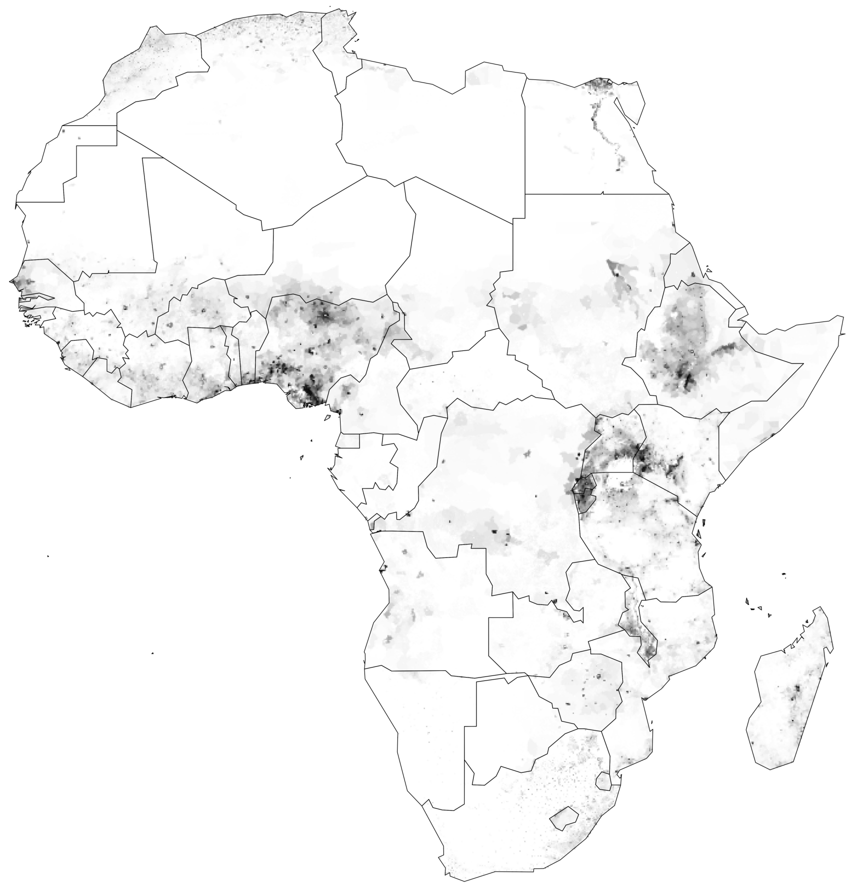
Supplementary Appendix for Online Publication

Does Household Electrification Supercharge Economic Development?

Kenneth Lee, Edward Miguel, and Catherine Wolfram

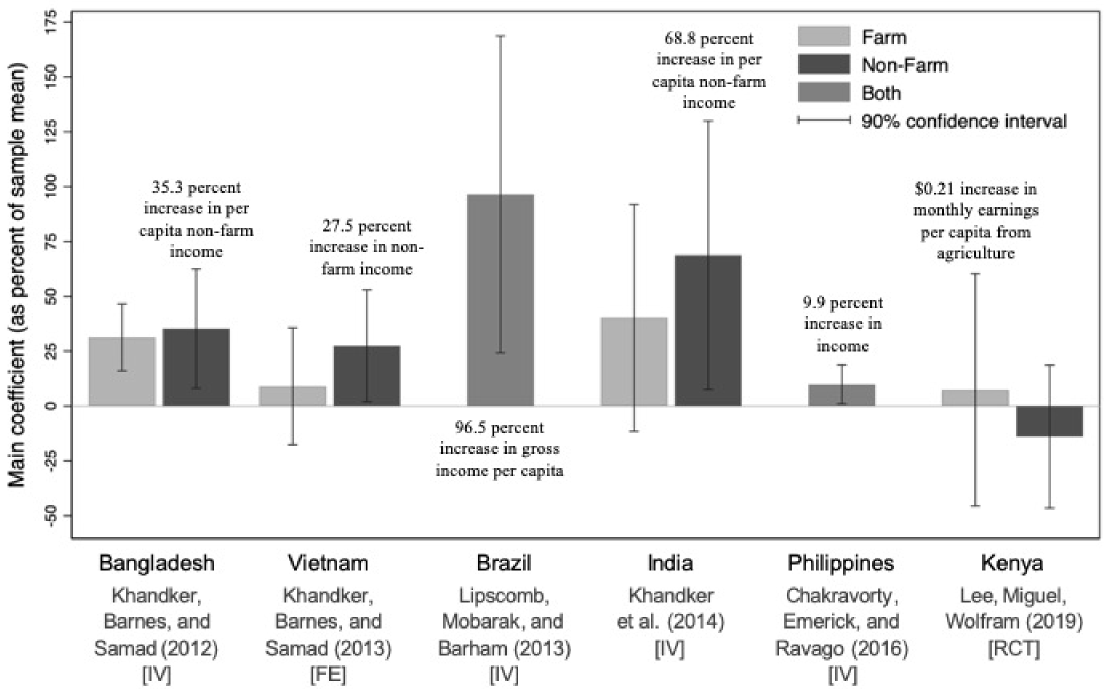
November 2019

Appendix Figure 1: Distribution of "missing" night lights in Africa



Notes: We combine data from satellite images of night lights in 2013, obtained from the National Oceanic and Atmospheric Administration's Defense Meteorological Satellite Program–Operational Line Scan (NOAA DMSP-OLS), with data from the Gridded Population of the World, Version 4 (GPWv4), provided by the Center for International Earth Science Information Network (CIESIN) at Columbia University, to predict where the largest gains in nighttime brightness would occur if everyone were able to enjoy the same levels of brightness as OECD countries. The simple procedure is as follows: (1) estimate $Log(N_{5km}) = \alpha + \beta Log(P_{5km})$ for grid cells in OECD countries; (2) using the estimated parameters, predict $\hat{N_{5km}}$ for grid cells in Africa; and (3) subtract N_{5km} from $\hat{N_{5km}}$ for grid cells in Africa to estimate "missing" night lights.

Appendix Figure 2: Key estimates of the impacts of rural electrification on income



Notes: For each study, coefficient estimates have been expressed as a percentage of the mean of the dependent variable.

Appendix Table 1: Local average treatment effects for different complier subgroups

	Control (1)	Adopter only when price is low (2)	Adopter when price is high (3)	p-value of diff. (4)
Share of sample (%)	100	67	22	
Panel A: Primary energy outcomes				
A1. Grid connected (%)	5.6	–	–	–
	(23.0)			
A2. Monthly electricity spending (USD)	0.14	2.00***	2.47***	0.28
	[0.91]	(0.21)	(0.31)	
Panel B: Additional energy outcomes				
B1. Electricity as main lighting (%)	5.2	86.8***	96.8***	0.01
	[22.2]	(2.4)	(2.6)	
B2. Number of appliance types owned	1.8	-0.2	1.5***	< 0.01
	[1.3]	(0.2)	(0.3)	
B3. Owns mobile phone (%)	84.3	-12.7***	18.5**	< 0.01
	[36.4]	(4.0)	(8.9)	
B4. Owns radio (%)	54.2	-5.1	23.5*	0.09
	[49.8]	(5.3)	(13.0)	
B5. Owns television (%)	17.9	-2.3	47.1***	< 0.01
	[38.4]	(4.6)	(10.7)	
B6. Owns iron (%)	4.1	-0.1	6.9	0.37
	[19.9]	(2.5)	(5.9)	
B7. Monthly kerosene spending (USD)	2.81	-1.21***	-1.66**	0.64
	[2.86]	(0.26)	(0.76)	
B8. Monthly total energy spending (USD)	11.66	4.62*	-16.64***	< 0.01
	[28.47]	(2.57)	(4.89)	
B9. Solar home system as main lighting (%)	11.8	-13.7***	-10.4	0.74
	[32.3]	(2.3)	(8.0)	

(Table continued on next page)

(Table continued from previous page)

	Control	Adopter only when price is low	Adopter when price is high	*p*-value of diff.
	(1)	(2)	(3)	(4)
Panel C: Primary economic outcomes				
C1. Household employed or own business (%)	36.8	1.1	14.6	0.31
	[38.8]	(4.2)	(10.3)	
—— Household women-only (%)	34.5	6.0	18.2	0.44
	[44.5]	(5.1)	(12.1)	
—— Household men-only (%)	40.2	-2.5	11.3	0.43
	[45.6]	(5.5)	(13.3)	
C3. Total hours worked last week	47.0	-2.2	-4.2	0.81
	[24.7]	(2.7)	(6.6)	
C4. Total asset value (USD)	914	3	630**	0.09
	[961]	(123)	(280)	
C5. Per capita cons. of major items (USD)	133	-3	-23	0.67
	[142]	(14)	(37)	
Panel D: Primary non-economic outcomes				
D1. Recent health symptoms index	0	-0.19*	0.29	0.14
	[1]	(0.10)	(0.26)	
D2. Normalized life satisfaction	0	0.12	0.23	0.76
	[1]	(0.11)	(0.28)	
D3. Avg. student test Z-score	0	-0.08	-0.11	0.92
	[1]	(0.12)	(0.24)	
D5. Political and social awareness index	0	-0.09	-0.03	0.86
	[1]	(0.10)	(0.26)	
Panel E: Mean treatment effects on grouped outcomes				
E1. Economic Index (C outcomes)	0	-0.03	0.30	0.31
	[1]	(0.11)	(0.26)	
E2. Non-Economic Index (D outcomes)	0	-0.10	0.06	0.64
	[1]	(0.10)	(0.26)	

Notes: Heterogeneous treatment effects estimated using only round 1 data (roughly 16 months post-connection). The following variables (reported in table 3, Lee, Miguel, and Wolfram 2019) were not collected in round 1: "C2. Monthly household earnings (USD)," "D4. Avg. student KCPE test Z-score," and "D6. Perceptions of security index." Column 1 reports mean values in the control group, with standard deviations in brackets. Using the established sample shares, columns 2 and 3 display weighted-average local average treatment effects for each adopter group. Robust standard errors, displayed in parentheses, are estimated using a stacked regression approach.

Appendix Table 2: Differences between adopter groups at baseline, when all households were unconnected

	Adopter only when price is low $(0 \leq p < \$171)$ (1)	Adopter when price is high $(171 \leq p \leq \$284)$ (2)	p-value of difference (3)
Panel A: Household head (respondent) characteristics			
Female (%)	60.5	62.6	0.77
Age (years)	53.4	53.4	> 0.99
Senior citizen (%)	2.08	26.2	0.78
Attended secondary schooling (%)	9.5	21.0	0.05
Married (%)	64.6	69.5	0.47
Not a farmer (%)	18.5	26.6	0.20
Employed (%)	34.1	41.9	0.28
Basic political awareness (%)	9.9	15.7	0.25
Has bank account (%)	14.7	32.4	< 0.01
Monthly earnings (USD)	11.55	24.39	0.10
Panel B: Household characteristics			
Number of members	5.1	6.0	0.02
Youth members (age $\leq$ 18)	2.9	3.4	0.17
High-quality walls (%)	12.6	21.3	0.16
Land (acres)	1.9	2.1	0.71
Distance to transformer (m)	363.2	356.9	0.75
Monthly (non-charcoal) energy (USD)	5.16	6.73	0.05
Panel C: Household assets			
Bednets	2.2	2.7	0.03
Sofa pieces	5.5	8.1	< 0.01
Chickens	5.9	9.3	0.01
Radios	0.4	0.4	0.42
Televisions	0.1	0.3	< 0.01
Share of sample (%)	67	22	

Notes: Columns 1 and 2 report sample means for "adopters when the price is low" and "adopters when the price is high," respectively, at the time of the baseline survey. Column 3 reports p-values of the difference between the means. The basic political awareness indicator captures whether the household head was able to correctly identify the presidents of Tanzania, Uganda, and the United States. Monthly earnings (USD) includes the respondent's profits from businesses and self-employment, salary and benefits from employment, and agricultural sales for the entire household.

Appendix Table 3: Correlations between various characteristics observed at baseline

	Attended schooling	Not a farmer	Employed	Has bank account	Monthly earnings	Total asset value
Attended schooling	1					
Not a farmer	0.111	1				
Employed	0.057	0.249	1			
Has bank account	0.278	0.068	0.177	1		
Monthly earnings	0.245	0.179	0.332	0.288	1	
Asset value	0.148	-0.002	0.180	0.207	0.220	1

Notes: This table presents correlations between six respondent and household characteristics that were observed at baseline, and appear to be important differences between "adopters only when the price is low" and "adopters when the price is high," as shown in Appendix Table 2. These include whether the household respondent (1) attended secondary schooling; (2) is not a farmer; (3) is employed; and (4) has a bank account; (5) estimated monthly earnings (USD), which includes the respondent's profits from businesses and self-employment, salary and benefits from employment, and agricultural sales for the entire household; and (6) estimated value of assets at baseline. These six variables are combined to construct a baseline measure of "Social and Economic Status." We then construct a binary variable, SES, indicating whether a household falls into the upper quartile of this measure. SES is then used as an interaction variable in instrumental variables regressions estimating the impacts of household electrification, in order to explore heterogeneity. We can compare the results of this approach to the alternative approach presented in the paper, and shown in Appendix Table 1.

Appendix Table 4: Local average treatment effects with high "Social and Economic Status" interaction

	Control	E	SES	$E \times SES$
	(1)	(2)	(3)	(4)
Panel A: Primary energy outcomes				
A1. Grid connected (%)	5.6	–	–	–
	[23.0]			
A2. Monthly electricity spending (USD)	0.14	1.86***	-0.05	1.19***
	[0.91]	(0.13)	(0.09)	(0.38)
Panel B: Additional energy outcomes				
B1. Electricity as main lighting (%)	5.2	88.4***	-0.5	2.8
	[22.2]	(2.4)	(0.8)	(3.9)
B2. Number of appliance types owned	1.8	0.4***	0.6***	-0.2
	[1.3]	(0.1)	(0.1)	(0.2)
B3. Owns mobile phone (%)	84.3	-5.0*	4.9***	7.4*
	[36.4]	(2.7)	(1.9)	(4.4)
B4. Owns radio (%)	54.2	7.7**	12.5***	-13.1*
	[49.8]	(3.8)	(3.5)	(7.5)
B5. Owns television (%)	17.9	13.3***	17.0***	-5.6
	[38.4]	(3.8)	(3.2)	(7.9)
B6. Owns iron (%)	4.1	3.3**	5.7***	-3.5
	[19.9]	(1.4)	(1.9)	(3.7)
B7. Monthly kerosene spending (USD)	2.81	-1.29***	-0.43**	0.43
	[2.86]	(0.17)	(0.19)	(0.35)
B8. Monthly total energy spending (USD)	11.66	0.62	2.28	-5.06
	[28.47]	(2.34)	(1.85)	(3.20)
B9. Solar home system as main lighting (%)	11.8	-9.3***	12.5***	-13.2***
	[32.3]	(1.3)	(2.5)	(3.3)

(Table continued on next page)

(Table continued from previous page)

	Control (1)	E (2)	SES (3)	$E \times SES$ (4)
Panel C: Primary economic outcomes				
C1. Household employed or own business (%)	36.8	6.2*	12.5***	-4.3
	[38.8]	(3.7)	(2.6)	(5.9)
—— Household women-only (%)	34.5	11.3***	15.8***	-6.9
	[44.5]	(4.1)	(3.1)	(7.6)
—— Household men-only (%)	40.2	0.8	6.5*	2.0
	[45.6]	(4.2)	(3.6)	(7.6)
C3. Total hours worked last week	47.0	-2.1	2.6	1.5
	[24.7]	(1.5)	(1.7)	(2.9)
C4. Total asset value (USD)	914	150	352***	232
	[961]	(136)	(85)	(203)
C5. Per capita cons. of major items (USD)	133	0	45***	-30
	[142]	(12)	(10)	(20)
Panel D: Primary non-economic outcomes				
D1. Recent health symptoms index	0	-0.06	0.23***	0.01
	[1]	(0.08)	(0.08)	(0.18)
D2. Normalized life satisfaction	0	0.21***	0.08	-0.27**
	[1]	(0.08)	(0.06)	(0.14)
D3. Avg. student test Z-score	0	0	0.39***	0
	[1]	(0)	(0.09)	(0)
D5. Political and social awareness index	0	0.05	0.32***	-0.28*
	[1]	(0.06)	(0.07)	(0.16)
Panel E: Mean treatment effects on grouped outcomes				
E1. Economic Index (C outcomes)	0	0.10	0.48***	-0.01
	[1]	(0.09)	(0.07)	(0.17)
E2. Non-Economic Index (D outcomes)	0	0.03	0.41***	-0.29
	[1]	(0.07)	(0.07)	(0.18)

Notes: We report coefficients from separate instrumental variables regressions in which household electrification status (E) and the interaction between E and SES are instrumented with the three subsidy treatment indicators (as well as their interactions with SES). SES is a binary variable indicating whether a household is in the upper quartile of the Social and Economic Status index constructed using observable characteristics at baseline (see Appendix Table 3 for a list of components). Regressions are estimated using only round 1 data (roughly 16 months post-connection). All specifications include pre-specified household, student, and community covariates, excluding those captured in SES. See Appendix Table B6A in Lee, Miguel, and Wolfram (2019) for additional results. Robust standard errors clustered at the community level in parentheses. Asterisks indicate coefficient statistical significance level (2-tailed): $^{*}P < 0.10$; $^{**}P < 0.05$; $^{***}P < 0.01$.

Appendix Note 1. Sources for historical rural electrification initiatives

In Table 1, we summarize some of the historical rural electrification initiatives evaluated in the recent microeconomics literature. In general, electrification rates are obtained from various studies, including those cited in the paper, as well as World Bank Open Data. GDP figures are calculated using GDP and CPI statistics from World Bank Open Data. Annual average exchange rates are obtained from the International Monetary Fund. The costs of electrification initiatives are obtained from various sources. Unless otherwise indicated, the reported cost is assumed to be in current prices of the year in which it is reported. All costs are then converted to 2017 USD. The following is a list of the references that were consulted.

United States of America

- Barnes, Douglas F. 2005. "Meeting the Challenge of Rural Electrification in Developing Nations: The Experience of Successful Programs." *Energy Sector Management Assistance Programme, The World Bank.*

- Lewis, Joshua and Edson Severnini. 2017. "Short- and Long-Run Impacts of Rural Electrification: Evidence from the Historical Rollout of the U.S. Power Grid." *Journal of Development Economics*, forthcoming.

- Maddison Project Database, version 2018. Bolt, Jutta, Robert Inklaar, Herman de Jong and Jan Luiten van Zanden. 2018. "Rebasing 'Maddison': new income comparisons and the shape of long-run economic development." *Maddison Project Working Paper 10.*

- Pellegrini, Lorenzo and Luca Tasciotti. 2013. "Rural Electrification Now and Then: Comparing Contemporary Challenges in Developing Countries to the USA's Experience in Retrospect." *Forum for Development Studies* 40(1): 153-176.

Bangladesh

- Asian Development Bank. "Rural Electrification in Bangladesh – Structure, Technical Performance, Achievements, and Areas for Improvement.", accessed here on November 4, 2019.

- Samad, Hussain and Fan Zhang. 2017. "Heterogeneous Effects of Rural Electrification: Evidence from Bangladesh." *Policy Research Working Paper 8102*

India

- Planning Commission, Government of India. "National Five Year Plans.", accessed here on November 4, 2019.

- Rural Electrification Corporation, Government of India. "Deendayal Upadhyaya Gram Jyoti Yojana Guidelines.", accessed here on November 4, 2019.

Brazil

- World Bank. 1994. "Eletrobras I & II Power Distribution Projects, Power Project Completion Report, Report No. 13173."

- Goldemberg, Jose, Emilio Lebre La Rovere, and Suani Teixeira Coelho. 2004. "Expanding access to electricity in Brazil." *Energy for Sustainable Development*

- Garcia-Escribano, Mercedes, Carlos Goes, and Izabela Karpowicz. 2015. “Filling the Gap: Infrastructure Investment in Brazil.” *IMF Working Paper No. 15/180*

- Borges da Silveira Bezerra, Paula, Camila Ludovique Callegari, Aline Ribas, André F P Lucena, Joana Portugal-Pereira, Alexandre Koberle, Alexandre Szklo, and Roberto Schaeffer. 2017. “The power of light: socio-economic and environmental implications of a rural electrification program in Brazil.” *Environmental Research Letters*

Ghana

- Adu, George, John Bosco Dramani, and Eric Fosu Oteng-Abayie. 2018. “Powering the powerless: Economic impact of rural electrification in Ghana” *International Growth Centre, E-33415-GHA-2*

- Ministry of Energy, Government of Ghana. 2010. “National Electrification Scheme, Master Plan Review (2011-2020).”, here on November 4, 2019.

South Africa

- Department of Minerals and Energy, Republic of South Africa. 2001. “National Electrification Programme 1994-1999: Summary Evaluation Report.”

Vietnam

- World Bank. 2007. “Implementation, Completion, and Results Report, Report No: ICR0000485.”

- World Bank. 2011. “Vietnam - State and People, Central and Local, Working Together - The Rural Electrification Experience.”

The Philippines

- Department of Energy, Republic of The Philippines. 2016. “Terminal Report on Sitio Electrication Program.”, accessed here on November 4, 2019.

Kenya

- Rural Electrification Authority. “Strategic Plan 2008-2012.”

- Rural Electrification Authority. “Strategic Plan 2016/17-2020/21.”

Appendix Note 2. Estimating heterogeneous treatment effects

We are interested in estimating Local Average Treatment Effects (LATE) separately for a set of complier subgroups. Since any LATE (and any treatment effect) can be represented as a weighted average of Marginal Treatment Effects (Heckman and Vytlacil (1999, 2001a, 2001b, 2001c, 2005, 2007a, 2007b)), it can also be represented as weighted average of sub-LATE's. This means that if we obtain a LATE using an instrumental variable, it can be represented as a weighted average of sub-LATE's for more narrowly defined complier subgroups. Therefore, the first step to calculate LATE's for different complier groups is to estimate the share that each complier group represents in the sample. Under randomization, we only need to make a standard assumption of monotonicity (Angrist, Imbens 1994) to get unbiased estimates of these shares.

In our experimental situation in which there are different randomized subsidy levels across treatment villages, we can define complier group $c1$ as those individuals who take up a grid connection when offered the low subsidy but not when offered no subsidy; those who take up power even without a subsidy are the "always takers". Those who only take up a connection when offered at least the medium subsidy are called $c2$, and those who only connect under the full subsidy (free treatment) are $c3$. The monotonicity assumption is equivalent to assuming that if individual i takes up treatment under a low subsidy, they would also take it up if offered the medium or full subsidy. This leads to the following logic: if $x\%$ of the control group take up a connection (without a subsidy), then $x\%$ of the entire sample are always takers. Then if $y\%$ of those offered the low subsidy take up treatment, we assume there is the same share of always takers in this treatment group, and therefore $y\% - x\%$ corresponds to the share of $c1$ compliers. Next, if $z\%$ of those offered the medium subsidy connect, then $z\% - y\% - x\%$ corresponds to the share of $c2$ compliers, since always takers and $c1$ compliers would also take up treatment under the medium subsidy. Following the same logic we can estimate the share of $c3$ compliers and never takers. We will denote these shares as $\pi_{at}, \pi_{nt}, \pi_{c1}, \pi_{c2},$ and π_{c3}.

Along with these shares, estimating the LATE for $c1$ compliers (LATEc1) is straightforward, since it simply corresponds to the 2SLS regression of an outcome on the low subsidy (using only control and low subsidy observations). If we estimate a 2SLS regression of an outcome on the medium subsidy, we are estimating a weighted average of LATEc1 and LATEc2, since the compliers at the medium subsidy level include those who would not have complied under the low subsidy, but also those who would have connected under a low subsidy if it had been offered. We call this weighted average LATEc1,c2. This same logic applies to the full subsidy.

Exploiting the monotonicity condition yields the following key expressions, which are similar in notation to those presented in Kowalski (2016):

For $c2$ compliers:

$$\frac{\pi_{c1}}{\pi_{c1} + \pi_{c2}} \text{LATE}_{c1} + \frac{\pi_{c2}}{\pi_{c1} + \pi_{c2}} \text{LATE}_{c2} = \text{LATE}_{c1,c2}$$

$$\implies \boxed{\text{LATE}_{c2} = \frac{\pi_{c1} + \pi_{c2}}{\pi_{c2}} \text{LATE}_{c1,c2} - \frac{\pi_{c1}}{\pi_{c2}} \text{LATE}_{c1}}$$

$$\boxed{V(\text{LATE}_{c2}) = \Big(\frac{\pi_{c1} + \pi_{c2}}{\pi_{c2}}\Big)^2 V(\text{LATE}_{c1,c2}) + \Big(\frac{\pi_{c1}}{\pi_{c2}}\Big)^2 V(\text{LATE}_{c1}) - 2\frac{\pi_{c1} + \pi_{c2}}{\pi_{c2}} COV(\text{LATE}_{c1,c2}, \text{LATE}_{c1})}$$

For $c3$ compliers:

$$\frac{\pi_{c1}}{\pi_{c1}+\pi_{c2}+\pi_{c3}}\text{LATE}_{c1}+\frac{\pi_{c2}}{\pi_{c1}+\pi_{c2}+\pi_{c3}}\text{LATE}_{c2}+\frac{\pi_{c3}}{\pi_{c1}+\pi_{c2}+\pi_{c3}}\text{LATE}_{c3}=\text{LATE}_{c1,c2,c3}$$

$$\implies \text{LATE}_{c3}=\frac{\pi_{c1}+\pi_{c2}+\pi_{c3}}{\pi_{c3}}\text{LATE}_{c1,c2,c3}-\frac{\pi_{c1}}{\pi_{c3}}\text{LATE}_{c1}-\frac{\pi_{c2}}{\pi_{c3}}\text{LATE}_{c2}$$

$$\implies \boxed{\text{LATE}_{c3}=\frac{\pi_{c1}+\pi_{c2}+\pi_{c3}}{\pi_{c3}}\text{LATE}_{c1,c2,c3}-\frac{\pi_{c1}+\pi_{c2}}{\pi_{c3}}\text{LATE}_{c1,c2}}$$

$$\boxed{\begin{aligned} V(\text{LATE}_{c3}) = &\Big(\frac{\pi_{c1}+\pi_{c2}+\pi_{c3}}{\pi_{c3}}\Big)^2 V(\text{LATE}_{c1,c2,c3})+\Big(\frac{\pi_{c1}+\pi_{c2}}{\pi_{c3}}\Big)^2 V(\text{LATE}_{c1,c2}) \\ &-2\Big(\frac{\pi_{c1}+\pi_{c2}}{\pi_{c3}}\Big)\Big(\frac{\pi_{c1}+\pi_{c2}+\pi_{c3}}{\pi_{c3}}\Big)COV(\text{LATE}_{c1,c2,c3},\text{LATE}_{c1,c2})\end{aligned}}$$

Baseline characteristics

The LATE's estimated implicitly come from the identification of the different complier groups using the monotonicity assumption. For example, LATEc2 is just the difference between the average outcome for the $c2$ compliers who were treated and the $c2$ compliers who were not. When it comes to baseline characteristics, it means we can identify the average for each complier group and treated/untreated combination. Following Abadie (2002), a number of studies have aimed to empirically identify the characteristics of different complier subgroups (for examples, see Card 2019; Mountjoy 2019; and Kline and Walters 2016). Here, we are able to calculate average baseline values corresponding to the combination of treated and untreated compliers within each subgroup using the assumption of monotonicity and the basic formula for weighted averages. As an example, consider the average characteristics X for those connected under the low subsidy. This average is a weighted average of the characteristics of the always takers and the treated $c1$ compliers. Since we can identify the always takers in the control group, we can back out the average characteristics for the treated $c1$ compliers. Similarly, the average characteristics X for those who remain unconnected under the medium subsidy is a weighted average of the never takers and those who only comply under the full subsidy, $c3$ compliers. Since the never takers are identified through those unconnected under the full subsidy, we can back out the average characteristics of the untreated $c3$ compliers. The following equations use the shares we previously described, $T = 1$ to represent treated (connected) households, and $Z_1, Z_2,$ and Z_3 to represent the low, medium, and full subsidy respectively:

For treated compliers:

$$\boxed{\begin{aligned} E[X|c1]_t &= \frac{\pi_{at}+\pi_{c1}}{\pi_{c1}}E[X|T=1,Z_1=1]-\frac{\pi_{at}}{\pi_{c1}}E[X|T=1,Z=0] \\ E[X|c2]_t &= \frac{\pi_{at}+\pi_{c1}+\pi_{c2}}{\pi_{c2}}E[X|T=1,Z_2=1]-\frac{\pi_{c1}}{\pi_{c2}}E[X|c1]_t-\frac{\pi_{at}}{\pi_{c2}}E[X|T=1,Z=0] \\ E[X|c3]_t &= \frac{\pi_{at}+\pi_{c1}+\pi_{c2}+\pi_{c3}}{\pi_{c3}}E[X|T=1,Z_3=1]-\frac{\pi_{c1}}{\pi_{c3}}E[X|c1]_t-\frac{\pi_{c2}}{\pi_{c3}}E[X|c2]_t-\frac{\pi_{at}}{\pi_{c3}}E[X|T=1,Z=0]\end{aligned}}$$

For untreated compliers:

$$\boxed{\begin{aligned}
E[X|c3]_u &= \frac{\pi_{nt}+\pi_{c3}}{\pi_{c3}}E[X|T=0,Z_2=1]-\frac{\pi_{nt}}{\pi_{c3}}E[X|T=0,Z_3=1]\\
E[X|c2]_u &= \frac{\pi_{nt}+\pi_{c3}+\pi_{c2}}{\pi_{c2}}E[X|T=0,Z_1=1]-\frac{\pi_{c3}}{\pi_{c2}}E[X|c3]_u-\frac{\pi_{nt}}{\pi_{c2}}E[X|T=0,Z_3=1]\\
E[X|c1]_u &= \frac{\pi_{nt}+\pi_{c3}+\pi_{c2}+\pi_{c1}}{\pi_{c1}}E[X|T=0,Z=0]-\frac{\pi_{c3}}{\pi_{c1}}E[X|c3]_u-\frac{\pi_{c2}}{\pi_{c1}}E[X|c2]_u-\frac{\pi_{nt}}{\pi_{c1}}E[X|T=0,Z_3=1]
\end{aligned}}$$

In Appendix Table 2, we compare baseline characteristics for each complier subgroup. Specifically, we present the minimum variance weighted average for each subgroup, instead of presenting them separately for treated and untreated households. Furthermore, we present a weighted average (using their shares) of the baseline characteristics and LATE's of complier subgroups c1 and c2, primarily because π_1 and π_2 are quite small and thus pooling the data leads to more statistical power. There is also a meaningful conceptual distinction between those willing (and able) to pay "something" for a connection, versus those who only connect when it is completely free. As noted above, the weighted average for c1 and c2 is just the LATE and the average characteristics for medium subsidy treatment group compliers, and in this case the relevant sub-LATE can be obtained directly from that IV estimate.

To be able to estimate the standard errors for the LATE's and the baseline characteristics, we need to estimate a full covariance matrix, given the covariance terms that appear in the expressions above. To do this, we employ a stacked regression approach, which is numerically equivalent to a seemingly unrelated regression (SUR) approach. This allows for a straightforward estimation of the analytical standard errors, rather than having to rely on bootstrapped estimates. All of the data and code that generated the results are in this article's replication files.

Additional References:

Abadie, Alberto. 2002. "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models." *Journal of the American Statistical Association* 97(457): 284-292.

Card, David, Alessandra Fenizia, and David Silver. 2019. "The Impacts of Hospital Delivery Practices on Infant Health," unpublished manuscript.

Heckman, James, and Edward Vytlacil. 2001. "Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect." *Econometric Evaluations of Active Labor Market Policies in Europe, ed. by M. Lechner and F. Pfeiffer. Heidelberg and Berlin: Physica.*

Heckman, James, and Edward Vytlacil. 2001. "Local Instrumental Variables." *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya, ed. by K. M. C. Hsiao and J. Powell. Cambridge University Press.*

Heckman, James, and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73(3): 669–738.

Heckman, James, and Edward Vytlacil. 2007. "Chapter 70 Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation." *Handbook of Econometrics* Vol. 6, Part 2.

Heckman, James, and Edward Vytlacil. 2007. "Chapter 71 Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." *Handbook of Econometrics* Vol. 6, Part 2.

Kline, Patrick, and Christopher Walters. 2016. "Evaluating public programs with close substitutes: The case of Head Start." *Quarterly Journal of Economics* 131(4): 1795-1848.

Mountjoy, Jack. 2019. "Community Colleges and Upward Mobility," unpublished manuscript.