EI @ Haas WP 270A

Supplementary Appendix for Online Publication

"Experimental Evidence on the Demand for and Costs of Rural Electrification"

Kenneth Lee, Edward Miguel, and Catherine Wolfram

May 2016

http://ei.haas.berkeley.edu

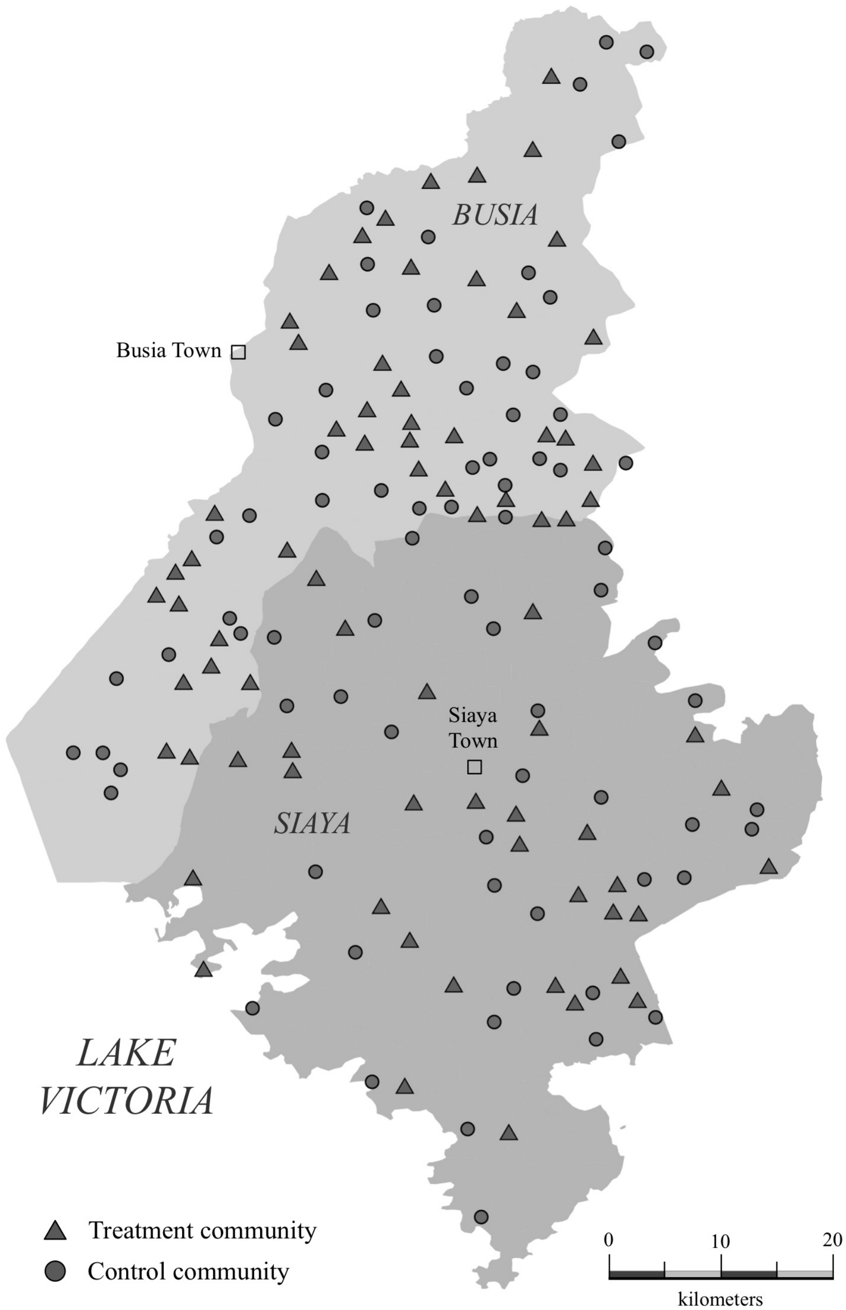
Supplementary Appendix for Online Publication

"Experimental Evidence on the Demand for and Costs of Rural Electrification"

Kenneth Lee, Edward Miguel, and Catherine Wolfram

May 2016

Figure A1—150 sample communities in Busia and Siaya counties in Kenya



Notes: The final sample of 150 communities includes 85 and 65 transformers in Busia and Siaya counties, respectively.

Figure A2—Example of REA offer letter for a subsidized household electricity connection

Rural Electrification Authority

NYANZA REGION OFFICE
Mamboleo junction, Opposite Lake Basin Development Authority, Kisumu – Kakamega Highway
P.O. Box 2604 – 40100

ipa INNOVATIONS FOR POVERTY ACTION

Tom Mboya Drive,
Milimani Kisumu.
P.O. Box 313

ELECTRICITY SUPPLY TO YOUR PREMISES (ESP)

UNIQUE CUSTOMER LOCATION DETAILS:

Project name:	|______|	Reference number:	|______|
Transformer:	|______|	Substation number:	|______|

Customer name: |________________________|

Customer address: |________________________|

Household coordinates: |____________| LATITUDE

|____________| LONGITUDE

Dear Sir/Madam,

Reference is made to the ongoing *Rural Electric Power Project* research project that is being carried out by Innovations for Poverty Action (IPA) in partnership with REA in this community. As part of this research project, we are pleased to advise you that a single-phase service cable can be installed to your premises at the following price:

SINGLE-PHASE SUPPLY CONNECTION: KES [EFFECTIVE PRICE]

If you would like to accept this offer, kindly arrange to pay this amount to the Rural Electrification Authority account at Kenya Commercial Bank, Milimani Branch, Account Number: 1103201557.

NOTICE: When making your payment, please remember to quote your unique ID number: [NUMBER] in the memo line. Please keep a copy of your bank payment slip for official receipting. This receipt is very important and must be shown to IPA in order to complete this process. No payment should be made to any individual. Your payment should only be made to the REA bank account above. This offer is valid for 8 weeks:

DATE OF OFFER: DD-MM-YY **DATE OF EXPIRY:** DD-MM-YY

For further enquiries, please contact the REA Regional Coordinator, Nyanza Region, [NAME] at [MOBILE] or the IPA Project Associate in Busia Town, [NAME] at [MOBILE].

For: RURAL ELECTRIFICATION AUTHORITY

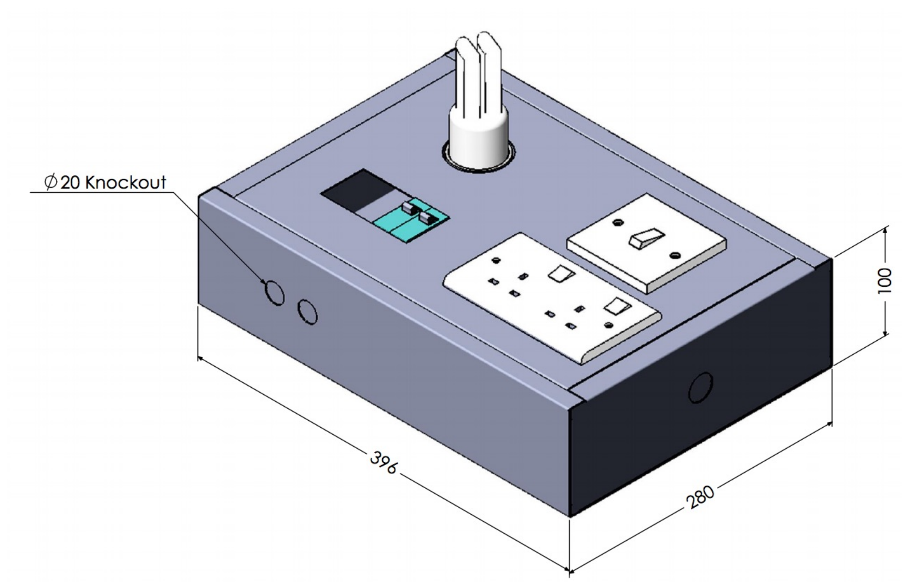
____

[NAME]
REGIONAL CO-ORDINATOR, NYANZA REGION.

Unique ID: [NUMBER]

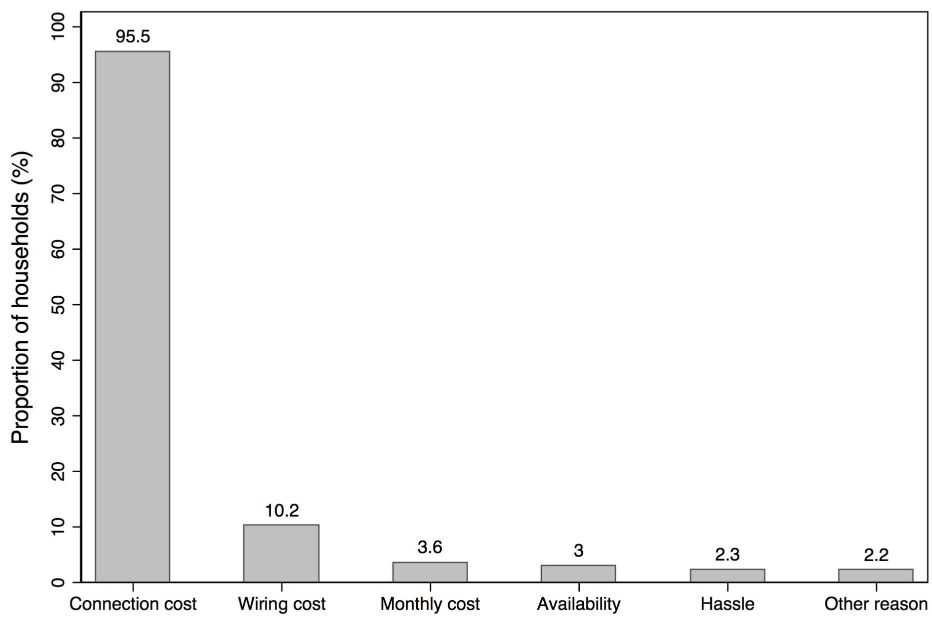
Notes: Each offer letter was signed and guaranteed by REA management. Project field staff members visited each treatment community and explained the details of the offer to a representative from each household in a community meeting. The meeting was held to give community members a chance to ask questions.

Figure A3—*Umeme Rahisi* "ready-board" designed by Power Technics



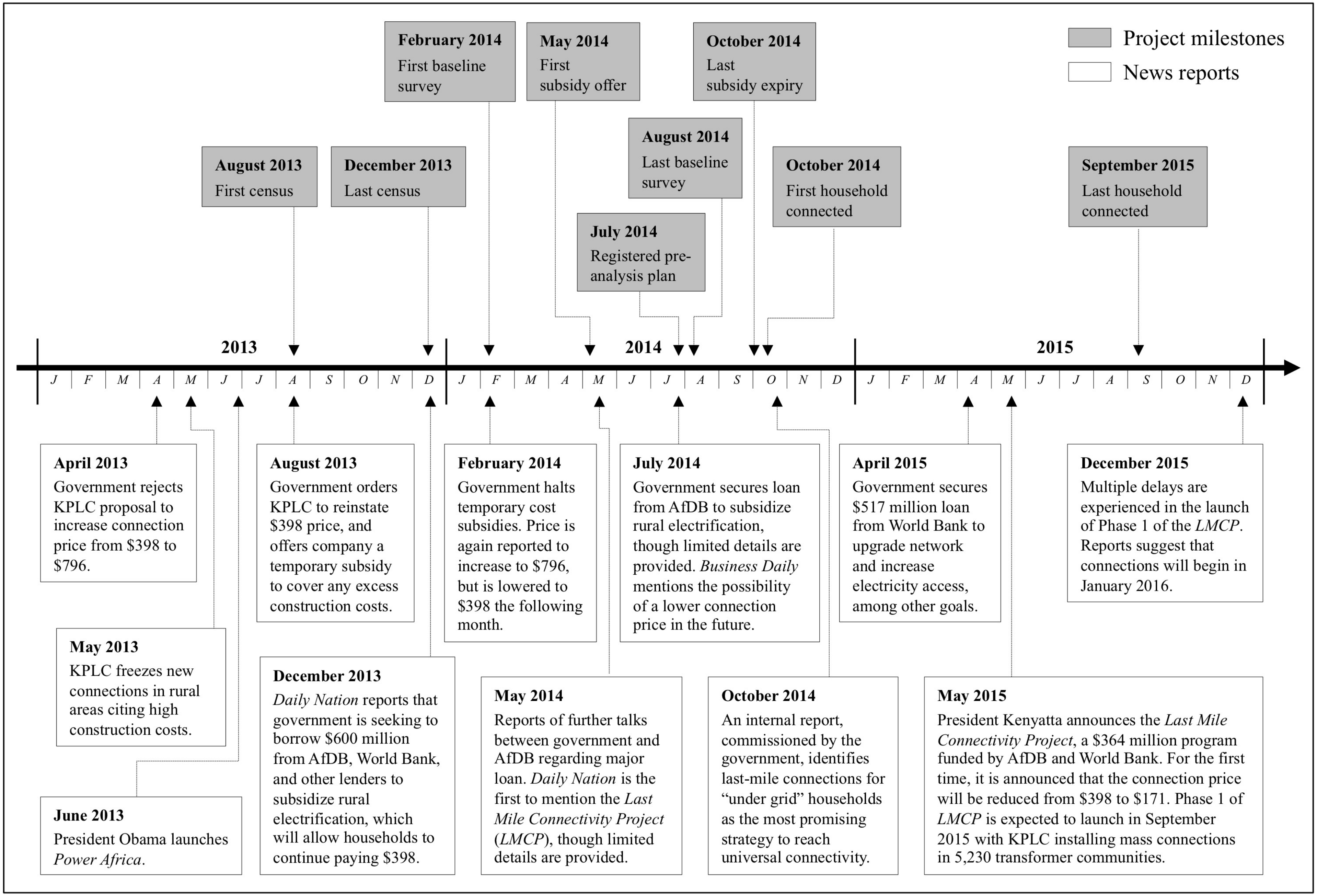
Notes: Treatment households received an opportunity to install a certified household wiring solution in their homes at no additional cost. 88.5 percent of the households connected in the experiment accepted this offer, while 11.5 percent provided their own wiring. Each ready-board, valued at roughly $34 per unit, featured a single light bulb socket, two power outlets, and two miniature circuit breakers. The unit is first mounted onto a wall and the electricity service line is directly connected to the back. The hardware was designed and produced by Power Technics, an electronic supplies manufacturer in Nairobi.

Figure A4—Stated reasons why households remain unconnected to electricity at baseline



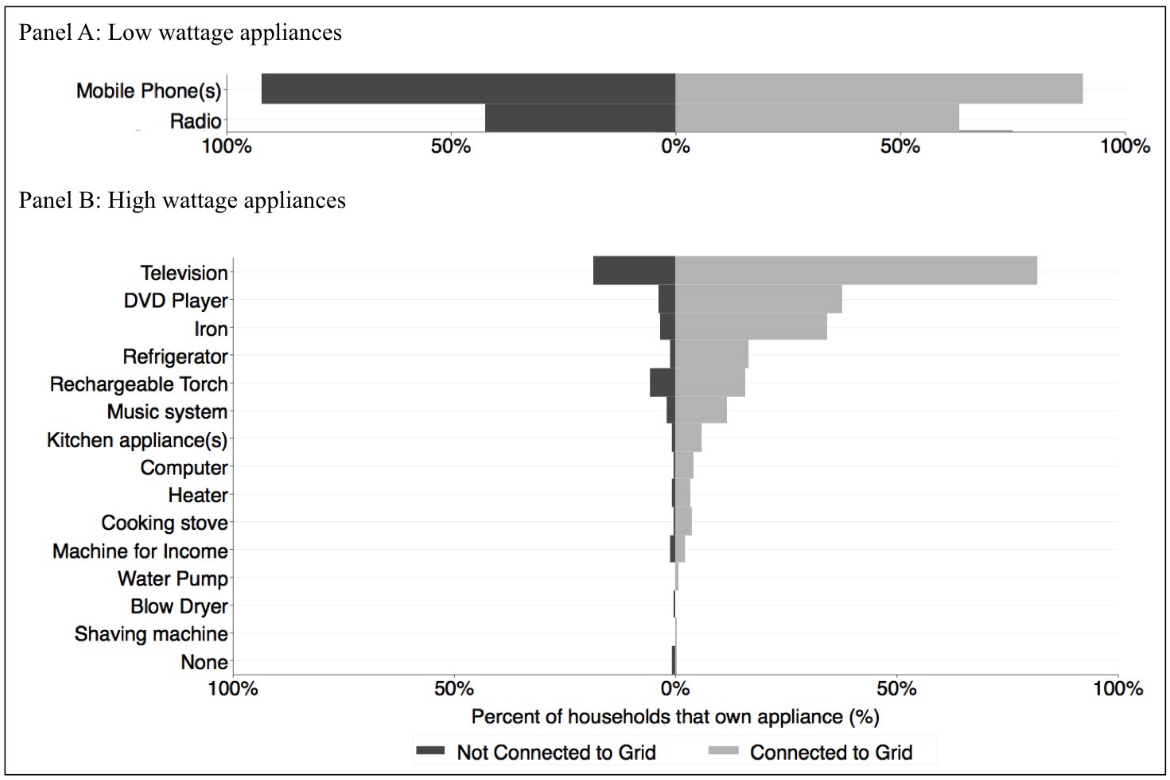
Notes: Based on the responses of 2,289 unconnected households during the baseline survey round.

Figure A5—Timeline of project milestones and connection price-related news reports over the period of study



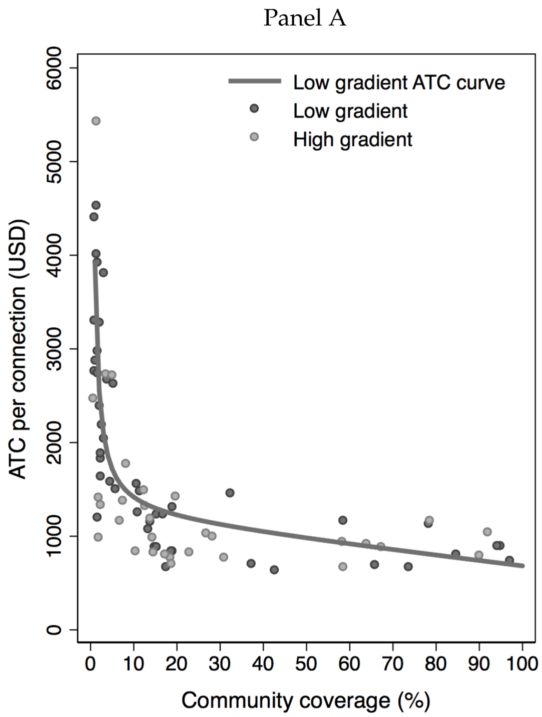
Notes: Sources for news reports related to the grid connection price include *Daily Nation*, Kenya's leading national newspaper, and *Business Daily*.

Figure A6—Electrical appliances owned by unconnected and connected households at baseline



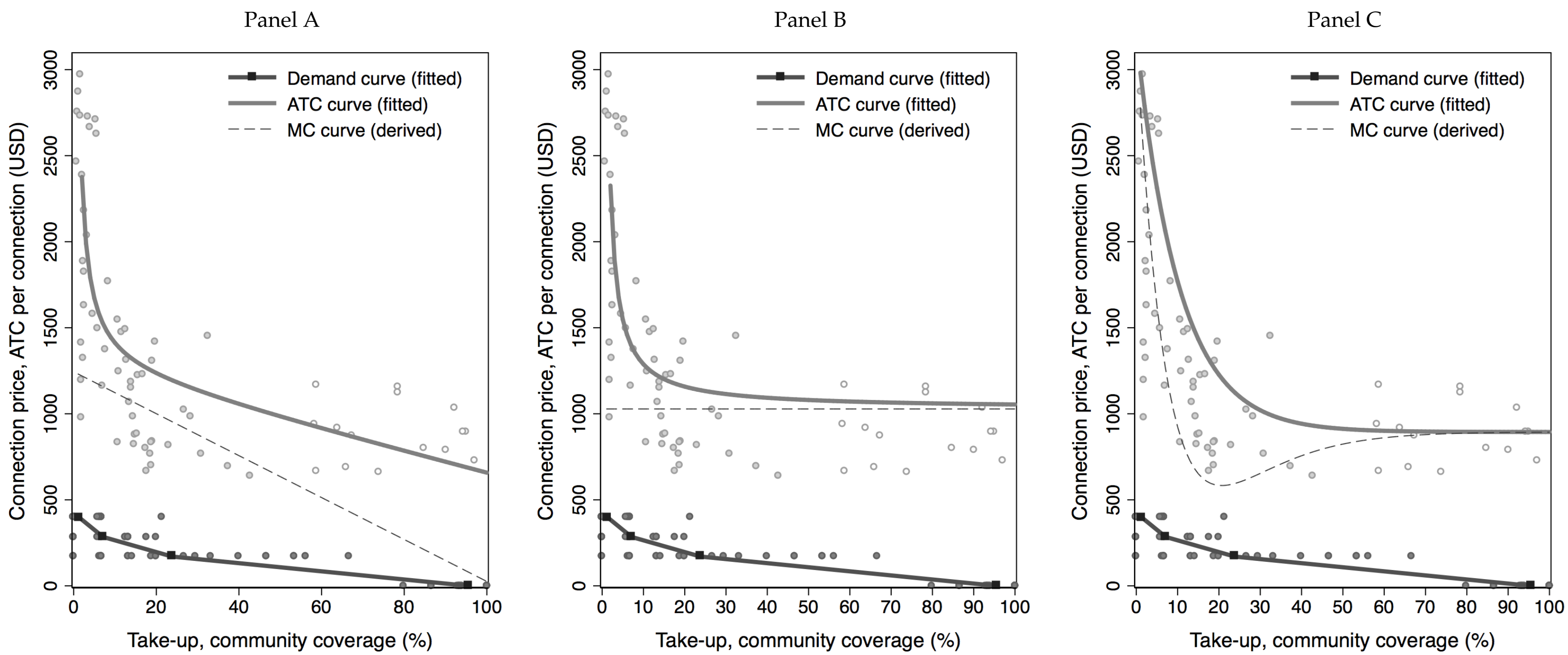
Notes: Based on the responses of 2,289 unconnected households and 215 connected households during the baseline survey round. See Lee, Miguel, and Wolfram (2016) for a discussion.

Figure A7—Average total cost (ATC) per connection by land gradient



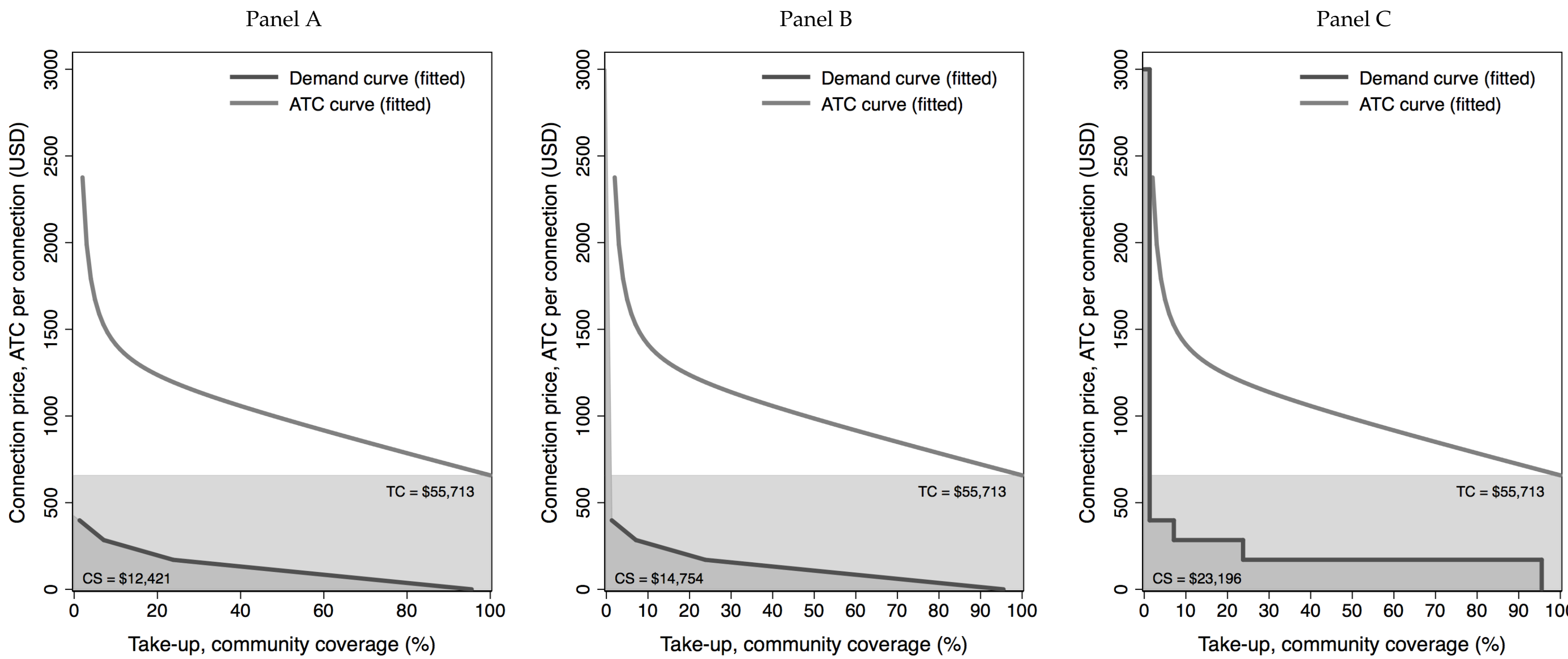
Notes: In the sample of communities, average land gradient ranges from 0.79 to 7.76 degrees with a mean of 2.15 degrees. We divide the sample into communities with "low" average land gradient (i.e., below median) gradient and communities with "high" average land gradient (i.e., above median). In Panels A and B, we plot fitted lines from nonlinear estimations of $ATC = b_0/x + b_1 + b_2x$ for the low and high gradient subsamples, respectively (they lie nearly on top of each other so we present them here in separate panels for clarity).

Figure A8—Experimental estimates of a natural monopoly: Alternative functional forms



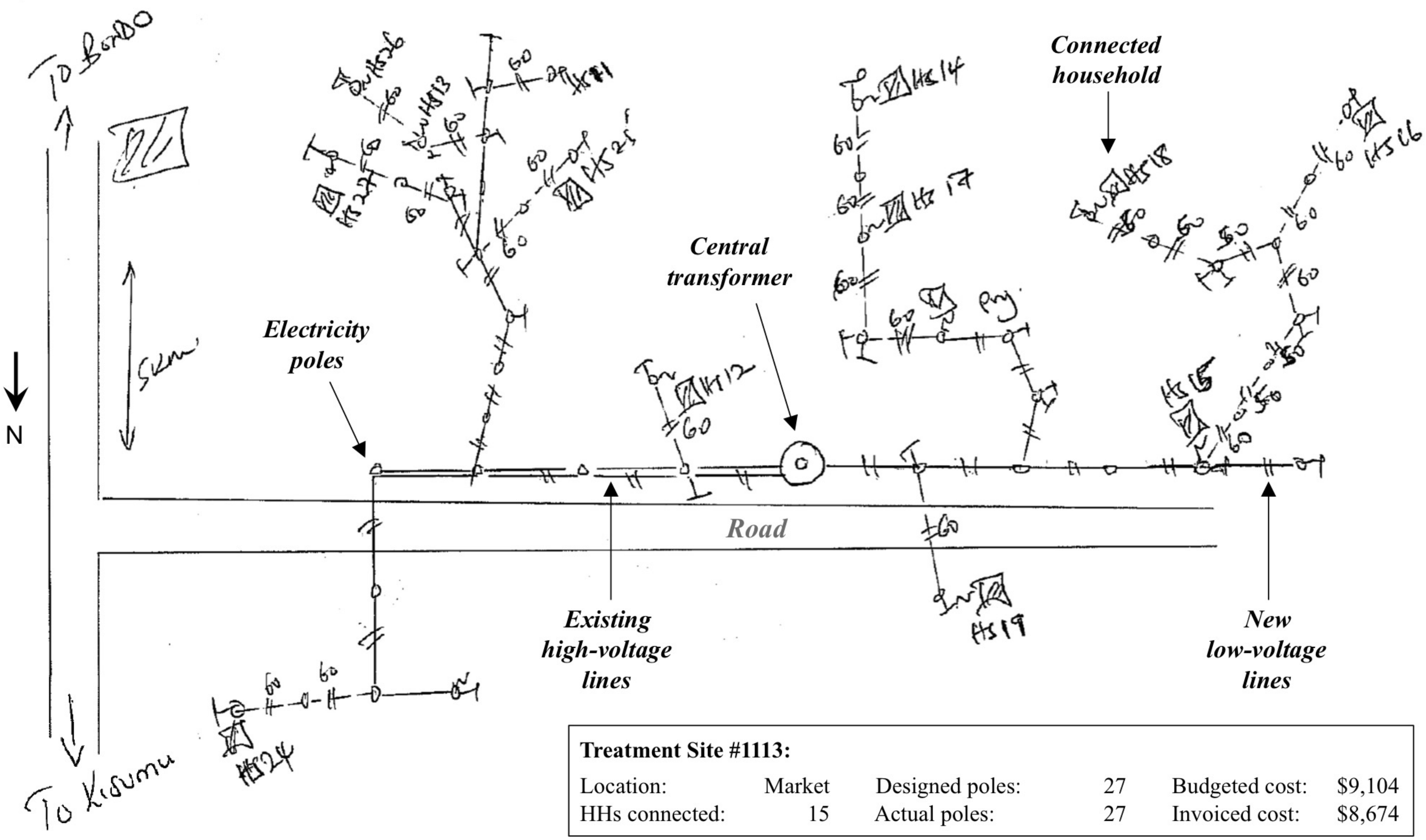
Notes: Panel A reproduces Figure 6, Panel A. In Panel B, we estimate an average total cost curve with constant variable costs. In Panel C, we estimate an exponential function to derive a marginal cost curve.

Figure A9—Welfare loss associated with rural electrification under various demand curve assumptions



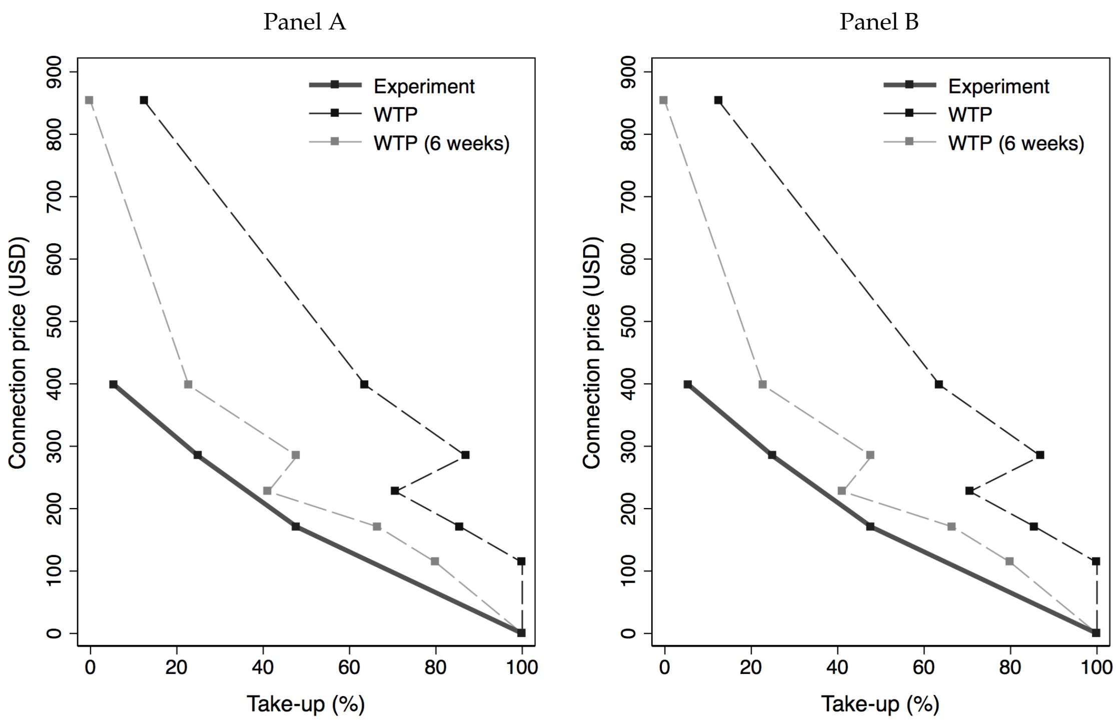
Notes: Panel A reproduces Figure 6, Panel B. In this scenario, the welfare loss associated with a mass electrification program is $43,292 per community. In Panel B, we estimate the area under the unobserved [0, 1.3] domain by assuming that the demand curve intercepts the vertical axis at $3,000, rather than $424 (as in Panel A). In this more conservative case, the welfare loss is $41,611 per community. In order to overturn this result (i.e. costs exceeding the consumer surplus), the intercept would need to be an astronomical $32,300. In Panel C, the most conservative case, we assume that demand is a step function and calculate the welfare loss to be $32,517 per community. The required discounted future welfare gains needed for consumer surplus to exceed total costs across the three scenarios range from $384 (in Panel C) to $511 (in Panel A) per household.

Figure A10—Example of a REA design drawing in a high subsidy treatment community



Notes: After receiving payment, REA designers visited each treatment community to design the local low-voltage network. The designs were then used to estimate the required materials and determine a budgeted estimates of the total construction cost. Materials (e.g. poles, electricity line, service cables) represented 65.9 percent of total installation costs. The community in this example is the same as that shown in Figure 2.

Figure A11—Stated willingness to pay with and without time constraints, by wall quality and bank accounts



Notes: We plot the experimental results (solid black line) against the responses to a set of contingent valuation questions included in the baseline survey. Households were first asked whether they would accept a hypothetical offer (i.e., randomly assigned price) to connect to the grid (dashed line, black squares). Households were then asked whether they would accept the same hypothetical offer if required to complete the payment in six weeks (dashed line, grey squares). In Panel A, the sample contains all households with low-quality walls and no bank account (n=1,647). In Panel B, the sample contains all households with high-quality walls and a bank account (n=236).

Table A1—Comparison of social and economic indicators for study region and nationwide counties

		Nationwide county percentiles		
	Study region	25th	50th	75th
Total population	793,125	528,054	724,186	958,791
per square kilometer	375.4	39.5	183.2	332.9
% rural	85.7	71.6	79.5	84.4
% at school	44.7	37.0	42.4	45.2
% in school with secondary education	10.3	9.7	11.0	13.4
Total households	176,630	103,114	154,073	202,291
per square kilometer	83.6	7.9	44.3	78.7
% with high quality roof	59.7	49.2	78.5	88.2
% with high quality floor	27.7	20.6	29.7	40.0
% with high quality walls	32.2	20.3	28.0	41.7
% with piped water	6.3	6.9	14.2	30.6
Total public facilities	644	356	521	813
per capita (000s)	0.81	0.59	0.75	0.98
Electrification rates				
Rural (%)	2.3	1.5	3.1	5.3
Urban (%)	21.8	20.2	27.2	43.2
Public facilities (%)	84.1	79.9	88.1	92.6

Notes: The study region column presents weighted-average and average (where applicable) statistics for Busia and Siaya counties. Specifically, total population, total households, and total public facilities represent averages for Busia and Siaya. We exclude Nairobi and Mombasa, two counties that are entirely urban, from the nationwide county percentile columns. Demographic data is obtained from the 2009 Kenya Population and Housing Census (KPHC). Public facility electrification data is obtained from the Rural Electrification Authority (REA). High quality roof indicates roofs made of concrete, tiles, or corrugated iron sheets. High quality floor indicates floors made of cement, tiles, or wood. High quality walls indicates walls made of stone, brick, or cement. Rural and urban electrification rates represent the proportion of households that stated that electricity was their main source of lighting during the 2009 census. Based on the 2009 census data, the mean (county-level) electrification rates in rural and urban areas were 4.6 and 32.6 percent, respectively. Nationally, the rural and urban electrification rates were 5.1 and 50.4 percent, respectively, and 22.7 percent overall. An earlier version of this table is presented in Lee et al. (2016).

Table A2—Baseline summary statistics and randomization balance check

		Regression coefficients on subsidy treatment indicators			
	Control	Low	Medium	High	p-value of F-test
	(1)	(2)	(3)	(4)	(5)
Panel A: Household head (respondent)					
Female=1	0.63	0.021	-0.026	-0.015	0.624
	[0.48]	(0.031)	(0.029)	(0.032)	
Age (years)	52.04	-1.098	1.029	1.701	0.287
	[16.29]	(1.242)	(1.064)	(1.388)	
Completed secondary school=1	0.14	-0.012	0.025	-0.033	0.297
	[0.34]	(0.024)	(0.027)	(0.025)	
Married=1	0.66	-0.005	0.013	-0.024	0.856
	[0.47]	(0.033)	(0.033)	(0.033)	
Not a farmer=1	0.23	0.002	-0.029	-0.004	0.793
	[0.42]	(0.036)	(0.031)	(0.026)	
Basic political awareness=1	0.13	-0.051***	-0.009	-0.026	0.039
	[0.33]	(0.018)	(0.020)	(0.016)	
Has bank account=1	0.19	-0.032	0.003	-0.022	0.452
	[0.39]	(0.022)	(0.028)	(0.025)	
Panel B: Household characteristics					
Number of members	5.30	-0.265*	0.100	-0.294*	0.071
	[2.71]	(0.145)	(0.160)	(0.177)	
Youth members (age $\leqq$ 18)	3.03	-0.101	0.075	-0.214	0.247
	[2.17]	(0.122)	(0.120)	(0.144)	
High-quality walls=1	0.15	0.051**	0.039	-0.014	0.092
	[0.36]	(0.025)	(0.030)	(0.026)	
Land (acres)	1.85	0.299	0.229	0.054	0.414
	[2.14]	(0.201)	(0.214)	(0.149)	
Distance to transformer (m)	348.6	14.85	9.50	22.10**	0.173
	[140.0]	(9.85)	(12.15)	(10.64)	
Monthly (non-charcoal) energy (USD)	5.55	-0.234	0.495*	-0.432	0.026
	[5.20]	(0.266)	(0.267)	(0.283)	

(*Table continued on next page*)

(*Table continued from previous page*)

	Control	Low	Medium	High	p-value of F-test
		Regression coefficients on subsidy treatment indicators			
	(1)	(2)	(3)	(4)	(5)
Panel C: Household assets					
Bednets	2.27	-0.032	0.056	0.000	0.887
	[1.50]	(0.091)	(0.092)	(0.096)	
Bicycles	0.66	-0.027	0.076	0.016	0.353
	[0.74]	(0.042)	(0.052)	(0.052)	
Sofa pieces	5.92	-0.039	0.477	-0.008	0.66
	[5.21]	(0.366)	(0.399)	(0.403)	
Chickens	7.03	0.420	-0.421	-0.218	0.739
	[8.74]	(0.690)	(0.616)	(0.680)	
Cattle	1.74	0.069	0.190	0.232	0.514
	[2.32]	(0.151)	(0.192)	(0.185)	
Radios	0.34	-0.015	0.047	-0.002	0.41
	[0.48]	(0.028)	(0.033)	(0.040)	
Televisions	0.16	0.001	-0.003	-0.054**	0.132
	[0.37]	(0.022)	(0.025)	(0.024)	
Panel D: Community characteristics					
Electrification rate (%)	5.25	1.57	-0.03	-0.08	0.674
	[4.61]	(1.34)	(0.99)	(0.86)	
Community population	534.69	42.10	26.37	9.85	0.793
	[219.02]	(45.01)	(41.70)	(39.13)	

Notes: Column 1 reports mean values for the control group, with standard deviations in brackets. Columns 2 to 4 report the coefficients from separate regressions in which a dependent variable is regressed on the full set of treatment indicators and stratification variables (including county, market status, and whether the transformer was funded and installed early on, between 2008 and 2010). Standard errors are in parantheses. Column 5 reports the p-values of F-tests of whether the treatment coefficients are jointly equal to zero. Robust standard errors clustered at the community level. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$. Sample sizes range from 2,279 to 2,289 depending on missing values except in the specification with Age as the dependent variable where the sample size is 2,205. The variables selected for the randomization check were not specified in the pre-analysis plan.

Table A3—Characteristics of households taking-up electricity by treatment arm

	High subsidy Price: \$0 (1)	Medium subsidy Price: *\$171* (2)	Low subsidy Price: *\$284* (3)	Control Price: \$398 (4)
Panel A: Respondent characteristics				
Female (%)	61.71	58.89	59.26	60.00
Age (years)	53.69	52.82	50.56	51.62
Attended secondary school (%)	9.92	27.78***	33.33***	26.67**
Married (%)	64.19	74.44*	70.37	66.67
Not a farmer (%)	22.31	28.89	29.63	28.57
Basic political awareness (%)	9.64	16.67*	14.81	6.67
Has bank account (%)	17.13	31.11***	40.74***	35.71*
Panel B: Household characteristics				
Number of members	5.02	6.19***	6.22**	5.80
Youth members	2.81	3.52***	3.89**	3.29
High-quality walls (%)	12.95	25.56***	51.85***	33.33**
Land (acres)	1.93	2.20	2.56	2.14
Distance to transformer (m)	369.74	357.41	369.06	360.67
Monthly (non-charcoal) energy (USD)	5.16	7.62***	8.24***	5.88
Panel C: Household assets				
Bednets	2.29	2.77***	3.44***	2.53
Sofa pieces	5.88	8.96***	9.44***	8.87**
Chickens	6.90	9.07**	10.31*	6.43
Radios	0.34	0.48**	0.48	0.53
Televisions	0.11	0.28***	0.48***	0.40***
Take-up of electricity connections	363	90	27	15

Notes: Columns 1, 2, and 3 report sample means for unconnected households that chose to take-up a subsidized electricity connection. Column 4 reports sample means for control group households that chose to connect on their own. Basic political awareness indicator captures whether the household head was able to correctly identify the heads of state of Tanzania, Uganda, and the United States. The asterisks in columns 2, 3, and 4 indicate statistically significant differences compared to column 1: * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$.

Table A4A—Impact of connection subsidy on take-up: Interactions with community-level variables

		Interacted variable			
		Busia county	Transformer funded early on	Market center	Baseline population
	(1)	(2)	(3)	(4)	(5)
T1: Low subsidy—29% discount	5.94***	2.65	4.97***	6.34***	2.29
	(1.50)	(1.71)	(1.89)	(1.71)	(3.97)
T2: Medium subsidy—57% discount	22.88***	20.93***	26.81***	23.54***	18.48*
	(4.02)	(5.78)	(6.23)	(4.80)	(10.27)
T3: High subsidy—100% discount	94.97***	95.19***	93.74***	94.92***	100.05***
	(1.27)	(1.69)	(1.73)	(1.64)	(4.49)
Interacted variable		0.20	0.24	0.88	-0.00
		(0.85)	(0.81)	(1.01)	(0.00)
T1 × interacted variable		5.59**	2.06	-1.64	0.01
		(2.65)	(3.06)	(3.32)	(0.01)
T2 × interacted variable		3.48	-8.19	-2.71	0.01
		(8.01)	(7.90)	(8.98)	(0.02)
T3 × interacted variable		-0.42	2.67	0.15	-0.01
		(2.59)	(2.52)	(2.43)	(0.01)
Take-up in control group	1.30	1.30	1.30	1.30	1.30
Observations	2176	2176	2176	2176	2176
R-squared	0.69	0.69	0.69	0.69	0.69

Notes: The dependent variable is an indicator variable (multiplied by 100) for household take-up. The mean of the dependent variable is 21.6. Robust standard errors clustered at the community level in parentheses. All specfications include the household and community covariates specified in the pre-analysis plan. Household covariates include the age of the household head, indicators for whether the household respondent attended secondary school, is a senior citizen, is not primarily a farmer, is employed, and has a bank account, an indicator for whether the household has high-quality walls, and the number of chickens (a measure of assets) owned by the household. Community covariates include indicators for the county, market status, whether the transformer was funded and installed early on (between 2008 and 2010), baseline electrification rate, and community population. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$. The number of observations in the above regressions is somewhat smaller than the total number of households in our sample (2,289) due to missing data. The coefficients do not change appreciably when the households with missing data are included in the regression with no covariates (as in column 1).

Table A4B—Impact of connection subsidy on take-up: Interactions with household-level variables

	Interacted variable				
	Household size	Age of household head	Senior household head	Number of chickens	Has bank account
	(1)	(2)	(3)	(4)	(5)
T1: Low subsidy—29% discount	0.61	5.49	5.55***	4.68***	4.50***
	(2.67)	(4.96)	(1.71)	(1.42)	(1.38)
T2: Medium subsidy—57% discount	9.81*	26.19***	23.66***	17.21***	20.30***
	(5.71)	(7.06)	(4.17)	(3.78)	(4.10)
T3: High subsidy—100% discount	94.16***	95.22***	95.50***	93.79***	94.91***
	(2.75)	(3.52)	(1.24)	(1.83)	(1.43)
Interacted variable	0.00	0.04	1.19	-0.07*	1.07
	(0.16)	(0.04)	(1.31)	(0.04)	(1.24)
T1 × interacted variable	1.01*	0.01	1.70	0.18	8.43
	(0.53)	(0.09)	(4.25)	(0.12)	(5.87)
T2 × interacted variable	2.41***	-0.06	-3.05	0.84***	13.54*
	(0.88)	(0.11)	(3.56)	(0.27)	(7.25)
T3 × interacted variable	0.13	-0.01	-2.01	0.16	-0.04
	(0.44)	(0.07)	(2.28)	(0.11)	(2.55)
Take-up in control group	1.30	1.30	1.30	1.30	1.30
Observations	2176	2176	2176	2176	2176
R-squared	0.69	0.69	0.69	0.70	0.69

Notes: The dependent variable is an indicator variable (multiplied by 100) for household take-up. The mean of the dependent variable is 21.6. Robust standard errors clustered at the community level in parentheses. All specfications include the household and community covariates specified in the pre-analysis plan. Household covariates include the age of the household head, indicators for whether the household respondent attended secondary school, is a senior citizen, is not primarily a farmer, is employed, and has a bank account, an indicator for whether the household has high-quality walls, and the number of chickens (a measure of assets) owned by the household. Community covariates include indicators for the county, market status, whether the transformer was funded and installed early on (between 2008 and 2010), baseline electrification rate, and community population. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$. The number of observations in the above regressions is somewhat smaller than the total number of households in our sample (2,289) due to missing data. The coefficients do not change appreciably when the households with missing data are included in the regression with no covariates (as in column 1).

Table A5—Impact of scale on average total cost (ATC) per connection

	Sample & Designed—OLS			
	(1)	(2)	(3)	(4)
Number of connections (M)	-87.77***	-81.11***	-83.42***	-114.05***
	(15.11)	(16.45)	(17.01)	(18.84)
M^2	0.83***	0.76***	0.79***	1.31***
	(0.18)	(0.20)	(0.21)	(0.25)
Land gradient		-173.90***	-599.29***	
		(58.05)	(164.10)	
Land gradient × M			36.68***	
			(13.86)	
Land gradient × M^2			-0.33*	
			(0.17)	
Households				-6.07
				(11.49)
Households × M				0.17
				(0.47)
Households × M^2 / 100				-0.91*
				(0.53)
Busia=1		247.69	461.80	304.85
		(388.75)	(391.79)	(448.41)
Market transformer=1		-148.84	-32.07	-170.02
		(195.41)	(185.58)	(177.70)
Transformer funded early on=1		109.29	128.41	41.56
		(218.63)	(216.06)	(205.59)
Community electrification rate		15.89	15.16	9.18
		(15.41)	(14.09)	(13.90)
Community population		-0.66	-0.90	0.48
		(0.66)	(0.68)	(1.92)
Round-trip distance to REA (km)		1.60	-2.45	-0.08
		(3.61)	(3.66)	(3.80)
Community controls	No	Yes	Yes	Yes
Observations	77	77	77	77
R^2	0.43	0.48	0.54	0.52

Notes: The dependent variable is the budgeted average total cost (ATC) per connection in USD. The sample is expanded to include the 15 additional designed communities. Robust standard errors are clustered at the community level. Community covariates were specified in the pre-analysis plan. Columns 1 and 2 report the same results as the regression specifications in columns 4 and 6 in Table 3, respectively. Average land gradient ranges from 0.79 to 7.76 degrees with a mean of 2.15 degrees. Column 4 includes interaction terms with an additional variable—the (demeaned) number of households (i.e., residential compounds) in each community. Asterisks indicate coefficient significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$.

Table A6—Breakdown of labor and transport costs for nine projects (three contracts)

	Contract #1	Contract #2	Contract #3
Panel A: Labor costs (e.g., digging holes, installation, clearing bush, dropping service lines, etc.)			
Budgeted LV poles	40	107	62
Invoiced LV poles	38	98	76
Actual (counted) LV poles	39	92	60
Difference (Actual - Invoiced)	*+1*	*-6*	*-16*
Avg. labor cost per LV pole	*27.59*	*27.59*	*27.59*
Total LV poles labor	1,048	2,704	1,655
Budgeted stays	–	–	35
Invoiced stays	32	68	43
Avg. labor cost per stay	*19.22*	*19.22*	*19.22*
Total stays labor	615	1,308	827
Budgeted HV poles	–	–	6
Invoiced HV poles	12	5	6
Avg. labor cost per HV pole	*35.59*	*35.59*	*35.59*
Total HV poles labor	427	178	214
Additional labor	832	1,552	2,199
Total labor	2,922	5,742	4,895
Panel B: Transport costs (e.g., wood pole and other materials)			
Large lorries	2	4	4
Invoiced round-trip distance (km)	320	300	300
Google round-trip distance (km)	218	256	218
Difference (Actual - Invoiced)	*-102*	*-44*	*-82*
Avg. cost per km	*3.75*	*3.75*	*3.75*
Total large lorry transport	2,402	4,503	4,503
Small lorries	1	3	2
Invoiced round-trip distance (km)	250	250	250
Avg. cost per km	2.98	2.98	2.98
Total small lorry transport	745	2,234	1,490
Total transport	3,146	6,738	5,993
Budgeted labor and transport costs	6,126	12,708	8,956
Invoiced labor and transport costs	7,040	14,477	12,516
Difference (Invoiced - Budgeted)	*14.9%*	*13.9%*	*39.8%*
Projects	3	3	3
Households connected	18	38	22
Construction days	36	31	35

Notes: Based on the detailed invoice submitted to REA. "LV" denotes low-voltage and "HV" denotes high-voltage. Additional labor includes costs of bush clearing, tree cutting, signage, dropping service cables, and other expenses. Each large lorry is capable of transporting 30 poles. Each small lorry is capable of transporting 2.3 km of line materials.

Table A7—Reasons for unexpected delays in household electrification

Phase	Description	Reasons for unexpected delays
A2	Wiring	• In order to begin using electricity, households are required to have a valid meter and a certificate of wiring safety. A large proportion of households were not able to register for a meter because they lacked a PIN (*Personal Identification Number*) certificate from the Kenya Revenue Authority. In our sample, 42 percent of households applying for electricity needed assistance in applying for a PIN certificate.
B1	Design	• Competing priorities at REA due to the 2014/15 nationwide initiative to connect primary schools to the national grid. This resulted in a persistent shortage of REA designers and planners. • Low motivation to perform design duties. In addition, since REA designers were required to physically visit each community, there were numerous challenges in scheduling field visits.
B2	Contracting	• Competing priorities (described above) delayed the bureaucratic paperwork required to prepare contracts. • REA staff members had strong preferences to assign certain projects to specific contractors. This resulted in delays because REA wanted to wait until specific contractors were free to take on new projects.
B3	Construction	• Insufficient materials (e.g., poles, cables) requiring site revisits. • Poor weather (i.e., rainy conditions) made roads impassable and digging holes (for electricity poles) impossible. • Issues in securing wayleaves (i.e., right of ways) to pass through neighboring properties. • Low-quality construction work that needed to be fixed. • Missing materials. • Faulty transformers requiring contractors to revisit sites to complete the final step of the process (e.g., connecting the new low-voltage network to the existing line). • Incorrect households were connected to the network, requiring site revisits. • Contractor issues installing "ready-boards" due to lack of experience.
B4	Metering	• Insufficient materials (e.g., prepaid meters, cables) contributed to lengthy delays at Kenya Power. • Lost meter application forms at local Kenya Power offices. • Changes in internal Kenya Power processes requiring applications to be approved in Nairobi as well as local offices in Siaya, Kisumu, and Busia. • Unexpected requests by local Kenya Power representatives for additional documents (e.g., photocopies of payment receipts). • Local Kenya Power representatives unable to perform metering duties due to competing priorities. • Scheduling difficulties due to the necessity for Kenya Power to make multiple trips to remote village sites, which increased the costs (metering costs are not documented in our cost estimates).

Notes: Each phase of the construction process corresponds to the timeline bar illustrated in Figure 8.

Table A8—Transformer problems in study communities during the 14-month study period (between September 2014 and October 2015)

Row	Site ID	Group	Wave	Treated HHs	Connected	Metered	Blackout	Primary issue
1	1204	Treatment	2	15	Feb-15	May-15	4 months	Burnt out
2	1403	Treatment	1	15	Mar-15	Jul-15	1 month	Commissioning
3	1505	Treatment	2	1	Mar-15	May-15	1 month	Commissioning
4	2101	Treatment	1	0	n/a	n/a	8 months	Burnt out
5	2103	Treatment	1	0	n/a	n/a	4 months	Technical failure
6	2106	Treatment	1	15	Nov-14	Nov-14	8 months	Commissioning
7	2114	Treatment	1	8	Dec-14	Dec-14	12 months	Relocated by Kenya Power
8	2116	Treatment	1	14	Sep-14	May-15	2 months	Technical failure
9	2202	Treatment	1	1	Sep-14	Oct-14	1 month	Technical failure
10	2217	Treatment	1	13	Oct-14	Dec-14	1 month	Technical failure
11	2222	Treatment	1	3	Oct-14	Dec-14	4 months	Leaking oil
12	2303	Treatment	2	7	May-15	Jun-15	4 months	Technical failure
13	2406	Treatment	2	15	Apr-15	Jun-15	1 month	Burnt out
14	2503	Treatment	1	1	Oct-14	Oct-14	6 months	Burnt out
15	2506	Treatment	1	15	Dec-14	Feb-15	9 months	Commissioning
16	1103	Control	n/a	0	n/a	n/a	2 months	Technical failure
17	1109	Control	n/a	0	n/a	n/a	6 months	Burnt out
18	1203	Control	n/a	0	n/a	n/a	1 month	Technical failure
19	1205	Control	n/a	0	n/a	n/a	1 month	Technical failure
20	1405	Control	n/a	0	n/a	n/a	6 months	Burnt out
21	1410	Control	n/a	0	n/a	n/a	2 months	Relocated by Kenya Power
22	2103	Control	n/a	0	n/a	n/a	4 months	Burnt out
23	2115	Control	n/a	0	n/a	n/a	2 months	Technical failure
24	2212	Control	n/a	0	n/a	n/a	5 months	Burnt out
25	2220	Control	n/a	0	n/a	n/a	8 months	Burnt out
26	2304	Control	n/a	0	n/a	n/a	3 months	Stolen
27	2315	Control	n/a	0	n/a	n/a	3 months	Burnt out
28	2504	Control	n/a	0	n/a	n/a	4 months	Technical failure
29	2515	Control	n/a	0	n/a	n/a	4 months	Damaged by weather

Note: “Commissioning” refers to a situation in which the transformer (and related equipment) is installed but electricity is not being delivered.

Table A9A—Impact of randomized offers on take-up

	WTP 1	WTP 2	Experiment
	(1)	(2)	(3)
$853 offer	-19.74***	-8.21***	
	(3.68)	(2.11)	
$284 offer / T1: Low subsidy—29% discount	16.26***	6.02**	5.94***
	(3.39)	(2.50)	(1.50)
$227 offer	14.25***	7.33***	
	(3.56)	(2.74)	
$171 offer / T2: Medium subsidy—57% discount	24.08***	18.50***	22.88***
	(3.37)	(2.66)	(4.02)
$114 offer	25.20***	19.71***	
	(3.51)	(2.88)	
Free offer / T3: High subsidy—100% discount	62.04***	87.47***	94.97***
	(2.90)	(2.20)	(1.27)
Number of household members	1.26***	0.43	0.62***
	(0.39)	(0.33)	(0.21)
High-quality walls=1	9.07***	11.63***	3.53
	(2.66)	(2.34)	(2.14)
Number of chickens=1	0.70***	0.40***	0.12**
	(0.12)	(0.11)	(0.06)
Age (years)	-0.39***	-0.18**	0.03
	(0.09)	(0.07)	(0.04)
Attended secondary school=1	15.61***	5.41**	3.77**
	(2.69)	(2.36)	(1.71)
Over 65 years old=1	0.90	1.30	0.53
	(3.48)	(3.04)	(1.40)
Not a farmer=1	0.37	0.10	1.86
	(2.40)	(1.94)	(1.63)
Has bank account=1	11.13***	11.03***	2.57
	(2.51)	(2.49)	(1.70)
Employed=1	2.27	1.18	1.06
	(2.17)	(1.88)	(1.27)
Take-up in status quo (i.e., $398) group	36.24	9.81	1.30
Mean of dependent variable	53.73	25.54	21.60
Observations	2,157	2,157	2,176
R^2	0.23	0.35	0.69

Notes: In column 1, the dependent variable is an indicator for whether the household accepted the hypothetical offer (i.e. randomly assigned price) to connect to the grid. In column 2, the dependent variable is an indicator for whether the household accepted the hypothetical offer if required to complete the payment in six weeks. In column 3, the dependent variable is an indicator for experimental take-up. All dependent variables are multplied by 100. Robust standard errors clustered at the community level in parentheses. All specfications include the community covariates specified in the pre-analysis plan. Community covariates include indicators for the county, market status, whether the transformer was funded and installed early on (between 2008 and 2010), baseline electrification rate, and community population. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$.

Table A9B—Impact of WTP offer on stated take-up of electricity connections

		Interacted variable		
	Baseline	High-quality walls	Has bank account	Attended secondary schooling
	(1)	(2)	(3)	(4)
\$853 offer	-8.21***	-6.00***	-7.95***	-5.40**
	(2.11)	(2.24)	(1.92)	(2.17)
\$284 offer / T1: Low subsidy—29% discount	6.02**	5.02**	4.86*	6.01**
	(2.50)	(2.42)	(2.57)	(2.40)
\$227 offer	7.33***	6.59**	7.15**	7.73***
	(2.74)	(2.82)	(2.80)	(2.72)
\$171 offer / T2: Medium subsidy—57% discount	18.50***	15.96***	16.59***	17.14***
	(2.66)	(2.69)	(2.85)	(2.69)
\$114 offer	19.71***	18.40***	15.03***	20.02***
	(2.88)	(3.21)	(2.94)	(2.91)
Free offer / T3: High subsidy—100% discount	87.47***	89.64***	89.58***	89.32***
	(2.20)	(2.27)	(2.08)	(2.15)
Interacted variable		7.95	5.63	7.23
		(5.26)	(4.76)	(5.94)
\$853 offer × interacted variable		-8.95	-4.06	-17.96***
		(6.43)	(7.56)	(6.14)
\$284 offer × interacted variable		6.40	5.47	0.02
		(8.51)	(7.28)	(8.82)
\$227 offer × interacted variable		4.62	0.60	-2.70
		(8.51)	(7.39)	(8.54)
\$171 offer × interacted variable		15.71*	9.90	11.38
		(8.29)	(7.79)	(9.88)
\$114 offer × interacted variable		8.50	25.10***	-2.09
		(8.55)	(8.38)	(9.16)
Free offer × interacted variable		-11.53*	-15.06**	-17.21**
		(5.92)	(5.91)	(6.62)
Take-up in status quo (i.e., \$398) group	9.81	9.81	9.81	9.81
Mean of dependent variable	25.54	25.54	25.54	25.54
Observations	2,157	2,157	2,157	2,157
R^2	0.35	0.36	0.36	0.35

Notes: The dependent variable is an indicator (multiplied by 100) for whether the household accepted the hypothetical offer if required to complete the payment in six weeks. All specfications include the household and community covariates specified in the pre-analysis plan. Household covariates include the age of the household head, indicators for whether the household respondent attended secondary school, is a senior citizen, is not primarily a farmer, is employed, and has a bank account, an indicator for whether the household has high-quality walls, and the number of chickens (a measure of assets) owned by the household. Community covariates include indicators for the county, market status, whether the transformer was funded and installed early on (between 2008 and 2010), baseline electrification rate, and community population. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$.

Table A9C—Predictors of financial constraints in WTP questions

	(1)	(2)
$853 offer	90.32***	91.37***
	(5.24)	(5.53)
$398 offer / Existing fixed price	72.93***	75.86***
	(4.10)	(4.05)
$284 offer / T1: Low subsidy—29% discount	70.30***	72.20***
	(3.33)	(3.44)
$227 offer	65.92***	68.21***
	(3.72)	(3.80)
$171 offer / T2: Medium subsidy—57% discount	52.73***	55.05***
	(3.30)	(3.39)
$114 offer	52.89***	54.20***
	(3.35)	(3.40)
Number of household members		-0.05
		(0.53)
High-quality walls=1		-12.51***
		(3.27)
Number of chickens=1		-0.23*
		(0.14)
Age (years)		0.14
		(0.12)
Attended secondary school=1		0.09
		(3.07)
Over 65 years old=1		-3.53
		(5.18)
Not a farmer=1		0.27
		(3.16)
Has bank account=1		-10.74***
		(3.21)
Employed=1		0.37
		(2.91)
Mean of dependent variable	52.36	52.46
Observations	1,184	1,159
R^2	0.25	0.27

Notes: In both columns, the dependent variable is an indicator (multiplied by 100) for whether the household first accepted the hypothetical offer (i.e. randomly assigned price) to connect to the grid, and then declined the hypothetical offer if required to complete the payment in six weeks. Robust standard errors clustered at the community level in parentheses. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$.

Table A9D—Comparison of stated willingness to pay and experimental curves

	WTP1	WTP2
	(1)	(2)
\$284 offer / T1: Low subsidy—29% discount	5.80***	5.80***
	(1.37)	(1.37)
\$171 offer / T2: Medium subsidy—57% discount	22.44***	22.44***
	(4.01)	(4.01)
Free offer / T3: High subsidy—100% discount	94.22***	94.22***
	(1.17)	(1.17)
WTP indicator	34.94***	8.50***
	(2.73)	(1.71)
\$284 × WTP indicator	9.10**	-0.42
	(3.69)	(2.83)
\$171 × WTP indicator	-1.09	-5.03
	(5.28)	(4.69)
Free offer × WTP indicator	-34.42***	-7.99***
	(3.31)	(2.48)
Mean of dependent variable	34.22	24.46
Observations	3,635	3,635
R^2	0.50	0.57
F-statistic (test for equality between WTP and experimental results)	103.22	10.54

Notes: The dependent variable is an indicator (multiplied by 100) for whether the household accepted the hypothetical (i.e., willingness to pay) or experimental offer. Columns 1 and 2 pool the results of the first hypothetical offer (i.e., without time constraints) with the experimental results, and the second hypothetical offer (i.e., with time constraints) with the experimental results, respectively. WTP indicates whether the observation is generated by the stated willingness to pay portion of the experiment. Robust standard errors clustered at the community level in parentheses. Asterisks indicate coefficient statistical significance level (2-tailed): * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$. *F*-statistics correspond to the test that $\beta_{\text{WTP}} = \beta_{\text{WTP}} + \beta_{284\times\text{WTP}} = \beta_{\text{WTP}} + \beta_{171\times\text{WTP}} = \beta_{\text{WTP}} + \beta_{\text{Free}\times\text{WTP}} = 0$.

Table A10—Summary of randomly-assigned, hypothetical credit offers

Offer	Months	Upfront	Monthly	NPV at discount rate of 5%	15%	25%	n	Take-up: Time unlimited	6 week deadline
1	36	79.60	11.84	475.23	425.67	387.38	406	50.6%	38.3%
2	36	59.70	12.58	480.03	427.38	386.69	379	53.5%	38.9%
3	36	39.80	13.32	484.83	429.09	386.01	369	52.7%	39.6%
4	36	59.70	13.45	509.29	452.98	409.46	353	49.7%	39.1%
5	24	59.70	17.22	452.57	418.07	389.91	419	52.4%	40.2%
6	36	127.93	26.94	1028.26	915.48	828.34	363	52.7%	28.2%
Offer 1 to 5 (average)		59.70	13.68	480.39	430.64	391.89		52.0%	39.3%

Notes: During the baseline survey, each household was randomly assigned a hypothetical credit offer consisting of an upfront payment (ranging from \$39.80 to \$79.60), a monthly payment (ranging from \$11.84 to \$17.22), and a contract length (either 24 or 36 months). Respondents were first asked whether they would accept the offer, and then asked whether they would still accept if required to complete the upfront payment in six weeks. Figure 9, Panels B and C plot the net present value and take-up results corresponding to offer 6 and the average for offers 1 to 5 (which are very similar), assuming a discount rate of 15 percent.

Table A11—Summary of selected historical national rural electrification initiatives

			Electrification rate					
			National		Rural		GDP per capita	
Country	Period	Government authority	Start	End	Start	End	Start	End
Kenya	2008 - present	Rural Electrification Authority	<20%		<5%		939	
China (I)	1949 - 1978	Maoist era of central planning	n/a	63%	n/a	53%	n/a	292
China (II)	1978 - 1998	Ministry of Water and Power, Rural Electricity Department	63%	98%	53%	97%	292	1,512
Vietnam	1975 - 2009	Vietnam Electricity (EVN)	10%	96%	3%	95%	376	1,235
Tunisia	1972 - 2001	National Rural Electrification Commission (CNER)	37%	95%	6%	88%	1,524	3,092
South Africa	1995 - 2001	SA National Electrification Programme (NEP)	30%	66%	21%	49%	5,321	5,811
USA (I)	1935 - 1940	Rural Electrification Administration	68%	79%	13%	33%	9,102	11,847
USA (II)	1940- 1955	Rural Electrification Administration	79%	98%	33%	94%	11,847	19,974

Notes: There is very little reliable data available from China in 1949. At that point, rural electrification rates were likely to be < 1 percent.

Note A1—Community selection process

In August 2013, local representatives of REA provided us with a master list of 241 unique REA projects, consisting of roughly 370 individual transformers spread across the ten constituencies of Busia and Siaya. Each project featured the electrification of a major public facility (market, secondary school, or health clinic), and involved a different combination of high and low voltage lines and transformers. Projects that were either too recent, or classified as not commissioned, were not included in this master list. Since the primary objective was to estimate local electrification rates, projects that were funded after February 2013 were excluded to ensure that households in sample communities had had ample opportunity to connect to the grid.

In September 2013, we randomly selected 150 transformers using the following procedure: 1) in each constituency, individual transformers were listed in a random order, 2) the transformer with the highest ranking in each constituency was then selected into the study, and 3) any remaining transformers located less than 1.6 km (or 1 mile) from, or belonging to the same REA project, as one of the selected transformers, were then dropped from the remaining list. We repeated this procedure, cycling through all ten constituencies, until we were left with a sample of 150 transformers for which: 1) the distance between any two transformers was at least 1.6 km, and 2) each transformer represented a unique REA project. In the final sample, there are 85 and 65 transformers in Busia and Siaya counties, respectively, with the number of transformers in each of the ten constituencies ranging from 8 to 23. This variation can be attributed to differences across constituencies in the number of eligible projects. In Budalangi constituency, for example, all of the 8 eligible projects were included in the sample. As a result of this community selection procedure, the sample is broadly representative of the types of rural communities targeted by REA in rural Western Kenya.

Note A2—Experimental design notes

Sampling—Households were identified at the level of the residential compound, which is a unit known locally as a boma. In Western Kenya, it is common for related families to live in different households within the same compound.

Surveying—The majority of the baseline surveys were conducted between February and May 2014. However, 3.1 percent of surveys were administered between June and August 2014 due to scheduling conflicts and delays.

Sample of connected households—Since electrification rates were so low, the sample of connected households covers only 102 transformer communities; 17 communities did not have any connected households at the time of census, and we were unable to enroll any connected households in the remaining 31 communities, for instance, if there was a single connected compound in a village and the residents were not present on the day of the baseline survey.

Randomization—For the stratification variable market status, we used a binary variable indicating whether the total number of businesses in the community was strictly greater than the community-level mean across the entire sample.

Note A3—Unexpected delays during the wiring phase

Kenya Revenue Authority (KRA) identification certificates—Households applying to Kenya Power are required to have (1) a National Identity Card (NIC), (2) a KRA Personal Identification Number (PIN) certificate, and (3) a completed Kenya Power application form. 42.0 percent of household heads requesting a connection did not already have a KRA PIN certificate, which could only be generated on the KRA website. Since most rural households do not regularly access the Internet, project enumerators provided registration assistance for 96.6 percent of the households lacking KRA PINs. At the time of the experiment, KRA PIN registration services were typically offered at local Internet cafes at a cost of $5.69 (500 KES).

Spelling mistakes on wiring certificates—Households connecting to the grid are required to have certificates that the wiring is safe. The ready-board manufacturer provided wiring certificates that needed to be signed by contractors after installation. We encountered delays when the spelling of the name on the certificate did not precisely match its spelling on the NIC or KRA PIN certificate.

Pre-analysis plan

"The demand for and costs of supplying grid connections in Kenya"

AEA RCT Title: "Evaluation of Mass Electricity Connections in Kenya"

RCT ID: AEARCTR-0000350

Principal Investigators: Eric Brewer, Kenneth Lee, Edward Miguel, and Catherine Wolfram

Date: 30 July 2014

Summary: This document outlines the plan for analyzing the demand for and costs of supplying household electricity connections in rural Kenya. The proposed analysis will take advantage of a field experiment in which randomly selected clusters of rural households were offered an opportunity to connect to the national grid at subsidized prices. This pre-analysis plan outlines the regression specifications, outcome variables, and covariates that will be considered as part of this analysis. We anticipate that we will carry out additional analyses beyond those included in this plan. This document is therefore not meant to be comprehensive. The overall research project will also include an impact evaluation of electricity connections that will be carried out in 2015 or 2016, upon completion of the endline survey round. For this portion of the project, we will register an additional pre-analysis plan at a later date, in either 2015 or 2016.

I. Introduction

Electrification has long been a benchmark of development, yet over two-thirds of the population of Sub-Saharan Africa lives without access to electricity. In June 2013, President Obama announced the Power Africa initiative, making energy access a top priority among six partner countries in Africa, including Kenya. In light of this initiative, and others being implemented by the World Bank and the UN General Assembly, there is considerable need for rigorous research to inform the effective scale-up of energy access programs in developing countries.

In this project, we have identified a unique opportunity to increase access to on-grid energy in Kenya. Since 2007, Kenya's Rural Electrification Authority (REA) has rapidly expanded the national grid, installing electricity distribution lines and transformers across many of the country's rural areas. Connectivity, however, remains low. While roughly three-quarters of the population is believed to live within 1.2 kilometers of a low voltage line, the official electrification rate is under 30%. In related work, we find that in regions that are technically covered by the grid, half of the unconnected households are no more than 200 meters from a low-voltage line.

We believe that the primary barrier to connecting these "under grid" households is the prohibitively high connection fee faced by rural households. The current connection price of KSh 35,000 ($412) may not be affordable for poor, rural households in a country where the GNI per capita (PPP) is $1,730. Despite this fact, Kenya's monopoly distribution company, Kenya Power, has recently proposed increasing the price to KSh 75,000 due to cost considerations.[1]

In general, little is known about the demand for electricity in rural areas, both initially and over time. Specifically, how many more households would opt to connect if the fee were,

[1] In March 2014, Kenya Power, the national utility, stated that it will continue to charge eligible customers KSh 35,000 for single-phase power connections, as long as the cost of connection does not exceed KSh 135,000 ($1,588), inclusive of VAT.

for example, KSh 25,000 ($294), KSh 15,000 ($176), or even KSh 0? How much power would households consume if they did connect, now and in the future? And once households are connected, do the social and economic benefits of access to modern energy in rural areas outweigh the costs?

In the coming years, REA will explore the feasibility of initiating a long-term, last-mile household connection program involving discounted connection fees for households and small businesses located close to existing REA electricity transformers. In order to evaluate this potential program, we have partnered with REA to conduct a randomized evaluation of grid connections involving roughly 2,500 households in rural Western Kenya.

The principal objectives of this study are twofold:

1. To trace out the demand curve for electricity connections, and in addition, to estimate the economies of scale in costs associated with spatially grouping connections together.
2. To measure the social and economic impacts of electrification, including schooling outcomes for children, energy use, income and employment, among other outcomes.

This pre-analysis plan outlines our strategy to address the first objective. The analysis on the impacts of the intervention will be carried out in 2015 and 2016, upon completion of the midline and endline survey rounds. The pre-analysis plan for the second stage of this project will therefore be registered at a later date, in either 2015 or 2016.

The remainder of this document is organized as follows. Section II provides a brief background on the existing literature on the demand for electricity connections. Section III provides a brief overview of the experimental design. Finally, Sections IV and V outline the main estimating equations that will be used in our analysis of both the demand for and costs of supplying electricity connections.

II. Brief literature review

In recent years, there has been a growing literature examining the demand for electricity connections in developing countries. The methods utilized in these studies range from contingent valuation approaches (see, e.g., Abdullah and Jeanty 2011) to randomized encouragement designs, where households are offered vouchers or subsidies to connect to the electricity network at a discounted price. Bernard and Torero (2013), for example, distribute two levels of randomized vouchers (10% and 20% discounts) to encourage household grid connections in Ethiopia, where the connection price ranges from $50 to $100, depending on the household's distance to the nearest electrical pole. Similarly, Barron and Torero (2014) utilize two levels of randomized vouchers (20% and 50% discounts) in El Salvador, where the connection price (in the study setting) is $100.

There is also an engineering literature simulating the costs of extending the grid to rural areas in developing countries. Parshall et al. (2009), for example, apply a spatial electricity planning model to Kenya and find that "under most geographic conditions, extension of the national grid is less costly than off-grid options." Zvoleff et al. (2009) examine the costs associated with extending the grid across various types of settlement patterns, demonstrating the potential for non-linearities in costs.

While our study is closely related to the earlier randomized encouragement designs, our objective is to evaluate the demand for electricity connections at randomized prices, as well as provide experimental evidence on the cost economies of scale associated with grouping connections together spatially.

III. Overview of project

1. Experimental design

Our experiment takes place across 150 "transformer communities" in Western Kenya. Each transformer community is defined as the group of all households located within 600 meters of a central electricity distribution transformer. In Kenya, all households within

600 meters of a transformer are eligible to apply for an electricity connection. In each transformer community, we have enrolled roughly 15 randomly selected unconnected households. In total, our study will involve roughly 2,250 unconnected households.

On 23 April 2014, our sample of transformer communities was randomly divided into treatment and control groups of equal size (75 treatment, 75 control). Each of the 75 treatment communities were then randomly assigned to one of three treatment arms (i.e. subsidy groups). These subsidies were designed to allow households to connect to the national power grid at relatively low prices (compared to the current connection price of KSh 35,000 or $412). In addition, each household accepting an offer to be connected as part of the study would receive a basic household wiring solution ("ready-board") at no additional cost. Each ready-board provides a single light bulb socket, two power outlets, and two miniature circuit breakers (MCBs).

The treatment and control groups are characterized as follows:

A. High-value treatment arm

25 communities. KSh 35,000 ($412) subsidy and KSh 0 ($0) effective price. This represents a 100% discount on the current price.

B. Medium-value treatment arm:

25 communities. KSh 20,000 ($235) subsidy and KSh 15,000 ($176) effective price. This represents a 57% discount on the current price.

C. Low-value treatment arm:

25 communities. KSh 10,000 ($118) subsidy and KSh 25,000 ($294) effective price. This represents a 29% discount on the current price.

D. Control group:

75 communities. No subsidy and KSh 35,000 ($412) effective price. There is no discount offered to households in the control group.

Within each treatment community, all enrolled and unconnected households would receive the same subsidy offer. After receiving the subsidy offer, treatment households would be given eight weeks to accept the offer and deliver the required payment to REA. At the end of this eight-week period, field enumerators would visit each household to verify that the required payment has been made to REA. Electricity connections are delivered once these verifications are complete. The collection of take-up responses comprises the main data set for the analyses outlined in this pre-analysis plan.

Once payments are verified, REA would hire its own contractors to deliver the connections within a period of four to six weeks. In order to economize on its own delivery costs, REA would connect all of the required connections in each community at the same time. REA would also group anywhere from two to four neighboring communities together, in order to further economize on transportation costs.

The first set of randomized offers were delivered in early-May and expired in early-July. The second set of randomized offers will be delivered in late-July and will expire in late-September. Our field enumerators began collecting take-up data on 4 July 2014. The full round of data collection will continue through the end of October 2014. As a result, it is expected that the final version of the data set for this analysis will be available in November 2014.

Data collection began before this document was uploaded to the AEA RCT registry website. In anticipation of this delay, we posted a document to our registered trial on 2 July 2014 titled "A note on pre-analysis plans" in order to describe how the investigators would be prohibited from accessing any data until a pre-analysis plan had been uploaded to the registry website.

2. Power calculations

At the beginning of this project, we knew little about the demand for electricity connections at various prices. We therefore made a set of assumptions on how take-up would vary at four different levels of prices. Taking into account our budgetary constraints, we designed the study to detect differences in take-up at these pricing levels, based on our set of ex-ante assumptions. In addition, we took into consideration the level of take-up that we would need in our future analysis on the social and economic impacts of electrification. These assumptions are outlined in Table 1.

Table 1: Ex-ante take-up assumptions

	Communities	Households (n)	Assumed take-up range
A. High-value arm ("High")	25	375	90 - 95%
B. Medium-value arm ("Medium")	25	375	40 - 50%
C. Low-value arm ("Low")	25	375	15 - 25%
D. Control group ("Control")	75	1,125	0 - 5%
Total	150	2,250	

Table 2: Communities required in each arm to detect differences with 80% power

Comparison	Description	Required size of each arm	Actual size of each arm
A vs. B	High vs. Med.	3 - 5	25
A vs. C	High vs. Low	2	25
A vs. D	High vs. Control	1 - 2	25 (High), 75 (Control)
B vs. C	Med. vs. Low	6 - 27	25
B vs. D	Med. vs. Control	3 - 5	25 (Med), 75 (Control)
C vs. D	Low vs. Control	6 - 26	25 (Low), 75 (Control)

In Table 2, we report the total number of communities required to detect differences ($\alpha = 0.05$) between groups with 80% power. For example, in the comparison of groups B (medium-value treatment arm) and C (low-value treatment arm), we expect that we will need 6 to 27 communities in each treatment arm (the actual size of each arm is 25 communities).[2] We assume an intracluster correlation coefficient of 0.1 within communities. In our design, we included a large number of high-value treatment communities in order to increase our statistical power to estimate the social and economic impacts of electrification (our second objective). Based on these assumptions, we expect that we are

[2]Since we had assumed a range of values for our assumptions on take-up, we report a range of values for the required size of each arm. For example, if take-up is 50% and 15% for groups B and C, respectively, we would require only 6 communities in each arm. However, if take-up is 40% and 25% for groups B and C, respectively, we would require 27 communities.

sufficiently powered, based on our ex-ante assumptions on take-up.

3. Data

This analysis will utilize four data sets: (1) Data on household take-up decisions; (2) Data on actual costs of supplying household connections; (3) Data on community-level characteristics; and (4) Household-level baseline survey data from the Living Standards Kenya (LSK) survey. The survey instrument is included in the Appendix.

IV. Analysis plan - Demand

The primary objective of this analysis is to estimate the demand for electricity connections, or in other words, the willingness of individual households to pay for a quoted price of an electricity connection. We will follow the procedure: (1) Estimate a non-parametric regression of household take-up on various subsidy levels. (2) Test for linearity: If we cannot reject linearity, we will estimate a linear regression of take-up on the effective connection price. If we can reject linearity, we will focus on the non-parametric estimation for the remainder of the analysis. (3) Estimate heterogeneous effects. (4) Plot the demand curve and compare these results to our contingent valuation results.

1. Non-parametric regression

We will begin by estimating the main equation:

$$y_{ic} = \alpha_0 + \alpha_1 T_c^{low} + \alpha_2 T_c^{mid} + \alpha_3 T_c^{high} + X_c' \gamma + \epsilon_{ic} \qquad (1)$$

where y_{ic} is a binary variable reflecting the take-up decision for household i in transformer community c.[3] The binary variables T_c^{low}, T_c^{mid}, and T_c^{high} indicate whether community c was randomly assigned into the low-value, medium-value, or high-value treatment arms, respectively. Following Bruhn and McKenzie (2009), we include a vector of community-level characteristics, X_c, containing the variables used for stratification dur-

[3]Refer to Section IV Part 3 for further details on the dependent variable.

ing randomization.[4] Standard errors will be clustered at the community level.

Equation (1) will be the primary equation that we estimate in our demand-side analysis. As a robustness check, we will also estimate the equation:

$$y_{ic} = \alpha_0 + \alpha_1 T_c^{low} + \alpha_2 T_c^{mid} + \alpha_3 T_c^{high} + X_c' \gamma + X_{ic}' \lambda + \epsilon_{ic} \tag{2}$$

where X_{ic} is a vector of household-level characteristics.[5] X_{ic} will include standard control variables that not only have predictive effects but may also serve as sources of heterogeneity in take-up.

We will also assess whether treatment and control households are balanced at baseline in terms of household characteristics. In addition to X_{ic}, we may also choose to control for any covariates that are both unbalanced at baseline and relevant for electricity take-up.

In equations (1) and (2), the baseline (i.e. $T_c^{low} = T_c^{mid} = T_c^{high} = 0$) estimates household take-up under the status-quo pricing policy (i.e. take-up when the price of an electricity connection faced by the rural household is KSh 35,000). α_1, α_2, and α_3 capture the incremental effects (over the baseline) on take-up of the low-value, medium-value and high-value subsidies, respectively. Since the randomized subsidies will lower the effective price of an electricity connection, we expect that our experiment will result in positive and statistically significant α-coefficients.

2. Testing for linearity

We are interested in testing for linearity in equation (1). We will use an F-test to assess the null hypothesis:

$$H_0\text{:}\ \frac{(\alpha_3 - \alpha_2)}{15} = \frac{(\alpha_2 - \alpha_1)}{10} = \frac{(\alpha_1 - \alpha_0)}{10}$$

[4] Refer to Section IV Part 4 for further details on the components of X_c.

[5] Refer to Section IV Part 4 for further details on the components of X_{ic}.

against the alternative hypothesis that the slope in between the various take-up points is unequal. If we cannot reject linearity in an F-test, we will also estimate the equation:

$$y_{ic} = \beta_0 + \beta_1 p_c + X_c' \gamma + \epsilon_{ic} \tag{3}$$

where p_c is the effective price of an electricity connection faced by households in community c.[6]. Standard errors will again be clustered at the community level. As in equation (2), we will similarly check robustness by including the vector X_{ic}.

If we can reject linearity in an F-test, it will be of interest to understand how take-up changes when moving across different subsidy levels. In a similar experiment conducted in El Salvador, Barron and Torero (2014) find that the effects of a relatively low subsidy (20%) and a relatively high subsidy (50%) are similar. This is taken to suggest that either the demand for connections is inelastic (in the price range offered), or that the subsidies affect take-up through alternative channels.[7] Given this unusual result, we will focus on equation (1) and test the hypothesis that:

$$H_0\colon \alpha_1 = \alpha_2$$

against the alternative that the higher-value subsidy has a larger effect on take-up compared to the lower-value subsidy (i.e. $H_1\colon \alpha_2 > \alpha_1$). We will conduct a similar test for each of the pairwise combinations listed in Table 2.

3. Two measures of take-up

We may find that some of the treatment households decided that they would like to accept the offer, but are unable to complete the full payment within the eight-week period. We may therefore have two measures of take-up:

[6]For example, in a high-subsidy treatment community, the subsidy amount is equal to the current price of an electricity connection and the effective price faced by households is 0 KSh (i.e. $p_c = 0$)

[7]For example, Barron and Torero propose that a subsidy may raise awareness that electrification is possible, resulting in higher take-up.

1. Actual take-up (y^1_{ic}): Binary variable indicating whether treatment household ic accepted the offer and completed the required payment within eight weeks.

2. Intended take-up (y^2_{ic}): Binary variable indicating whether treatment household ic intended to accept the offer, and began to make payments, but was unable to complete the full payment within eight weeks.

Our primary outcome of interest, however, will be the actual take-up captured by y^1_{ic}.

4. Covariate vectors X_c and X_{ic}

There are two sets of covariates in equations (1), (2), and (3). X_c is a vector of community-level characteristics and X_{ic}, which will mainly be used in robustness checks, is a vector of household-level characteristics. X_c will primarily include the stratification variables that were used during randomization.[8] The list of X_c variables will include:

1. County indicator: Binary variable indicating whether community c is in Busia or Siaya. This was used as a stratification variable during randomization.

2. Market status: Binary variable indicating whether the total number of businesses in community c is strictly greater than the community-level mean across the entire sample. We use this definition to define which communities could be classified as "markets" relative to the others. This was used as a stratification variable during randomization.

3. Transformer funding year: Binary variable indicating whether the electricity transformer in community c was funded "early" (i.e. in either 2008-09 or 2009-10). This was used as a stratification variable during randomization.

4. Electrification rate: Residential electrification rate in community c.

5. Community population: Estimated number of people living in community c.

X_{ic} will include a set of household-level variables that not only have predictive effects but may also serve as sources of heterogeneity in take-up. The survey from which we will obtain this data is attached in the Appendix. For example, it is possible that take-up will vary depending on household size, household wealth, or the education level and employment type of the survey respondent. In the majority of cases, the survey respondent

[8]The collection of this data is described in further detail in Lee et al. (2014).

is either the household head or the spouse of the household head. The list of X_{ic} variables will include (*LSK question numbers in parentheses*):

1. Household size ($a1$): Number of people living in household ic.
2. Household wealth indicator - Walling material ($c1c$): Binary variable indicating whether the walls of household ic can be considered "high quality" (i.e. made of brick, cement, or stone).
3. Household wealth indicator - Chickens ($d9a$): Number of chickens owned by household ic.
4. Age of respondent in years ($a4c$)
5. Education of respondent ($a5b$): Binary variable indicating whether respondent ic has completed some level of secondary education.
6. Farming as primary occupation of respondent ($a5c$): Binary variable indicating whether the primary occupation of respondent ic is farming.
7. Access to financial services of respondent ($g1a$): Binary variable indicating whether respondent ic uses a bank account.
8. Business or self employment activity of respondent ($e1$): Binary variable indicating whether the respondent (or the respondent's spouse) in household ic engages in any business or self-employment activities.
9. Senior household ($a4c$): Binary variable indicating whether respondent ic is over 65 years old.

5. Heterogeneous effects

We are interested in understanding how take-up varies across several important socio-economic dimensions. For example, will take-up depend on community characteristics? Will it be higher for households that are located in more electrified communities or in market centers? Alternatively, will take-up depend on individual characteristics? Will it be higher for the more educated households, or those that are engaged in more "entrepreneurial activities"? In order to answer these questions, we will estimate heterogeneous effects along a number of dimensions, captured in the vectors X_c and W_{ic} (which is a subset of X_{ic}):

1. County indicator (X_c)
2. Market status (X_c)
3. Transformer funding year (X_c)
4. Electrification rate (X_c)
5. Community population (X_c)
6. Household wealth indicator - Walls (W_{ic})
7. Education of respondent (W_{ic})
8. Farming as primary occupation of respondent (W_{ic})
9. Access to financial services of respondent (W_{ic})
10. Business or self employment activity of respondent (W_{ic})
11. Senior household (W_{ic})

We will estimate heterogeneous effects by adding interactions between the treatment variables and the vectors X_c and W_{ic} to equations (1), (2), and (3). We will also carry out additional analyses, depending on the types of heterogeneous effects that we estimate. For example, if we find that take-up is higher in communities with higher electrification rates, we may explore whether there are any "bandwagon" effects, as in Bernard and Torero (2013), by focusing on the interaction between the treatment and community electrification variables. Since we do not know the nature of these heterogeneous treatment effects, it is not possible to fully specify all of the potential analyses in this document.

6. Comparison of contingent valuation to revealed preference results

During the LSK survey round, conducted between February and July 2014, we asked respondents from unconnected households whether they would be hypothetically willing to connect to the national grid at a randomly selected price (see questions $f16b$ and $f16c$ in Appendix). These amounts were randomly drawn from the following set of prices:

$$\text{Hypothetical Price} \in \{0, 10000, 15000, 20000, 25000, 35000, 75000\}$$

This question was followed by an additional hypothetical question asking the respondent whether they would accept an offer at this price if they were given six weeks to complete the payment.[9]

In comparison, there were four effective prices (randomized at the community-level) in our experimental design:

$$\text{Effective Price} \in \{0, 15000, 25000, 35000\}$$

By making comparisons between these two measures of take-up at similar levels of prices, we will test whether we could reject equal demand (in terms of contingent valuation and revealed preferences). In addition, we will plot various demand curves, with take-up plotted along one axis and the effective (or hypothetical) price plotted along the other. Finally, we will run contingent valuation regressions using the same specifications and covariates as those described in Section IV, Parts 1, 2, and 6.

V. Analysis plan - Costs

The secondary objective of this analysis is to characterize how connection costs decrease with the number of neighboring households that choose to connect at the same time.[10]

<u>1. Potential for economies of scale in costs</u>

Given that rural households are often located in remote areas, the cost of supplying an electricity connection to an individual household can be very high. This is due to the high cost of transportation and the necessity of building additional low-voltage lines. However, significant economies of scale could be achieved by connecting multiple households

[9] In our experimental design, treatment households were given eight weeks to complete the payment. This change was made at the request of REA, after we had already launched our baseline survey round. In this hypothetical question, we do not believe that providing an additional two weeks would have influenced the responses.

[10] We make a distinction between the *price* of an electricity connection, which is the fixed price of an electricity connection faced by households, and the *cost* of an electricity connection, which is the physical cost of supplying the electricity connection faced by the utilities.

at the same time. In a related paper, we use the current costs of materials to estimate that the incremental cost of supplying an electricity connection to a single household 200 and 100 meters away from a low-voltage line is $1,940 and $1,058, respectively, inclusive of material and transportation costs, as well as a 25% contractor markup (Lee et al. 2014).

While this cost is extremely high, it is desirable from the perspective of the supplier to connect spatially-clustered groups of households at the same time. For example, when two neighboring households are connected along the same length of line, the above per household costs are projected to fall by roughly 47%, to $1,021 and $580, respectively.

2. IV approach to estimating economies of scale in costs

In our experimental design, randomized subsidies are assigned at the community level. In addition, there are three levels of subsidies. We expect that different levels of subsidies—low, medium, and high—will create variation in the number of households that choose to apply for electricity at the same time. For example, larger numbers of applicants should be observed in the high-subsidy communities (where households pay 0 KSh), and smaller numbers of applicants should be observed in the low-subsidy communities (where households pay 25,000 KSh).

We can therefore estimate the community-level construction cost, Γ_c, as a function of the number of connected households in the community, M_c, using the randomized community-level subsidy amounts, Z_c^{low}, Z_c^{mid}, and Z_c^{high}, as instruments for M_c.[11] In order to allow for the possibility of non-linearities in costs, we will include higher-order polynomials in our estimation of Γ_c. Specifically, we will estimate an instrumental variables regression using the equations:

$$M_c = \delta_0 + \delta_1 Z_c^{low} + \delta_2 Z_c^{mid} + \delta_3 Z_c^{high} + V_c' \mu + \nu_c \tag{4}$$

[11] Refer to Section V Part 3 for additional information on how we plan to construct the variable Γ_c.

$$M_c^2 = \delta_0 + \delta_1 Z_c^{low} + \delta_2 Z_c^{mid} + \delta_3 Z_c^{high} + V_c' \mu + \nu_c \quad (5)$$

$$M_c^3 = \delta_0 + \delta_1 Z_c^{low} + \delta_2 Z_c^{mid} + \delta_3 Z_c^{high} + V_c' \mu + \nu_c \quad (6)$$

$$\Gamma_c = \pi_0 + \pi_1 M_c + \pi_2 M_c^2 + \pi_3 M_c^3 + V_c' \mu + \eta_c \quad (7)$$

where the first-stage equations (4), (5), and (6) estimate the effects of the treatment variables on the number of applicants, and the second-stage equation (7) estimates the effect of higher-order polynomials of the number of connected households on the community-level cost. Since there are multiple endogeneous variables in this framework, equations (4), (5), and (6) will be estimated jointly. V_c is a vector of community-level characteristics that will be relevant in this regression.[12] ν_c and η_c are error terms.

We will take the derivative of our estimates in equation (7) in order to uncover different points along the marginal cost curve. We will plot these points to sketch out a marginal cost curve, with the number of connected households on the horizontal axis and the marginal cost on the vertical axis. We will also expand equations (4) through (7) by interacting the Z_c and M_c variables with the V_c vector to explore any potential heterogeneous effects.

We should note that this analysis is highly speculative. We have not carried out any power calculations because we do not have baseline data on the community-level costs of household electrification. Furthermore, our ability to identify the desired effects will depend on the specified functional forms. If we estimate linear relationships in both stages, we will focus only on estimating equation (4) in the first-stage and substitute equation (7) with the equation:

$$\Gamma_c = \pi_0 + \pi_1 M_c + V_c' \mu + \eta_c \quad (8)$$

[12] Refer to Section V Part 4 for further details on the components of V_c.

In addition, we may pursue additional analyses, depending on the nature of the cost data that we eventually receive.

3. Constructing the variable Γ_c

Through our partnership with REA, we will collect actual cost invoices related to the connections that are delivered as a part of this study. Specifically, we will be provided with an itemized list of costs (e.g. cost of low-voltage lines, cost of service lines, cost of transportation etc.), as well as the design drawings detailing the planned locations of electricity poles. Using these data, we will work with REA to determine the total construction cost for each community.

4. Covariate vector V_c

V_c will include variables that should have an impact on construction costs, including all of the community-level variables in X_c, in addition to a community distance and land gradient variables. The list of V_c variables will include:

1. County indicator
2. Market status: This may approximate community density or the pre-existing coverage of the local low-voltage network.
3. Transformer funding year
4. Electrification rate: This should approximate the pre-existing coverage of the local low-voltage network. Higher electrification rates (and more local low-voltage network coverage) should decrease construction costs.
5. Community population
6. Distance from REA warehouse: Travel distance (in kilometers) between community c and the primary REA warehouse located in Kisumu where the construction materials are stored. Longer travel distances should increase construction costs.
7. Terrain or land gradient: We will use two different measures of terrain or land gradient. Dinkelman (2011) identifies land gradient as a major factor contributing to the costs of electrification. In flatter areas, the soil tends to be softer, making it cheaper to lay power lines and erect transmission poles. Our primary community-level land gradient variable will therefore be constructed using the same methodology as Dinkelman (2011). Specifically, we will use the 90-meter Shuttle Radar Topog-

raphy Mission (SRTM) Global Digital Elevation Model (available at www.landcover.org) to access elevation data and then construct measures of the average land gradient for each transformer community.[13] Our secondary community-level land gradient variable will be the variance in the distribution of altitudes collected across the entire population of geo-tagged buildings for each transformer community.[14]

References

Abdullah, Sabah and P. Wilner Jeanty. 2011. Willingness to pay for renewable energy: Evidence from a contingent valuation survey in Kenya. *Renewable and Sustainable Energy Reviews* 15: 2974-2983.

Barron, Manuel and Maximo Torero. 2014. Short Term Effects of Household Electrification: Experimental Evidence from Northern El Salvador.

Bernard, Tanguy and Maximo Torero. 2013. Bandwagon Effects in Poor Communities: Experimental Evidence from a Rural Electrification Program in Ethiopia.

Bruhn, Miriam and David McKenzie. 2009. In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics* 1(4): 200-232.

Dinkelman, Taryn. 2011. The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review* 101(December 2011): 3078-3108.

Lee, Kenneth, Eric Brewer, Carson Christiano, Francis Meyo, Matthew Podolsky, Javier Rosa, Catherine Wolfram, and Edward Miguel. 2014. Barriers to Electrification for "Under Grid" Households in Rural Kenya. NBER Working Paper 20327. National Bureau of Economic Research, Cambridge, MA. http://www.nber.org/papers/w20327.

Parshall, Lily, Dana Pillai, Shashank Mohan, Aly Sanoh, and Vijay Modi. 2009. National electricity planning in settings with low pre-existing grid coverage: Development of a spatial model and case study of Kenya. *Energy Policy* 37: 2395-2410.

Zvoleff, Alex, Ayse Selin Kocaman, Woonghee Tim Huh, and Vijay Modi. 2009. The impact of geography on energy infrastructure costs. *Energy Policy* 37(10): 4066-4078.

[13] Each transformer community is defined as all of the buildings within a 600 meter radius of a central electricity distribution transformer, as defined in Lee et al. (2014).

[14] Usage of this secondary definition of land gradient will depend on whether we can verify that our altitude records (taken using the GPS application on Android tablets) are relatively accurate.