the monthly expenditures of residential and small-commercial mini-grid customers respectively. Another hypothesis could be that mini-grid customers have a low willingness to pay for electricity services, such that they spend a higher proportion of their budget on other household or business expenditures besides electricity. However, previous studies [238, 111] suggest that even poor households in SSA are willing to dedicate substantial parts of their budget.
to electricity. At the upper end of the spending distribution, we observe that there are slightly more residential mini-grid monthly electricity expenditures above $13 than SHS monthly electricity expenditures and about 10% more small commercial mini-grid monthly electricity expenditures above $10. This supports our earlier hypothesis that mini-grids connect a wider distribution of customers, enabling a broader distribution of expenditure and consumption among their customers.

**Figure 4.3:** Distribution of the monthly expenditures on electricity consumption for mini-grid customers and solar home system customers

We also observe that mini-grid customers make smaller, more frequent payments than SHS customers, with 76% of payments by residential mini-grid customers below $1 each compared to only 35% of payments by SHS customers (see Figure 4.4). Surprisingly, small-commercial mini-grid customers, whom we observed to spend more per month than the residential customers, also make frequent small payments; 68% of payments are below $1 each. A possible reason for this observation could be that the spending behavior of off-grid customers is largely driven by the nature of their incomes. We noted in Section 4.2.3 that a large percentage of mini-grid customers mainly derive their incomes from agricultural sources, which are highly seasonal. In contrast, the plurality of SHS customers reports a more predictable and regular source of income from salary.
or wage work. It is possible that mini-grid customers, who are likely more budget-constrained, make small electricity payments whenever they have disposable income, while SHS customers make larger, more regular electricity payments. This is an important factor that could be useful in planning business models for both off-grid platforms. In addition, strategies that mini-grid companies and development partners employ to facilitate increased connection of enterprises and integrating productive energy uses should be intentional about diversifying the kind of enterprises, and productive energy uses beyond those associated with irregular and seasonal incomes.

Figure 4.4: Distribution of payment per top-up for mini-grid customers and solar home system customers

Given the differences between the SHS and mini-grid unit power costs ($/kWh) and tariff structures, it is not helpful to directly compare the consumption values of the two customer groups. SHS are estimated to have a higher unsubsidized electricity retail cost on site ($/kWh) than mini-grids [106], therefore, SHS customers likely pay a higher unit energy cost. Our observations align with this estimate; As shown in Figure 4.5, both residential and mini-grid customers consume more electricity in a month than SHS customers. About 98% of SHS consumption values are under 5 kWh/month, which is consistent with the statistics on the energy limits of the SHS customers. While
we would expect that a much larger percentage of monthly consumption values for mini-grid customers, particularly small commercial customers, would be greater than 5 kWh/month, we observe that only 14% and 30% of monthly consumption values of residential and small commercial mini-grid customers respectively are above 5 kWh.

We observed in Table 4.4 that despite the capability of mini-grids to support larger electric loads, the bulk of the mini-grid customers in this study still own a limited set of basic appliances. Coupled with the earlier hypothesis that mini-grid customers are more financially-constrained, these reasons could explain the low consumption among mini-grid customers.

**Figure 4.5:** Distribution of the monthly consumption for mini-grid customers and solar home system customers

While the quantity of demand is important for the financial viability of off-grid systems, how electricity is consumed over the course of the day is equally important. Both solar hybrid mini-grids and SHS achieve better economic performance when the demand profile follows the PV generation profile. When demand is at or near the peak during the evening or early morning hours, which coincide with hours of low or non-existent solar generation, the requirement for storage and/or backup generation increases the Levelized cost of energy (LCOE) of mini-grids. For SHS, this type of demand profile
increases the battery and charge controller requirement and subsequently decreases
the performance of these components, thereby increasing the cost of the system. We,
therefore, explored how SHS customers, and residential and small-commercial mini-grid
customers consume electricity over the course of the day.

The clustering process results suggest three distinct load profile patterns in the
normalized monthly average load profiles of both SHS customers and residential mini-grid
customers, and two distinct load profile patterns in the normalized monthly average
load profiles of small-commercial mini-grid customers, as shown in Figure 4.6. We find
that most SHS customers (those with Profiles 1 and 3) primarily consume electricity
during evening hours, with very little daytime electricity use. Further, a more significant
proportion of the residential mini-grid customers than the SHS customers (43% compared
to 19%) have some daytime consumption between the hours of 5 am and 2 pm. However,
the SHS customers with this profile (Profile 2) consume more of their load during these
hours than the residential mini-grid customers. In their analysis, clustering over 60
measured mini-grid load profiles across 17 developing countries [160] found that the
mini-grid load profile was characterized by low nighttime demand, some daytime use,
and a more pronounced evening peak (similar to profiles 1 and 2 for our mini-grid
residential sample and profile 1 for the commercial sample) is well represented across
their data set, especially among privately owned mini-grids. While previous analysis
into load profiles of SHS in other countries on the continent is not as extensive, [246]
find that SHS customers across several villages in Rwanda exhibit an average load profile
characterized by a pronounced evening peak and some daytime consumption, similar to
profile 3 in our findings. Further, load profiles for single SHS households in villages in
Zimbabwe and Uganda [265] exhibit a small daytime peak and a much more pronounced
evening peak, similar to profile 2 in our findings.

In our analysis, we also find that more than half of the small-commercial mini-grid
customers primarily consume electricity during hours of peak PV generation, which is beneficial for the mini-grid. However, we note that small-commercial connections make up only 18% of the mini-grid customers. The remaining small-commercial mini-grid customers primarily consume electricity during evening hours, with a peak at 8 pm.

Figure 4.6: Daily load profile patterns of mini-grid and solar home system customers

Our results suggest that mini-grid companies must incentivize existing customers who can shift their evening consumption to daytime hours, as well as find and connect more enterprises that primarily operate during the day to fully utilize peak solar power generation. They should also find and connect flexible loads, such as pumps, water purifiers, or even cold storage, that can be switched on or off to follow the generation profile. The World Bank estimates that increasing daytime electricity consumption in a mini-grid can decrease the LCOE by 25 percent or more [76]. SHS companies should also incentivize electricity use during daytime hours, such as electricity use for phone charging, which could be a potential income-generating activity.

So, how do off-grid customers grow their electricity consumption? In developing strategies to facilitate organic consumption growth among their customer base, it would
be valuable for off-grid companies to have insights into this question. This could enable customized and targeted demand stimulation programs based on customers’ consumption behavior. We sought to answer this question by first evaluating the change in the timing of demand with increasing electricity spending, characterized by the change in the proportion of the day’s electricity use that takes place in the evening. Second, by evaluating the change in the duration of demand (number of hours in the day when demand is over 5% of the total daily kWh) with increasing electricity spending, which speaks to the sharpness or flatness of the daily load profiles. We carry out this analysis in three bins of electricity spending: spending under $5, between $5 and $10, and above $10. It is important to note that mini-grid customers pay per unit of electricity consumed (price per kWh) without a ceiling on the amount of electricity they can consume in a day. This would mean that their ability to afford the unit cost of power influences their consumption behavior, with less influence from their appliance ownership. On the other hand, electricity spending among SHS is directly proportional to their appliance package, which dictates the maximum amount of electricity that can be consumed in a day (explained in 4.2.3). This would suggest that their consumption behavior would likely be influenced by the appliances they own, which, to an extent, speaks to their financial abilities and other external factors that affect the timing of electricity use.

Figure 4.7 shows how the proportion of consumption during peak evening hours of 5 pm to 10 pm correlates with monthly expenditure and how this compares between SHS customers and residential and small-commercial mini-grid customers. These hours are often analyzed separately because they are crucial for how consumers experience electricity. The bottom panels show the percentage of the sample in each bin. To begin, we observe that the highest proportion of total monthly expenditures over a year, that is 53% and 46% are under $5 for residential and small-commercial mini-grid customers, respectively, while for SHS customers, 48% of total monthly expenditures are between $5
- $10. In all three customer groups, about a similar proportion of total monthly spending, 20 to 25%, is above $10.

![Figure 4.7: Proportion of average daily consumption consumed during evening peak hours (5pm - 10pm). Bottom panels: Percentage of sample that is in each bin. Note: *p < 0.1; **p < 0.05; ***p < 0.01](image)

We find the most interesting observations during months when spending was less than $5; we observe a statistically significant positive and negative correlation between monthly spending and the proportion of consumption during evening hours for SHS customers and mini-grid customers respectively – a dollar increase in monthly electricity spending goes towards a 2% increment in consumption during off-peak hours among both residential and small-commercial mini-grid customers, while SHS customers increase their consumption during peak evening hours by 3%. As we observed in Figure 4.6, 50% of the monthly expenditures in this bin made by residential mini-grid customers belonged to those whose load profiles comprised primarily of evening consumption with a peak at 8 pm and about 38% were made by those whose load profiles have an evening peak at 5 pm, albeit with some daytime consumption. The monthly expenditures in this bin made by commercial mini-grid customers are almost split evenly between customers with evening peak load profiles and those with load profiles comprising primarily daytime...
consumption. Therefore, customers spending slightly more on electricity likely have the latter profiles, and those spending less have the former.

We elucidate this hypothesis by considering how the total number of hours of high consumption in a day changes with electricity spending, alluding to how the shape of the daily profile changes with electricity spending. As shown in Figure 4.8, there is a statistically significant positive correlation between monthly spending and load profiles with a higher load factor, which supports our hypothesis. These findings suggest that small increments in electricity among mini-grid customers spending under $5 per month on electricity tend to either go towards using the same or additional appliances for more hours off-peak, which implies that these customers would be more susceptible to policies that encourage mini-grid customers to shift peak consumption to off-peak hours. We note that we controlled for the tariff structure among mini-grid customers because customers on a time-of-use tariff may have more incentive to increase consumption during off-peak hours than customers on a block tariff.

For SHS customers, 50% of the monthly expenditures in this bin also belong to customers whose load profiles comprise primarily of evening consumption, albeit with a peak at 5 pm. The remaining monthly expenditures were made by customers whose load profiles consisted of some daytime or early morning consumption. Unlike the mini-grid customers, the SHS customers spending slightly more on electricity likely have the former profiles, and those spending less have the latter profiles. This hypothesis is supported by our observations in Figure 4.8, where we observe a statistically significant correlation between increased monthly spending, a sharper load profile, and a subsequent lower load factor. This means that any additional electricity spending in this bin goes towards additional appliance usage during peak evening hours. This

---

2load factor, which is the average load divided by the peak load in a specified time period, is a measure of the utilization rate; a high load factor indicates that load is using the electricity system more efficiently
could have implications on product design, as the exacerbation of the evening peak may require higher-capacity batteries, resulting in lower capacity utilization. Similarly, in their analysis, [24] found that SHS users in Rwanda had the highest diversification of appliance usage during peak evening hours.

During the rest of the months when electricity spending was more than $5, we continue to see a statistically significant correlation between monthly expenditure and the proportion of consumption during evening hours only among residential mini-grid customers spending over $10 a month. The consumption ratio during evening hours remains unchanged, regardless of changes in monthly spending among SHS and small-commercial mini-grid customers. This could mean that for these two types of off-grid customers spending over $5 a month on electricity, any incremental electricity spending goes towards increasing the quantity of consumption with little effect on the timing of use. SHS customers spending over $5 a month have bigger appliance packages, including televisions. These additional appliances are likely used during the same hours of the day – as shown in Figure 4.6, about 56% and 70% of monthly
expenditures between $5-$10 a month and above $10 a month respectively were made by SHS customers whose load profiles comprise primarily of evening consumption with a peak at 5 pm. In addition, our observations in Figure 4.8 suggest that, similar to residential mini-grid customers, there is a statistically significant positive correlation between monthly spending over $10 and a flatter daily load profile for small-commercial mini-grid customers – about 60% of the monthly spending over $10 by small-commercial mini-grid customers belongs to customers with load profiles comprising primarily of daytime consumption. The implications for these higher-spending customers are similar to the lower-spending customers, whereby additional spending by mini-grid customers correlates with consumption behavior that benefits the financial viability of the mini-grid. In contrast, additional spending by SHS customers correlates with lower capacity utilization of SHS units, which negatively impacts the systems’ technical and economic performance. Further work in understanding the reasons that drive the changes we have observed in customers’ consumption behavior, perhaps via additional surveys, ethnographic methods, or a similar data collection instrument, could be a valuable next step for providing insights that could be useful in improving the business models of off-grid electricity platforms, as well as promoting the broader development benefits of electrification.

4.5 Conclusions and policy implications

This chapter presents a first analysis of its kind examining the spending patterns and behavior of electricity consumption among SHS and mini-grid customers in the East Africa region. While these customers are largely similar demographically and reside in relatively socioeconomically similar areas, we found significant differences in spending and consumption habits. Specifically, some key findings are:
• About 70% of monthly revenues per mini-grid connection are less than $5 and consist of smaller, more frequent payments of less than a dollar each compared to about 30% of the monthly revenues per solar home system customer. Electricity demand among off-grid customers, especially mini-grid customers, is therefore constrained by the customers’ ability to pay for electricity services rather than the technology.

• Most (over 80%) of the residential mini-grid connections and SHS customers consume a significant proportion of their daily electricity during evening hours, with peak consumption between 5 pm and 8 pm, while about 55% of small-commercial mini-grid connections consume a significant proportion of their daily electricity during daytime hours, with peak consumption around noon.

• For every $1 increase in monthly electricity expenditures, residential and small-commercial mini-grid customers increase consumption during off-peak hours by about 2%, while SHS customers increase their consumption during peak evening hours by about 3%.

Since SHS and mini-grid companies select customers carefully to develop economically viable business models, we acknowledge that these customers may not be representative of the entire SHS and mini-grid market, and there is a chance that our findings may be unique to the sample in this study. That said, we believe that our results have several policy implications for the stakeholders working towards increasing electrification through adopting these decentralized systems. We, therefore, conclude by discussing these policy implications.
4.5.1 Implications for business model development

Widespread adoption of off-grid electricity systems hinges on developing commercially-viable business models for solar home systems and mini-grids. One of the significant challenges facing existing mini-grids is the ironic problem of low demand. In addition, the costs of current mini-grids are still high, with a typical Levelized cost of energy (LCOE) of well over $0.60 per kWh [43]. Recovering the cost of an investment, when about 70% of monthly revenues per connection are less than $5 and over 50% of monthly consumption is under 1 kWh, is simply impossible without subsidies and (or) implementing strategies to increase the revenue-generating potential of the mini-grid. Our observations, therefore, support the consensus among development organizations [31, 76] that breaking free of this cycle of high cost-reflective tariffs, low demand, and financially unsustainable mini-grids would require investments in technologies or business models that help increase the average revenue per connection. With the general uncertainty in whether subsidizing the cost of electricity is a sustainable strategy in both supporting consumption growth and increasing revenues for mini-grid developers [163], current demand stimulation efforts have focused on strategies to support investment in and integration of appliances, primarily used during the day that improve incomes by providing higher value to end-users than incumbent non-electric solutions [135]. However, our results suggest that ownership of productive use appliances among mini-grid enterprises may not necessarily translate into their presumed patterns of use – 46% of enterprises are consuming electricity primarily during peak evening hours. While a generic push for productive uses of energy is valuable, mini-grids must mainly focus on attracting customers whose periods of consumption coincide with their lowest-cost periods of supply, which are typically during the day for solar-powered systems. Further, most enterprises are also exhibiting spending behavior, that is, small albeit frequent pay-
ments, which suggests that they are subject to unpredictable and possibly low incomes. These observations have significant implications for mini-grid companies implementing strategies that anchor the success of their business models on enterprises and call for integrating innovative yet diverse productive uses that encompass different sectors of the economy and consume power beyond the popular evening hours.

As for solar home systems, business model innovations, including Pay-As-You-Go (PAYGo) solutions, are accelerating their widespread adoption and diffusion. However, there is still a debate on whether SHS are effective tools for enabling the broader benefits of electrification, particularly because there is an overwhelming tendency for SHS customers to consume electricity for basic household needs that make a home comfortable rather than for income-generating activities [250]. 90% of SHS customers in our study consume under 2 kWh/month, which suggests that a majority are utilizing only basic appliances – mainly lighting and phone charging – that leave these systems largely underutilized. Given that SHS customers are spending much more on electricity than mini-grid customers, these observations imply that there is still a cost barrier inhibiting both adoption and upgrading to larger SHS packages that include more appliances. A previous study [105] found that 50,Wp SHS customers’ choice of the system they sign up for heavily hinges on affordability. Well-designed subsidies may therefore be required to enable SHS customers to expand their access to systems with a broader range of appliances that could be used for income-generating activities. It is also imperative for SHS companies to develop strategies that incentivize their customers to utilize their systems better. An out-of-the-box approach to networking SHS customers to improve access and subsequent socio-economic outcomes for end-users, as well as financial outcomes for the SHS companies, has been proposed and evaluated by a previous study [60].
4.5.2 Implications for demand side management

Off-grid electricity system planners in SSA, where the available solar generating resource is unlikely to match the latent demand, have a strong incentive to implement strategies that reduce demand from peak hours. This can include shifting existing loads and facilitating the incorporation of new daytime loads. The results of our analyses suggest that strategies by mini-grid companies that seek to incentivize their customers to shift peak consumption to off-peak customers could be successful if they are targeted toward residential customers, particularly those spending below $5 a month or above $10 a month on electricity. Our results also imply that strategies and policies that support the growth of energy-intensive mini-grid-powered enterprises operating during the day could contribute to the financial viability of the mini-grids as the added spending on electricity does not negatively impact their favorable timing of consumption. On the contrary, it may result in a higher load factor, which reduces the LCOE of the mini-grid. Put simply, not all newly-stimulated demand is of equal value to the mini-grid operator.

Our results concerning SHS are consistent with a previous study [105] that found that households in East Africa powered by SHS undertake more productive activities by moving household activities from the daytime to the evening to work longer hours. They also found that the most common motivation for SHS customers to upgrade their systems is to own a television, which they mainly use for about 3-4 hours a day in the evenings. This would explain our observations that higher spending among SHS customers correlates with a steeper load profile and an evening peak. These findings suggest that solar home systems designed to support larger appliances or a bigger appliance package may become increasingly underutilized – bigger batteries are required to meet the peak consumption, yet customers consume less during off-peak hours. The SHS industry could benefit from policies that support the establishment of a market that makes
available a broader set of SHS-compatible electric appliances, such as super-efficient refrigerators, that promote additional income-generating uses of a SHS, particularly during daytime hours, beyond just phone charging and working an existing job later in the evening. Government incentives such as import tariff subsidies, common technology standards, and quality control mechanisms could aid in unlocking the potential for versatility in the SHS appliance market.

Succinctly, while some of the differences we observed between SHS and mini-grid customers are due to technology and business model characteristics, other factors include differences among the customer bases and how they interact with their systems. More broadly, our results reflect the diversity found both between and within these customer bases, highlighting opportunities for system expansions, energy intensity increases, and targeted policy interventions to improve the consumption and productivity of these customers as well as enhance the effectiveness of these electrification systems for future customers, all the while with a focus of translating increased electricity access into broadly improved livelihoods.
CHAPTER 5

STIMULATING ORGANIC ELECTRICITY CONSUMPTION GROWTH AMONG MINI GRID CUSTOMERS IN EAST AFRICA

5.1 Background and Motivation

Electricity is not useful without appliances. In order to benefit from electricity, users must be able to afford the cost of a connection, appliances, and the electrical energy that the appliances use. MG operators often keep initial connection costs low and recover their initial investment by charging higher unit rates. This results in higher electricity tariffs, but permits users to access a connection who otherwise would be unable to do so [211]. Indeed, Lee et al. found connection rates in newly-electrified areas on the grid are highly sensitive to connection costs [150]. In effect, these customers are being offered a form of credit. This makes sense in a context in which many rural electricity users are capital constrained and lack access to traditional forms of credit, such as banking services. SHS providers, who face the same barrier of credit constraints, address the issue by combining the electricity generation technology and compatible appliances into a single package or “SHS kit,” and allowing customers to pay for the kit over time.

Thus, for most customers served by MGs, the supply of electricity is effectively
financed by the pricing model—but the appliances by which customers might use power are not. This has led to relatively low average consumption per user (ACPU) and low average revenue per user (ARPU)—threatening the economic sustainability of MG business models. In order to stimulate demand for electricity, nascent MG operators in Africa have begun to experiment with programs that finance customers’ use of electricity by offering them credit to acquire electrical appliances. The goals of these programs are to stimulate demand for electricity, increase revenues, and at the same time to permit customers to access energy services that were previously unavailable to them [277]. For commercial users, this allows businesses to acquire income-generating equipment and use the increased revenues to cover their costs. Households may benefit from reduced expenditure on more expensive alternative forms of energy.

The idea of financing electrical appliances to enable increased use of electricity is not new. In the United States, the Rural Electrification Administration and Electric Home and Farm Authority provided for loans to acquire appliances in the 1930s [44]. Similar public efforts in SSA are limited or nonexistent. However, the MG sector has begun to experiment with financing models, with an eye toward improving their own financial sustainability. The Mini-Grid Innovation Lab has launched a series of prototypes to investigate the financial viability of new business interventions for MGs in different regions throughout Africa. The data generated by the first of these prototypes, conducted in East Africa, provide the empirical setting for this paper (as described in Section 5.2.1).

Academic research on demand stimulation programs is scant, though researchers acknowledge that the provision of electricity connections must be accompanied by complementary services to realize socioeconomic impacts [207, 133] and that MGs must support and stimulate demand for electricity to achieve impact and financial sustainability [172]. One of these complementary services, particularly for businesses

1A related stream of research, not specific to off-grid settings, considers the heterogeneity that
and would-be entrepreneurs, is access to credit [208]. Access to financial services allows businesses to invest in machinery or appliances and pay for them over time. In turn, the use of electricity for economically productive purposes increases average income. In some settings the increase in income arises from higher labor productivity (which enables higher wages) or higher employment, but in rural settings where most businesses are typically one-person operations where the owner is the residual claimant, higher income typically arises from being able to sell more of the same products, the same products for higher prices (e.g., because of better perceived quality), or new products. As income rises for owners or wage laborers who are also MG customers, this permits customers to afford more electricity. This in turn improves the financial performance of the MG [77].

Promotion of productive electricity use through this virtuous cycle is a common element of discussion both in the field and among academic researchers. Implicit in the theory of change is the assumption that customers will use their access to credit in part to finance the necessary electrical appliances. To our knowledge, no prior work addresses the effects of offering credit for the specific purpose of acquiring electrical appliances, at least in the context of rural Africa.

This chapter aims to address the following questions in our appliance financing evaluation: (1) Will people take up domestic and commercial appliances if they are offered on credit? (2) How will electricity consumption be affected among those customers who take up appliances? (3) Do customers repay their appliance loans in a timely and complete fashion?, and (4) Under what conditions do MG developers benefit from providing credit to their customers? In our far more limited tariff subsidy evaluation, we examine the effects on consumption from reduced tariffs, which is a topic of particular

characterizes rigorous empirical evaluation of rural electrification: Some prominent studies suggest that widespread rural electrification has no effect on average, while others suggest important benefits, and highlights the role of complementary conditions in explaining these heterogeneous effects [151, 85].

57
5.2 Methods and Data

5.2.1 Study Area

This study was conducted in 29 villages in Kenya and Tanzania, each with a MG operated by one of five participating private sector developers and no availability of the main grid. The MGs were installed at various times between September 2014 and November 2017. The number of MG customers per village range between about 20 and 190, with an average of about 85 customers.

As noted above, this study was conducted in partnership with the Mini-Grid Innovation Lab coordinated by CrossBoundary, which works directly with mini-grid developers to prototype and test innovations that improve the business model. Once proven, the Lab works with partners – developers, government, and funders – to scale the prototypes across other developers and markets, and shares evidence on successful prototypes’ impact on the business model to inform how partners can best support the prototypes to scale.

5.2.2 Study Design

Appliance financing program

The appliance financing program for residential and commercial (micro-enterprise) customers was implemented in 27 of the 29 villages. All customers at treatment sites were invited to apply for credit to procure one or more of the electrical appliances listed in Table 5.1. Appliances were selected based on the expected consumption of each appliance, suitability for MG customers’ needs, and developers’ ability to procure them.
Any customers who applied for credit, and could make the initial deposit requirement for a given appliance, received the appliance on credit and then proceeded to make regular payments to the developer. Appliances were delivered to customers from February 2018 through mid-July 2018.

Appliances were financed with a loan, secured against the appliances as collateral. Customers who received appliances were required to make monthly payments for each appliance based on monthly compounding at an effective annual interest rate of 35 percent and over a period of 9, 10, or 12 months. These loan terms represent interest rates that, according to the developers working in these villages, are generally consistent with what they have observed in these villages for other types of purchases by households or businesses; however, the developers also report that their customers generally do not have access to such credit for the purpose of purchasing appliances. market rates. Customers were informed that nonpayment would result in operators repossessing appliances or turning off power. Anecdotally, while some offgrid companies in the region have been known to repossess SHSs or turn off mini-grid power due to substantial payment noncompliance, repossession is generally rare. In practice, repossession is costly and developers face potential reputation damage if they switch off power.

Tariff subsidy program

Reduced tariffs were provided to existing customers at the remaining two mini-grid sites, each operated by a different MG operator. Tariffs were set at a level required to cover the operating expenses of the MG operator over the number of years required to pay off the majority of the project investment costs on a given site. This tariff at each of the two sites, based on 12 months of historical costs and consumption data, was

---

Developers made individual choices about the loan period; these were set as a policy across all loan applicants and did not differ by customer creditworthiness or other factors.
calculated as operating expenses over 12 months ($) divided by the site consumption over 12 months (kWh). The variable subsidy amount was then calculated as the difference between the historical tariff over the previous 12 months and the calculated subsidized tariff.

Customers at the first site received a tariff subsidy of 50% in June 2018 and those at the second site received a tariff subsidy of 75% in May 2018. The maximum length of time for the subsidy was chosen as 5 years to enable analysis of the impact of long-term subsidized power.

5.2.3 Data

Data sources

A baseline survey, which was conducted in each of the 29 villages participating in either the appliance financing or tariff subsidy program prior to the start of the program (February 2018), gathered information on customer demographics, socioeconomic status and appliance ownership. Seven months into the study, a midline survey was conducted for continued monitoring of socioeconomic and demographic indicators. Lastly, an endline survey was conducted in October 2019. While the baseline and midline surveys were conducted in person, the endline survey was a telephone interview with a shorter overall length. Perhaps due to the change in mode of delivery, only 574 (29%) of the 1,965 customers who completed the baseline and midline surveys completed the endline survey.

Participating developers received hourly consumption readings from previously installed smart meters at each customer site, and the data were simultaneously automatically uploaded to a central server. We were not able to obtain data on the electricity consumption patterns of individual appliances. Of the 3,388 customers with smart meter
data, we were able to match 1,953 of these to survey data, linking records by using customer meter numbers. Customer payment data for electricity consumed and (separately) appliance loan repayment were uploaded to the same server, and linked to consumption data by the customer meter number.

We excluded data from two appliance financing MG sites in which the developer launched an additional reduced-tariff intervention, and from two other sites where there were prolonged service disruptions during the study period. Thus, our appliance financing analysis uses data for 1,772 customers in 23 villages served by three of the five developers, while the tariff subsidy analysis uses data for 116 customers in 2 villages served by two of the five developers.

Descriptive analysis

Appliance uptake
Of the 1,772 customers in our sample, appliances were offered to 1,654 customers, of which 348 households (about 21 percent) bought at least one appliance. The remaining 118 households were part of the tariff subsidy sites. Table 5.1 shows that of those households that purchased at least one appliance, most (254) bought just one type of appliance, with the remaining households purchasing more than one type of appliance. The most popular appliances were televisions, speakers, refrigerators or freezers, and satellite dishes.

Demographic and socio-economic characteristics
First, we consider which households apply for and take up appliances. Table 5.2 summarizes the distribution of the demographic and socioeconomic characteristics of customers who took up appliances and those who did not, as well as their average consumption, prior to the start of the program. Households that purchased appliances and those that did not are observed to be largely similar in demographic characteristics. However, those who purchased appliances are observed to have a higher average monthly income ($191 compared to $154),
Table 5.1: Appliance purchases through the appliance financing program

<table>
<thead>
<tr>
<th>Appliance type</th>
<th>Power rating</th>
<th>Number of appliances bought</th>
<th>Number of households who purchased only the listed appliance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Television</td>
<td>30 to 58 W</td>
<td>211</td>
<td>94</td>
</tr>
<tr>
<td>Speaker</td>
<td>25 to 80 W</td>
<td>157</td>
<td>78</td>
</tr>
<tr>
<td>Satellite dish/decoder</td>
<td>18 to 20 W</td>
<td>77</td>
<td>25</td>
</tr>
<tr>
<td>Fridge or freezer</td>
<td>50 to 60 W $^a$</td>
<td>70</td>
<td>39</td>
</tr>
<tr>
<td>Blender</td>
<td>350 to 450 W</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>Hair clipper</td>
<td>11 W</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>Electric iron</td>
<td>1.1 kW</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Power tools $^b$</td>
<td>3 to 7.5 kW</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Egg incubator</td>
<td>60 to 160 W</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Rice cooker</td>
<td>700 W</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Fan</td>
<td>7 to 11 W</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Laptop</td>
<td>65 W</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>544</td>
<td>254</td>
</tr>
</tbody>
</table>

$^a$ Assuming 12 - 14 hours of operation a day.

$^b$ Power tools include grinder machine, welding machine, driller, wood sander and compressor.

which is consistent with the fact that appliances were only distributed to customers who could afford to pay a deposit on the appliance. A higher percentage of these customers also report having a bank account compared to those who did not purchase appliances (24% compared to 18%). Appliance ownership prior to the start of the program is higher among customers who purchased appliances, and this is somewhat reflected in their electricity consumption – on average, their monthly consumption is 76% higher than those who did not purchase appliances.

Next, we look at the distribution of demographic and socioeconomic characteristics among households in the sites that received 75% and 50% electricity tariff subsidies. The MG sites which received 75% and 50% tariff subsidies are small sites comprising 61 and 55 connections respectively. Although they are demographically similar, we note considerable socio-economic differences between these two sites. On average, the customers in the site that received a 50% tariff subsidy have substantially higher monthly incomes (over 150% higher), with a higher percentage reporting to have a bank account.
Table 5.2: Demographic and socioeconomic characteristics of mini-grid customers

<table>
<thead>
<tr>
<th>Variable</th>
<th>Offered appliances purchased N&lt;sub&gt;max&lt;/sub&gt; = 348</th>
<th>Offered appliances did not purchase N&lt;sub&gt;max&lt;/sub&gt; = 1424</th>
<th>75% tariff subsidy site N&lt;sub&gt;max&lt;/sub&gt; = 61</th>
<th>50% tariff subsidy site N&lt;sub&gt;max&lt;/sub&gt; = 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of household head</td>
<td>38.8 (10.2)</td>
<td>41.2 (12.2)</td>
<td>41.4 (7.7)</td>
<td>41.1 (10.2)</td>
</tr>
<tr>
<td>Household size</td>
<td>5.0 (2.7)</td>
<td>5.6 (10.9)</td>
<td>5.1 (2.2)</td>
<td>4.2 (2.0)</td>
</tr>
<tr>
<td>Number of rooms</td>
<td>3.5 (1.9)</td>
<td>3.6 (2.2)</td>
<td>4.0 (1.4)</td>
<td>3.3 (2.4)</td>
</tr>
<tr>
<td>Monthly consumption, kWh</td>
<td>5.3 (11.2)</td>
<td>3.0 (5.5)</td>
<td>1.1 (2.6)</td>
<td>10.7 (17.2)</td>
</tr>
<tr>
<td>Monthly income in USD</td>
<td>191.1 (233.7)</td>
<td>154.4 (231.9)</td>
<td>86.6 (48.4)</td>
<td>211 (152.1)</td>
</tr>
<tr>
<td>Wealth index</td>
<td>0.06 (1.8)</td>
<td>-0.02 (1.5)</td>
<td>-0.12 (0.44)</td>
<td>0.22 (1.24)</td>
</tr>
<tr>
<td>Bank account</td>
<td>24%</td>
<td>18%</td>
<td>8%</td>
<td>18%</td>
</tr>
</tbody>
</table>

Primary source of income for household

<table>
<thead>
<tr>
<th>Source</th>
<th>Offered appliances purchased</th>
<th>Offered appliances did not purchase</th>
<th>75% tariff subsidy site</th>
<th>50% tariff subsidy site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsistence farming</td>
<td>59%</td>
<td>63%</td>
<td>84%</td>
<td>24%</td>
</tr>
<tr>
<td>Commercial farming</td>
<td>42%</td>
<td>36%</td>
<td>3%</td>
<td>46%</td>
</tr>
<tr>
<td>Commerce</td>
<td>30%</td>
<td>27%</td>
<td>36%</td>
<td>53%</td>
</tr>
<tr>
<td>Salary work</td>
<td>14%</td>
<td>13%</td>
<td>13%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Appliance ownership before interventions

<table>
<thead>
<tr>
<th>Appliance</th>
<th>Offered appliances purchased</th>
<th>Offered appliances did not purchase</th>
<th>75% tariff subsidy site</th>
<th>50% tariff subsidy site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Television</td>
<td>41%</td>
<td>30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sound system</td>
<td>28%</td>
<td>17%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radio</td>
<td>26%</td>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satellite dish</td>
<td>27%</td>
<td>18%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refrigerator</td>
<td>6%</td>
<td>4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is also a substantial difference in consumption between the two sites prior to the tariff subsidy program; about 10 kWh/month in the 50% tariff subsidy site compared to 1.5 kWh/month in the 75% tariff subsidy site. The primary source of income for a majority of the households in the 75% tariff subsidy site is subsistence farming, while for the majority of households in the 50% tariff subsidy site, it is commerce.

5.2.4 Methods

Effect of the appliance financing program

If cashflow constraints, or the unavailability of credit, hinder customers’ ability to take up and utilize electrical appliances that are generally available in the open market, then providing financing should result in a sustained increase in power consumption. To evaluate if this is the case, we measure changes in energy consumption and payments for electricity among customers before and after they received appliances, using a
difference-in-differences design. The ideal experiment would involve randomly assigning appliances to customers, and then evaluating the behavior of recipients compared to non-recipient customers. However, this program design was determined to be infeasible due to logistical complexity and potential damage to developer-customer relationships. Thus, all customers in our sample were invited to purchase appliances under the program. For the purpose of our matching analysis, the customers who chose to purchase appliances are the ones we consider “treated”.

To evaluate changes in consumption, we used nearest-neighbor propensity score matching to identify one control customer (that did not receive an appliance under the program) comparable to each treatment customer. We select control customers within the same developer based on the estimated propensity to purchase an appliance, which we estimate as a function of average ex ante electricity consumption, the age of the head of household, household income and wealth, and whether the customer had a bank account.\(^3\) All of these variables were measured in the baseline survey, that is, prior to the placement of appliances under the program. For matched controls, we assign the “placebo appliance delivery date” as identical to that of the corresponding treatment household.\(^4\) Our preferred specification uses a one-to-one match of treatment to control customers; in additional specifications shown in C, Table C.2, we show that our results are robust to alternative forms of matching that permit many-to-one matches, including kernel and radius matching.

\(^3\)In the preferred specification, we calculate wealth as the first component in a principal components analysis (PCA) of a comprehensive set of assets that includes the number of livestock animals, access to various means of transportation, number of rooms, and ownership of a mobile phone. In a robustness check, shown in Table C.2, we match on these individual elements rather than the first component of the PCA.

\(^4\)About one-third of the customers that applied for at least one appliance applied for (and received) multiple appliances under the program. These appliances were not necessarily delivered on the same date. We use the earliest date of the appliance delivery for each customer to define the date of the intervention for that customer.
The propensity score matching method relies on assumptions of conditional independence—that is, that potential outcomes are independent of treatment assignment—and common support (that is, customers with the same covariates have a positive probability of being both participants and non-participants) [41]. The first assumption is only partially testable, in that we can test if the treatment assignment is independent with respect to observable covariates, but not to unobservables. Table 5.3 provides a comparison of mean values for treated customers, all untreated customers, and matched control customers, for socioeconomic and demographic characteristics of customers, at the baseline. The table shows that there are no significant differences between treatment customers and matched controls in terms of the covariates used for the matching assignment (which are also the top five rows in the table), nor in terms of most other observable characteristics. The table also indicates some differences: treatment households have fewer rooms, on average; own more appliances at baseline; and in particular are more likely to own a television, a sound system, and a satellite dish or decoder. It is worth noting that these higher rates of baseline appliance ownership do not come with significantly higher ex-ante electricity consumption; although treatment customers did consume about 15% more electricity than matched controls in the months prior to the intervention, the variation in ex-ante electricity consumption is substantial, and the difference in means is not significantly different from zero.

The “common support” assumption refers to the idea that treatment and matched control customers have a positive probability of being both participants and non-participants. Figure C.3 in C provides graphical evidence that this is the case. Although the quality of the matching approach cannot be fully tested, particularly with respect to unobservable characteristics that may affect treatment status, C provides the results of several
additional tests to verify the quality of the propensity score match.\(^5\) That appendix also documents the effects of the intervention under several alternative matching methods, as noted above. We discuss the key implications of these tests for our findings in Section 5.3.1.

The first question of interest concerns the effect of the appliance financing program on consumption. For this element, we aggregate the hourly consumption data to a total weekly measure per household (in some specifications, we also examine total monthly consumption per household). We limit our sample to the treatment customers (i.e., those that received an appliance under the program) and the matched control customers, and estimate the average treatment effect (ATE) using the following difference-in-differences model:

\[
y_{ijt} = \alpha_i + \beta_t + \sum_{\tau=-52}^{93} \delta_{\tau} 1(\text{appliance received in week } t - \tau)_{ijt} + \epsilon_{ijt} \quad (5.1)
\]

In (5.1), \(y_{ijt}\) denotes the total energy consumption by household \(i\) in village \(j\) in time \(t\), \(\alpha_i\) is a household fixed effect, and \(\beta_t\) is a time (week) fixed effect. This specification allows us to estimate separate effects on the treatment customers for each week leading up to the delivery of the (first) appliance as well as each week after, for 52 weeks before and 93 weeks after delivery. We include household fixed effects to control for time-invariant idiosyncratic variations at the household level (see [132]), and calendar-week fixed effects to control for seasonality of income or other time-related variations that span the region. We also estimate a specification without household fixed effects, but with household characteristics from the survey data, as well as village fixed effects and other elements. For this latter specification, we aggregate the time series to a monthly (rather than weekly) series, strictly to facilitate display of the results in a tabular (rather

\(^5\)We are grateful to two anonymous reviewers for suggesting several of these robustness checks.
Table 5.3: Differences between treated, control, and matched control customers

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Treated</th>
<th>(2) Non-treated</th>
<th>(3) Matched control</th>
<th>(4) Diff (T - NT)</th>
<th>(5) Diff (T - C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household wealth</td>
<td>0.06</td>
<td>-0.01</td>
<td>-0.10</td>
<td>0.08</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(1.76)</td>
<td>(1.42)</td>
<td>(0.97)</td>
<td>(0.10)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Bank account</td>
<td>0.24</td>
<td>0.17</td>
<td>0.25</td>
<td>0.06**</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.43)</td>
<td>(0.38)</td>
<td>(0.43)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Elec. consumption</td>
<td>5.24</td>
<td>3.23</td>
<td>4.59</td>
<td>2.01***</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>(11.25)</td>
<td>(6.43)</td>
<td>(8.28)</td>
<td>(0.63)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>Age of HH head</td>
<td>38.82</td>
<td>41.22</td>
<td>38.06</td>
<td>-2.40***</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>(10.22)</td>
<td>(11.98)</td>
<td>(10.70)</td>
<td>(0.63)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>HH income</td>
<td>190.83</td>
<td>153.84</td>
<td>183.50</td>
<td>36.98***</td>
<td>7.33</td>
</tr>
<tr>
<td></td>
<td>(233.94)</td>
<td>(224.91)</td>
<td>(229.76)</td>
<td>(13.98)</td>
<td>(17.73)</td>
</tr>
<tr>
<td>Educ. of HH head</td>
<td>7.82</td>
<td>7.56</td>
<td>7.54</td>
<td>0.26</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(3.00)</td>
<td>(3.49)</td>
<td>(3.38)</td>
<td>(0.19)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>HH size (persons)</td>
<td>4.98</td>
<td>5.52</td>
<td>5.14</td>
<td>-0.54*</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>(2.65)</td>
<td>(10.49)</td>
<td>(2.53)</td>
<td>(0.32)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Number of rooms</td>
<td>3.53</td>
<td>3.61</td>
<td>3.93</td>
<td>-0.08</td>
<td>-0.40*</td>
</tr>
<tr>
<td></td>
<td>(1.85)</td>
<td>(2.18)</td>
<td>(3.40)</td>
<td>(0.12)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Appliances owned (#)</td>
<td>1.46</td>
<td>1.00</td>
<td>1.19</td>
<td>0.46***</td>
<td>0.27**</td>
</tr>
<tr>
<td></td>
<td>(1.50)</td>
<td>(1.40)</td>
<td>(1.55)</td>
<td>(0.09)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Owns radio</td>
<td>0.27</td>
<td>0.20</td>
<td>0.22</td>
<td>0.06**</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.40)</td>
<td>(0.42)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Owns TV</td>
<td>0.42</td>
<td>0.29</td>
<td>0.32</td>
<td>0.12***</td>
<td>0.09**</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.45)</td>
<td>(0.47)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Owns fridge</td>
<td>0.06</td>
<td>0.04</td>
<td>0.08</td>
<td>0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.20)</td>
<td>(0.27)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Owns sound system</td>
<td>0.29</td>
<td>0.17</td>
<td>0.22</td>
<td>0.13***</td>
<td>0.07**</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.37)</td>
<td>(0.41)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Owns satellite decoder</td>
<td>0.27</td>
<td>0.18</td>
<td>0.21</td>
<td>0.09***</td>
<td>0.06*</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.39)</td>
<td>(0.41)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
</tbody>
</table>

Notes. All values represent characteristics at baseline, prior to intervention. Columns 1, 2, 3 report, respectively, means and standard deviations for customers in treatment group (i.e., purchased at least one appliance under the program), all non-treated customers, and matched control customers. Column 4 (5) reports the mean difference and standard error for treated vs. non-treated (treated vs. control) customers. Significance stars represent the results of t-tests (or \( \chi^2 \)-tests for variables that are proportions): *** \( p < 0.01 \), ** \( p < 0.05 \), * \( p < 0.1 \). Household wealth is the first component of the principal components analysis described in text. Results for appliances owned by less than 2% of households in any group (DVD player, computer, internet access, printer, microwave, iron, hair clipper, mill, blender) available from authors.

than graphical) format, and compare the estimates that use household characteristics to those that use household fixed effects. (The results are shown in Table B.1.)

For both the weekly and monthly series, we describe the week (month) in which the
first appliance was delivered to the household as 1 week (month) prior to the intervention and normalize the weeks (months) before and after appliance delivery accordingly.

As noted in the section on background, we also examine whether customers repay appliance loans in a timely and complete fashion, and explore the conditions under which MG developers benefit from providing credit to their customers.

Effect of the tariff subsidy program

The effect of the tariff subsidy is reflected in the changes in the monthly average revenue per user (ARPU), which is the metric most commonly used by mini-grid developers to evaluate revenues and the monthly average consumption per user (ACPU) before and after the per unit cost of power was lowered in the participating sites. Similar to the appliance financing program, the ideal experiment would involve randomly assigning a set of subsidised tariffs to a set of mini-grid sites without any other interventions. Customers at these sites would make up the treatment groups, and customers at other mini-grid sites of the same developer, also without any other interventions, would make up the control group. However, due to the limitations of the appliance financing experiment discussed in the previous section, for one of the developers that implemented the tariff subsidy program, all their other sites that were not part of the tariff subsidy program participated in the appliance financing program, and the other developer only has one site, which was part of the tariff subsidy program. It was therefore impossible to construct a control group for the tariff subsidy sites. We therefore present the changes in ARPU and ACPU in the treatment sites only. We also analyze the customers’ sensitivity to tariff changes by looking at the price elasticity of electricity demand, $\epsilon$, given by Eq. 5.2
\[ \epsilon = \frac{\Delta q}{q} \frac{\Delta p}{p} \]  

(5.2)

where \( \epsilon \) is the price elasticity, \( p \) is the electricity price, and \( q \) is the electricity demand.

### 5.3 Analysis

#### 5.3.1 Appliance financing program: Effects on consumption

Figure 5.1 shows our main results with respect to consumption: prior to the delivery of appliances, consumption among treated customers and matched controls is statistically equal. The week-level results, going back 52 weeks before the initial date of delivery of appliances, support the parallel trends assumption of the difference-in-differences approach. Upon delivery and in subsequent weeks, consumption increases by about 0.6 kWh/week on average (this amounts to about a 66% increase in consumption; the average weekly consumption for an average household prior to the intervention is 0.9 kWh). This increase is sustained for about 7 – 8 months (30 weeks) after delivery, but then falls gradually and the point estimates level off to about 0.3 kWh/week between about 47 weeks and 85 weeks after delivery, which is roughly a 33% increase in consumption. After 85 weeks, the point estimates begin to drop again to about 0.15 kWh/week.

It is worth reiterating that the estimated increase in consumption is attributable to the treatment only under fairly restrictive assumptions. As discussed in Section 5.2.4, all customers in our sample were invited to purchase appliances under the program, which means that the matched control customers actively declined to purchase appliances, which in turn implies that there are likely to be unobservable characteristics that differ between treated customers (i.e., those who took up the appliances) and matched controls. Furthermore, the bias in the estimated treatment effect is likely to be positive, if the
Figure 5.1: Average treatment effect of appliance financing program. Notes: Robust standard errors are clustered by village; error bars represent a 95% confidence interval. Includes household fixed effects, calendar-month fixed effects, and relative week fixed effects. *p < 0.1; **p < 0.05; ***p < 0.01.

households who voluntarily purchased appliances under the program are also those who were more likely to use them (e.g., have preferences for owning and using electric appliances, or believe they can use them productively) or more likely able to afford them (e.g., feel they have more stable incomes). If so, then these unobservable factors likely play a substantial role in the increase in consumption that we observe. Indeed, a Rosenbaum bounds analysis (Table C.1) suggests that even a relatively small amount of “hidden bias” would have to be present in order for us to conclude there is no significant change in consumption. If treated customers were 1.5 times more likely than control customers to select into treatment (Γ = 1.5 in the Rosenbaum bounds framework), then the p-values associated with most of the (month-level) estimated increases in consumption would exceed 0.1, from which we would conclude the treatment had no
statistically significant effect on electricity consumption.

The increase and subsequent drop off in consumption is difficult to explain. We look at the change in average consumption of the treatment and control customers separately before and after the program to investigate how these move separately, as well as how their difference changes. There is a slight decrease in consumption among the control households in weeks following delivery of appliances to treatment customers, however, there is a comparatively larger increase in consumption among customers who purchased appliances. During the period when consumption among treatment drops after the initial increase, we observe that there is no significant change in consumption among control households. We can therefore rule out spillover effects as a possible explanation for the drop in consumption.

Another possible explanation for the increase and subsequent drop off in consumption could be that MG customers are budget-constrained, therefore after seven months, treatment customers begin to lower their appliance usage in response to higher spending on electricity consumption. We explored this hypothesis further by analyzing whether the electricity payment behavior of the treatment customers could somehow indicate when they realize how costly the additional consumption is. The billing system is pre-paid, where customers purchase units of electricity (kWh) at a given tariff. We therefore looked at the distribution of the frequencies of electricity top-ups per month. We would expect that when customers realize how costly the additional consumption is, they would start topping-up their electricity units less frequently. However, we observed that the electricity payment behavior among treatment customers, as indicated by the trend in the average frequency of top-ups, does not differ from that of control customers, as shown in Figure B.1. Therefore, this hypothesis may not explain why we observe a reduction in consumption after the initial increase. In addition, based on customer responses on appliance usage over time during the program, with the caveat
that there was high attrition in the endline survey as discussed in Section 5.3.6, the majority of customers reported similar appliance use over time. Another reason could be that customers do not earn enough additional income from the appliances purchased to be able to consistently cover the costs of both electricity and loan repayment. However, we did not have data on changes in household income during the program to test this hypothesis.

Another reason could be that although we have tried to control for seasonal effects by using calendar-month fixed effects, there may be macroeconomic effects that are not controlled for that drive a decrease in disposable household income. To explore this further, we investigate whether the drops in consumption coincide with the lean season, which is the period between planting and harvesting when incomes plummet in agricultural areas. About 75% of the appliance financing customers reported that the main source of income for the household is from agriculture (farming or post-harvest processing). As shown in Figure B.2, the drop in consumption between weeks 33 - 52 after appliance delivery coincide with the 2018/2019 lean season, which starts in November to the end of February [91]. The second drop in consumption after 85 weeks also coincides with the start of 2019/2020 lean season. In addition, as discussed in Section 5.3.6, about 25% of the subset of appliance financing customers who responded to the endline survey reported difficulty in making loan repayments due to limited income during the lean season. It is likely that they also had difficulty making electricity payments as well. However, the fact that the average consumption does not increase after the lean season suggests that there could be confounding factors that we are not able to account for that are also impeding consumption growth. Furthermore, when we compared the effect of the program on customers who reported that their main source of income is from agriculture and those who reported that their main source of income is from commerce, services, or salaried work – i.e., those who may be less affected by seasonality – we
observed that in both cases there is a drop in consumption after 30 weeks from the initial increase albeit the drop in consumption between the two groups of customers does not occur simultaneously as shown in Figure B.3.

When the effect of the program was analyzed at an appliance level, consumption among customers who purchased the three most popular devices (televisions, speakers, and refrigerators or freezers) was estimated to increase by up to 1.0 kWh/week, 0.5 kWh/week and 2.5 kWh/week respectively as shown in Figure 5.2. This is about a 108%, 98% and 82% increase in consumption from the respective average control customers prior to the delivery of appliances to treatment customers. This increase in consumption is sustained longer among the refrigerator or freezer customers and speaker customers (that is, about 10 months), compared to 8 months among television customers. We also observe that the drop in consumption after the initial increase is more pronounced among television customers, whose point estimates drop to about 0.1 kWh/week compared to fridge/freezer customers whose consumption at the end of this study period is about 0.5 kWh/week higher than consumption prior to the intervention. The quicker and more significant drop in consumption among television customers compared to fridge/freezer customers underpins the second reason of inadequate additional income generation from the new appliances to support increased consumption: Fridges and freezers offer a better opportunity for sustained higher income, which may be put towards consuming more electricity, whereas televisions are mainly for residential use, with limited uses for income generation.

Although customers were asked whether they used the appliances they purchased to generate an income as part of the endline survey, the high attrition of this survey, which resulted in a very small sample size, prevented us from comparing the treatment effect

---

6 The 124 customers who purchased multiple appliances under the program are excluded from this analysis.
of productive use and consumptive use appliances. However, from the responses that were received, a higher percentage of customers reported using their fridges/freezers for income generation compared to the other appliances. Additionally, information on whether customers were enterprises or households was only collected as part of the endline survey; therefore, we are unable to compare the average treatment effect of the program between households and enterprises. However, when we compared the average weekly consumption among the subset of customers for whom we were able to classify as residential or business customers, we observed similar behavior after appliance delivery, that is the initial increase and subsequent drop in consumption, albeit the drop starting earlier among residential customers as shown in Figure B.4.

We also consider how the effect of the program on the consumption of customers who purchased an individual type of appliance compares with customers who purchased a blend of appliances. As shown in Figure B.5, after appliances are delivered, both sets of customers are observed to have an initial growth in consumption, which erodes over time. This growth appears to be larger among customers who purchased multiple types of appliances. These customers are likely to be higher income customers as they were able to afford the deposit on more appliances. Section 5.3.4 describes the implication of appliance offerings to these customers on the economic viability of appliance financing programs in this type of setting.

Our preferred analysis involves using week-level data, as it provides a granular measure of changing consumption patterns over time, and household fixed effects, which control for idiosyncratic variations in household behavior. Nevertheless, as a form of a robustness check, we also analyze the effect on consumption using (1) a simple difference-in-difference analysis without covariates and (2) an analysis using household-level covariates rather than household fixed effects. To facilitate comparisons with the main consumption results we present these results in tabular form, alongside
Figure 5.2: Treatment effect of individual appliance uptake. Notes: Error bars represent a 95% confidence interval; robust standard errors are clustered by village. Includes household fixed effects, calendar-month fixed effects and relative week fixed effects. *p < 0.1; **p < 0.05; ***p < 0.01

results from a specification with household fixed effects, and to make the table legible we perform this analysis at the level of month rather than week. Table B.1 shows the results of this analysis. As the table demonstrates, the results of these various alternative approaches are broadly consistent with the results of our main specification: the parallel trends assumption holds for periods prior to the intervention, and subsequent monthly
consumption among treatment households increases substantially immediately after the intervention—then the gap between treatment and control households begins to decrease approximately 7 months after the intervention date.

5.3.2 Appliance financing program: Time-of-Use Effects

Operators of solar MGs, like any grid, must manage aggregate load profiles to smooth excess consumption over the course of the 24-hour period. Storage requirements—or increased operating costs due to the need for backup generation, often through diesel motors—represent a substantial portion of the levelized cost of energy (LCOE), so minimizing demand in excess of current supply helps to reduce the LCOE. Evening, when solar insolation is low or nonexistent, but demand is at or near a peak, often represents the time of the greatest excess demand. To analyze the effects on peak demand in this setting, we explore shifts in daily load profiles from adopters of the three most popular appliances (televisions, speakers, and fridges/freezers). We begin by considering the change in load factor, which is the ratio of average to peak consumption in a given time period and measures the efficiency of electricity usage. Prior to the appliance financing program, the consumption of customers who purchased televisions only, and fridges/freezers only, is already somewhat balanced across the day, with a load factor of 47% (Figure 5.3). The respective control customers have similar baseline load factors of 45% and 47% respectively. Speaker buyers tend to have higher consumption in the late afternoon to early evening hours, with peak consumption at 5pm resulting in a load factor of 33%. The control customers also have a baseline load factor of 33%. After implementation of the program, we find that television buyers increase evening peak consumption relative to matched control customers, resulting in a reduction in the load factor by 4%. For speaker buyers, the increase in consumption relative to the control customers is more spread out throughout the afternoon and evening hours, resulting in
a load factor increase of 8%. For fridge/freezer buyers, relative to the control, customers show a statistically significant change in the load profile during morning hours resulting in an increase in load factor by 10%.

Figure 5.3: Effect of appliance uptake on daily load profiles. Notes: Top panels are normalized daily load profiles; bottom panels show the average effect of individual appliances, distinguished by hour of the day. Error bars represent a 95% confidence interval, calculated from robust standard errors clustered by village. The regressions that generated the coefficients illustrated here include household fixed effects, calendar-month fixed effects, and relative hour fixed effects. *p < 0.1; **p < 0.05; ***p < 0.01.

These results may suggest that a large percentage of appliances purchased through this program were primarily for residential use, particularly during evening hours. As such, the program was not effective in shifting peak evening load to daytime hours, when there is peak solar generation. This means that if MG developers are interested in shifting consumption to match hours of generation in order to amortize the high fixed costs of provisioning energy storage, they need to identify and offer appliances that are primarily used during the day and, at the same time, meet the needs of their customers.
5.3.3 Appliance financing program: Loan repayment

One of the particularly novel features of our data set is that it contains information on customer loan repayment. This is especially notable because data on loan repayment and loan default from private sector providers of off-grid electricity solutions (including SHS and MGs) are often unavailable in the literature; operators often consider this some of their most sensitive information about customers. In this setting, two of the four participating MG operators set a loan term of 12 months in the financing agreement with their customers, while the remaining two set loan terms of 9 and 10 months each. The cumulative distribution function in Figure 5.4 shows how customer repayments progress over the course of the loan term and after. It shows the proportion of appliance buyers on the y-axis that have less than and up to the corresponding loan balance as a percentage of the expected repayment amount on the x-axis at a specified point in the loan repayment period, indicated by each of the lines on the figure.

![Figure 5.4: Cumulative distribution functions of loan balances for customers who bought appliances](image)

We observe that in the initial phases of the loan term, customers tend to be on time with payments, but the repayment rate steadily deteriorates over the term. One-quarter
of the way through the loan term, 61% of customers were on track with repayments; in progressive quarters this falls to 50% then 37%, and by the end of the loan term just 24% of customers have fully repaid the initial loan. The average repayment rate at the end of the term is about 66%; that is, at the end of the term, the average customer had repaid 66% of her loan. Customers do continue to repay loans after the stipulated term; at 125% of the loan term (i.e., past the end of the term), about 35% of customers had repaid loans in full and the average repayment rate increases to 78%. This evidence of good-faith behavior helps to explain why developers infrequently activate the threat of repossession when customers fail to repay on the contract terms, and we hypothesize that this relationship between developers and their customers (and the potential to limit access to electricity) leads to continued payments after the conclusion of the loan term. Yet the relatively high rate of nonpayment suggests MG operators must choose carefully to whom to offer financing, and may have to charge higher rates to all customers in order to mitigate the risk of nonpayment or late payment. This is especially true when MG operators must bear the upfront costs of appliance purchases alongside the capital costs for mini-grid development itself. We note that the MG operators in this study did not bear the upfront costs of the appliance purchases, therefore it’s possible that the problem of moral hazard could also explain why the efforts to enforce loan repayments were more measured.

To our knowledge, only a handful of studies have reported information on repayment for household goods purchased on credit in developing-country settings. In the most directly comparable study, researchers working in collaboration with a micro-lender in Orissa, India, offered 12-month loans for rural householders to purchase insecticide-treated bednets (intended to reduce the incidence of mosquito-borne disease including malaria) at market prices [255]. In that study—in which the market price amounted to three to five times the average daily agricultural wage, and householders paid 20
percent annual interest on loans—researchers found an average repayment rate of 64 percent after 18 months (i.e., 6 months after the expiration of the contract term.) Table 5.4 provides a summary of the repayment rate from this study, as well as four others that we identified that quantify repayment rates for consumer products in comparable settings. Of note, all the other studies shown in that table have involved zero-interest loans and shorter contract terms than in our setting. Thus, the repayment rate we observe here is comparable to, and somewhat higher than, that documented in the limited prior literature on comparable (non-zero-interest) household loan repayment in rural low-income settings. We reflect further on the implications of repayment rates in the economic feasibility analysis; see Section 5.3.4.

Table 5.4: Repayment rates for other consumer goods

<table>
<thead>
<tr>
<th>Item</th>
<th>Location</th>
<th>Repayment rate</th>
<th>Monitoring period</th>
<th>Loan period</th>
<th>Interest rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITNs</td>
<td>India</td>
<td>64%</td>
<td>18 months</td>
<td>12 months</td>
<td>20%</td>
</tr>
<tr>
<td>Cookstoves</td>
<td>Uganda</td>
<td>97%</td>
<td>4 weeks</td>
<td>4 weeks</td>
<td>0%</td>
</tr>
<tr>
<td>Cookstoves</td>
<td>Senegal</td>
<td>&gt;95%</td>
<td>10 weeks</td>
<td>10 weeks</td>
<td>0%</td>
</tr>
<tr>
<td>SHSs</td>
<td>Rwanda</td>
<td>77%</td>
<td>11 months</td>
<td>1 week to 5 months</td>
<td>0%</td>
</tr>
<tr>
<td>Water filters</td>
<td>Kenya</td>
<td>93%</td>
<td>6 months</td>
<td>8 weeks</td>
<td>0%</td>
</tr>
</tbody>
</table>

Notes. SHS = solar home systems; ITNs = insecticide-treated nets. Sources: India ITNs from [255]; Uganda cookstoves from [153]; Senegal cookstoves from [22]; Rwanda SHS from [111]; Kenya water filters from [165].

We further explored whether the demographic and socioeconomic characteristics of the appliance financing customers could predict timely repayment of appliance loans. We consider three logistic regression models: the first, without any fixed effects and an additional explanatory variable that considers the length of the loan term, the second with developer fixed effects only, and the third with village fixed effects. The results are presented in Table 5.5. We expected income to be a strong indicator of customers’ ability to repay their loans on time; however, we find that none of the characteristics we
considered, including income, are statistically significant indicators of the propensity for customers to fully repay their loans on time.

Table 5.5: Customer characteristics predicting on-time appliance loan repayment

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household size</td>
<td>-0.137*</td>
<td>-0.027</td>
<td>-0.043</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.077)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>Number of rooms</td>
<td>0.037</td>
<td>0.044</td>
<td>0.137</td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.113)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Income (USD)</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Wealth index</td>
<td>-0.209</td>
<td>-0.017</td>
<td>-0.178</td>
</tr>
<tr>
<td></td>
<td>(0.143)</td>
<td>(0.174)</td>
<td>(0.276)</td>
</tr>
<tr>
<td>Loan amount (USD)</td>
<td>-0.002**</td>
<td>-0.001</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Less than 12 month loan term</td>
<td>1.077*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.578)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Village fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Developer fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td>0.10</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Observations</td>
<td>282</td>
<td>282</td>
<td>282</td>
</tr>
<tr>
<td>Full repayment by end of loan term</td>
<td>85</td>
<td>85</td>
<td>85</td>
</tr>
</tbody>
</table>

Notes: Regression uses a logit model, in which the dependent variable is coded as 1 if a customer repaid their loan in full (including interest) by the end of the loan term, and 0 otherwise. ***, **, and * indicate statistical significance at 1%, 5%, and 10%. Robust standard errors, clustered at village level, are in parentheses.

Another question of interest is how consumption changes after customers’ appliance loans are fully paid off. For instance, we might expect to see an increase in electricity consumption when loans are paid off, if we believe that customers’ demand for the services these appliances provide does not change before and after full repayment, but disposable income increases after the loans are fully repaid. If we see such an effect, this suggests that customers may be credit-constrained in a way that hinders their ability to purchase appliances, but not in a way that hinders their ability to purchase electricity, conditional on owning the appliances they desire. On the other hand, if we see no
substantial change in electricity consumption after loans are fully repaid, this would suggest that customers may face constraints on their ability to purchase electricity, over and above the constraints on initial appliance purchases. Figure 5.5 shows the average day-to-day change in consumption for customers who fully repaid their loans, for 10 months before and after completion of loan repayment. Customers’ consumption begins to drop before they complete their loan repayments, and we see this drop continue for a couple of months even after they complete their repayments. This suggests that customers face additional constraints that adversely affect their ability to purchase electricity (and realize the benefits of access), even after loans are repaid.

![Graph showing average daily consumption](image)

**Figure 5.5:** Average daily consumption prior to and after full loan repayment. Includes those customers who paid loans in full, regardless of whether the payment was within the loan term \([n = 131 (35\%)]\).

### 5.3.4 Appliance financing program: Economic analysis

Should developers implement appliance financing programs similar to the one we study here? To shed some light on this question, we consider a hypothetical mini-grid operator who could borrow funds (or divert funds from other investments) to purchase appliances and then sell them to customers, with financing. Developers’ net revenue
from offering an appliance financing program is equal to

\[ \Delta R = P\Delta Q + A(1 + APR_c) \times REPAY - A(1 + APR_d), \]

where \( P \) is the price per unit (e.g., kWh) and \( \Delta Q \) is the change in consumption; \( A \) is the cost of the appliance; and \( APR_c \) and \( APR_d \) represent the cost of capital for the customer and the developer, respectively. As we show in Section 5.3.3, not every customer repays their loan in full; \( REPAY \) is the average amount of the appliance loan that is repaid. We assume that developers face a cost of capital of 20% per year and charge customers 35% per year, which is the rate several developers used in the present study; we also assume developers sell power at a price of $1.50 per kWh. For our hypothetical scenario, we use the observed appliance cost \( A \), as well as the average values of \( \Delta Q \) and \( REPAY \), from our empirical analysis. For each individual appliance shown in the figure, we limit the sample to those treatment households that purchased only that appliance, and corresponding matched control households, and run the regression in (5.1) on that sub sample. We then calculate \( \Delta Q \) as a simple average of the difference in consumption (i.e., treatment minus matched controls), over all weeks following appliance delivery. It should be noted that to the extent that the estimated increase in consumption is biased upwards due to self-selection among the treatment customers (see Section 5.3.1), this financial analysis may have limited external validity—for instance, the specific conclusions about net payoffs may not hold for another setting, or if developers scale up an intervention to a much broader set of customers. Nevertheless, the findings about relative profitability across different appliances, and perhaps the approach to assessing potential profitability, may be instructive for some developers considering similar programs.

Figure 5.6 shows the result. To show the appliances on an equal footing, the vertical
axis shows the change in ARPU divided by the appliance cost $A$.\footnote{For instance, the refrigerator/freezer in our setting costs USD 207, and customers who bought only this appliance increased consumption relative to controls by about 0.73 kWh per week; thus, the ratio of ARPU to $A$ is $0.73 \times 52 + 1.5/207 = 0.276$.} The downward-sloping line shows where net developer revenues equal zero; points above this line represent appliances that would return positive net profits, while points below the line represent unprofitable investments. We also plot a point for the blend of customers and appliance offerings that comprise the program that we analyzed (the “any appliance” marker); note that this blend additionally includes the customers who took up multiple appliances. While our blended appliance program did yield a positive net profit, the overall findings suggest that developers may see the strongest results by implementing appliance financing programs only for select appliances (e.g., refrigerator/freezers). Alternatively, developers may be able to realize profits from other appliances if they can access low-interest sources of capital, thus driving down $APR^d$, or if they have reason to believe their customers would repay loans at a higher rate (increasing $REPAY$) or increase consumption ($\Delta Q$) by a greater degree than we found in our setting. Charging a higher rate to customers ($APR^c$) or a higher unit price ($P$) could be feasible depending on policy or market conditions, but would likely come at a tradeoff of decreasing $REPAY$ or $\Delta Q$ or both.

### 5.3.5 Appliance financing program: Robustness analysis

To test the quality of the estimates obtained using the nearest one-to-one neighbor matching, we compute estimates using two additional matching algorithms using propensity score matching: radius matching with replacement and kernel matching. As before, the variables used to match households are the average daily energy consumption in the three months preceding the first appliance delivery to any household in the sample, household size, and the household asset index. In both these algorithms,
there may be many-to-one matching between the control and treated households. We bootstrap the standard errors for these estimates [149].

The radius matching algorithm generates counterfactuals for each of the treated households within the common support using control households whose propensity scores are within a given caliper. A caliper is a maximum permissible distance between the propensity score of the treated and counterfactuals [54]. Each of these households within this caliper is assigned the same weight \( w_{ki} = \frac{1}{N_i^C} \) such that \( N_i^C \) is the number of counterfactual households matched with the \( ith \) treated household. Consequently, the number of control group households that are matched to each of the treated households may vary. The quality of matching is superior relative to kernel matching since only those control group households are used as matches that have propensity scores similar to the treated household.

The kernel matching algorithm matches each treated household in the common support to a weighted average of all the control group households. We use weights that are derived from kernel weights using a normal distribution. These weights are a
function of the distance between the propensity scores of the households in the treated households and control group households [42]. This method of matching contains more information about the control group as all households in this group are used to match each time. This lowers the variance of the estimator.

The difference-in-difference estimates computed after matching using various propensity score matching techniques are presented in Table C.2. These results indicate that our estimates obtained using the nearest one-to-one neighbor matching method are consistent across different matching algorithms.

5.3.6 Appliance financing program: Customer perspective

To gain insight on how customers perceived the appliance financing intervention, we consider the responses of customers who were part of the endline survey on program satisfaction. However, as a result of attrition in the endline survey, our observations are based on 178 appliance financing customers (51% of the total customers who obtained an appliance). 78% of these customers reported that if they could go back in time, they would still buy an appliance through the program. Almost none reported that they stopped using their appliances. When asked about the ease or difficulty in making monthly loan repayments, 25% reported medium to high difficulty, stating the reason why as limited income.

5.3.7 Tariff subsidy program: Effects on consumption

Figure 5.7 shows our main results with respect to the effect of the tariff subsidy on average revenue and consumption per user. The tariff structure in the 75% subsidy site is a time-of-use tariff, where a lower tariff is charged during off-peak hours than during peak usage hours, while the structure in the 50% subsidy site is a block tariff, where the price per kilowatt-hour changes at different levels of consumption.
Immediately after the 75% tariff subsidy took effect, the ACPU in this site increases by about 72%, with an accompanying decrease in ARPU by about 37%. In subsequent months, we observe a steep upward trend in the change in ACPU, reaching roughly a 240% increase that leads to an 18% recovery in ARPU, which translates into a 19% decrease in ARPU 9 months after the program took effect. The increase in consumption could be attributed to the fact that following the tariff cut, some customers at this site used the additional disposable income to purchase appliances through a community member who facilitated sales from a local vendor. However, we have no knowledge of whether there was a change in the usage patterns of their existing appliances.

In the site that received a 50% tariff subsidy, a 46% increase in ACPU in the month after the subsidy took effect is accompanied by only a 2% decrease in ARPU. In the subsequent months, the growth in consumption is more measured compared to the site with a 75% tariff subsidy, culminating in a 68% increase in consumption 9 months after the subsidy took effect. At 9 months, the ARPU at this site is about 5% less than the value at the start of the program. We have no information indicating whether the increase in consumption is due to customers purchasing additional appliances or increasing the usage of their existing appliances. A majority of community members at this MG site are pastoralists, whose economy is centered around cattle rearing. We observe a notable dip and peak in the ARPU trend at 2 months prior and 3 months after the tariff cut respectively, which fall on April, which is the start of the rainy season and August, which is the peak of the dry season respectively. A possible explanation could be that during the dry season, when there is scarcity of pasture, the MG customers sell their cattle and thus have increased disposal income which they use to purchase more electricity units, while the expectation of an abundance of pasture during the rainy season may cause the customers to direct more of their disposable income to the replenishing and maintenance of their herd [39].
Since the customers in the 50% tariff subsidy site were already significantly higher consumers, with higher monthly incomes than the customers in the 75% tariff subsidy site prior to implementation of the tariff subsidy program as shown in Table 5.2, we cannot definitively attribute the differences in the effect of the program on ARPU and ACPI between these sites solely to the difference in the tariff subsidy provided.

![Figure 5.7: Effects of tariff reduction program](image)

Given the three tariff levels, that is, at 100%, 50% and 25%, and their corresponding average consumption per user values, as shown in Figure 5.8, we calculate the price elasticity of electricity demand as -1.2. This means that a 10% reduction in electricity tariff leads to a 12% increase in the average consumption per user. This shows that mini-grid customers in this admittedly limited sample are very sensitive to changes in the electricity price and there is room for developers to impact the consumption of their customers with moderate changes to the electricity price index. However, we caution that these conclusions should be taken lightly, as the sample MG customers are distinctly different (see Section 5.2.3) and the sample size is quite constrained.

Lastly, we explore the impact of the tariff subsidy program on the hourly load profiles of residential and small commercial customers. Data on the classification of customers as
either a residential connection or a business connection were only available for the site that received a 75% subsidy. In both cases there is an increase in consumption intensity in the load profile, rather than a broader range of consumption hours, as shown in Figure 5.9. This indicates that the increase in the magnitude of consumption of both residential and business customers is not accompanied by a significant change in the timing of consumption. In the case of an MG developer trying to encourage customers to shift their consumption away from peak evening hours, these results suggest that applying a uniform tariff subsidy may not be the right approach. However, given that mini-grid customers are highly sensitive to changes in the price of electricity, a more effective approach may be to apply different tariff subsidy levels to time-of-use tariffs.

5.4 Discussion

Developers, governments, donors, and communities are increasingly interested in the potential for MGs to provide power to hundreds of millions of people who lack it,
particularly in South Asia and sub-Saharan Africa. Given that grid infrastructure is expensive and time-consuming to construct and is often subject to routine load-shedding, it is evident that off-grid solutions will form part of the solution to achieve universal access to affordable, reliable, and modern energy, especially in rural communities far from the grid. While increasingly popular SHS can provide power for common uses such as lighting, mobile phone charging, and perhaps even refrigeration, these are insufficient for many commercial applications. MGs offer the potential of a combination of affordability, reliability, and capacity to service areas that need more power than a home solar panel can provide, but do not have enough load density for the central grid. At the same time, MG developers face their own challenges, chief among them whether their business models are economically sustainable—and, therefore, that the potential benefits of MGs for communities will be realized.

In this context, interventions that aid developers, donors, and researchers to better understand the constraints on demand among MG customers are especially helpful. We study the effects of an appliance financing intervention conducted among roughly 2,000
households in 27 microgrid-powered villages in East Africa, using a novel and unique data set on hourly electricity consumption, payments, and customer demographics to analyze the effects on consumption, repayment dynamics, and the economic returns of the program for developers. The results support the idea that customers face credit constraints that hinder demand growth, and relieving those constraints by providing market-rate financing to purchase appliances increases consumption—at least for several months. However, the increase in consumption does not appear to be sustained, relative to a matched control sample of customers who were offered appliance financing but did not take it up. We also report important results regarding loan repayment rates; to our knowledge, these results represent the only rigorous analysis (and indeed, the only publicly-available analysis) of loan repayment timeliness in the sector.

5.5 Conclusion

This chapter has analyzed two approaches for stimulating electricity consumption among residential and commercial customers of privately-operated mini-grids in East Africa. Our results show that the program offering a range of appliances to customers on market-rate credit terms yielded appreciable yet uneven gains in consumption growth, with a notable initial increase—the measurement of which proves to be rather sensitive to potential selection bias arising from unobservable characteristics of customers who selected into the program—eroding over time. In particular, refrigerator/freezer units showed two key benefits relative to other appliances offered: (1) significant changes in consumption that imply overall profitability of an appliance financing program based solely on these appliances and (2) time-of-use consumption patterns that complement rather than exacerbate existing evening-heavy daily consumption profiles on mini-grids.

These findings are relevant beyond the mini-grid sector. Grid operators in sub-
Saharan Africa also struggle with low consumption in rural areas that, combined with high infrastructure costs per connection, result in significant financial losses [88]. Low levels of electricity use also suggest that public funds invested in grid extension may not be achieving significant economic development gains, at least in the short term. Appliance finance programs could also be implemented in rural areas served by the grid. Lower grid tariffs may in fact permit beneficiaries to increase their electricity use more than on mini-grids.

With a smaller sample size, the tariff subsidy program we evaluated indicated mixed signals for whether overall revenue could be maintained at a lower tariff. We believe that this small-scale experiment calls for further research to find the optimal balance of increased consumption for livelihood development while driving a profitable business model for electricity service companies in settings like East Africa.

Overall, the importance of developing new and effective strategies for demand stimulation is not only important for the sustainability of emerging electricity service companies, but also, and especially so, for the mostly rural citizens gaining access to life-changing, foundational electricity services. Our work can serve as a step towards enabling those crucial gains.
6.1 Background and Motivation

Despite the growing popularity of mini-grids as an attractive alternative to grid expansion for providing power to rural and underserved communities, the long-term operation and management of mini-grids to provide electricity to the poor faces considerable financial and operational challenges. A sustainable mini-grid business model would require that the capital expenditure (CapEx) and the operating expenses (OpEx) be recovered from either initial connection costs, cost-reflective tariffs, or subsidy schemes. However, to date, there are few examples of established mini-grids that are operating sustainably in Africa [207]. Since most non-electrified households are poor and located in rural areas, mini-grid business models typically must involve low connection costs. With the challenge of high capital costs and limited to non-existent financing and subsidy schemes for mini-grids, this leaves only one pathway for mini-grid developers – charging significantly higher “cost-reflective” tariffs to recover their investments, at levels that only around 10 to 15% of rural customers can afford [212]. Given the already diminished consumption levels of newly-connected customers, the high tariffs exacerbate the problem of low capacity utilization.
This chapter focuses on a novel strategy to stimulate electricity use for human development in addition to making private mini-grid business models viable: demand stimulation on a mini-grid via electrifying fishing boats for a hybrid 600 kWp solar-battery-diesel mini-grid on an island in Lake Victoria. While our systems study deeply examines a particular setting and its attendant design and deployment challenges, we believe that our work has generalized utility. mini-grids have long sought anchor loads (e.g., telecom towers or irrigation pumps [214, 224]) to provide predictable demand and increased revenue. We extend this line of inquiry by studying tradeoffs among the multiple load classes that a financially-sustainable mini-grid may encounter. Using this lens, our study maps to a variety of demand stimulation strategies that can be applied to the great range of settings where mini-grids are found. Additionally, we characterize electric boats, a previously unstudied electric mobility load class that offers substantial promise for strengthening the electric systems of coastal communities worldwide.

Our study proceeds as follows: 1) We conduct surveys among fishing boat operators and outfit a set of fishing boats with custom tracking devices to understand the potential for adoption of this relatively new electricity technology and better understand boat usage patterns. 2) We use the insights from this in-the-field activity to size and identify an electric outboard motor and battery pack candidate, which we then incorporate into a model of electric mobility that embodies the range of usage patterns derived from our dataset. 3) We then evaluate electric mobility both technically – understanding the ability of this system to meet user needs – as well as financially – characterizing the payback of such a system within the context of a privately-operated mini-grid in our target environment, as both of these perspectives are crucial for adoption in such a setting. 4) We examine tradeoffs in incorporating this demand stimulation technique with the rest of the mini-grid, which includes a range of domestic and small commercial customers as well as an ice manufacturing operation. 5) Having modeled the operation
of the entire mini-grid, we also consider the benefits of demand response via scheduled charging of boat batteries and the implications of an alternative target depth of discharge after charging. 6) We then discuss design considerations for a boat monitoring system given our observations from the target environment and conclude the study.

6.2 Data and Methodology

In this section, we describe our modeling approach, shown in Fig. 6.1. We begin by describing our data collection process and describe the datasets collected in Section 6.2.1. Next, in Section 6.2.2, we discuss our methodology in sizing an electric outboard motor and battery based on fishing boat movement patterns. We then describe the components of the electric load on the island in Sections 6.2.3, and 6.2.4, from residential and small commercial connections and the ice factory respectively. We then present an iterative mini-grid operation model with a stochastic electric boat charging load algorithm developed to determine the maximum electric boat charging load each day over a year, while minimizing the charging infrastructure based on the capacity constraints of the mini-grid, as well as an economic analysis of the system in Section 6.2.5.

6.2.1 Data collection and description

Our study takes place on Lolwe Island, which is situated in Lake Victoria in Eastern Uganda and has an estimated population of 14,841 people. Fishing is the major economic activity on the island, home to a vibrant fishing hub of over 1,000 boats. Figure 6.2 shows an example fishing boat. Currently, the mini-grid developer for Lolwe Island is planning for a mini-grid with an installed capacity of 600 kWp of solar PV, 650 kWh of lithium-ion battery bank capacity, and a 120 kW backup diesel generator [210]. The

1The planned mini-grid is slated for commissioning in 2020.
developer is also setting up an industrial park on the island to mitigate the challenges of fish storage. It will include an ice factory to provide affordable ice for preserving Nile perch and a fish drying factory to enable efficient drying of Silver fish, which is currently done under the sun or using firewood [210]. For this study, we limited our power demand model of the industrial park to the ice factory only.

**Survey data**

We conducted a survey with 69 respondents in three villages on the island as part of our study to learn more about the socioeconomic and demographic characteristics of the fishing community, as well as their fishing habits\(^2\). Fishing boat owners on the island on average have a fleet of 3 – 4 boats and a majority of them handle the management of

---

\(^2\)We obtained approval for our study from our university’s Institutional Review Board.
their fleets and the marketing of their catch, as opposed to fishing themselves. Instead, they hire young men to fish and operate their boats. The island depends on diesel to power fishing boats. On average, each boat consumes about 20 liters of fuel per trip, which translates to approximately 20,000 liters of fuel for the entire island for every boat to make a fishing trip. On average, fishing takes place during 6 days each week. This is expensive, unreliable, and has a negative impact on the environment.

The fishing boats in use on the island are V-shaped bottom boats (see Figure 6.2), locally constructed with wood, which has a density of $440 \text{ kg/m}^3$. We measured three random boats, whose lengths ranged between $9m$ and $13m$, with a width of $1.85m$. Based on these dimensions, the shape of the boat and density of the wood, we estimated that the boats weigh between $1,209kg$ and $2,156kg$ unloaded. At the beginning of each fishing trip, we observed that each boat was loaded with fishing gear and two male fishing boat operators and on return, there was additional weight from the fish caught. Fishing boat operators in two of three villages we surveyed fish Nile perch, whose peak season is from July to December. They catch an average of $16 \text{ kg/day}$ on a day with sparse catch and an average of $98 \text{ kg/day}$ when there is bountiful catch. Fishing boat operators in the third village catch Silver fish. Peak Silver fish season is between March and June, during which operators reported catching on average about 35,000 basins of Silver fish and as few as 900 basins during low season. We therefore estimated a maximum weight, $W_{\text{boat}}$, of about $3000kg$ for a loaded fishing boat.

Boat owners in the Nile perch villages reported purchasing between 30 and 1500 $kg$ of ice a day. We therefore estimated an average of 10,000 - 15,000 $\text{kg/day}$ per day of ice demand, $Q_{\text{ice,d}}$, on the island. Transporting ice from the mainland to the island and storing it before use results in substantial ice loss and embodied energy consumption. Typically, 20% of any ice that is purchased is lost before use for fish preservation.
Boat movement data

To collect information about boat movements and the communications environment, we constructed 20 custom boat tracking devices. Each device consisted of a custom mobile application on a low-cost mobile phone (Motorola G5) to sense and log different metrics including GPS location coordinates, accelerometer and gyroscope readings, and cellular signal strength. The mobile phone was sealed in a water-tight enclosure for deployment. We favored a low-cost mobile phone for its configurability over an off-the-shelf microcontroller device, the Particle Electron IoT suite, because the Particle system is designed to store data in the cloud as it is collected. However, the cellular network on the island and in the lake is slow and unreliable, and thus a constant uplink to upload data was unavailable. While we could have stored data locally, the expansion slots on the Electron do not provide a direct way of expanding the memory storage. For our deployment, we also planned to extract logged data on a daily basis, which was easier to do from the mobile device than the Electron. A third reason is that the Electron typically uses an external 1,800mAh battery, which we projected would not have been able to last the duration of a 12-16 hour fishing trip. On the other hand, the Motorola
G5 uses a 2,800mAh battery, which can last up to 24 hours when the phone’s functions are limited to the core requirements. Finally, the mobile phone provided all the required sensors already assembled and a readily available application programming interface (API) to interact with them, compared to the Particle Electron. Note that we made these decisions given our limited deployment purpose, size, and timeline – we discuss design considerations for a longer-term deployment in Section 6.4.

The data collection device was attached to a fishing boat at the start of a fishing trip with the mobile application (app) running in the foreground, as seen in Figure 6.2. Putting into consideration the hardware constraints and software delays, the app continuously logged sensor data by sampling at 30 second intervals. We limited data logging to only the required sensors to maintain the privacy of boat operators. Over the course of six days, boat owners and operators were randomly approached each day to host the devices on their boats for the duration of their upcoming fishing trip. On each day, we deployed the devices at the start of the fishing trips in the afternoon. We then returned the next morning to detach and collect devices at the end of each fishing trip to recharge and deploy them in a different village in the afternoon. Deployment took place twice in each village. While we initially aimed to deploy all 20 devices each day, the number of devices deployed each day varied depending on a number of logistical factors such as the turnaround time of charging the phones, which was limited to about three hours each day when the island’s diesel generator was running. We logged the departure time and level of fuel in the boat’s fuel containers at the beginning of each fishing trip and the arrival time, as well as the corresponding level of fuel on arrival at the end of the trip.

We tracked 77 fishing trips in total. Figure 6.3 shows the probability distribution functions of the departure times, \( f_d(\mu_d, \sigma_d) \), and arrival times \( f_a(\mu_a, \sigma_a) \) of the 77 fishing trips. We can see the distinct differences of times based on the type of fish pursued:
Silver fish boats tend to operate for shorter journeys in the pitch dark of the night. Additionally, we can see that the window for departure times is narrow, while arrival times are more spread out. This has positive implications by enabling fewer charge stations to charge all the boats.

Figure 6.3: Probability distributions of the arrival and departure times of all 77 recorded fishing trips.

Due to challenges with environmental conditions, logistics, and a software bug, we only captured near-complete GPS traces of 27 fishing trips. Figure 6.4 shows a select number of traces that we captured. The three distinct origin points on the island are the shores of the three participating villages.

We utilize the GPS data to calculate the distance travelled by each boat between subsequent coordinates using Vincenty’s solution for the distance between points on an ellipsoidal earth model [257]. We calculated that the longest fishing trip, $d_{\text{max}}$ covered 62km round trip and the shortest trip covered 12km. Despite the Nile perch fishing trips lasting longer than the Silver fish fishing trips, we do not observe much difference in the distances covered. The mean distances of the Nile perch and Silver fish fishing trips are 24.8km and 24.5km respectively, with 80% of the fishing trips covering under 35km. We also see a distinct portion of the trip before the return trip begins which we surmise is
Figure 6.4: Select daily traces of fishing boats. Three departure villages are each denoted with an X.

when fishing is taking place. We unfortunately were not be able to determine whether the engine was on or off during this period due to noisy inertial measurement unit data.

We calculated the speeds of the boat over the course of the fishing trips based on the GPS coordinates data at a one-minute resolution using Eq. 6.1.

\[
s(t) = \sum_{i=1}^{n_i} (d_{\lambda_1,\phi_1,\lambda_2,\phi_2})_i \times 60 \tag{6.1}
\]

During the times when the boats were traveling to and from the fishing grounds, we calculated a range of maximum speeds, \( S_{max} \) between 4.5 m/s and 5.2 m/s. In between these two portions of the trips, we calculated speeds lower than 0.5 m/s, which could be indicative of a stationary or drifting boat.

6.2.2 Sizing of an electric outboard motor and battery system

The first step in the sizing an electric outboard motor is determining the required power output of the electric motor. A number of resources [179, 221] explain in detail how to calculate the force/thrust required to propel a boat based on the ship resistance model. However, for the purpose of this study, we used a simplified estimation of the power requirement of the propulsion system. We utilize a simplified thrust-to-weight
relationship [30] to estimate the thrust requirement based on the calculated maximum weight of the boats.

Though the thrust of the electric motor is important, the range of the battery is equally important and often overlooked when choosing an electric motor. We therefore also consider the range estimates of batteries compatible with the list of potential electric motors. The final selection of an electric outboard motor has sufficient thrust to propel the boat, and is compatible with a battery that has sufficient capacity and range to cover the distance of the longest fishing trip measured in our dataset.

6.2.3 Residential and commercial demand estimation

A crucial step in the deployment of any mini-grid is the assessment of electricity demand prior to implementation. Although electricity demand is hard to predict, especially in a village that has never had access to electricity, different methodologies are used to carry out electricity demand assessment of mini-grids to provide a baseline for a well-founded project design. Two fairly common practices have been used in previous studies; using primary data collected through pre-electrification surveys [115] and second, using existing demand data from other mini-grid projects in a similar socioeconomic and cultural context [144, 276].

We utilize the second approach to estimate the load profiles of prospective mini-grid customers on the island using electricity consumption data from customers of existing mini-grids in East Africa because we did not have the primary data collected for the island. These customers are categorized based on three main connection types: residential customers, small commercial customers, and residential customers who run businesses in their homes. We begin by determining the daily load profile patterns of the existing mini-grid customers, by applying a $k$-means clustering approach to the normalized hourly consumption readings of customers in each of the three mentioned
categories. The consensus among a number of studies [215, 87, 283] is that it is the best known and most frequently applied partitioning clustering technique to analyze daily load profile patterns of electricity consumers. The objective of clustering is to improve the accuracy of the predicted load profile of the prospective mini-grid customers. The clustering technique divides the customers, \((x_1, x_2, ..., x_n)\) in each connection type, \(j\), into \(k\) clusters such that similar load profile patterns are placed in the same cluster \(x_i; x_j \in C_k\) and dissimilar load profile patterns are grouped into different clusters. We determine the optimal number of clusters, \(k\), in each dataset using the NbClust approach [46]. Next, we get the average load profile of each cluster, \(P_{k,\text{avg}}\) using Eq. 6.2.

\[
P_{k,\text{avg}} = \frac{\sum_{i,j \in C_k} (x_i; x_j)}{N_{C_k}}
\]  

(6.2)

We then use a weighted allocation method to allocate the number of prospective customers in each connection, \(N_j\) to each cluster and then calculate the total hourly load of each cluster, \(C_k\), in each connection, \(j\), which we sum up to generate an aggregate load profile, \(P_{\text{cust}}\) for all the prospective mini-grid customers using Eq. 6.3.

\[
P_{\text{cust},1} ... P_{\text{cust},24} = \sum_{j=1}^{3} \sum_{C_k \in j} \frac{P_{k,\text{avg}} \cdot N_{C_k}}{N_j \cdot N_{pc,j}}
\]  

(6.3)

### 6.2.4 Modeling ice factory power demand

We modeled the ice factory as a series of similar small ice machines, which operate during hours of PV supply – that is, between the hours of 9 am and 6 pm. Assuming perfect knowledge of the next day’s demand for ice \(Q_{\text{ice,d}}\), we calculate the number of machines started up for the day to meet this demand. We also assume that there is enough storage for all the ice produced for at least 24 hours. We use the technical data of one ice machine, including its capacity \(Q_{\text{tot}}\), in kg/day, and power drawn by
the compressor, $P_c$, in addition to the number of hours in a day the machines run, $h_m$ to determine the number of machines required, $N_{mach}$, to meet demand, $Q_{ice,d}$. Lastly, we generate the demand profile of the ice machine for various levels of ice demand as summarized in Eq. 6.4

$$\left( P_9, P_{2...}, P_{18} \right) = P_c \ast \left[ N_{mach} = \frac{Q_{ice,d}}{Q_{tot}} \ast \frac{24}{h_m} \right]$$

(6.4)

6.2.5 mini-grid operation model considering stochastic electric boat charging load

We present an iterative dispatch and control model for the proposed mini-grid over a 24-hour period for each day over a year, considering the stochastic nature of electric boat charging and PV supply in Algorithm 1. While Markov Chain models have been widely used in the stochastic generation of EV charging load [110, 245], it is a decision-based time and state model that models vehicle traffic flows and as such, not applicable to the boat movement patterns in this study. We therefore propose a stochastic method based on Monte Carlo simulations that considers the boat movement patterns we observe in our data. This method has many advantages, such as possibility of simultaneous consideration of many probabilistic factors and ease of implementation. The boat movement data collected is processed to identify the probability distribution functions of battery State of Charge (SoC), $f_{soc}(\mu_{soc}, \sigma_{soc})$, hours of boat arrival, $f_a(\mu_a, \sigma_a)$ and hours of boat departures, $f_d(\mu_d, \sigma_d)$.

We initialize the algorithm with three charging stations, $N_{ch}$ one at each shore in each village. We assume that the charging stations are fast charging and the power output from the chargers, $P_{ch}$ is constant. We run the simulation for an entire year. During each day, $k$, we estimate the charging demand of the electric boats during the
charging window determined by the boat arrival and departure times. We use 1 hour as the sampling interval time. At each time step of the day, we consider the total load from the ice factory, $P_{\text{ice}}$, customer connections, $P_{\text{cust}}$ and the charging stations, $P_{\text{charge}}$ as well as the total generation from the PV array, $P_{\text{array}}$, mini-grid battery bank storage, $Q_{\text{batt}}$ and backup diesel generator, $P_{\text{gen}}$. We used HOMER software’s algorithm [73] to generate synthetic hourly solar data for Lolwe Island for an entire year by combining monthly averaged solar insolation data and the clearness index for the coordinates corresponding to Lolwe averaged over a 35-year period, from 1983 to 2018, available from NASA’s Prediction of Worldwide Energy Resource (POWER) project [3]. The PV supply is first used to meet the total load at each hour, and any excess PV supply is used to recharge the battery bank if needed. Any excess PV supply after this is curtailed. During hours of insufficient PV supply, the battery bank is discharged to meet the remaining load. The Depth-of-Discharge (DoD) is limited to 80%. When the battery bank reaches its DoD, the backup generator then ramps up to meet the remaining load.

If there is still additional mini-grid capacity during the least sunny day of the year, the number of charging stations at each shore is incremented by one during each iteration to allow for additional boats to charge until the maximum capacity of the mini-grid is reached. The model therefore minimizes the charging infrastructure required to maximize boats charging within the constraints of the charging window and capacity of the mini-grid.

In addition to the technical operation of the mini-grid, we estimate the profitability of the system characterized by Net Present Value (NPV), a parameter that expresses the initial capital investment and all future cash flows arising from operating the system over its lifetime as an equivalent amount at present time, summarized by Eq. 6.5.
Algorithm 1: mini-grid dispatch and control algorithm with stochastic boat charging load

**Result:** Maximum number of boats charged in a day, $N_{b, \text{max}}$. Annual diesel consumption from backup generator, $\text{Annual excess PV generation}, Q_{\text{excess}}$

**foreach day of year, $k$ do**

1. Initialize $N_b = N_{b, i}$;
   **foreach charging station, $N_{ch, i}$ do**
   1. Arrival time, $T_{a, i} = \text{np.random}(\mu_a, \sigma_a, n) = \text{Connection hour of first boat}, T_{\text{conn}, i}$;
   2. Battery state of charge on arrival, $B_{soc, i} = \text{np.random}(\mu_{soc}, \sigma_{soc}, n)$;
   3. Departure time, $T_{d, i} = \text{np.random}(\mu_d, \sigma_d, n)$;
   4. Charging duration, $T_{ch, i} = [0.8 - B_{soc, i}] \times Q_{\text{boat}} / P_{ch}$;
   5. Disconnection hour, $T_{d\text{isc}, i} = T_{\text{conn}, i} + T_{ch, i}$;
   6. Available charging time, $T_{\text{avail}} = T_{d, i} - T_{d\text{isc}, i}$;
   while $T_{\text{avail}} \neq 0$ do
     a. $T_{\text{conn}, i} + 1 = T_{d\text{isc}, i}$;
     b. $T_{d, i} + 1 = \text{np.random}(\mu_d, \sigma_d, n)$;
     c. $B_{soc, i} + 1 = \text{np.random}(\mu_{soc}, \sigma_{soc}, n)$;
     d. $T_{d\text{isc}, i} = T_{\text{conn}, i} + T_{ch, i}$;
     e. $T_{\text{avail}} = T_{d, i} - T_{d\text{isc}, i}$;
     f. Populate connection and disconnection matrices to keep track of the number of electric boats in charging status at every given hour, $N_{b, h, r}$.
   end
   7. Total charging load, $P_{\text{charge}, h, r} = P_{ch} \times N_{b, h, r}$

end

Sum all boats charged: $N_b = N_b + N_{b, \text{conn}}$

**foreach hour, $i$ do**

1. $P_{\text{tot}, h, r} = P_{\text{ice}, h, r} + P_{\text{cust}, h, r} + P_{\text{charge}, h, r}$;
   if $P_{\text{array}, h, r} > P_{\text{tot}, h, r}$ then
     $P_{\text{excess}} = P_{\text{array}, h, r} - P_{\text{tot}, h, r} - [650 - Q_{\text{batt}}]$;
   else
     if $[P_{\text{tot}, h, r} - P_{\text{array}, h, r}] < (\eta_{\text{batt}} Q_{\text{batt}})$ then
       $Q_{\text{batt}} = Q_{\text{batt}} - [P_{\text{tot}, h, r} - P_{\text{array}, h, r}]$;
     else
       $[P_{\text{tot}, h, r} - P_{\text{array}, h, r} - Q_{\text{batt}}] = P_{\text{gen}}$
     end
   end

end

If $P_{\text{gen, required}} > P_{\text{gen, max}}$, decrease number of boats charged per day, $N_b$ and rerun algorithm.

end

$$\text{NPV} = \sum_{t=0}^{n} \frac{A_t}{(1 + d)^t}$$ (6.5)

where $A_t$ is the project’s revenues minus costs in time $t$, from year 0 to year $n$ and $d$ is the discount rate. We calculate the NPV over a period of 20 years, which is the average lifetime of a PV system and discounted at a rate of 14%, which is the reported discount rate for Uganda [192]. For this study, we assumed that the system without infrastructure for ice production and electric boat charging has a zero NPV, which means that the project breaks even.
6.3 Analysis

In this section we present a techno-economic feasibility analysis of adding electric boat charging and ice factory load to the proposed mini-grid, the impact of infrastructure planning on the maximum electric boat charging load, the financial benefit to the boat owners and finally the impact of demand response on the operation of the mini-grid.

6.3.1 Residential and small commercial demand profile

We began by analyzing the hourly electricity consumption data of customers of 18 mini-grids in East Africa, including those described in previous work [276]. 56% of these customers have a residential connection, 36% have a small business connection, and the remaining 8% run a business from their residential premises. From the results of the clustering algorithm, we found three distinct load patterns for each category as shown in Figure 6.5.

The mini-grid developer estimates a total of 3000 prospective residential connections and 700 small commercial connections on Lolwe island [210]. Of the residential connections, using data from the Kenya Integrated Household Budget Survey [199], which indicates a proportion of rural households that run business out of their homes, we estimate 68% prospective residential connections and the remaining 12% as residential connections with businesses on their premises. Figure 6.6 shows the resulting demand profile of all prospective customers. We observe high demand during the evening hours of 8pm - 10pm, with a peak of $75kW$ at 9pm.

6.3.2 Electric outboard motor and battery sizing

Based on the estimated weight of a loaded fishing boat we calculated the required thrust to be $458lbs$, which results in about $11kW$ ($15HP$) of propulsive power at the
maximum speed of 5.2m/s. The 15 - 30 HP electric outboard motors were not compatible with batteries with sufficient capacity to last the duration of the longest trip, 62km. The 40HP Torqeedo Deep Blue 25 RL electric outboard motor, with a propulsive power of 16kW, met our requirements [260]. It is compatible with two 9.1kWh BMW i8 lithium ion battery packs connected in parallel. This battery setup has a range of between 32km
and 86 km at a speed of 12 m/s and 2 m/s respectively. It draws 3.7 kW at a 240V fast charging station. Based on these characteristics, we calculated that the total battery capacity is sufficient to last the duration of all monitored trips, with 90% of the trips using less than 65% of the battery capacity as shown in Figure 6.7. It is necessary to slightly oversize the battery to prevent stranded boats.

![Figure 6.7: Cumulative distribution of boat battery usage from two 9.1 kWh BMW i8 lithium ion battery packs connected in parallel.](image)

6.3.3 Impact on mini-grid operation

The ice factory comprises of multiple 5000 kg/day flake ice machines with compressor power of 17.5 kW, each costing $15,000 [157]. We assume that the machines operate everyday between the hours of 9am and 6pm. In our simulation of mini-grid operations, we assume that ice production takes precedence over recharging the electric fishing fleet batteries. It is becoming common practice to limit EV battery depth-of-discharge to about 20% (i.e., SoC to 80%), which reduces battery degradation and increases longevity. Our simulations therefore limit the electric boat battery charging to 80%, which we observed in Figure 6.7 to be sufficient for over 95% percent of the trips.

The average daily ice demand on the island was estimated as 13,000 kg/day. At this
level of ice production, the minimum number of charging stations required to maximize boat charging on the day of the year with minimum PV supply was determined to be 15. This is the most risk averse charging infrastructure plan, where the mini-grid operator would install just enough charging stations such that on any given day of the year, all the charging stations are in use during the entire available charging window without resulting in inadequate supply to meet the charging demand. During the charging window, a maximum of 102 boats are able to charge in a day, adding 466 kWh of load to the system (17% of total load), in addition to 1,350 kWh from ice production (51% of total load). As shown in Figure 6.8, the capacity utilization of the mini-grid increases, reducing the amount of curtailed PV supply by about 20%. However, generation from the backup diesel generator increases.

![Figure 6.8: mini-grid supply and demand curve on an average day of the year (a) without ice factory and electric boat charging load and (b) with ice production of 13,000 kg/day of and 102 boats charged at 15 charging stations.](image)

We also carried out a sensitivity analysis, to observe how the maximum daily charging load that the mini-grid can serve at different quantities of ice production between 5,000 kg and 30,000 kg is impacted by charging infrastructure planning. As shown in Figure 6.9,
we find that a more audacious charging infrastructure plan that increases the number of charging stations to 51, increases the maximum daily boat charging demand by about 240%. We also observe that below 60 charging stations, the number of charging stations limit the maximum daily charging demand on days with high PV supply, therefore we see very little variation with the quantity of ice produced. Above this, the quantity of ice production is pivotal to the number of boats that can recharge in a day. For example, we observe that when 126 charging stations are installed, almost 3 times as many boats can be charged when ice production is decreased from 30,000$kg$ to 5,000$kg$. These results help to elucidate some of the load tradeoffs between ice manufacturing and electric mobility, as well as charging infrastructure planning.

![Figure 6.9: Change in maximum daily electric boat charging demand a function of number of charging stations and ice produced](image)

6.3.4 **Impact on economics of mini-grid project**

Technical performance by itself is insufficient for the viability of this system; financial sustainability is also crucial for system viability. We quantify the impact of ice factory
and charging demand on the economics of the system, as well as capacity utilization. For the NPV calculation, we made a number of assumptions:

- The base case system, i.e., without ice production and boat charging, has an NPV of zero

- A standard tariff of $0.40 per kWh is charged for electric boat charging

- Each charging station costs $200 to install

- We assumed that all the ice produced is sold at $0.068 per kg, which is the average cost of ice quoted by the boat owners.

- Lastly, we did not account for the operating costs of maintaining the ice machines and the charging stations.

The aforementioned cost assumptions are competitive numbers based on discussions with mini-grid developers in the area. The price of diesel on the island is reported to fluctuate between $1.08 per liter and $1.32 per liter. We used the highest price ($1.32 per liter) to calculate the costs associated with the fuel consumption of the backup generator.

As shown in Figure 6.10, there is a tradeoff between maximizing the NPV of the system, and minimizing both the diesel fuel consumption from running the backup generator and the amount of PV supply that is curtailed. We observe that any level of ice production, planning the charging infrastructure to accommodate the charging of more boats per day would increase the NPV and improve capacity utilization but would also require higher usage of the backup generator. The planning of charging infrastructure ultimately depends on the goal of the mini-grid developer. Preference for a higher NPV, while ensuring adequate charging infrastructure to service a large fleet would mean compromising by adding diesel consumption, which is vulnerable to price fluctuations, not to mention the environmental costs of fossil fuel use.
Figure 6.10: Net Present Value, annual diesel consumption and annual excess PV supply as a function of number of charging stations and daily ice production. Note that the number of boats are maximized in each scenario.

6.3.5 Economic impact for boat owners

To consider the potential benefit of converting diesel-powered fishing boats to electric for the boat owners, we consider the payback period, which is the amount of time it takes to recover the cost of an investment. We consider the cost of purchasing the engine and battery system through an asset financing scheme with a loan term of 24 months, a 10% down payment and a monthly interest rate of 2.55% per month. These terms are comparable with other asset financing programs in East Africa. As part of the survey, we also examined fishing boat owners’ willingness to participate in asset financing. 90% of owners have access to either a mobile money platform or a bank account. 25% of owners reported to have requested for a loan within the past
year, mainly for expenses related to their fishing boat, with a majority of those reported
to have borrowed from family/friends or from a savings group. However, only 45% reported to have completed loan repayments. They all reported a willingness to take a loan in the near future for further investment in their fishing business.

We also consider the cost of recharging the battery at $0.40 per kWh. We compare these costs to the savings from not purchasing diesel at $1.32 per liter and the cost of maintaining a diesel engine, which we determined to be $50 per month from the survey. The repayment amount on a 40 HP Torqeedo Deep Blue engine with the corresponding 18.2 kWh lithium-ion battery system and charger, which costs $26,200 [168] would therefore be $38,766. As shown in Figure 6.11, a boat owner with an average round-trip of 25 km per fishing trip, using an average of 200 liters of fuel per week would recover the cost of their investments after about 3 years and by 5 years they could potentially see about $20,000 in savings. Shortening this payback period would entail exploring cheaper or pre-owned options for an electric engine/battery system that still meet the requirements of the boat owners.

Figure 6.11: Payback period on purchase of 40 HP Torqeedo Deep Blue electric outboard motor and 18.2 kWh lithium-ion battery system.
6.3.6 Demand response

Considering that the charging window of the electric fishing fleet does not coincide with peak demand hours of residential and commercial load, the approach we took with regard to DR is to restrict electric boat charging to the hours of PV production (7am - 6pm). This means that early boat arrivals delay connection of their batteries to charging stations until 7am. This DR strategy would require up to 56% increase in the charging infrastructure to serve the same number of boats as shown in Figure 6.12, which reduces the NPV by at most 1%. However, about 5% of fuel can be saved. This DR strategy would be useful to a mini-grid operator whose main goal is decreasing fuel consumption from running the backup generator.

Controlling the level to which EV batteries charge is also a common DR strategy. While we have established an 80% charging baseline for the electric boats in this study, we evaluated how the economic and operational performance of the grid would change if the charging limit was relaxed to allow boats to charge to 100%. We expect the number of boats that charge in a day to decrease, given all the other factors that influence the number of boats charged remain unchanged. This hypothesis is validated as shown in Figure 6.13. We observe that there is minimal difference in NPV between the two cases. However, we find that between 5000kg and 15000kg ice production, there is between 5 - 10% in fuel savings. Based on these results, it would seem favorable to the mini-grid developer to limit charging to 80% as this allows up to 60% more boats to be charged at a minimal cost to their financial and operation goals and with little (but nonzero) risk for stranded boats.
6.4 Discussion and Future Work

To test this project at scale, there is a need to explore the potential for charging station networks on other surrounding islands to address range anxiety. We are also yet to explore the potential for a boats-to-grid and boat-to-boat demand response strategy. The risk factors affecting adoption of this project, such as policy and political factors, have to be detailed and strategies to mitigate these risks explored. Real-time boat monitoring could be explored, as it allows for predictive management of boat charging, improving the operational efficiency of the mini-grid. To that end, we leverage our deployment
experience to discuss design considerations for a long-term electric boat monitoring system:

While a low-cost mobile phone is a better solution for collecting data in the short-term, it may not be feasible in the long run. A fishing expedition can last up to 72 hours. As such, even a 2800mAh battery that can hold up to 24 hours of battery charge would not suffice for this application. Therefore, an additional power bank that can potentially store up to threefold the phone battery power would be an ideal back-up. Second, a mobile device is susceptible to being tampered with and used otherwise, thus, building an embedded device would be a better solution in the long run. Third, while storing data on the device is a better solution because of network unreliability, data can easily be lost when the device comes into contact with water or gets stolen, thus, by routinely monitoring the cellular signal, the data collected can be backed up to a secure cloud service each time the device is within a better signal range. Fourth, while the rest of the world has embraced 4G and is gearing towards 5G, most places in rural sub-Saharan Africa oscillate between EDGE and 3G. Thus, the device to be deployed should cater for typically unreliable networks, and adapt accordingly. For example, from our deployment, Figure 6.14 shows that the signal strength quickly deteriorates with increasing distance from the island. There is an option of using LAN beacons atop buoys, which would then upload data to the cloud via a satellite link. These are more reliable, albeit expensive. Therefore, this approach would only work if the project is to be scaled beyond a single island. Finally, some mobile sensors (e.g., gyroscope and accelerometer) are very sensitive to small changes in the movement of fishing boats. The data collected by these sensors is noisy as the sensor readings are affected by weather changes, boat drifting, and other uncoordinated boat movements. Going forward, we could explore using commercial-grade, albeit expensive sensors, if there is a need to collect certain metrics whose sensors are prone to noise.
6.5 Conclusion

This chapter has presented a study on the potential for electric fishing boats to provide valuable and flexible load to a decentralized mini-grid system on an island in Lake Victoria, with the potential to improve outcomes for fishing boat owners and operators as well as mini-grid developers alike. We applied a survey, a low-cost boat tracking system, and substantial system modeling to create a large-scale model for understanding technical and financial tradeoffs in the mini-grid. Our work shows the significant scale of load possible from a modest deployment of electric boats, and the crucial value from adding relatively trivial control to the boat charging system. We also outline the considerations for a future boat tracking system, laying the groundwork for a much larger future operational deployment. We intend for this effort to serve as unique guidance to mini-grid developers for incorporating electric mobility to ensure the technical and financial viability of their systems, paving the way for sustainable progress towards universal electrification and its associated economic empowerment.
7.1 Introduction

The growing climate crisis is making a complete transformation of the world’s energy systems increasingly urgent. Transportation is one of the most significant emissions drivers, accounting for sixteen percent of global greenhouse gas (GHG) emissions and twenty-four percent if we consider only direct CO$_2$ emissions from fuel combustion [121]. Technology advancements and rapid reductions in technology costs, are driving the deployment and adoption of electric vehicles (EVs) globally, as a viable alternative to gasoline powered vehicles, in response to environmental and public health concerns. Presently, there are over 10 million electric vehicles on the world’s roads, predominantly in industrialized countries like China, countries in Europe and the United States [120]. In Sub-Saharan Africa (SSA), EVs are still very nascent and their adoption pathway is far less clear. Countries in the region present a set of unique challenges, including the dominance of used car imports, rapid growth in motorization, major power grid constraints, poor urban air quality, a unique transportation culture, and low vehicle...
affordability, that require a thoughtful approach to EV adoption to balance the trade-offs between economic growth, sustainable development and meeting climate and public health goals. Nonetheless, some countries in the region, like South Africa, Cape Verde, Rwanda, Zimbabwe, Egypt, and Kenya, have started to announce non-binding vehicle electrification targets and incentives for adoption [123]. Chiefly among them is Kenya, which reduced the excise duty for electric vehicle imports to 10% and is targeting a 5% share of EVs in total vehicle imports by 2025 [130].

In this work, we contribute to a sparse body of work that utilizes a data-driven approach to evaluating vehicle electrification pathways for an urban city in sub-Saharan Africa considering the unique local transportation culture and mobility patterns, as well as the constraints of the distribution grid. Our study is carried out for Nairobi, the most populous city in Kenya. Here, we explore the impacts of different electrified transportation options on the electricity distribution grid, taking into account a number of non-binding vehicle electrification targets that have been outlined locally and globally. We seek to answer the following research questions:

- What are the grid limitations of electric vehicle charging load in Nairobi?
- What is the impact of electric vehicle charging load on system load factor and capacity utilization, considering local consumer behavior and transportation systems?
- What is the potential impact of electric vehicle adoption on generation in Kenya?
- How can Nairobi’s grid infrastructure and the built environment change to better accommodate the growth of EVs?

We explore these research questions by developing a model, whose framework is presented in Figure 7.1, that fills in the gap of facilitating a city-scale analysis of
strategies for future transportation electrification planning in an urban city in Sub-Saharan Africa that considers the local transportation culture and social movement patterns, as well as the constraints of the electricity distribution system. Specifically, we investigate how various types of EVs (private, commercial, and paratransit vehicles), different modes of charging (fast and slow) and different consumer behaviors (low and high range anxiety) could affect the loading on distribution transformers under different levels of EV penetration. We leverage different independent data sources and stochastic models to fill in the data gap of transportation data that is fundamental for this work. Hourly simulations of electrified transportation are carried out for Nairobi, incorporating granular models for driving patterns, charging decision models and comprehensive modeling of the load on distribution transformers.

7.2 Literature review

EV adoption requires a dependable EV charging ecosystem, which includes sufficient charging points, power generation and a reliable electricity network. Therefore, planning for EVs will be crucial for Sub-Saharan Africa. Yet, the breadth of recent literature that explore vehicle electrification planning are focused on developed countries, mainly in the Global North. Specifically, with the anticipation of the swift growth in EV penetration, the impact of mass vehicle electrification on power generation expansion planning [167, 269, 253] and transmission expansion planning [109, 219] has been extensively studied. Further, with an increased focus on the role of EVs in climate action plans, and the fact that environmental impacts of EVs are more positive if their electricity consumption comes from renewable energy sources, the interplay between electric vehicle integration and renewable energy sources integration is a vastly growing body of work. Research topics on this front include the potential contribution of grid-connected
Figure 7.1: Flowchart showing the model and analysis framework
EVs to balance generation from RES through vehicle-to-grid systems [20, 252, 64, 84, 175], scheduling and smart charging strategies [134, 80, 81, 186], and fundamentally, the impact of EVs on emissions reductions [49, 128, 154, 256, 158].

Beyond consideration of power system capacity, planning for EV mobility needs from the context of power grid reliability and stability is of particular importance, creating the need for interdisciplinary approaches to study the complex interaction between transportation networks, electricity infrastructure, consumer behavior and power demand. Several studies have assessed the energy requirements for EVs and subsequent impact of large scale deployment of EVs on urban distribution networks based on real-world traffic and charging data using both deterministic and stochastic methods [282, 213, 183, 34, 13, 229, 183, 189, 190]. Further, to embody other real-world factors that influence EV electricity consumption, a growing body of work is focusing on EV drivers’ range concern, charging decisions and charging behavior [116, 259, 92, 217, 279]. Additionally, the availability of adequate public charging infrastructure is often cited as effective strategy to reduce consumer range anxiety and increase EV adoption. As such, optimal charging infrastructure deployment strategies have been thoroughly studied [155, 272, 164, 62].

A critical factor that has enabled this extensive body of literature in the context of industrialized countries is the availability of high quality and granular transportation, electricity system, economic and environmental data. Such data on mobility patterns of different transport modalities and their geographic distribution, grid infrastructure and granular electricity consumption is not well documented, available, or accessible in most countries in SSA. Data collected by private actors is often behind a paywall or entirely inaccessible and that collected by public actors is often in formats that render the data unusable. This data gap is evidenced by the dearth body of work on EV adoption planning in SSA, which has primarily focused on country level or city level analyses.
that inform national climate change policies [71, 55, 94, 6], or analyze the barriers to large-scale EV adoption [248]. Only one study so far assesses the grid impact of EVs using unique urban mobility patterns [32]. However, it only considers minibus taxis.

The research community that is focused on analyzing the pathways for EV adoption in SSA agree that the data and models used to study vehicle electrification planning in the context of industrialized countries cannot be shoehorned into research informing data-driven vehicle electrification planning and decision-making for SSA. One of the main reasons is that the characteristics of transport systems in SSA are fundamentally different from those in industrialized countries. For example, unlike transportation systems in most cities in industrialized countries that comprise privately owned vehicles and public buses, 50 - 98% of transportation demand in SSA cities is met by privately-owned and informally run “public” transport vehicles, known as paratransit [131]. They are known by different names in different countries e.g. matatus in Kenya, danfos in Nigeria, daladala in Tanzania, trotros in Ghana, and gbakas in Cote d’Ivoire, to refer to minibuses; bodabodas in Kenya and Uganda and Okada in Nigeria to refer to motorbike taxis; and bajaji in Tanzania to refer to auto-rickshaws [131]. As such, innovate approaches to modeling SSA-specific paratransit systems will need to be included in analyses of EV adoption pathways. Another key reason for a different approach to large-scale EV penetration in SSA is the grid capacity limitations, which are more pronounced than in industrialized countries, as well as the unreliable electricity supply and low power quality. In fact, a recent 2019 study suggests that less than half of Africans enjoy a reliable supply of electricity [51]. Therefore vehicle electrification and electricity system planning will have to be considered in tandem.
7.3 Transport electrification pathway scenarios

We tested different scenarios envisioning how adoption of electric vehicles in Nairobi may change in the future. Our analysis simulates five scenarios that account for EV adoption rates of different fleet types. The first two scenarios are Kenya specific scenarios. The first, the stated policy scenario, captures Kenya’s EV adoption target of 5% of new vehicle registrations by 2025 outlined in the Kenya National Energy Efficiency and Conservation Strategy [177] and the second captures the Kenya National Strategic Plan goal of 15% penetration of 2 and 3-wheelers [187]. The last two scenarios are global EV adoption scenarios. The first of the two captures the International Energy Agency’s (IEA) Sustainable Development scenario which projects that the stock of electric two/three-wheelers reaches 40% of all two/three-wheelers by 2030, and an almost 15% stock share of passenger cars, light duty commercial vehicles and buses [120]. The last scenario captures the Electric Vehicle Initiative scenario which targets at least 30% of new electric vehicle sales by 2030 [122].

A common behavior of drivers in Kenya is taking trips to the gas station when the vehicle’s tank is nearly empty with the quantity of their gas purchases being heavily dependent on the cash on hand. Yet, passenger vehicle owners surveyed in Kenya expressed skepticism that EVs can meet their mobility needs and noted range anxiety - the fear of being stranded due to an EV’s battery having insufficient charge to reach one’s destination - as one of their primary barriers to adoption [58]. Therefore, to better understand the importance of behavioral realism, in each of the above scenarios, we allow for two scenarios of consumer behavior with respect to range anxiety. Details of the above scenarios are summarized in Table 7.1
Table 7.1: Description of the scenarios considered for this study

<table>
<thead>
<tr>
<th>Scenario</th>
<th>EV penetration and Fleet composition</th>
<th>Consumer behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stated Policy Scenario</td>
<td>5% of total vehicles registered by 2025</td>
<td>Low range anxiety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High range anxiety</td>
</tr>
<tr>
<td>Kenya National Strategic Plan Scenario</td>
<td>15% of annual 2 and 3 wheeler registrations by 2028</td>
<td>Low range anxiety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High range anxiety</td>
</tr>
<tr>
<td>IEA Sustainable Development Scenario</td>
<td>40% of 2 and 3-wheelers; 15% of passenger cars, LDCVs and buses</td>
<td>Low range anxiety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High range anxiety</td>
</tr>
<tr>
<td>Electric Vehicle Initiative Scenario</td>
<td>30% of all vehicles by 2030</td>
<td>Low range anxiety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High range anxiety</td>
</tr>
</tbody>
</table>

7.4 Methods

7.4.1 Model overview

Our analysis process follows the steps below, and are visually summarized in Figure 7.1:

- We begin by estimating spatial vehicle ownership across Nairobi at a high granularity.

- We estimate the daily driving patterns of different vehicle fleet types by employing a probabilistic approach using person and vehicle mobility probability distribution functions (arrival and departure times, average commute times, vehicle travel times and distances) in Nairobi.

- We determine if and when EVs decide to charge based on their range anxiety and state of charge. We use a stochastic Markov chain charging behavioral algorithm adapted from [92].

- Using the estimated driving patterns and charging decisions, we estimate proba-
bilistic charging behavior (arrival and departure times, the state of charge, duration of charging events) for vehicles of different fleet types under different charging conditions.

• The probabilities are employed in Monte Carlo simulations to project the collective charging demand for vehicles for each spatial unit of analysis throughout the work week at a 1-min time resolution. We optimize the number of charging points to minimize queuing.

• Lastly, we estimate the potential impacts of different vehicle electrification strategies on Nairobi’s distribution transformers, and demand profile under a set of scenarios.

Each of the steps are discussed in detail below:

*Modeling spatial vehicle ownership*

To sample EV owners from all vehicle owners in each of the 400 zones of Nairobi, we first extract vehicle owners from the entire population using information from Kenya’s 2019 Census data on household vehicle ownership rates and household sizes for constituencies in Nairobi [141], and geospatial population data from WorldPop [281]. Previous studies [152, 65] have shown the rate of vehicle ownership in relation to income can be modeled by an S-shaped curve. Also, income and overall wealth have been shown to have a positive and significant correlation with electricity consumption [203]. We validate that this correlation holds for Nairobi’s data by evaluating the correlation between constituency level electricity consumption and the constituency wealth status variables (household ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and
sanitation facilities) obtained from the Census data used to construct the Demographic Health Survey Wealth Index [226].

Given that we have granular electricity consumption data and only constituency level wealth data, we use the average electricity consumption of each zone as a proxy for wealth and construct wealth indices for each zone. We fit a log-logistic function represented by Eq. 7.1 [156], to the zone wealth indices and average vehicle ownership rate.

\[
V = \frac{V_{\text{max}}}{1 + e^{-\lambda(\ln(Y)-\theta)}}
\]  

(7.1)

where \( V \) is the vehicle ownership rate, \( V_{\text{max}} \) is the saturation level of vehicle ownership, \( Y \) is the wealth index and \( \lambda \) and \( \theta \) are two positive parameters defining the shape of the curve. We adjust the curve parameters so that the cumulative number of vehicles from the model are about the same as the total number of vehicles in the study area mentioned in section 7.4.2.

We utilize a similar function to estimate the zone level distribution of LDCVs. In this case, we hypothesize that the number of LDCVs are a function of the level of commercial activity, for which we also use the small commercial customers’ electricity consumption as a proxy. For each scenario, the number of EV owners randomly sampled for each fleet type in each zone correspond to the scenario’s EV adoption rate.

A Monte Carlo simulation draws from the registered vehicle curb weight probability distribution function (Figure 7.2) to randomly assign each EV owner electric drive train technical parameters (range, battery capacity, energy consumption and charge power).
Simulating EV driving patterns

A key limitation of our study is the lack of granular transportation data, specifically real-world data on driving patterns in Nairobi. We overcome this challenge by stochastically generating synthetic driving patterns using probability distributions of commuter departure, arrival and commute times from surveys and vehicle travel times and speeds from the Uber movement data. The departure times, travel times or vehicle kilometers traveled (VKTs) and initial SOCs for each EV in each origin zone were generated using Monte Carlo simulations specific to fleet type, subject to the appropriate probability distribution.

Private vehicle mobility: For scenarios exploring private vehicles, we first determined the trip destination (work place or in public) stochastically. We then sample the departure times and duration of commute stochastically based on the trip destination. Based on the departure and commute times, we determine the destination zones of each EV by sampling the hourly inter-zone travel time probability distributions that reflect the traffic conditions within the city at any given hour of the day and the level of commercial activity in each zone. The probability of traveling to a destination zone is given by:
\[ P_{\text{dest}} = \arg\min_{P_{\text{th}}, P_{\text{c}}} \]  \hspace{1cm} (7.2) 

where the likelihood of driving to a destination zone, \( P_{\text{dest}} \), is given by jointly maximizing the probability of no time difference between the stochastically drawn commute time and the travel time to a destination zone at hour, \( h \), and the probability of high commercial activity in a destination zone, \( P_{\text{c}} \).

Once the destination zone is determined, we estimate the travel speeds, which we use as inputs to the EV energy demand estimation model described in section 7.4.1. We simulate the return trip to the origin destination (home) at the end of the day by sampling departure times from the probability distribution of the \( \text{home} \) trip destination.

**Commercial vehicle mobility:** For scenarios exploring light duty commercial vehicles, we first determined the daily vehicle kilometers traveled (VKT/day) stochastically based on the LDCV probability distribution of VKT/day. We modeled the number of stops per trip per e-LDCV as a linear function of the VKT/day fitted to data from a 2017 pilot study of 4 electric LDCVs in Nairobi [93]. The equation is given by:

\[ N_{\text{stops}} = 0.0606 \text{VKT}_{\text{day}} + 0.5606 \]  \hspace{1cm} (7.3) 

We determine the average trip distance stochastically subject to the sum of average trip distances equating to the e-LDCV’s daily VKT. Further, we randomly determine the stop duration, between 5 and 120 minutes, with uniform probability. We sample the initial departure time of each e-LDCV from a uniform probability distribution covering the morning peak hours (6 a.m - 9 a.m) and sample the end of day arrival time back at the origin zone from the probability distribution function of the \( \text{home} \) trip destination (Figure 7.6b) derived from the Nairobi person trip survey. Like the case of private vehicles,
we determine the destination zones of each stop by sampling the hourly inter-zone travel time probability distributions based on the departure times from each stop and subsequently calculate the average speeds which we use as inputs to the EV energy demand estimation model.

**Paratransit mobility**: 14-seater matatus and minibuses which accommodate between 25 and 35 passengers are the most common mode of transportation in Nairobi, used by approximately two-thirds of daily commuters [230]. For scenarios exploring the 14-seater matatus and 25-seater minibuses, we assume that charging only takes place at the outbound transit terminals, not located within the central business district. The spatial locations of outbound terminals of all the matatu and bus rapid transit routes are highlighted in Figure 7.9a. We model the departures of the paratransit vehicles on each route at the start of each day (operations begin at 6:00 am) following the gamma distribution of a poison process as shown in the following equation:

\[
  f(x) = \frac{1}{(n-1)!} \lambda^n x^{n-1} e^{-\lambda x}
\]  

(7.4)

where \( f(x) \) is the density function for the waiting time, \( t \) until the departure of the \( n^{th} \) EV with an average rate of departures of \( \lambda \) per unit time. Based on data from the digital matatus project, during peak operating hours, the paratransit vehicles depart every 2 minutes [278]. We determine the road segments from the speeds data set that correspond to those that join together to make up each of the paratransit routes shown in Figure 7.9a. We determine the duration of travel along each road segment by sampling the hourly speeds probability distributions based on the time when each EV arrives at the particular road segment, as well as the length of the road segments in km. We use the speeds obtained as inputs to the EV energy demand estimation model. We make the assumption that the electric matatus and minibuses are only charged during off-peak...
Motorcycle mobility: Motorcycles, bodabodas, though an unconventional and unregulated mode of transport, continue to make a significant contribution to urban mobility in many SSA cities. Yet, their mobility patterns are vastly understudied and very little information exists on how they move within urban cities. Nonetheless, there have been previous efforts aimed at broadly characterizing the market share, trip distances and trip fares of motorcycle taxis operating in some African cities like Accra, Douala, Kampala and Lomé [262, 201, 9, 70] but none for Nairobi. A study by [78] that both quantitatively and qualitatively analyzes bodaboda mobility within Kampala found that most bodabodas journeys tend to be between the central business district, the nucleus of commercial activity and residential areas at the periphery of the city, especially low income residential areas. For scenarios exploring two-wheelers in this study, we make a similar assumption and simulate motorcycle journeys between residential and commercial zones within Nairobi similar to private and commercial fleets. We therefore simulate motorcycles a proportion of the private and commercial fleets.

We use data on modal share of privately owned motorcycles versus four-wheeled vehicles in different zones of the city from the 2013 person trip survey [278]. We use these ratios of motorcycle ownership to four-wheeled vehicle ownership to model the spatial distribution of the privately owned motorcycles. For the commercial motorcycle fleet, we use a similar model to the commercial four-wheeled vehicles, that is, modeling ownership based on the level of commercial activity in each zone.

We acknowledge that motorcycles within Nairobi move differently than four-wheeled vehicles. They have the benefit of being able to move more quickly through congested roads as they can maneuver through small spaces. A recent study analyzed traffic counts on a major road within Nairobi and found that motorcycles on average move almost twice as fast as private and commercial fleets. Further they find that private and
commercial fleets take four times longer than motorcycles to complete a certain trip during peak hours and three times longer during peak hours [240]. We use the findings of this study to adjust the travel times and speeds drawn stochastically drawn from the probability distributions of the Uber travel time functions in the case of simulating motorcycles.

**EV energy demand model**

We estimate the energy demand of each EV trip with a drivetrain model that builds on the relationship between the energy consumption (kWh/km), the average travel speed (km/hr) and the trip distance. That is:

\[
E_{\text{trip}} = f(V_{\text{avg}}) \times d_{\text{trip}}
\]

where \(E_{\text{trip}}\) is the energy consumed by the EV during the trip in kWh, \(d_{\text{trip}}\) is the route distance and \(f(V_{\text{avg}})\) is the energy consumed per km (kWh/km) when the EV is traveling at speed \(V_{\text{avg}}\) (km/h). The \(f(V_{\text{avg}})\) is dependent on the EV battery type, therefore different EV models show different shapes of \(f(V_{\text{avg}})\). In this work, we leverage algorithms developed by [95] for energy consumption and speed for eight commonly sold EVs using publicly available lab dynamometer tests [12]. From the eight curves, we get the average energy-speed curve represented by the equation below:

\[
f(V_{\text{avg}}) = 0.2022 - 0.0032V_{\text{avg}} + 0.00003V_{\text{avg}}
\]

The energy consumption values in the technical specifications of each EV model in the data obtained from the EV database, presented in Table D.1, are given for an assumed speed of 110 km/hr. Using these energy consumption values and the average
energy-speed curve, we generate energy consumption - speed curves for each of the 87 EV models. Some example energy-speed curves generated from the ev options used in this study are presented in Figure 7.3.

![Energy consumption curves with speed](image)

**Figure 7.3: Energy consumption curves with speed**

*Determining EV charging decisions and simulating charging*

We simulate EV movement patterns over 5 weekdays to create an average 24 hour driving and charging profile for each EV. We assume that all the batteries are fully charged at the beginning of the simulation i.e. at the beginning of the week. The battery soc at the beginning of the subsequent days is determined by the charging decisions made in the previous day. For each trip, in the case of all fleet types, we determine the arrival battery SOC based on the SOC at the time of departure, $SOC_{dep}$, the energy consumed during the trip, $E_{trip}$ and the available EV battery capacity, $EV_{cap}$ as follows:
\[ \text{SOC}_{\text{arr}} = \text{SOC}_{\text{dep}} - \frac{E_{\text{trip}}}{E_{\text{cap}}} \quad (7.7) \]

We develop a user charging decision model based on a stochastic Markov chain charging behavioral algorithm adapted from [92]. Inputs to the model are logistic functions which define and adjust the driver’s behavioral related parameters that represent the probabilities of connecting to and disconnecting from charging when the EV battery reaches a certain state of charge. These functions are given by:

\[ p_i = \frac{1}{1 + e^{-k_p(i-x_p)}} \quad (7.8) \]

\[ q_i = \frac{1}{1 + e^{-k_q(-i+x_q)}} \quad (7.9) \]

where \( p_i \) is the probability of disconnecting when charging and \( q_i \) is the probability of connecting when driving or on arrival based on state of charge, \( i \). \( k_p \) and \( k_q \) are the coefficients of the gradient change in the logistic functions and control the level of range anxiety. \( x_p \) and \( x_q \) are the central points of the gradient and depend on the stage of charge. For example, driver exhibiting a high level of range anxiety, will be described with a larger \( x_p \) and/or \( x_q \) and a larger \( k_p/k_q \), while a driver exhibiting a low level of range anxiety will be described with a larger \( x_p \) and/or \( x_q \) and a larger \( k_p/k_q \).

For the low range anxiety scenario, we define \( x_p = 60, x_q = 15, k_p = 0.3 \) and \( k_q = 0.35 \) such that there is almost a zero percent chance of an EV driver choosing to charge when the EV state of charge is greater than 30% and for electric vehicles charging there is almost a 100% chance of disconnecting when the battery state of charge is greater than 70% as shown in Figure 7.4. For the high range anxiety scenario, we define \( x_p = 90, x_q = 50, k_p = 0.5 \) and \( k_q = 0.5 \) such that there is almost a 100% chance of an EV driver
choosing to charge when the EV state of charge is less than 45% and almost a 0% chance of an EV driver who is charging to disconnect when the battery state of charge is less than 80% as shown in Figure 7.5.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure7.4.png}
\caption{Low range anxiety probability distributions}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure7.5.png}
\caption{Low range anxiety probability distributions}
\end{figure}

Given information on simulated arrival/departure times, the user charging decision and the charging requirements for each vehicle, the charging decisions and requirements of the EVs are aggregated to schedule charging such that each EV that requests to charge begins charging instantaneously at either its rated level 2 power or the maximum power allowed at the level 2 charging point, whichever is bottleneck, continuing until the state of charge from the decision to disconnect is achieved. This strategy optimizes the number of charging points required such that there is no queuing in any zone. In all the simulations performed in this study, uncoordinated slow charging is assumed, meaning that each EV starts charging as soon as it is connected to the grid.
Impact of EV charging on the power system

We assess the impact of EVs on Nairobi’s power infrastructure in two ways: first, the impact of increase in the quantity on demand and timing of demand from electricity consumed by EVs on the loading of distribution transformers. Second, the impact of increased demand on the amplitude and shape of Nairobi’s cumulative demand profile and the potential influence on electricity generation. We consider the worst case scenario that EVs are charged in an uncontrolled way meaning that the electric utility has no control over when charging occurs.

While we have data on the transformer capacities and transformer loading, we do not have data on the transformer load profiles, which will be crucial in determining how the EV demand profiles impact transformer loading. Therefore using both the customer and transformer location data, we map each residential, commercial and industrial electricity customer to their respective distribution transformer. We make an assumption that each residential and small commercial customer in Nairobi follows a similar demand profile as the aggregate demand profile. For each transformer, we scale the aggregate demand profile based on the number of residential and small commercial customers connected to the transformer as well as the transformer loading. We use the smart meter data for 400 industrial customer to develop an average demand profile for industrial customers. We add industrial customer demand profiles to the resulting transformer profiles of the transformers they are connected to. We aggregate transformer load profiles at zone level and calculate baseline transformer loading, $T_{x_i,\text{load}}$ as:

$$T_{x_i,\text{load}} = \frac{T_{x_i,\text{peak}}}{T_{x_i,\text{max}} \ast pf}$$  \hspace{1cm} (7.10)

where $T_{x_i,\text{peak}}$ is the peak demand of transformer $i$ in kW and $T_{x_i,\text{max}}$ is the maximum transformer capacity in kVA and $pf$ is the average power factor.
The revised transformer loading values for each zone are recalculated after adding the aggregate level EV charging demand profile to the baseline transformer load profile.

7.4.2 Data sets and sources

Electricity demand and generation: To model Nairobi’s electricity demand profiles at the distribution level, we relied on closed-source data obtained from Kenya’s utility company, Kenya Power, which included monthly billing information, electricity meter information and GPS coordinates of over 5 million residential customers, over 200,000 small commercial customers, and about 3,000 industrial customers for 2015. Of these customers, about 1.85 million residential customers, about 200,000 small commercial customers and 140 industrial customers are located within our study area. In addition, we obtained GPS coordinates of about 20,000 distribution transformers, as well as their rated capacities. Of these transformers, about 8,100 are within Nairobi. We also use data on Kenya’s generation mix for 2017 and 2018, which comprises of hydro, geothermal, solar and thermal supply.

Travel behavior: Travel surveys are widely used in metropolitan areas to collect information on the travel behaviors of individuals in a given day. For this work, we make use of a person trip survey conducted by the Japan International Cooperation Agency (JICA) within the city of Nairobi in 2013, in which 10,000 households were interviewed [198]. We used information on the daily trip rate per person, categorized by mode of transport and trip purpose, information on the mode of travel for different zones of the city, as well as information on the hourly distribution of travel times, departure times and arrival times by trip purpose in our simulation model of driving patterns of private vehicles. For light duty commercial vehicles (LDCVs), we make use of a questionnaire-based quantitative transport survey conducted in Nairobi in 2015 [171]. We generated a probability distribution function of vehicle activity represented by vehicle kilometers
traveled (VKT) per day from data collected from 58 LDCVs. For paratransit transport, we used route information, stops and trip schedules for 135 routes of the city’s semi-formal paratransit system called *matatus* in a General Transit Feed Specification (GTFS) format obtained from The Digital Matatus project. Detailed information on the structure of the data can be found in ref [278]. We also obtained route information for the five bus rapid transit corridors proposed as part of the Nairobi Mass Rapid Transit System program.

**Vehicle movement:** A fundamental input to our vehicle mobility model is granular spatio-temporal traffic activity data for Nairobi. These data on vehicle travel times and speeds within the city of Nairobi was obtained from the Uber Movement Project [264]. Uber’s movement data provides statistical data about travel times between different zones of a city, and another data set provides hourly measurements of street speeds across a city. We downloaded aggregated measurements of travel time during weekdays and weekends for the city of Nairobi for each quarter of the period 2016 - 2019. The data includes both the central and suburban regions of the city which is divided into 400 zones and includes entries of the hourly mean and standard deviation of travel times between zones. Detailed information on how the statistics were derived can be retrieved from Uber’s official methodology paper [263]. We also use Uber’s street speeds data set to simulate driving patterns of the paratransit (14-seater *matatu* and 25-seater minibuses) fleets on their given routes. The Uber speeds data for Nairobi includes the mean and standard deviation of hourly speeds for 82,275 road segments within the city for each day of 2018 and 2019.

**Vehicle registration:** Information on the number of vehicles registered in Kenya from 2007 - 2018 is publicly available from the CEIC database [66]. We use vehicle registration data dis-aggregated by fleet type provided by the Kenya Revenue Authority (KRA) spanning 2010 to 2016 to estimate the number of vehicles in Kenya in 2018 by fleet type. We estimate that of the 3,280,934 vehicles registered in Kenya in 2018, 70%
were passenger vehicles, 14% were light duty commercial vehicles, 11% were heavy duty commercial vehicles and the remaining 5% were public transit vehicles. We then estimate the proportion of Kenya’s registered vehicles that are owned within our study area by considering Nairobi’s economic activity captured by the Gross County Product (GCP) as a proportion of that of the entire country [117]. We estimate that there are about 625,000 privately owned vehicles, 193,000 light duty commercial vehicles, 11,000 public service minivans (matatus) and 600 public service buses owned and mainly operated within Nairobi. Each vehicle in the KRA data set also contains information on it’s tare (curb) weights. We found that mean curb weight for registered private vehicles is about 1,400 kg with a standard deviation of about 400 kg. The mean curb weight for registered light duty commercial vehicles is about 1,700 kg with a standard deviation of about 400 kg. The vehicle curb weight is one of the main factors that influences the EV battery capacity and energy consumption, which in turn influences the range of the vehicle [182]. We obtain the technical information of 87 electric vehicle models, that is electric range (km), battery capacity (kWh), charge power (kW AC and kW DC), vehicle energy consumption (Wh/km) and tare weight (kg), also known as unladen weight from the electric vehicle database [67]. This information is summarized in Table D.1.

All data sources can be found in Table 7.2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Region</th>
<th>Year of dataset</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity consumption data</td>
<td>CSV</td>
<td>Kenya</td>
<td>2017</td>
<td>Kenya Power</td>
</tr>
<tr>
<td>Electricity generation data</td>
<td>CSV</td>
<td>Kenya</td>
<td>2015 - 2017</td>
<td>Kenya Power</td>
</tr>
<tr>
<td>Distribution transformers data</td>
<td>CSV</td>
<td>Kenya</td>
<td>2015 - 2017</td>
<td>Kenya Power</td>
</tr>
<tr>
<td>Vehicle movement data</td>
<td>Raster</td>
<td>Nairobi</td>
<td>2018-2019</td>
<td>Uber Inc. [264]</td>
</tr>
<tr>
<td>Road network</td>
<td>Raster</td>
<td>Nairobi</td>
<td>2018</td>
<td>Uber Inc. [264]</td>
</tr>
<tr>
<td>Person trip survey</td>
<td>PDF</td>
<td>Nairobi</td>
<td>2013</td>
<td>JICA [198]</td>
</tr>
<tr>
<td>Vehicle travel survey</td>
<td>PDF</td>
<td>Nairobi</td>
<td>2019</td>
<td>Mbandi et.al. [171]</td>
</tr>
<tr>
<td>Paratransit (matatu) data</td>
<td>GTFS</td>
<td>Nairobi</td>
<td>2015</td>
<td>Digital Matatus project [278]</td>
</tr>
<tr>
<td>Vehicle registration data</td>
<td>CSV</td>
<td>Kenya</td>
<td>2012, 2015</td>
<td>Kenya Revenue Authority</td>
</tr>
<tr>
<td>Total vehicle registrations</td>
<td>online</td>
<td>Kenya</td>
<td>2007-2018</td>
<td>CEIC database [66]</td>
</tr>
<tr>
<td>EV technical information</td>
<td>online</td>
<td>n/a</td>
<td>2022</td>
<td>EV database [67]</td>
</tr>
</tbody>
</table>
7.5 Analysis

7.5.1 Estimating the mobility of different EV fleets

We hypothesize that early adopters of passenger electric vehicles in Nairobi will be in the more affluent areas of the city. Currently, there are just over 300 electric vehicles registered in Kenya, primarily in Nairobi and a few public charging stations located in the relatively wealthier neighborhoods in Nairobi [142]. We therefore model the potential spatial distribution of private electric vehicles based on the socio-economic status of different zones within the city (see Section 7.4.1). As illustrated in Figure 7.7c, our analysis suggests that early adoption on private EVs will be concentrated to the north-west and south-west regions of the city, that is, the Westlands, Kilimani and Lang’ata constituencies. A previous study analyzing urban travel in Nairobi using household survey data also found higher vehicle use in zones that are relatively wealthy in the northern, western and south-western parts of the city [230]. The results of our trip generation model, shown in Figure 7.7a suggests that during the week private EV trips are likely to be concentrated in zones surrounding the central business district (CBD) and to the east of the CBD such as those in the Embakasi, Makadara and Kamukunji constituencies. Based on our data of the location of commercial and industrial electricity customers, there is significant commercial and industrial activity in these zones. We compare the results of our vehicle mobility model to those of the JICA trip generation and attraction model (Figure 7.7b) based on their 2013 Person Trip Survey [198]. We observe that our results of where there would likely be significant vehicle activity of passenger vehicles within the city are comparable to theirs.

We also analyze the resulting weekday daily distances traveled by the private EVs simulated in our model. Data from the person trip survey indicates that the average trip rate is about 2 trips per day, meaning that the private vehicles in Nairobi are mainly used
for commuting between homes and work places during the week. We therefore make
the assumption that each EV makes one round-trip, that is from home to its destination
and back home. The distribution of departure and arrival times, which are based on data
obtained from [198] are shown in Figure 7.6. The majority of commuters leave home
between 6:00 am and 9:00 am and return home from work between 5:00 pm and 8:00 pm.

As shown in Figure 7.7d, we estimate that most daily EV trips would be between
15 and 40 km/day. This distribution of distances is slightly higher than those estimated
from the JICA person trip survey responses – majority of trips from home to work
were found to range between 1 and 15 km and a similar range of distances for work
to home trips [198]. Similarly, in their study, [230] also find that private vehicle trips
within Nairobi are quite short, also ranging between 1 and 13 km between destinations. However, based on their vehicle fleet survey of a representative sample of about 200 private vehicles in Nairobi, [171] found high private vehicle usage of about 50 ± 15 km/day.

Light-duty commercial vehicle (LDCV) fleets in Nairobi are typically informal vans and trucks for hire within the city to transport goods and run errands [171]. We therefore simulate electric LDCV mobility on the assumption that they will also be used for trip purposes such as running company errands, doing technical field work or delivering goods (see Section 7.4.1). As shown in Figure 7.8a, we find that most of the electric LDCVs

---

**Figure 7.7:** Simulated origin-destination private vehicle movements
will likely originate from zones within and around the central business district since these zones have a high level of commercial activity. Comparable to the average daily vehicle kilometers traveled of LDCVs surveyed in Nairobi [171], we estimate electric LDCV usage of $55 \pm 25$ km/day. To cover this daily distance, we find that the electric LDCVs would have to mainly service zones in the peri-urban regions of the city. For example, zones in Roysambu, Juja, Thika, Kasarani, Kikuyu and Limuru constituencies, as shown in Figure 7.3. Further, each e-LDCV trip would include 3 to 5 stops/destinations.

Over 50% of Nairobi residents rely on some form of public transportation to meet their day to day mobility needs [198, 231]. Further, over 98% of residents report having access to Nairobi’s paratransit network [231]. As shown in Figure 7.9a, the paratransit

Figure 7.8: Simulated commercial vehicle movements
network is quite vast, linking the central parts of the city, with the peri urban areas, where a lot of the lower income residents live [230]. Based on the data obtained from the *DigitalMatatus* project, *matatus* and minibuses are normally in service between 6:00 am and 11:00 pm. Further, they stop for about 5 min at the terminals to load passengers during peak operating hours of 6:00 am - 9:00 am and 4:00 pm to 8:00 pm, and stop for about 15 min during off-peak hours [278]. From the results of our simulation, we find that *matatus* and minibuses operating on shorter routes, within 5 km of the CBD make between 20 and 30 trips a day, with those operating on the longer routes, over 15 km from the CBD, making about 5 trips a day as shown in Figures 7.9b and 7.9e. Further, we find that most *matatus* cover between 110 and 170 km a day, while most of the minibuses cover between 120 and 180 km a day. Our results are comparable to the results of the 2015 Nairobi travel survey, which found that 14-seater *matatus* on average cover between 125 and 175 km a day, with the 25-seater minibuses covering a similar range of daily distances [171]. Given that the expected driving distance of an electric *matatus* and minibuses in this study is 140 km and 250 km, the electric *matatus* and minibuses must be charged at least once a day at night, with *matatus* likely charging more than once (also during the day).

**7.5.2 Potential implications of future vehicle electrification on Nairobi’s grid under different scenarios**

The potential implications of EV charging on the distribution network under different scenarios was evaluated on a 30-minute based over a five day (work week) period based on the simulation results described above. We consider EV charging load aggregated at a zone level and its impact on transformer loading, also aggregated at a zone-level resolution (see section 7.4.1). First, we employed the user charging decision model to each EV’s simulated driving pattern. We found that on average drivers chose to re-charge
when the battery state of charge fell below 20% in the low range scenario compared to 40% in the high range anxiety. This is reflected in our results in Figure 7.10, which show that on an average week, there were as high as four times more private EVs connected simultaneously in the high range anxiety scenario compared to the low range anxiety scenario. In the case of the paratransit fleet, EVs in the high range anxiety scenario also choose to connect much sooner during the day than in the low range anxiety scenario, as shown in Figure 7.10c.

In a business-as-usual (BAU) scenario, without electric transportation, we observe that zones in the eastern part of the city have the highest peak demand of between 100
and 200 MW peak, with zones in the peri-urban areas of the city having under 50 MW peak demand as shown in Figure 7.11a. This is likely due to the fact that zones in the eastern part of the city have the highest number of commercial and industrial electricity customers. Not surprisingly, these zones also have the highest cumulative transformer capacities (Figure 7.11b). Therefore, we find that transformer loading is highest in zones in the peri-urban regions of the city such as those in Kikuyu, Kajiado North, Kasarani and Juja constituencies, as shown in Figure 7.11c.

We observe that estimated EV charging demand under all scenarios results in more load than the maximum aggregate transformer capacity available in at least one zone, as summarized in Figure 7.12. Specifically, in the case of low range anxiety, where ev
drivers charge less frequently, about 1% of zones would need to upgrade transformers to serve EVs under all four scenarios. On the other hand, in the case of high range anxiety, we find that in a higher percentage of zones, the capability of the distribution network to support EV charging would be restricted by transformer capacities. Specifically 8% of zones in the sustainable development scenario which includes electrification of 40% of motorcycles and 15% of private vehicles, commercial vehicles and paratransit vehicles; 6% of zones in the EV initiative scenario which includes electrification of 30% of all vehicles and 2% of zones in the national strategic plan which includes electrification of 15% of motorcycles.

We spatially visualize the aggregate transformer loading of each zone after adding
the estimated ev half-hourly charging demand to the baseline demand from residential and commercial electricity customers under all four scenarios. As shown in Figure 7.13 observe that in the sustainable development scenario and EV initiative scenario most of the zones experiencing transformer loading are concentrated in the south-west region of the study area, specifically in Lang’ata constituency, with some zones in Kikuyu, Embakasi East and Westlands constituencies also exhibiting transformer overloading. The likely reason for this observation is that Lang’ata and Westlands constituencies being one of the most affluent areas of the study area will likely have a high adoption of EVs (Figure 7.7c) which will result in significant evening charging in residences that own EVs. Further, zones in Embakasi East have high commercial activity, which would explain overloading in a number of zones in this region. In the case of the stated policy and national strategic plan scenarios, albeit resulting in less zones with transformer loading, we observe that these zones also fall within Kikuyu, Embakasi East and Westlands constituencies. These results suggest that there may be a need for the utility to prioritize the upgrading of distribution infrastructure in regions where there will likely be high initial uptake of EVs such as Lang’ata and Westlands constituencies. Further, it would be imperative to explore smart charging strategies to ensure that residential EV charging does not exacerbate the evening peak.
From an equity perspective, the above findings are problematic since they are indicative of prioritizing EV investments in the more affluent neighborhoods in the city, where most of the vehicle owners live. We therefore explore an additional equity scenario of solely electrifying the paratransit network. The decision to explore an all electric paratransit system as an equity scenario is because majority of the commuters within the city and in the peri-urban areas around the city rely on the paratransit system to meet their mobility needs. Further, in their study, [230] found that as residents in Nairobi get wealthier they increasingly choose to own private vehicles, therefore an overwhelmingly large proportion of commuters who use the paratransit system are low income residents.
7.5.3 Potential implications of future vehicle electrification on the city’s load profile and national generation under different scenarios

We also consider the potential impacts of the estimated EV charging load on the load factors of Nairobi’s load profile. By considering the changes in load factor given by Equation 7.11, we explore whether EV charging improves capacity utilization of the system. A higher load factor suggests that the system is being utilized more efficiently. The implications of improving capacity utilization of a power system is a lower levelized cost of electricity (LCOE), which could enable utilities to lower the cost of electricity for the consumer.

\[ f_{\text{load}} = \frac{\sum_{t=1}^{24} P_t}{P_{\text{peak}} \cdot \sum_{t=1}^{24} t} \]  \hspace{1cm} (7.11)

The results of the changes in load factor for the different scenarios are summarized in Table 7.3. We find that in all scenarios there are no negative impacts on the capacity utilization of the system from added EV charging load. However, in some scenarios, specifically the low range anxiety scenarios, there is little to no improvement in the load factor. On the other hand, the high range anxiety scenarios, particularly in the sustainable development scenario and the EV initiative scenario exhibit about a 10% increase in the load factor. This is as a result of significant night-time charging load as shown in Figure 7.14. Further increases in the capacity utilization of the system could be achieved by employing smart charging strategies, as well as other strategies such as offering pricing incentives to EV owners to charge during favorable hours of the day according the system operator.
Table 7.3: Change in Nairobi’s load factor under different scenarios.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Baseline load factor</th>
<th>New load factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stated policy scenario</td>
<td>low range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td></td>
<td>high range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td>National strategic plan scenario</td>
<td>low range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td></td>
<td>high range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td>Sustainable development scenario</td>
<td>low range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td></td>
<td>high range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td>IEA EV initiative scenario</td>
<td>low range anxiety</td>
<td>77%</td>
</tr>
<tr>
<td></td>
<td>high range anxiety</td>
<td>77%</td>
</tr>
</tbody>
</table>

Figure 7.14: Changes to Nairobi demand profile with added EV charging load for different scenarios

7.6 Discussion

This study suggests that electrification of transportation in Nairobi could be beneficial for the technical performance of the system, as evidenced by improvement in the capacity utilization of the system under certain scenarios on adding load from EV charging. However, our results support the argument that the distribution network will be a major bottleneck to the broad diffusion of EVs in Nairobi – we find that under all scenarios at
least 1% and up to 8% of the zones will experience transformer overloaded with added EV charging load. Further our results provide insight into which parts of city would likely need to be prioritized for grid infrastructure upgrades to support early adoption of EVs. It is worthwhile pointing out that these regions are the more affluent parts of the city, which raises some concern about the equity implications of encouraging early adoption of EVs among private and commercial vehicle owners. Considering early investments in the electrification of the paratransit network could offer a potentially more equitable means of transitioning the transportation system while still improving the technical performance of the system.

7.6.1 Caveats of the model and limitations of our work

We acknowledge that this work has some limitations, which we summarize below:

• We estimate EV driving patterns stochastically based on self reported survey data, which is prone to error. Replicating our study with GPS-tracked driving data would allow for a more accurate representation of individual mobility needs.

• Another caveat of our model is the inability to spatially dis-aggregate the Uber movement data by different transportation modes. Augmenting our study with real-word CCTV footage of traffic in different parts of the city would allow for a more precise estimation the speeds and travel times of different transportation modes.

• We aggregate transformer loading a zone level due to lack of granular data on potential EV driving patterns. This technique allowed us to only make broad conclusions about the impact of EV load on the distribution network. Replicating our study with granular and spatially diverse mobility data, speeds data and
distribution network data would enable much more precise estimation of the potential bottlenecks of the distribution network in supporting EV adoption.

7.6.2 Future work

Further work needs to be done to obtain real-world data to validate our models. However, we note that collecting data can be quite expensive. Partnerships with stakeholders entering the EV space to fund and run pilot studies of different electric fleets could be a game-changer in providing much needed real-world data to develop much more accurate data driven insights in transport electrification pathways in the country.

Second, there is need for further work into how the potential electrification of transportation in the city could impact people’s mobility choices as well as the cost of transportation. It is imperative that strategies supporting electric vehicle adoption are grounded in the development priorities of the country.

Third, further research into the potential social, economic and environmental implications of electrifying transportation in the city, as well as policies and strategies that could encourage better techno-economic and social outcomes from electric vehicles needs to be done.

Lastly, further work needs to be done to ensure that the models and techniques used in this study are scalable and transferable to other cities, and countries in sub-Saharan Africa.
8.1 Background and Motivation

Over the past decade, the rate of electrification in sub-Saharan Africa has been steadily increasing, with the proportion of people with access to electricity growing from 33% in 2010 to about 47% in 2019, outpacing population growth. Despite this marked improvement, sub-Saharan Africa still remains the region with the largest electrification deficit. Ethiopia is one of the three largest deficit countries in the world, accounting for 58 million people without access to electricity [124]. In response, the Government of Ethiopia, with support from the World Bank, launched the National Electrification Plan (NEP) in 2017, which includes a comprehensive plan to reach 100% electricity access by 2025, 65% of the population through the grid and the rest through off-grid solutions [197].

While increasing access to an electricity connection is crucial, there is growing consensus that electricity connections need to be accompanied by affordable, reliable, and sustainable electricity consumption, as well as financially sustainable power systems,
to realize the full impacts of electrification [247, 82]. The unfortunate reality is that in most countries in sub-Saharan Africa, increasing electricity access rates has not been met by the same level of electricity consumption growth [89]. As a consequence, people often still rely on traditional fuels, especially for cooking, and utilities and mini grid companies are struggling for financially viability. In fact, a 2016 study of utilities in the region found that only two sub-Saharan African countries have financially viable electricity utilities [143].

To ensure that electrification stimulates economic growth, it is crucial that electricity demand constraints are addressed at every stage of electrification, not least the planning stage. This concept of demand stimulation during electrification planning has historical precedence: Vietnam, which is an electrification success story, prioritized the electrification of areas with high potential for growth in productive uses of electricity, especially irrigation of agricultural areas, which in turn increased government revenue and household incomes, leading to higher consumption and thus promoting the overall financial viability of rural electrification [97].

Demand stimulation has increasingly become an integral part of electrification planning in sub-Saharan African countries as well, including Ethiopia [197]. Ethiopia’s NEP specifically prioritizes grid access to areas with the highest potential for irrigation and agricultural processing, considering the particular importance of agriculture for rural livelihoods [197]. From a utility’s standpoint, the financial viability of rural electrification rests on their ability to generate sufficient revenue from the sale of electricity that outweighs the cost incurred by grid extension. Extending the grid to places where there is sufficient demand from anchor customers, who offer utilities a consistent and substantial source of revenue to supplement the low demand from rural customers is therefore crucial. In turn, cost-recovery for the utility means that cost-reflective tariffs for rural electricity consumers can be set at an affordable rate and that existing capital
banners can stretch to provide access to more communities. Instead, if capital is spent extending the grid to communities where consumption remains low, then utility finances suffer and as-yet-unconnected communities wait ever longer for electricity connections.

How can communities with significant potential electricity consumption be identified? Currently, about 2% of agricultural land in Ethiopia are irrigated, with smallholder farmers largely relying on diesel-powered motorized pumps and manually operated pumps for irrigation. An estimated 200,000 diesel-powered motorized irrigation pumps were in use nationally in 2019 [102]. Reliance on these fuels is unsustainable in that it is expensive, as well as detrimental to human health and the environment. Their combustion releases pollutants into the atmosphere, mainly nitrogen oxides ($NO_x$), carbon monoxide (CO) and particulate matter (PM) [86]. Moving from diesel-powered to electric pumps therefore has the potential to increase the development impact of electrification.

In this chapter, we propose a novel approach to identify areas with existing diesel-powered irrigation in Ethiopia by combining ground data from an agricultural survey with satellite-measured pollution data and crop cover, elevation, and surface water data. We apply supervised machine learning classification techniques that leverage the coincidence between irrigation seasons and seasonal variability of pollution and crop cover datasets. The goal of our study is thus to develop a model that can later detect fossil fuel-powered irrigation activity in Ethiopia based on publicly-available data alone. These areas can then be targeted for electrification and serve as a first productive use load for the grid, improving environmental and financial sustainability for both farmers and the utility, enabling acceleration of rural electrification.

A notable previous effort to develop scalable tools for mapping areas with high potential for electric irrigation pumps in sub-Saharan Africa largely focused on mapping potential for off-grid solar irrigation pumps through a multi-decision criteria model.
that considers factors such as the groundwater levels, aquifer productivity, crop and land suitability, and population [235]. At the same time, publicly-available satellite-measured pollution data is becoming a reliable measure of surface level pollution where real-time air pollution monitoring on the ground is scarce [137, 118, 289]. This enables the development of transferable and scalable models as ours to estimate ground-level pollution.

Numerous studies have used satellite pollution measurements to identify spatial and temporal changes and patterns in surface emissions [169, 101, 159]. However, most of these studies have identified emissions patterns at a country-level scale or over large areas and have not distinguished individual or highly localized sources of emissions. Studies that have used satellite-measured pollution data to detect individual sources of surface emissions have largely focused on fossil fuel power plants, oil tanks, and ships. They specifically apply Gaussian plume air pollution models to $\text{CO}_2$ or $\text{NO}_2$ data to detect and quantify emissions from individual fossil fuel plants [188, 36, 234] and ships [99]. These models have only been applied to sources with visible smoke plumes and, to our knowledge, have not been applied to smaller sources of emissions, like diesel-powered pumps. In addition, this technique faces the challenge of distinguishing between multiple sources of emissions that are in close proximity due to the effect of winds. To overcome this challenge, some studies have applied deep learning techniques to satellite imagery to detect fossil fuel power plants [286, 284], oil tanks [287, 271], and ships [285]. These techniques however, can only be applied to emissions sources that are distinguishable from their surrounding environment in satellite imagery, e.g. through the cooling towers of power plants. Some studies have leveraged the seasonal variability in time series tropospheric pollution measurements to identify the individual source of emissions. These studies, for example, match seasonal $\text{NO}_2$ patterns to harvesting seasons to identify biomass burning especially in Africa and South-east Asia [1, 129, 274]. However,
no studies to our knowledge have explored matching seasonal pollution patterns to irrigation seasons to detect fossil fuel powered irrigation pumps. A recent study applied a supervised machine learning classification technique to remote sensed rainfall and surface water data to identify functioning and non-functioning groundwater pumps in Kenya and Ethiopia, leveraging the relationship between surface water availability and groundwater pump use [258]. However, this study is limited to identifying ground water pumps that are in use but does not distinguish the type of pumps or the activity for which the pump is used, so that pumps commonly used for drinking water supply cannot be considered separately, for example.

Our specific research approach involves three steps: (1) We conduct on-the-ground surveys to collect a first-of-its kind comprehensive dataset on the locations and measurements of cultivated plots in two regions of Ethiopia with high levels of irrigation, as well as their crop cultivation and irrigation practices, including the source and method of obtaining water for irrigated plots. (2) We assess two approaches that both classify cultivated areas into three classes: not irrigated, irrigated using diesel pumps, and irrigated with other non-diesel-based methods. The first approach directly classifies all observations into the three classes. The second approach is a two-step binary classification, first detecting irrigation activity, and then classifying irrigated areas into two classes: irrigated using diesel pumps and irrigated with other methods. (3) We evaluate the performance of four supervised classification algorithms and compare the efficacy of our classification approaches. In doing so, we discuss the limitations of this analysis.

### 8.2 Methodology

In this section, we introduce our approach for detecting areas irrigated with diesel-powered water pumps summarized in Figure 8.1. We begin by describing the study area,
followed by a description of the ground reference primary data and a set of features selected according to insights into characteristics of diesel-powered pumps and irrigated land. Next, we select supervised binary classification models based on the characteristics of our datasets and, finally, describe our classification approach.

8.2.1 Ground Truth Data

Our study area consists of two regions in Ethiopia, the area East of Lake Tana in Amhara region and central parts of Oromia region, each covering around 3,000 km\(^2\) (Figure 8.2). Ethiopia is located in the North-Eastern part of the African continent between 3°N and 15°N latitude and 33°E and 48°E longitude. Amhara and Oromia regions have the largest area of irrigated agriculture in Ethiopia, comprising over 70% of the existing irrigation schemes in the country [74]. In consultation with the Ministry of Agriculture, six districts in Amhara and six districts in Oromia were selected for data collection. Satellite imagery over the study districts was then divided into square pixels of 5 km resolution, of which 36 square pixels were selected as study sites, 18 in each region. In order for the sample to capture the local variety in irrigation intensity, the study site selection was based on a stratified random sampling approach, where the five
strata referred to different levels of signal strength determined through an analysis of satellite imagery.

Figure 8.2: Location of cultivated plots in (a) East of Lake Tana in Amhara region and (b) East of Lake Ziway in central Oromia region, Ethiopia.

Our analysis is based on irrigation data collected in the study area during the months of April and May 2021 using two methods of primary data collection. First, at each of the 36 sites we conducted an extensive farmer survey in the community lying closest to the center of the square pixel selected as the study site. We thereby interviewed a random sample of about 1,000 farmers and elicited information on their socio-economic and demographic status, crop cultivation, and irrigation practices (household survey component). If accessible, the farmers’ locations and measurements of cultivated plots (four GPS coordinates on the boundary and one at the center of the plot) were also captured together with the status of irrigation and crop cultivation on the plots at the
Table 8.1: Number and total area of irrigated plots by method of irrigation

<table>
<thead>
<tr>
<th>Irrigation status/method</th>
<th>Number</th>
<th>Area (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrigated, diesel water pump</td>
<td>561</td>
<td>2.69</td>
</tr>
<tr>
<td>Irrigated, mechanical water pump</td>
<td>31</td>
<td>0.09</td>
</tr>
<tr>
<td>Irrigated, gravity</td>
<td>326</td>
<td>2.07</td>
</tr>
<tr>
<td>Irrigated, hand held method</td>
<td>28</td>
<td>0.07</td>
</tr>
<tr>
<td>Not irrigated</td>
<td>4162</td>
<td>20.63</td>
</tr>
</tbody>
</table>

time of data collection (plot measurement component). A second team of enumerators visited a total of 140 randomly-sampled pixels of 250m resolution at each study site. Here, the same plot measurement was conducted, with the difference that features had to be independently observed. Enumerators were directed to the plot at the center of each pixel, and to up to two additional plots in the case of multiple plots within a pixel. We note that these ground truth pixels were over-sampled in irrigated areas to provide more positive observations for algorithm training. The distribution of our ground truth dataset therefore, does not mirror the underlying distribution of irrigation in the regions of study. A total of about 6,100 plots were measured, among which 570 plots were measured as part of the farmer survey exercise. Information on the method used to obtain water for irrigation was collected for about 1,100 plots which were being irrigated at the time of data collection – data for about 450 of these plots were collected as part of the farmer survey exercise and the remaining plots as part of the independent observation by enumerators. The remaining 5,000 plots were labeled as non-irrigated. We cleaned the dataset by removing the plot instances with invalid geometries or that intersect with other plot polygons. The number and area coverage of resulting plots by the status and method of irrigation are shown in Table 8.1.

The coordinates of about 260 irrigation wells was also collected together with the respective plots to which they supply water. Irrigation pumps are placed near the irrigation wells, which were dug to access water. Based on these data, we determined
that the average distance of the irrigation wells from the plots is about 35 m. To account for placement of the motorized pumps with respect to the plots, we include a 35 m buffer around the plots. The spatial distribution of the plots by label in the study regions is shown in Figure 8.2.

We define a 250 m resolution grid cell covering both of the study areas, selected to align with the sampling pixel resolution. Then, we label and select pixels such that the pixels with a specific class label only contain plots with the same class label. We specify three classes: diesel irrigated, non-diesel irrigated, and non-irrigated (Table 8.2). The resolution of the pixels is selected such that the spatial resolution of the satellite-measured pollution data is improved without compromising the consistency of the satellite measurements with \textit{in situ} surface measurements, as well as our ability to detect spatial heterogeneities in the pollution measurements [45, 108].

\subsection*{8.2.2 Defining Features}

We identify and select features of the pixels based on knowledge of the characteristics of cultivated land, factors that would influence the choice of irrigation method, and the characteristics of diesel-powered water pumps.

\textit{Time series pollution data}

We use time series satellite-measured pollution data as one of the main input features to our model. Our hypothesis is that for primarily agricultural areas, there is statistically

\begin{table}[h]
\centering
\caption{Number of $250 \times 250 \, m^2$ pixels per class}
\begin{tabular}{ll}
\hline
Class & Number of pixels \\
\hline
Diesel irrigated & 558 \\
Non-diesel irrigated & 358 \\
Non-irrigated & 4763 \\
\hline
\end{tabular}
\end{table}
significantly higher pollution in areas with plots irrigated via diesel-powered pumps than in areas with plots irrigated via other methods during the irrigation seasons. In Ethiopia, these seasons are the so-called Belg and Meher season, which we determined to fall between September to March based on our survey data (Figure 8.3).

The Tropospheric Monitoring Instrument (TROPOMI), which was launched on 13 October 2017 aboard the European Space Agency’s Sentinel-5 Precursor, measures ultraviolet (UV), ultraviolet–visible (UV–VIS), ultraviolet near-infrared (NIR), and shortwave infrared (SWIR) spectral bands. From these bands, a wide range of pollutant gases including nitrogen dioxide ($\text{NO}_2$), carbon monoxide ($\text{CO}$), sulfur dioxide ($\text{SO}_2$), ozone, and formaldehyde, can be retrieved at a spatial resolution of $7 \times 3.5 \text{ km}^2$ (reduced to $3.5 \times 5.6 \text{ km}^2$ on 6 August 2019) at nadir at a swath width of 2600 km, and daily overpasses at approximately 13:30 local solar time at the equator. We downloaded TROPOMI level-2 $\text{NO}_2$ and $\text{CO}$ tropospheric vertical column density measurements from the ESA Copernicus Open Access Hub over the Amhara and Oromia regions of Ethiopia for the period of January 2018 to July 2021. A tropospheric vertical column density is the vertically integrated number of $\text{NO}_2$ (or $\text{CO}$) molecules per unit area between the surface and the tropopause in units of $\text{molec./cm}^2$. We filtered the data to remove cloud-covered scenes, errors, and problematic retrievals (i.e., measurements with a quality assurance value of less than 0.75). Since the angle of polar orbiting satellites on a given area is slightly different for each overpass, we first accumulate daily observations from consecutive overpasses over each month in each study region to improve spatial sampling. We then arrive at average monthly $\text{NO}_2$ and $\text{CO}$ values in each pixel using inverse distance weighting.

Previous studies [268, 289, 35, 119] have shown that TROPOMI satellite measurements are a good estimate of $\textit{in situ}$ surface concentrations of both $\text{NO}_2$ and $\text{CO}$ in most parts of the world. However, none of these have performed ground-based validation.
in sub-Saharan Africa likely due to the paucity of ground-deployed air quality sensors. We obtained data for August 2020 from the $NO_x$ sensor on a Real-time Affordable Multi-Pollutant (RAMP) monitor recently deployed in Accra, Ghana, [5] to validate the correlation between surface $NO_2$ concentrations and tropospheric $NO_2$ measurements. The RAMP was developed in a collaboration between Carnegie Mellon University and SenSevere. It incorporates Alphasense electrochemical sensors to measure $CO$, $NO_2$, $SO_2$ and $O_3$, and a non-dispersive infrared (NDIR) sensor to measure $CO_2$ [291]. We find that the TROPOMI $NO_2$ measurements of the $7 \times 3.5 \ km^2$ pixel that encompasses the $NO_x$ sensor in Accra has both temporal agreement as well as a strong correlation (0.71) to measurements taken by the $NO_x$ sensor (Figure 8.4), providing some confidence in the ability of this satellite to reflect air quality dynamics at ground level.

**Time series vegetation indices**

To capture crop cover dynamics over time over these cultivated pixels, we obtained remotely sensed vegetation data, that is the Normalized Difference Vegetation Index
Figure 8.4: Comparing tropospheric NO₂ column measurements and ground NO₂ concentrations in Accra, Ghana

(NDVI) and Enhanced Vegetation Index (EVI) measurements, which are used to quantify vegetation greenness from measurements of light intensity coming off the Earth’s surface in visible and near-infrared wavelengths. EVI improves upon the quality of the NDVI by correcting for some atmospheric conditions and canopy background noise and is more sensitive in areas with dense vegetation. These data products are collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard NASA’s Terra spacecraft at a 16-day temporal resolution and a 250-meter spatial resolution. We extracted monthly averaged data from Google Earth Engine over the study regions for the same period as the pollution data (January 2018 to July 2021) and re-sampled the data to the 1 km² grid cells.

Time constant features

Elevation: The choice of which method of irrigation to use was found to be heavily influenced by topography. Our hypothesis is that areas irrigated with methods of irrigation that rely on gravity and slope will be at higher elevations than
those irrigated with diesel-powered pumps (Figure 8.5a). We therefore use digital elevation data for our analysis, obtained from the NASA Shuttle Radar Topography Mission (SRTM) dataset at a 30 m spatial resolution. The SRTM was a primary component of the payload on the Space Shuttle Endeavour, which launched on February 11, 2000, and flew for 11 days [267]. We averaged the elevation in meters for each 1 km² pixel.

**Water availability:** Availability of water sources for irrigation impacts the type of irrigation method used by farmers. We therefore hypothesize that areas irrigated with diesel pumps will likely be closer to large surface water bodies such as lakes and rivers, than those irrigated with other methods, especially non-mechanized methods (Figure 8.5b). Therefore using geospatial data on water bodies in Ethiopia from the RCMRD GeoPortal [98], we define the distance of each pixel centroid to the nearest major water source, that is a lake or main river, as a feature to capture water availability.

**Settlement patterns:** Lastly, we consider settlement patterns as a possible feature of the method of irrigation. We use data on population densities and proximity to road infrastructure to capture the variation in settlement patterns. Our hypothesis is that cultivated plots with non-motorized and even more non-mechanized irrigation methods which rely on human operation will likely have higher population and will be closer to roads than plots irrigated with diesel-powered pumps (Figure 8.5c and d). We use 1 arc-second (approximately 30m) resolution population estimates of Ethiopia obtained from satellite imagery as part of the High Resolution Settlement Layer (HRSL) population density dataset [79]. We sum the population values for each pixel, as well as define the distance of each pixel centroid to the nearest major road.
Figure 8.5: Empirical cumulative distribution functions of time constant features for pixels irrigated with and without diesel.

8.2.3 Time series feature extraction and selection

A common approach to time series classification is to treat each time point as a separate feature and directly apply a standard classifier. However, this approach is problematic because the classifier ignores information contained in the time order of the data. In our case, the classifier will ignore the seasonality of irrigation cycles in the pollution and crop cover features. To ensure that all the information contained in the time series is captured, we extract features that capture the overall properties of the time series and the correlation between the different measurements in the time series [53]. The algorithm calculates over 200 features for each time series including the minimum,
maximum, mean, median, 25th percentile, 75th percentile, standard deviation (stdev),
the linear regression $y = ax + b$ coefficients $a$ and $b$, and the area under the curve (AUC).
A comprehensive list of the features can be found in [52]. To avoid increased model
and computation complexity, and poor model accuracy from the inclusion of irrelevant
inputs, the algorithm implements an additional step of feature significance testing
and feature selection. After feature extraction, each feature vector is individually and
independently evaluated with respect to its significance for the classification problem,
and its significance quantified as a $p$-value. Finally, a vector of all $p$-values is evaluated
on the basis of the Benjamini-Yekutieli procedure [21] in order to decide which features
to keep.

8.2.4 Data pre-processing

The features in our dataset vary in scale, range, and units. Therefore, to ensure that
our models do not make assumptions about the distribution of our data, prior to training,
we first employ a standardization technique to re-scale our features to have a standard
normal distribution (mean of 0 and standard deviation of 1). Then we employ a min-max
normalization technique to re-scale the features into a range of $[0,1]$.

8.2.5 Classifiers

The choice of classifiers for our analysis was mainly based on the nature of our dataset,
that is the dataset size, as well as nature and distribution of features. We consider four
classifiers that work well with small, complex datasets and support nonlinear distribution
of features: Random Forest, Support Vector Machine (SVM), $k$-Nearest Neighbor ($k$-NN),
and Logistic Regression.

**Random forest** is an ensemble of tree-structured classifiers such that each tree is
trained on the values of a random vector sampled independently and with the same
distribution for all the trees in the forest. Using a random selection of features to split each node in a tree decreases the correlation between decision trees in the forest, and thus decreases the possibility of overfitting [38]. It uses a bootstrapping technique which enables it to work well on relatively small datasets. In addition, it is simple to implement, and robust to outliers. After training, the classification result is determined by averaging the most frequent prediction.

**Support vector machine (SVM)** is a widely used classifier due to its flexibility and robustness. It is based on maximizing the gaps between two classes by defining a hyperplane that splits the two classes [61]. In the case of multi-class classification SVM uses a ”one-versus-one” classification approach, whereby \( n_{\text{classes}} \times ((n_{\text{classes}} - 1)/2) \) classifiers are constructed and each one trains data from two classes [236]. It is effective in high-dimensional spaces and is able to compute decision boundaries without assuming specific distributions of the input data. It also performs very well with limited training data. To adapt SVM for nonlinear classification to avoid overfitting, we use a Radial Basis Function (RBF) kernel to map the data into a higher dimensional space.

**k-Nearest Neighbor (k-NN)** is one of the simplest classifiers, often used for its simplicity of interpretation and low computation time. It hinges on the assumption that similar observations exist in close proximity in a multidimensional space. Therefore it works by calculating the distance between observations and assigning each observation to the class most common among its \( k \) nearest neighbors [114]. Besides its simplicity, another advantage is that is is agnostic to the distribution of the data. However, its drawbacks are that it has low efficiency and its performance highly depends on the value for \( k \).

**Logistic regression** is a widely used classification technique, which uses a logistic function to model a binomial target variable. However, it can also be extended to model a multinomial target variable [166]. It is easy to implement, efficient, and is a high-bias
model, which means it works well for small datasets. However, its major limitation is the assumption of linearity between the dependent variable and the independent variables.

8.2.6 Diesel-powered irrigation detection

In this section we present our approach to developing and training a model which detects areas with diesel-powered irrigation activity using the above list of input features.

Class balancing

Given the imbalanced nature of our dataset that reflects the ground reality of irrigation methods used in Ethiopia, we use a class balancing technique during model training. Datasets with imbalanced classes cause poor performance with traditional machine learning models and evaluation metrics that assume a balanced class distribution. Previous approaches such as random under sampling [270] have been used to address class imbalance by randomly removing some observations of the majority classes. However, this technique is only a good choice when working with a very huge dataset, which is not the case for our study where each ground datapoint is both expensive and onerous to collect. Therefore to address class imbalance in our dataset, we use the Synthetic Minority Oversampling Technique (SMOTE) [47]. This approach over-samples the minority classes by creating synthetic instances of the class. This is achieved by selecting a minority class at random, finding its five nearest minority class neighbors, randomly selecting $k$ of the neighbors based on the amount of oversampling required and generating a synthetic example a randomly selected point between the two examples in feature space. This results in a balanced dataset that aims to reflect the characteristics of the underlying unbalanced classes.
Cross validation

Since we have a limited dataset, removing part of our dataset for validation risks introducing a problem of under-fitting and bias. We therefore use a tenfold cross-validation approach, which works by dividing the data into ten subsets and during each iteration, one of the subsets is used as the validation set and the other subsets are combined into a training set ensuring that each observation is a validation set once and a training set nine times [223]

Hyperparameter tuning

We use a grid search strategy for different model hyperparameter values during training. We tune the kernels and C hyperparameters for the SVM classifier, solver and C hyperparameters for the logistic regression classifier, the nearest neighbors, metric and weights hyperparameters in the k-NN classifier, and the number of estimators and maximum features for the random forest classifier.

Classification approach

We propose to compare three classification approaches, a single-stage ternary classification approach, a single-stage binary classification approach, and a two-stage binary classification approach, to detect diesel-powered irrigation. In both cases, we train and cross-validate the classifiers on 70% of the data and reserve the remaining 30% as a test dataset.

Single-stage ternary classification approach: In this approach (Figure 8.6), we use all the labeled observations to train multi-class classification models to directly categorize the feature vectors into three classes: diesel irrigated, non-diesel irrigated, and non-irrigated.
Two-stage binary classification approach: This approach breaks the classification technique into two stages: first, a binary classification of irrigated and non irrigated pixels. In an ideal case, we would then use the resulting true positive labels (correctly classified irrigated pixels) from the first step to train a binary classification model to categorize the feature vectors into the two classes: diesel irrigated and non-diesel irrigated (Figure 8.7). However, given the relatively small sample size in our study, for the first step of our approach, we validate the irrigation status of the pixels labeled as irrigated in our sample with an irrigation detection model [56] which pairs Sentinel-2 imagery with labeled plot polygons to train an irrigation detector that achieves 95% accuracy in our study area. We then train a binary classification model to categorize the feature vectors of the irrigated pixels into the two classes: diesel irrigated and non-diesel irrigated.

Figure 8.6: Single-stage ternary classification approach

Figure 8.7: Two-stage binary classification approach
8.2.7 Performance Metrics

We evaluate the performance of the classifications based on the widely-used metrics for classification tasks: accuracy, precision, recall, F1 score, the Matthews correlation coefficient (MCC), and the area under the Receiver Operating Characteristic (ROC) curve (AUC) [244]. Accuracy is the proportion of correctly classified results among the total number of observations and is denoted by:

\[
\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{8.1}
\]

where in the case of a binary classification, true positives (TP) would denote the number of correctly classified observations in Class A, true negatives (TN) would denote the number of correctly classified observations in Class B, false positives (FP) would denote the number of observations in Class B incorrectly classified as Class A and false negatives (FN) would denote the number of observations in Class A incorrectly classified as Class B.

Precision is the proportion of observations classified as Class A actually belong to Class A and is denoted by:

\[
\text{Precision} = \frac{(TP)}{(TP + FP)} \tag{8.2}
\]

Recall is the proportion of Class A observations that are classified correctly and is denoted by:

\[
\text{Recall} = \frac{(TP)}{(TP + FN)} \tag{8.3}
\]

These performance metrics are summarized in a confusion matrix. For instance, in the case of a binary classification, it has four quadrants: true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN).

Given that our objective is to detect diesel irrigation activity, our study optimizes for precision and recall over accuracy. Therefore, we also consider F1 score, which is
the harmonic mean of precision and recall. It is a number in the [0,1] range and it is denoted by:

\[ F_1 = 2 \cdot \frac{(\text{precision} \cdot \text{recall})}{(\text{precision} + \text{recall})} \] (8.4)

We also consider the area under the ROC curve, which tells us how well observations from the positive class are distinguished from the negative class. The ROC is a probability curve, plotted as true positive rate against false positive rate. It is denoted by:

\[ AUC = \frac{1}{2} \left( \frac{TP}{(TP + FN)} + \frac{TN}{(TN + FP)} \right) \] (8.5)

Lastly, we compute the Matthews correlation coefficient (MCC), which is a balanced measure of the quality of binary and multi-class classifications, taking into account all of the four confusion matrix categories (true positives, false negatives, true negatives, and false positives) proportionally. It ranges from minus 1 to 1, with minus 1 meaning perfect misclassification and 1 a perfect classification. It is very useful in that it is a good measure of predictions even with an unbalanced dataset [50]. It is demoted by:

\[ MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \] (8.6)

8.3 Analysis

We evaluate the performance achieved by the models in classifying pixels with diesel irrigation activity using a 10-fold cross validation approach. We present results in terms of precision, recall, $F_1$ score, and the Matthews correlation coefficient (MCC). We also evaluate sensitivity of the classification using area under the ROC curve.

8.3.1 Single-stage ternary classification

In this classification approach, we consider both the pollution and vegetation index time-series. We extracted 3,156 features from the $NO_2$, $CO$, NDVI, and EVI time series
Table 8.3: Comparison of model performance and classification approaches

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Single-stage ternary</th>
<th>Two-stage binary</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>$F_1$</td>
<td>MCC</td>
<td>Precision</td>
<td>Recall</td>
<td>$F_1$</td>
<td>MCC</td>
</tr>
<tr>
<td>Random Forest</td>
<td>77%</td>
<td>80%</td>
<td>0.77</td>
<td>0.75</td>
<td>86%</td>
<td>92%</td>
<td>0.87</td>
<td>0.70</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>64%</td>
<td>76%</td>
<td>0.66</td>
<td>0.67</td>
<td>85%</td>
<td>80%</td>
<td>0.84</td>
<td>0.57</td>
</tr>
<tr>
<td>SVM</td>
<td>75%</td>
<td>80%</td>
<td>0.76</td>
<td>0.76</td>
<td>81%</td>
<td>96%</td>
<td>0.84</td>
<td>0.67</td>
</tr>
<tr>
<td>$k$-NN</td>
<td>67%</td>
<td>81%</td>
<td>0.70</td>
<td>0.71</td>
<td>91%</td>
<td>86%</td>
<td>0.90</td>
<td>0.72</td>
</tr>
</tbody>
</table>

data. Feature significance testing and selection based on the Benjamini-Yekutieli procedure results in 1291 time series features, 374 selected from EVI, 359 from NDVI, 307 from $CO$, and 251 from $NO_2$. This results in a total 1295 features, including elevation, population, distance to a major water source, and distance to a major road.

We find that Random Forest achieves the highest $F_1$ score of 0.77 for the diesel irrigated class, outperforming the other three classifiers in identifying diesel irrigated pixels, as shown in Table 8.3. The least performance is from the logistic regression model. We note however, that the SVM slightly outperforms the Random Forest model when we consider the MCC, with a value of 0.76 compared to Random Forests’s MCC of 0.75. It is noteworthy that the $F_1$ scores do not vary significantly from the MCCs, suggesting that the models are not incorrectly favoring one class over the others.

The confusion matrices, presented in Figure 8.8, show where incorrect classification is occurring. We observe that for all the classifiers, a large proportion of incorrectly classified diesel irrigated observations are being categorized as non-irrigated. This may suggest that the two-stage binary classification approach is likely to perform better as it separates the classification of irrigation activity and the classification of the type of irrigation method. It is also noteworthy that the models in this approach are better at classifying the non-irrigated observations than the other two classes. This finding is supported by the results of the AUC analysis.

The ROC curve of the Random Forest model is shown in Figure 8.9. It is plotted using the One vs ALL methodology, which means that the ROC for each class is classified
Figure 8.8: Confusion matrices of models in the single-stage ternary classification approach against the other two. We find that the model has the highest ability to separate the non-irrigated class, with an AUC value of 0.72. The diesel irrigated and non-irrigated classes both have AUC values under 0.7, suggesting that the model is struggling to distinguish these two classes.

8.3.2 Two-stage binary classification

We evaluate whether separating the classification task into two separate binary classification tasks outperforms the direct ternary classification approach presented in the previous subsection in identifying diesel irrigated activity. In the irrigation detection
model that we employ in the first step of our classification task, non-irrigated and irrigated labels are predicted with 98.3% and 95.5% accuracy [56]. Subsequently, we use the irrigated pixels in our sample that align with the predictions from the irrigation detection model as inputs to the second step of the classification task.

In this step, of 1,578 time series features extracted from the $NO_2$ and $CO$ time-series, 81 from $NO_2$, and 105 from $CO$ were selected as being significant for the classification problem. Therefore, together with the time constant features, resulting in a total 190 features. Note that we do not use the vegetation indices time-series in this step, as it is the main feature used in the irrigation detection model in the first step of this approach.

We find that overall, every model in the binary classification approach outperforms the models in the ternary approach. The $k$-NN model outperforms the other three models considering both its $F_1$ score of 0.9, which captures how well the diesel-irrigated class is predicted and its MCC of 0.72, which captures how well both classes are predicted. We can see the balanced performance of the $k$-NN in its confusion matrices presented in Figure 8.10, showing correct predictions about 86% of the observations of both classes.

We find that the measure of separability of the models significantly improves with
the two-stage binary classification approach. The $k$-NN model achieves a mean AUC of 0.93 (Figure 8.11) with 10-fold cross validation, suggesting that there is a 93% chance that the model will correctly distinguish a diesel irrigated observation from a non-diesel irrigated observation.

8.3.3 Feature Contribution

We evaluate which features played a key role in enabling our features to distinguish between diesel-irrigated and non-diesel irrigated pixels by comparing two techniques which estimate feature importance: Random Forest and Gradient Boosting. We find that
both techniques commonly identify a number of time series features extracted from CO, elevation and proximity to major water source as the key features that contribute to the success of our model as shown in Table 8.4. It is noteworthy that features from CO pollution time-series contribute more to the performance of our model that features from NO₂ pollution time-series.

### 8.3.4 Spatial Analysis

We also consider the areas where our model was unsuccessful in classification. We find that the largest proportion of incorrectly classified pixels are in areas where the
pixels of different classes are in close proximity, mainly to the East of Lake Tana in the Amhara region, as shown in Figure 8.12 (compared with Figure 8.2). One possible reason is that the pollution measurements, which have been re-gridded at a 250 m resolution, are hard to distinguish among pixels that are close together.

To that end we explore the impact of the spatial placement of pixels of different classes on the performance of our $k$-NN model. We consider (1) only on pixels of the two classes that are in close proximity to each other and (2) only on pixels of the two classes that are far from each other. We determine this proximity threshold by considering the distribution of distances between diesel-irrigated pixels and non-diesel irrigated pixels and vice versa and categorizing those below the median distance value as being in close proximity. The median distance of diesel-irrigated pixels from non-diesel irrigated pixels is 714 m and that of non-diesel irrigated pixels from diesel-irrigated pixels is 357 m. As shown in Table 8.5, we find that that our model performs much
Table 8.5: Performance of k-NN model considering the spatial proximity of class pixels

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>$F_1$ score</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full dataset</td>
<td>91%</td>
<td>86%</td>
<td>0.90</td>
<td>0.72</td>
</tr>
<tr>
<td>Class pixels in close proximity</td>
<td>70%</td>
<td>68%</td>
<td>0.70</td>
<td>0.23</td>
</tr>
<tr>
<td>Class pixels far from each other</td>
<td>92%</td>
<td>90%</td>
<td>0.92</td>
<td>0.79</td>
</tr>
</tbody>
</table>

worse when we only consider the pixels of two classes in close proximity to each other, achieving an MCC of only 0.23 compared to 0.72 when trained on the full dataset. On the other hand, we also note that there is better performance when we only consider the pixels of two classes that are further than the median distance from each other, achieving an improved MCC score of 0.79. While our model struggles to separate the diesel and non-diesel irrigated pixels in areas with multiple pixels of different classes as evidenced by the MCC, electrification decisions are usually made atomically over areas that comprise of heterogeneous activities. As such, a good precision and recall score for the diesel-irrigated class in these areas is sufficient.

8.3.5 Validation of Results

In our analysis, we make use of a sample of own-collected primary data given the absence of comprehensive data for Ethiopia that contains information on the location and type of water pumps for irrigation. We acknowledge that the small size of our dataset is challenging for classification, but note the paucity of any similar datasets in the literature. While our dataset is a good representation of the overall classification problem, there is a likelihood that it may not be enough to capture the complexity of the classification problem at large scale. In the following, we therefore present two efforts to give confidence to our results.
Table 8.6: Bootstrap Matthews correlation coefficient estimates and confidence intervals for classifiers in the two-stage classification approach

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Bootstrap estimate</th>
<th>Confidence interval</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>0.75</td>
<td>[0.67,0.83]</td>
<td>0.030</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>0.56</td>
<td>[0.48,0.64]</td>
<td>0.031</td>
</tr>
<tr>
<td>SVM</td>
<td>0.72</td>
<td>[0.64,0.80]</td>
<td>0.031</td>
</tr>
<tr>
<td>$k$-NN</td>
<td>0.74</td>
<td>[0.66,0.82]</td>
<td>0.032</td>
</tr>
</tbody>
</table>

Bootstrap Confidence Intervals

We construct confidence intervals around the Matthews correlation coefficient of our models in the two-stage binary classification approach to evaluate its variability. We use a widely used statistical method, the bootstrap method, which randomly draws samples from the original dataset (with replacement) to obtain estimates of the MCC thus creating a distribution. We generate a 99% confidence interval of coefficient values from 1000 bootstrap samples for the models in the two-stage binary classification approach. We find that there is a 99% chance that the interval [0.66, 0.82] contains the MCC of our best performing model, the $k$-NN, and the MCC of the poorest performing model, the logistic regression is contained in the [0.48,0.64] interval with high certainty.

Permutation testing

Permutation testing helps to investigate whether or not the performance score obtained from the models has been obtained by chance. The algorithm generates a null distribution of the performance score of the classifier on 1000 different permutations of the dataset, where features remain the same but labels undergo different permutations [232]. We carried out permutation testing on the $F_1$ score of the $k$-NN model in the two-stage classification approach. We find that the weighted average $F_1$ score obtained on the original data, 0.86, is statistically significantly higher than the scores obtained
using the permutated data, as shown in Figure 8.13. This indicates that our dataset contains real dependency between features and labels, which our model was able to use to identify the observations of the diesel irrigated class.

![Figure 8.13](image)

**Figure 8.13**: Distribution of average weighted $F_1$ scores from permutation test of $k$-NN model in two-stage classification approach.

When we apply our $k$-NN model to the predicted irrigated areas of the Amhara region [56], it classifies about 20% of the irrigation activity as using diesel-powered pumps. As shown in Figure 8.14, a significant proportion of the diesel-powered irrigation activity is in the western part of the region.

### 8.4 Discussion and Future work

The high precision and recall of our model suggests that our approach could be valuable in identifying areas with potential anchor loads for grid extension from the replacement of existing diesel pumps for irrigation with electric pumps in Ethiopia. While grid extension planning has myriad considerations – among them the locations of populations without electricity, unelectrified productive use loads including mining
and industry, proximity to the existing grid, and as always political considerations – information on where diesel-powered irrigation already is – representing sites more likely to have stable revenue from electricity sales – can be of significant value in settings like Ethiopia where irrigation is central to the electrification strategy and little other economic activity exists in rural areas [197]. While capturing the relative importance of each consideration is difficult enough in planning grids, it becomes even harder in integrated electrification planning scenarios that also consider decentralized electrification via minigrids and solar home systems. Nonetheless, as a simple thought experiment to show the value of accurate predictions of diesel-powered irrigation areas, imagine that two grid planners each have enough budget to extend an existing grid to 100 additional sites, with each gaining an additional site for every two diesel-powered
irrigation sites that is electrified (since these sites return more consistent revenues). Both planners have perfect maps of all irrigation sites for the country, but one has no information about diesel-powered irrigation sites and the other can interrogate any site and receive a correct answer 75% of the time (similar to our algorithm’s performance). Given that only 20% of irrigation sites are diesel-powered, our second planner is nearly four times more likely to correctly select sites with diesel-powered irrigation. This results in the first planner using their capital budget to electrify 112 sites total, while our second planner is able to electrify 250 total sites (2.2x). While this exercise is highly simplified and any assumptions may be discounted, it aims to convey the enormous value to electricity planners in having improved insight into which areas may yield higher stable revenues.

Future directions will consider the equity implications of this work by investigating the prioritization electrification in more affluent areas where consumers can already afford diesel pumps. We are also yet to consider the applicability of this technique in built environments that could have confounding pollution activity.

8.5 Conclusion

While universal electrification is beginning to come into view in Sub-Saharan Africa, the attendant livelihood gains from widespread adoption of electricity for economic benefits remains stubbornly far. Financially sustainable electricity service providers are a key ingredient to an electrified economy and identifying potential sources of sustainable revenue for utilities that also provide economic benefits to customers is a central challenge. In this work, we present a novel technique that leverages a raft of remote sensing datasets – including pollution, vegetation, and population as derived from satellites – to identify whether irrigation sites in Ethiopia are powered by diesel or
non-diesel sources. We evaluate our technique via a unique household survey among farmers in two regions of Ethiopia that we collected. Our results show a more than 3.5x improvement over the random chance threshold for this problem (20% to \( \approx 75\% \)) that can enable enhanced business models for electricity service providers, environmentally-sustainable production for farmers, and accelerated electrification for the people of Ethiopia. The broad application of this technique can substantially aid in the successful expansion of electricity systems throughout agriculture-led developing regions.
This thesis presents the evaluation of strategies that are aimed at facilitating sustainable energy use for human development in sub-Saharan Africa, either through stimulating organic consumption growth or converting existing fossil fuel energy uses into electric, mainly for income generating purposes. First, we investigate the electricity consumption and financial expenditure patterns of decentralized energy system users in East Africa to better understand how these customer bases consume and spend on electricity. Second, we evaluate the efficacy of two programs, an appliance financing program and tariff subsidy program, in stimulating organic consumption growth among mini-grid customers in East Africa. Third, we evaluate the potential for converting diesel-based fishing boats in Lake Victoria to electric motor and battery-based systems that can provide a crucial anchor load for the island mini-grid while improving the socio-economic outcomes of the fishing community. Fourth, we study the potential power system challenges and opportunities of supporting electric vehicle adoption in Nairobi under different scenarios. Finally, we present a novel approach to mapping areas with existing diesel-powered irrigation in Ethiopia by combining ground survey data with satellite measured pollution data and machine learning techniques to inform electrification planning.
SOCIOECONOMIC STATUS COMPARISON OF OFF-GRID AREAS USING NIGHTTIME LIGHTS LUMINOSITY

We evaluate whether we can detect differences in nighttime luminosity between areas with groups of SHS customers and mini-grid sites. We consider changes in the average nightlights radiance within solar home system and mini-grid sites before and after their installation dates using a difference-in-difference approach, where we use a nearest-neighbor ball-tree matching algorithm to identify one control site (without a solar home system or mini-grid) for each solar home system and mini-grid site. We then cluster customers within a 1 km radius of each school, while grouping schools that are within 1 km of each other. The 36 clusters, with the largest number of solar home systems, were classified as the treatment clusters. The potential control clusters were obtained by defining a 1 km radius around the school locations within the same counties as the treatment clusters that did not match any of the schools near solar home systems. Each mini-grid treatment site was defined as a 1 km radius around the centroid of the site. The potential control sites were defined as any 1 km radius site within the same district or country as the treatment sites that did not intersect with any mini-grid sites. We obtain the population density of each treatment and potential control solar home system cluster and mini-grid site, data on the location of substations to obtain...
the distance of each cluster/site to the nearest substation and the wealth indices of each cluster/site from the demographic health surveys. We identify a one-to-one match of each solar home system treatment cluster and mini-grid treatment site to a control cluster/site based on population density, proximity to the national grid and wealth index. We then use a difference-in-difference model to evaluate the average effect of installing decentralized systems on nighttime luminosity in SHS and mini-grid sites, given in Eq.A.1.

\[ Y_{it} = \beta_1 X_{tr} + \beta_2 X_q + \beta_3 X_{tr} X_q + \epsilon_{it} \]  

where \( Y_{it} \) is the average nighttime lights radiance band (\( nW/cm^2/sr \)) in site \( i \) for the month \( t \) in the period of 2014-2019, \( X_{tr} \) is a dummy variable coded 0 and 1 for control and treatment sites respectively, \( X_q \) is the month from the installation date of mini-grid sites (lagged by a month to allow for the effect of system installation to be captured) and in the case of SHS, it is the month from the date when 75% of customers signed contracts with the SHS company. The interaction term captures the possible differences in nighttime luminosity between the treatment and control clusters.

As shown in Figure A.1 we did not detect any significant changes in the average nighttime lights radiance band after installation of both decentralized systems. Furthermore, the values of the average radiance band are quite low, likely to be classified as lower bound noise – a study by [83] validating electricity access using nighttime lights set a lower-bound noise floor of 250 \( nW/cm^2/sr \). This would therefore suggest that nighttime lights by themselves would likely not be an effective method or proxy for measuring economic activity in these regions, as has successfully been done in urban areas or electrified regions with higher population densities. We therefore have to rely on the other ground based data sets to compare the socioeconomic status of the study
Figure A.1: Average effect of installing decentralized systems on average nighttime lights radiance band areas.
## EFFECTS OF APPLIANCE FINANCING PROGRAM

### Table B.1: Month level dynamic treatment effects of appliance financing program: Alternative regression specifications

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: Monthly electricity consumption (kWh)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-12 months prior:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>-0.122</td>
<td>-0.042</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.648)</td>
<td>(0.370)</td>
<td>(0.434)</td>
</tr>
<tr>
<td>-11 months prior:</td>
<td></td>
<td>0.022</td>
<td>0.100</td>
<td>0.228</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.770)</td>
<td>(0.536)</td>
<td>(0.584)</td>
</tr>
<tr>
<td>-10 months prior:</td>
<td></td>
<td>-0.027</td>
<td>0.370</td>
<td>0.476</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.853)</td>
<td>(0.643)</td>
<td>(0.688)</td>
</tr>
<tr>
<td>-9 months prior:</td>
<td></td>
<td>0.381</td>
<td>0.511</td>
<td>0.784</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.133)</td>
<td>(0.778)</td>
<td>(0.931)</td>
</tr>
<tr>
<td>-8 months prior:</td>
<td></td>
<td>0.247</td>
<td>0.234</td>
<td>0.498</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.935)</td>
<td>(0.639)</td>
<td>(0.824)</td>
</tr>
<tr>
<td>-7 months prior:</td>
<td></td>
<td>0.094</td>
<td>0.050</td>
<td>0.241</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.015)</td>
<td>(0.665)</td>
<td>(0.825)</td>
</tr>
<tr>
<td>-6 months prior:</td>
<td></td>
<td>0.391</td>
<td>0.314</td>
<td>0.605</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.941)</td>
<td>(0.674)</td>
<td>(0.783)</td>
</tr>
<tr>
<td>-5 months prior:</td>
<td></td>
<td>1.258</td>
<td>1.197</td>
<td>1.534</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.063)</td>
<td>(0.839)</td>
<td>(0.963)</td>
</tr>
<tr>
<td>-4 months prior:</td>
<td></td>
<td>0.909</td>
<td>0.889</td>
<td>1.295</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.134)</td>
<td>(0.932)</td>
<td>(1.048)</td>
</tr>
<tr>
<td>-3 months prior:</td>
<td></td>
<td>0.694</td>
<td>0.762</td>
<td>1.204</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.948)</td>
<td>(0.739)</td>
<td>(0.910)</td>
</tr>
<tr>
<td>-2 months prior:</td>
<td></td>
<td>0.656</td>
<td>0.752</td>
<td>1.151</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.918)</td>
<td>(0.710)</td>
<td>(0.894)</td>
</tr>
<tr>
<td>-1 months prior:</td>
<td></td>
<td>0.846</td>
<td>0.983</td>
<td>1.374*</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(0.965)</td>
<td>(0.694)</td>
<td>(0.822)</td>
</tr>
<tr>
<td>1 months after:</td>
<td></td>
<td>2.915***</td>
<td>3.029***</td>
<td>3.355***</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.021)</td>
<td>(0.937)</td>
<td>(1.082)</td>
</tr>
<tr>
<td>2 months after:</td>
<td></td>
<td>2.615**</td>
<td>2.735***</td>
<td>3.064**</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.193)</td>
<td>(1.042)</td>
<td>(1.298)</td>
</tr>
<tr>
<td>3 months after:</td>
<td></td>
<td>3.447***</td>
<td>3.577***</td>
<td>3.896***</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td>(1.284)</td>
<td>(1.175)</td>
<td>(1.488)</td>
</tr>
<tr>
<td>4 months after:</td>
<td></td>
<td>3.318**</td>
<td>3.524***</td>
<td>3.764**</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.326)</td>
<td>(1.263)</td>
<td>(1.509)</td>
<td></td>
</tr>
<tr>
<td>5 months after:treatment</td>
<td>3.246**</td>
<td>3.484***</td>
<td>3.797**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.410)</td>
<td>(1.330)</td>
<td>(1.641)</td>
<td></td>
</tr>
<tr>
<td>6 months after:treatment</td>
<td>3.093**</td>
<td>3.321***</td>
<td>3.576**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.360)</td>
<td>(1.265)</td>
<td>(1.586)</td>
<td></td>
</tr>
<tr>
<td>7 months after:treatment</td>
<td>3.306**</td>
<td>3.530**</td>
<td>3.827**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.554)</td>
<td>(1.495)</td>
<td>(1.724)</td>
<td></td>
</tr>
<tr>
<td>8 months after:treatment</td>
<td>2.780**</td>
<td>2.940***</td>
<td>3.219**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.223)</td>
<td>(1.092)</td>
<td>(1.324)</td>
<td></td>
</tr>
<tr>
<td>9 months after:treatment</td>
<td>2.111*</td>
<td>2.231**</td>
<td>2.595**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.273)</td>
<td>(1.029)</td>
<td>(1.181)</td>
<td></td>
</tr>
<tr>
<td>10 months after:treatment</td>
<td>2.167*</td>
<td>2.214**</td>
<td>2.776**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.307)</td>
<td>(1.045)</td>
<td>(1.263)</td>
<td></td>
</tr>
<tr>
<td>11 months after:treatment</td>
<td>1.685</td>
<td>0.903</td>
<td>1.193</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.336)</td>
<td>(0.746)</td>
<td>(0.968)</td>
<td></td>
</tr>
<tr>
<td>12 months after:treatment</td>
<td>1.255</td>
<td>1.159</td>
<td>1.751*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.188)</td>
<td>(0.765)</td>
<td>(1.009)</td>
<td></td>
</tr>
<tr>
<td>13 months after:treatment</td>
<td>1.204</td>
<td>1.128</td>
<td>2.077*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.289)</td>
<td>(0.822)</td>
<td>(1.121)</td>
<td></td>
</tr>
<tr>
<td>14 months after:treatment</td>
<td>1.065</td>
<td>1.003</td>
<td>1.993*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.361)</td>
<td>(0.844)</td>
<td>(1.115)</td>
<td></td>
</tr>
<tr>
<td>15 months after:treatment</td>
<td>1.102</td>
<td>1.052</td>
<td>2.045*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.397)</td>
<td>(0.853)</td>
<td>(1.134)</td>
<td></td>
</tr>
<tr>
<td>16 months after:treatment</td>
<td>1.407</td>
<td>1.309*</td>
<td>2.323**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.290)</td>
<td>(0.788)</td>
<td>(1.115)</td>
<td></td>
</tr>
<tr>
<td>17 months after:treatment</td>
<td>1.220</td>
<td>1.356</td>
<td>2.269*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.326)</td>
<td>(0.887)</td>
<td>(1.263)</td>
<td></td>
</tr>
<tr>
<td>18 months after:treatment</td>
<td>1.550</td>
<td>1.674**</td>
<td>2.578**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.293)</td>
<td>(0.850)</td>
<td>(1.246)</td>
<td></td>
</tr>
<tr>
<td>19 months after:treatment</td>
<td>1.116</td>
<td>1.202</td>
<td>2.094</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.414)</td>
<td>(0.954)</td>
<td>(1.324)</td>
<td></td>
</tr>
<tr>
<td>20 months after:treatment</td>
<td>1.364</td>
<td>1.535</td>
<td>2.222*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.409)</td>
<td>(0.975)</td>
<td>(1.315)</td>
<td></td>
</tr>
<tr>
<td>21 months after:treatment</td>
<td>1.016</td>
<td>1.263</td>
<td>1.898</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.311)</td>
<td>(0.891)</td>
<td>(1.273)</td>
<td></td>
</tr>
<tr>
<td>22 months after:treatment</td>
<td>0.579</td>
<td>0.583</td>
<td>1.311</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.369)</td>
<td>(0.785)</td>
<td>(1.095)</td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>0.0002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household wealth index</td>
<td>-0.311**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of rooms</td>
<td>0.187</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.395)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household size</td>
<td>-0.119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative month fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Calendar month fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Village fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Household fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.19</td>
<td>0.19</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>20,318</td>
<td>20,318</td>
<td>20,318</td>
<td></td>
</tr>
<tr>
<td>Treatment sample</td>
<td>348</td>
<td>348</td>
<td>348</td>
<td></td>
</tr>
</tbody>
</table>
Figure B.1: Distribution of frequency of electricity top-ups per month among treatment and control customers before and after appliance delivery. Note: Bottom panel shows the average treatment effect of the program. Error bars represent a 95% confidence interval; *p < 0.1; **p < 0.05; ***p < 0.01
Figure B.2: Weeks from appliance delivery that coincides with the lean season. Notes: Top panel shows the months of the year that the weeks from appliance delivery in the bottom panel fall on for majority of the sample during each week.

Figure B.3: Average Treatment effect of program among customers whose reported primary source of income is from agriculture and those whose reported primary source of income is from commerce, services or salaried work. Notes: Error bars represent a 95% confidence interval; robust standard errors are clustered by village. Includes household fixed effects, calendar-month fixed effects and relative week fixed effects. Note: *p < 0.1; **p < 0.05; ***p < 0.01
Figure B.4: Average weekly consumption among residential and business customers prior to and after appliance delivery. Note: Includes only a subset of the appliance financing customers for whom we have information on their connection type.

Figure B.5: Treatment effect of taking up individual appliances and treatment effect of taking up multiple types of appliances. Notes: Error bars represent a 95% confidence interval; robust standard errors are clustered by village. Includes household fixed effects, calendar-month fixed effects and relative week fixed effects.
As noted in Section 5.2.4, the identification of treatment effects based on matching treatment and control observations relies on two key assumptions: conditional independence (i.e., potential outcomes are independent of treatment assignment), and common support (i.e., that there is sufficient overlap in the characteristics of treated and non-treated units to find sufficient matches) [41]. While it is impossible to test these assumptions fully, particularly the conditional independence of unobservable characteristics; however, this appendix presents the results of several tests to supplement the analyses provided in the main text.

First, to further investigate whether covariates are balanced across the two groups, we conduct a test of joint orthogonality [173], which evaluates whether these variables (even if not individually significant) co-vary in a way that jointly predicts treatment assignment. We fail to reject ($p = 0.56$) the hypothesis of joint insignificance. In a similar test, over all of the demographic and socioeconomic variables shown in Table 5.3, we also fail to reject the null hypothesis of joint insignificance ($p = 0.17$). These tests complement the individual variable tests (Table 5.3) in suggesting that the demographic variables used to identify matched control customers do not vary significantly (in this case, jointly) over assignment into treatment and matched control groups.
Second, we evaluate match quality by regressing treatment status on the characteristics used to identify matched controls in two samples: (i) the full set of treated and non-treated customers, and (ii) the set of treatment customers and matched controls [41]. We then compare two statistics across the regressions on these two samples: (a) pseudo-$R^2$ values, and (b) the likelihood ratio for the joint significance of all regressors (i.e., the regression with the characteristics used to identify matched controls, versus regression on a constant). We find that before matching, the pseudo-$R^2$ for regressing treatment status on covariates is 0.018 and the likelihood ratio is 29.17 ($p = 0.0000$). After matching, the pseudo-$R^2$ is 0.004, and the likelihood ratio is 4.00 ($p = 0.5488$). The finding that the post-matching pseudo-$R^2$ is very close to zero, and that the likelihood ratio test suggests that the match covariates do not significantly increase the likelihood compared to an empty model, suggests there are no systematic differences in the distribution of covariates between the treatment group and matched controls. This in turn supports the notion that our propensity score matching approach succeeded in producing a group of matched controls that is substantially similar to treated customers, except for the treatment itself.

Third, we perform a sensitivity analysis, using the Rosenbaum bounds approach, to quantify how the estimated treatment effect would change under a violation of the assumption of conditional independence. This approach allows us to determine how strong the “hidden bias” from unobserved confounding variables must be in order to undermine the implications of the matching analysis. We use a range of values for the parameter $\Gamma$ (expressed in terms of the odds ratio between assignment to the treatment and control groups) and calculate both the upper bound p-value (i.e., the significance level for the test of the null hypothesis), and the upper and lower Rosenbaum bounds for the Hodges-Lehmann point estimate.\textsuperscript{1} Table C.1 provides the resulting Rosenbaum

\textsuperscript{1}Note that the Hodges-Lehmann point estimate is roughly equivalent to the difference in medians
bounds for values of $\Gamma$ ranging from 1.0 to 1.5. Within this range of odds ratios, both the upper bound p-value and the Rosenbaum bounds for the Hodges-Lehmann point estimate are relatively stable, at least for the months immediately following the intervention; above 1.5, we quickly see the upper bound p-values rise to greater than 0.1 for all months (except the very first month after appliance delivery, which is still significant at the 10% level until $\Gamma = 1.9$). From this we conclude that if an unobserved variable caused the odds ratio of treatment assignment to differ between treatment and control groups by more than 1.5 and if this variable’s effect on post-treatment consumption was sufficiently strong that it almost perfectly determines whether the post-treatment consumption is bigger for the treatment or the control case in each pair of matched cases in the data, then the effect of the treatment would be considered not statistically significant [72]. Considering that our matched control customers were offered the appliance financing program and actively declined to take it up, it seems possible that unobservable factors could drive a difference of this magnitude. This leads us to conclude that our results are moderately sensitive to hidden bias, even as the other tests in this section present somewhat encouraging evidence for the stability of our results to deviations from the conditional independence assumption.

Fourth, we estimate the time-varying treatment effects of the appliance financing program under a range of alternative matching assumptions, with the results shown in Table C.2. Comparing column (1), which shows our preferred matching method, to the other columns, shows that although the magnitude of the month-by-month treatment

---

2There is no hard and fast rule to inform the appropriate range of values over which to consider the potential for hidden bias. One author [139] suggests that 1.0 to 1.5 is a typical range of values applicable to most social science analyses, but the right range is surely driven largely by the specific context of a given study.
effects is somewhat different for matched controls identified with alternative methods (and usually smaller), the overall pattern of the treatment effects—notably, the finding that treatment effects are significant and positive for the first six to nine months after treatment, and then positive but not significant in subsequent months—are similar across all other matching methods. Figure C.1 shows the week-by-week treatment effects for treatment customers compared to controls identified using the kernel matching method, and Figure C.2 shows the same for the radius matching method.

Finally, Figure C.3 shows the distribution of propensity scores for the treated and untreated observations so as to illustrate the region of common support. As the figure shows, the propensity scores among treatment and matched control customers overlap relatively well, including in relative frequency, especially for the developers identified as "Developer 1" and "Developer 2". The overlap for Developer 4 is not as strong, particularly for the higher values of propensity scores for treatment customers (above 0.4 or so). Treatment customers in this region were not matched to control customers in the 1:1 propensity score match approach, and are excluded from the analysis.

**Table C.2:** Month level dynamic treatment effects of appliance financing program: Alternative matching methods

<table>
<thead>
<tr>
<th>Dependent variable: Monthly electricity consumption (kWh)</th>
<th>(1) 1-to-1 match (PSM with wealth index)</th>
<th>(2) 1-to-1 match (PSM with PCA)</th>
<th>(3) Kernel matching (PSM with elements)</th>
<th>(4) Radius matching (PSM with elements)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 months prior:treatment</td>
<td>-0.022 (0.434)</td>
<td>-0.063 (0.410)</td>
<td>-0.136 (1.203)</td>
<td>-0.070 (1.093)</td>
</tr>
<tr>
<td>11 months prior:treatment</td>
<td>0.228 (0.584)</td>
<td>-0.589 (0.802)</td>
<td>-0.525 (1.070)</td>
<td>-0.370 (0.953)</td>
</tr>
<tr>
<td>10 months prior:treatment</td>
<td>0.476 (0.688)</td>
<td>-0.242 (0.689)</td>
<td>-0.132 (0.965)</td>
<td>-0.324 (0.929)</td>
</tr>
<tr>
<td>9 months prior:treatment</td>
<td>0.784 (0.931)</td>
<td>0.224 (1.095)</td>
<td>0.005 (0.980)</td>
<td>0.163 (0.917)</td>
</tr>
<tr>
<td>8 months prior:treatment</td>
<td>0.498 (0.824)</td>
<td>0.157 (1.191)</td>
<td>-0.259 (0.985)</td>
<td>-0.187 (0.928)</td>
</tr>
<tr>
<td></td>
<td>1-to-1 match (PSM with wealth index PCA)</td>
<td>1-to-1 match (PSM with individual elements)</td>
<td>Kernel matching</td>
<td>Radius matching</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>7 months prior: treatment</td>
<td>0.241 (-0.825)</td>
<td>-0.551 (1.211)</td>
<td>-0.267 (0.927)</td>
<td>-0.243 (0.876)</td>
</tr>
<tr>
<td>6 months prior: treatment</td>
<td>0.605 (0.783)</td>
<td>-0.046 (1.030)</td>
<td>-0.366 (0.954)</td>
<td>-0.313 (0.901)</td>
</tr>
<tr>
<td>5 months prior: treatment</td>
<td>1.534 (0.963)</td>
<td>0.111 (0.883)</td>
<td>0.346 (0.979)</td>
<td>0.258 (0.944)</td>
</tr>
<tr>
<td>4 months prior: treatment</td>
<td>1.295 (1.048)</td>
<td>-0.604 (0.931)</td>
<td>-0.267 (0.974)</td>
<td>-0.400 (0.915)</td>
</tr>
<tr>
<td>3 months prior: treatment</td>
<td>1.204 (0.910)</td>
<td>0.018 (0.811)</td>
<td>-0.184 (0.977)</td>
<td>-0.233 (0.931)</td>
</tr>
<tr>
<td>2 months prior: treatment</td>
<td>1.151 (0.894)</td>
<td>-0.026 (0.945)</td>
<td>0.100 (0.985)</td>
<td>0.191 (0.917)</td>
</tr>
<tr>
<td>1 month prior: treatment</td>
<td>1.374* (0.822)</td>
<td>0.760 (0.988)</td>
<td>0.323 (0.959)</td>
<td>0.311 (0.909)</td>
</tr>
<tr>
<td>1 month after: treatment</td>
<td>3.355** (1.082)</td>
<td>2.65*** (1.019)</td>
<td>2.108** (0.945)</td>
<td>2.003** (0.907)</td>
</tr>
<tr>
<td>2 months after: treatment</td>
<td>3.064** (1.298)</td>
<td>2.279 (1.474)</td>
<td>1.781* (0.956)</td>
<td>1.735* (0.902)</td>
</tr>
<tr>
<td>3 months after: treatment</td>
<td>3.896** (1.488)</td>
<td>3.347** (1.610)</td>
<td>2.762** (1.131)</td>
<td>2.774** (1.081)</td>
</tr>
<tr>
<td>4 months after: treatment</td>
<td>3.764** (1.509)</td>
<td>3.471* (1.852)</td>
<td>2.398** (1.190)</td>
<td>2.250** (1.141)</td>
</tr>
<tr>
<td>5 months after: treatment</td>
<td>3.797** (1.641)</td>
<td>2.699 (1.674)</td>
<td>2.274** (1.121)</td>
<td>2.230** (1.093)</td>
</tr>
<tr>
<td>6 months after: treatment</td>
<td>3.576** (1.586)</td>
<td>3.572** (1.739)</td>
<td>2.246** (1.124)</td>
<td>2.033* (1.085)</td>
</tr>
<tr>
<td>7 months after: treatment</td>
<td>3.827** (1.724)</td>
<td>2.956* (1.737)</td>
<td>2.118* (1.223)</td>
<td>1.916 (1.177)</td>
</tr>
<tr>
<td>8 months after: treatment</td>
<td>3.219** (1.324)</td>
<td>2.058** (0.846)</td>
<td>1.299 (0.947)</td>
<td>1.125 (0.895)</td>
</tr>
<tr>
<td>9 months after: treatment</td>
<td>2.595** (1.181)</td>
<td>1.867* (1.038)</td>
<td>0.857 (0.956)</td>
<td>0.613 (0.891)</td>
</tr>
<tr>
<td>10 months after: treatment</td>
<td>2.776** (1.263)</td>
<td>1.227 (1.098)</td>
<td>1.343 (0.956)</td>
<td>1.078 (0.877)</td>
</tr>
<tr>
<td>11 months after: treatment</td>
<td>1.193 (0.968)</td>
<td>0.236 (1.120)</td>
<td>0.392 (0.956)</td>
<td>0.117 (0.915)</td>
</tr>
<tr>
<td>12 months after: treatment</td>
<td>1.751* (1.009)</td>
<td>0.393 (0.890)</td>
<td>0.755 (1.013)</td>
<td>0.231 (0.921)</td>
</tr>
<tr>
<td>13 months after: treatment</td>
<td>2.077* (1.121)</td>
<td>1.166 (1.069)</td>
<td>0.959 (0.993)</td>
<td>0.480 (0.877)</td>
</tr>
<tr>
<td>14 months after: treatment</td>
<td>1.993* (1.115)</td>
<td>1.019 (1.108)</td>
<td>0.787 (0.974)</td>
<td>0.388 (0.875)</td>
</tr>
<tr>
<td>15 months after: treatment</td>
<td>2.045* (1.134)</td>
<td>1.099 (1.102)</td>
<td>1.424 (0.988)</td>
<td>1.001 (0.893)</td>
</tr>
<tr>
<td>16 months after: treatment</td>
<td>2.322** (1.115)</td>
<td>1.240 (1.118)</td>
<td>0.921 (0.929)</td>
<td>0.599 (0.883)</td>
</tr>
<tr>
<td>17 months after: treatment</td>
<td>2.269* (1.263)</td>
<td>1.618 (1.090)</td>
<td>0.906 (0.937)</td>
<td>0.673 (0.890)</td>
</tr>
</tbody>
</table>
### Table

<table>
<thead>
<tr>
<th></th>
<th>1-to-1 match (PSM with wealth index PCA)</th>
<th>1-to-1 match (PSM with individual elements)</th>
<th>Kernel matching</th>
<th>Radius matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 months after: treatment</td>
<td>2.578** (1.246)</td>
<td>1.297 (1.139)</td>
<td>0.822 (0.956)</td>
<td>0.562 (0.903)</td>
</tr>
<tr>
<td>19 months after: treatment</td>
<td>2.094 (1.324)</td>
<td>1.178 (1.089)</td>
<td>1.030 (0.975)</td>
<td>0.731 (0.915)</td>
</tr>
<tr>
<td>20 months after: treatment</td>
<td>2.222* (1.315)</td>
<td>1.486 (1.177)</td>
<td>1.148 (0.985)</td>
<td>0.763 (0.937)</td>
</tr>
<tr>
<td>21 months after: treatment</td>
<td>1.898 (1.273)</td>
<td>1.143 (1.118)</td>
<td>0.796 (0.994)</td>
<td>0.438 (0.938)</td>
</tr>
<tr>
<td>22 months after: treatment</td>
<td>1.311 (1.095)</td>
<td>0.608 (0.998)</td>
<td>0.512 (0.980)</td>
<td>0.188 (0.925)</td>
</tr>
<tr>
<td>Relative month fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Calendar month fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Household fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.68</td>
<td>0.71</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Observations</td>
<td>20,318</td>
<td>20,287</td>
<td>51,011</td>
<td>51,011</td>
</tr>
<tr>
<td>Treatment sample</td>
<td>348</td>
<td>348</td>
<td>348</td>
<td>348</td>
</tr>
<tr>
<td>Control sample</td>
<td>342</td>
<td>342</td>
<td>1347</td>
<td>1347</td>
</tr>
</tbody>
</table>

Notes: ***, **, and * indicate statistical significance at 1%, 5%, and 10%. Robust standard errors, clustered at village level for 1-to-1 matching and bootstrapped for kernel and radius matching, are in parentheses.
### Table C.1: Rosenbaum sensitivity analysis for average treatment effect of appliance financing program

<table>
<thead>
<tr>
<th>Period</th>
<th>$\Gamma = 1.0$</th>
<th>$\Gamma = 1.1$</th>
<th>$\Gamma = 1.3$</th>
<th>$\Gamma = 1.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 months prior:treatment</td>
<td>0.999 -1.09</td>
<td>-1.09</td>
<td>1.000 -1.29</td>
<td>-0.89</td>
</tr>
<tr>
<td>12 months prior:treatment</td>
<td>1.000 -1.27</td>
<td>-1.27</td>
<td>1.000 -1.47</td>
<td>-1.07</td>
</tr>
<tr>
<td>11 months prior:treatment</td>
<td>0.999 -0.78</td>
<td>-0.78</td>
<td>1.000 -0.98</td>
<td>-0.58</td>
</tr>
<tr>
<td>10 months prior:treatment</td>
<td>0.997 -0.59</td>
<td>-0.59</td>
<td>1.000 -0.79</td>
<td>-0.39</td>
</tr>
<tr>
<td>9 months prior:treatment</td>
<td>0.997 -0.56</td>
<td>-0.56</td>
<td>1.000 -0.76</td>
<td>-0.36</td>
</tr>
<tr>
<td>8 months prior:treatment</td>
<td>1.000 -0.75</td>
<td>-0.75</td>
<td>1.000 -0.95</td>
<td>-0.55</td>
</tr>
<tr>
<td>7 months prior:treatment</td>
<td>1.000 -0.84</td>
<td>-0.84</td>
<td>1.000 -1.04</td>
<td>-0.64</td>
</tr>
<tr>
<td>6 months prior:treatment</td>
<td>1.000 -0.86</td>
<td>-0.86</td>
<td>1.000 -1.06</td>
<td>-0.66</td>
</tr>
<tr>
<td>5 months prior:treatment</td>
<td>0.996 -0.41</td>
<td>-0.41</td>
<td>1.000 -0.61</td>
<td>-0.21</td>
</tr>
<tr>
<td>4 months prior:treatment</td>
<td>1.000 -0.62</td>
<td>-0.62</td>
<td>1.000 -0.82</td>
<td>-0.42</td>
</tr>
<tr>
<td>3 months prior:treatment</td>
<td>1.000 -0.50</td>
<td>-0.50</td>
<td>1.000 -0.70</td>
<td>-0.30</td>
</tr>
<tr>
<td>2 months prior:treatment</td>
<td>1.000 -0.68</td>
<td>-0.68</td>
<td>1.000 -0.88</td>
<td>-0.48</td>
</tr>
<tr>
<td>1 month prior:treatment</td>
<td>0.994 -0.38</td>
<td>-0.38</td>
<td>1.000 -0.58</td>
<td>-0.18</td>
</tr>
<tr>
<td>1 month after:treatment</td>
<td>0.000 1.25</td>
<td>1.25</td>
<td>0.000 1.05</td>
<td>1.45</td>
</tr>
<tr>
<td>2 months after:treatment</td>
<td>0.000 0.67</td>
<td>0.67</td>
<td>0.001 0.47</td>
<td>0.87</td>
</tr>
<tr>
<td>3 months after:treatment</td>
<td>0.000 0.81</td>
<td>0.81</td>
<td>0.000 0.61</td>
<td>1.01</td>
</tr>
<tr>
<td>4 months after:treatment</td>
<td>0.000 0.94</td>
<td>0.94</td>
<td>0.000 0.74</td>
<td>1.14</td>
</tr>
<tr>
<td>5 months after:treatment</td>
<td>0.004 0.49</td>
<td>0.49</td>
<td>0.026 0.29</td>
<td>0.69</td>
</tr>
<tr>
<td>6 months after:treatment</td>
<td>0.000 0.68</td>
<td>0.68</td>
<td>0.004 0.48</td>
<td>0.88</td>
</tr>
<tr>
<td>7 months after:treatment</td>
<td>0.000 0.77</td>
<td>0.77</td>
<td>0.001 0.57</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Notes: $\Gamma$ refers to the ratio of odds of treatment assignment. The $p$-critical value shown is the Rosenbaum upper bound $p$-value, that is, the significance level for the test of the null hypothesis. LB and UB denote the lower and upper Rosenbaum bounds, respectively, for the Hodges-Lehmann point estimate.
Figure C.1: Average treatment effect of appliance financing program from kernel matching technique. Standard errors are bootstrapped; error bars represent a 95% confidence interval. Includes household fixed effects, calendar-month fixed effects, and relative week fixed effects. Note: *p < 0.1; **p < 0.05; ***p < 0.01.

Figure C.2: Average treatment effect of appliance financing program from radius matching technique. Standard errors are bootstrapped; error bars represent a 95% confidence interval. Includes household fixed effects, calendar-month fixed effects, and relative week fixed effects. Note: *p < 0.1; **p < 0.05; ***p < 0.01.
Figure C.3: Propensity scores for households in treatment (appliance uptake) and matched control (no uptake) groups indicating domain of common support
APPENDIX D

ELECTRIC VEHICLE TECHNICAL OPTIONS
<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Model</th>
<th>Tare weight (kg)</th>
<th>Battery capacity (kWh)</th>
<th>Range (km)</th>
<th>Consumption (Wh/km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audi</td>
<td>e-tron 50</td>
<td>2595</td>
<td>86.5</td>
<td>365</td>
<td>262</td>
</tr>
<tr>
<td>Audi</td>
<td>e-tron 55</td>
<td>2445</td>
<td>64.7</td>
<td>280</td>
<td>219</td>
</tr>
<tr>
<td>Audi</td>
<td>e-tron Sportback 55</td>
<td>2695</td>
<td>86.5</td>
<td>335</td>
<td>251</td>
</tr>
<tr>
<td>Audi</td>
<td>e-tron Sportback 50</td>
<td>2445</td>
<td>64.7</td>
<td>295</td>
<td>207</td>
</tr>
<tr>
<td>BMW</td>
<td>i3</td>
<td>1345</td>
<td>37.9</td>
<td>235</td>
<td>176</td>
</tr>
<tr>
<td>BMW</td>
<td>i3s</td>
<td>1365</td>
<td>37.9</td>
<td>230</td>
<td>302</td>
</tr>
<tr>
<td>BMW</td>
<td>iX3</td>
<td>2260</td>
<td>74</td>
<td>360</td>
<td>211</td>
</tr>
<tr>
<td>Byton</td>
<td>M-Byte</td>
<td>2300</td>
<td>72</td>
<td>325</td>
<td>248</td>
</tr>
<tr>
<td>Byton</td>
<td>M-Byte</td>
<td>2500</td>
<td>95</td>
<td>400</td>
<td>271</td>
</tr>
<tr>
<td>Chevrolet</td>
<td>Bolt EV</td>
<td>1626</td>
<td>60</td>
<td>383</td>
<td>169</td>
</tr>
<tr>
<td>Citroen</td>
<td>e-C4</td>
<td>1650</td>
<td>45</td>
<td>250</td>
<td>184</td>
</tr>
<tr>
<td>Dacia</td>
<td>Spring</td>
<td>950</td>
<td>26.8</td>
<td>170</td>
<td>179</td>
</tr>
<tr>
<td>DS 3</td>
<td>Crossback E-Tense</td>
<td>1598</td>
<td>45</td>
<td>250</td>
<td>196</td>
</tr>
<tr>
<td>Fiat</td>
<td>500e hatchback</td>
<td>1465</td>
<td>42</td>
<td>250</td>
<td>173</td>
</tr>
<tr>
<td>Fiat</td>
<td>500e hatchback</td>
<td>1350</td>
<td>23.8</td>
<td>145</td>
<td>170</td>
</tr>
<tr>
<td>Fiat</td>
<td>500e Cabrio</td>
<td>1405</td>
<td>37.3</td>
<td>230</td>
<td>182</td>
</tr>
<tr>
<td>Ford</td>
<td>Mustang Mach-E AWD SR</td>
<td>2086</td>
<td>68</td>
<td>330</td>
<td>226</td>
</tr>
<tr>
<td>Ford</td>
<td>Mustang Mach-E AWD ER</td>
<td>2218</td>
<td>88</td>
<td>420</td>
<td>230</td>
</tr>
<tr>
<td>Ford</td>
<td>Mustang Mach-E RWD SR</td>
<td>1993</td>
<td>68</td>
<td>345</td>
<td>219</td>
</tr>
<tr>
<td>Ford</td>
<td>Mustang Mach-E RWD ER</td>
<td>2111</td>
<td>88</td>
<td>440</td>
<td>222</td>
</tr>
<tr>
<td>Ford</td>
<td>Mustang Mach-E GT</td>
<td>2250</td>
<td>88</td>
<td>410</td>
<td>236</td>
</tr>
<tr>
<td>Honda</td>
<td>e</td>
<td>1588</td>
<td>28.5</td>
<td>170</td>
<td>184</td>
</tr>
<tr>
<td>Honda</td>
<td>e Advance</td>
<td>1595</td>
<td>28.5</td>
<td>170</td>
<td>184</td>
</tr>
<tr>
<td>Hyundai</td>
<td>Ioniq Electric</td>
<td>1602</td>
<td>38.3</td>
<td>250</td>
<td>200</td>
</tr>
<tr>
<td>Hyundai</td>
<td>Kona Electric</td>
<td>1535</td>
<td>39.2</td>
<td>255</td>
<td>170</td>
</tr>
<tr>
<td>Hyundai</td>
<td>Kona Electric</td>
<td>1760</td>
<td>64</td>
<td>400</td>
<td>175</td>
</tr>
<tr>
<td>Jaguar</td>
<td>I-PACE</td>
<td>2208</td>
<td>84.7</td>
<td>365</td>
<td>246</td>
</tr>
<tr>
<td>Kandi</td>
<td>K27</td>
<td>1030</td>
<td>17.69</td>
<td>95</td>
<td>244</td>
</tr>
<tr>
<td>Kia</td>
<td>Niro Electric</td>
<td>1667</td>
<td>39.2</td>
<td>235</td>
<td>242</td>
</tr>
<tr>
<td>Kia</td>
<td>Niro Electric</td>
<td>1812</td>
<td>64.8</td>
<td>365</td>
<td>188</td>
</tr>
<tr>
<td>Kia</td>
<td>e-Soul</td>
<td>1610</td>
<td>39.2</td>
<td>230</td>
<td>206</td>
</tr>
<tr>
<td>Kia</td>
<td>e-Soul</td>
<td>1757</td>
<td>64</td>
<td>365</td>
<td>206</td>
</tr>
<tr>
<td>Lexus</td>
<td>UX 300e</td>
<td>1860</td>
<td>50</td>
<td>260</td>
<td>214</td>
</tr>
<tr>
<td>Lucid</td>
<td>Air Dream</td>
<td>2300</td>
<td>110</td>
<td>645</td>
<td>190</td>
</tr>
<tr>
<td>Lucid</td>
<td>Air Touring</td>
<td>2150</td>
<td>85</td>
<td>530</td>
<td>166</td>
</tr>
<tr>
<td>Lucid</td>
<td>Air Grand Touring</td>
<td>2300</td>
<td>110</td>
<td>660</td>
<td>175</td>
</tr>
<tr>
<td>Lucid</td>
<td>Air Pure</td>
<td>2100</td>
<td>85</td>
<td>540</td>
<td>164</td>
</tr>
<tr>
<td>Mazda</td>
<td>MX-30</td>
<td>1720</td>
<td>30</td>
<td>170</td>
<td>200</td>
</tr>
<tr>
<td>Mercedes-Benz</td>
<td>EQC 400 4MATIC</td>
<td>2495</td>
<td>80</td>
<td>370</td>
<td>235</td>
</tr>
<tr>
<td>Mercedes-Benz</td>
<td>EQA-250</td>
<td>2040</td>
<td>66.5</td>
<td>355</td>
<td>222</td>
</tr>
<tr>
<td>MG</td>
<td>MG5 EV</td>
<td>1532</td>
<td>48.8</td>
<td>285</td>
<td>185</td>
</tr>
<tr>
<td>MG</td>
<td>ZS EV</td>
<td>1566</td>
<td>42.5</td>
<td>220</td>
<td>204</td>
</tr>
<tr>
<td>Mini</td>
<td>Cooper SE Hardtop 2 door</td>
<td>1440</td>
<td>28.9</td>
<td>185</td>
<td>175</td>
</tr>
<tr>
<td>Mitsubishi</td>
<td>i-MiEV</td>
<td>1080</td>
<td>16</td>
<td>160</td>
<td>186.4</td>
</tr>
<tr>
<td>Nissan</td>
<td>Leaf</td>
<td>1580</td>
<td>36</td>
<td>220</td>
<td>181</td>
</tr>
<tr>
<td>Nissan</td>
<td>Leaf e+</td>
<td>1750</td>
<td>56</td>
<td>325</td>
<td>190</td>
</tr>
<tr>
<td>Nissan</td>
<td>Ariya 63 kWh</td>
<td>1800</td>
<td>63</td>
<td>330</td>
<td>210</td>
</tr>
<tr>
<td>Nissan</td>
<td>Ariya 87 kWh</td>
<td>2000</td>
<td>87</td>
<td>440</td>
<td>218</td>
</tr>
<tr>
<td>Nissan</td>
<td>Ariya e-4ORCE</td>
<td>1900</td>
<td>63</td>
<td>325</td>
<td>252</td>
</tr>
<tr>
<td>Nissan</td>
<td>Ariya e-4ORCE</td>
<td>2300</td>
<td>87</td>
<td>420</td>
<td>229</td>
</tr>
<tr>
<td>Manufacturer</td>
<td>Model</td>
<td>Tare weight (kg)</td>
<td>Battery capacity (kWh)</td>
<td>Range (km)</td>
<td>Consumption (Wh/km)</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>------------------------</td>
<td>------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Opel</td>
<td>Corsa</td>
<td>1530</td>
<td>45</td>
<td>275</td>
<td>173</td>
</tr>
<tr>
<td>Opel</td>
<td>Mokka</td>
<td>1550</td>
<td>45</td>
<td>255</td>
<td>196</td>
</tr>
<tr>
<td>Peugeot</td>
<td>e-208</td>
<td>1530</td>
<td>45</td>
<td>450</td>
<td>173</td>
</tr>
<tr>
<td>Peugeot</td>
<td>e-2008 SUV</td>
<td>1548</td>
<td>45</td>
<td>250</td>
<td>196</td>
</tr>
<tr>
<td>Polestar</td>
<td>Polestar 2</td>
<td>2198</td>
<td>75</td>
<td>395</td>
<td>189</td>
</tr>
<tr>
<td>Porsche</td>
<td>Taycan Plus</td>
<td>2295</td>
<td>83.7</td>
<td>375</td>
<td>186</td>
</tr>
<tr>
<td>Porsche</td>
<td>Taycan Turbo</td>
<td>2380</td>
<td>83.7</td>
<td>400</td>
<td>217</td>
</tr>
<tr>
<td>Porsche</td>
<td>Taycan Turbo S</td>
<td>2370</td>
<td>83.7</td>
<td>385</td>
<td>226</td>
</tr>
<tr>
<td>Porsche</td>
<td>Taycan 4S</td>
<td>2215</td>
<td>71</td>
<td>375</td>
<td>189</td>
</tr>
<tr>
<td>Porsche</td>
<td>Taycan 4S Plus</td>
<td>2295</td>
<td>83.7</td>
<td>435</td>
<td>192</td>
</tr>
<tr>
<td>Renault</td>
<td>Zoe ZE40 R110</td>
<td>1577</td>
<td>52</td>
<td>310</td>
<td>186</td>
</tr>
<tr>
<td>Renault</td>
<td>Zoe ZE50 R110</td>
<td>1577</td>
<td>52</td>
<td>310</td>
<td>186</td>
</tr>
<tr>
<td>SEAT</td>
<td>Mi Electric</td>
<td>1235</td>
<td>32.3</td>
<td>195</td>
<td>180</td>
</tr>
<tr>
<td>Skoda</td>
<td>Enyaq iV 50</td>
<td>1950</td>
<td>52</td>
<td>280</td>
<td>193</td>
</tr>
<tr>
<td>Smart</td>
<td>EQ forfour</td>
<td>1200</td>
<td>16.7</td>
<td>95</td>
<td>186</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model 3 Long Range Dual Motor</td>
<td>1919</td>
<td>70</td>
<td>460</td>
<td>163</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model 3 Long Range Performance</td>
<td>1919</td>
<td>76</td>
<td>455</td>
<td>163</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model S Performance</td>
<td>2316</td>
<td>95</td>
<td>525</td>
<td>178</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model X Long Range Plus</td>
<td>2533</td>
<td>95</td>
<td>470</td>
<td>224</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model X Performance</td>
<td>2572</td>
<td>95</td>
<td>450</td>
<td>216</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model Y Long Range</td>
<td>2000</td>
<td>72.5</td>
<td>425</td>
<td>185</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model Y Long Range Performance</td>
<td>2000</td>
<td>72.5</td>
<td>410</td>
<td>195</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model 3 Standard Range Plus</td>
<td>1825</td>
<td>50</td>
<td>335</td>
<td>160</td>
</tr>
<tr>
<td>Tesla</td>
<td>Model S Long Range Plus</td>
<td>2290</td>
<td>95</td>
<td>540</td>
<td>178</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>ID.4 1st</td>
<td>2124</td>
<td>77</td>
<td>400</td>
<td>211</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>ID.4 Pro S</td>
<td>1934</td>
<td>77</td>
<td>450</td>
<td>205</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>e-Golf</td>
<td>1615</td>
<td>32</td>
<td>190</td>
<td>188</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>ID.3 Pro</td>
<td>1794</td>
<td>58</td>
<td>350</td>
<td>181</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>ID.3 1st</td>
<td>1794</td>
<td>58</td>
<td>350</td>
<td>181</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>e-up!</td>
<td>1235</td>
<td>32.3</td>
<td>195</td>
<td>223</td>
</tr>
<tr>
<td>Volvo</td>
<td>XC40 Recharge</td>
<td>2188</td>
<td>75</td>
<td>335</td>
<td>285</td>
</tr>
<tr>
<td>Audi</td>
<td>e-tron GT RS</td>
<td>2420</td>
<td>85</td>
<td>375</td>
<td>227</td>
</tr>
<tr>
<td>Fiat</td>
<td>500e Cabrio</td>
<td>1281</td>
<td>21.3</td>
<td>135</td>
<td>178</td>
</tr>
<tr>
<td>Kia</td>
<td>EV6 GT</td>
<td>2200</td>
<td>74</td>
<td>335</td>
<td>221</td>
</tr>
<tr>
<td>Renault</td>
<td>Twingo</td>
<td>1208</td>
<td>21.3</td>
<td>120</td>
<td>178</td>
</tr>
<tr>
<td>Toyota</td>
<td>bZ4X FWD</td>
<td>2020</td>
<td>71.4</td>
<td>340</td>
<td>210</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


[12] Argonne. Downloadable dynamometer database, generated at the Argonne national laboratory, advanced powertrain research facility (aprf), under the funding and guidance of the U.S. department of energy (DOE), 2015.


[194] NRECA. History - the story behind america’s electric cooperatives and nreca, 2021.


[266] W. Ukpe. Africa accounts for only 6% of global energy demand – afdb, 2 2021.


