Online Appendix for "How Do Electricity Shortages Affect Industry? Evidence from India"

By Hunt Allcott, Allan Collard-Wexler, and Stephen D. O'Connell

.

2005 World Bank Enterprise Survey: Barriers to Growth

* Allcott: New York University, NBER, and Poverty Action Lab. NYU Economics Department, 19 W. 4th St., New York, NY 10012. Email: hunt.allcott@nyu.edu. Collard-Wexler: Duke University and NBER. 230 Social Sciences Building, Durham, NC 27708. Email: collardwexler@gmail.com. O'Connell: City University of New York - Graduate Center. Department of Economics Room 5313, 365 5th Avenue, New York, NY 10016. Email: soconnell@gradcenter.cuny.edu. We thank Maureen Cropper, Jan De Loecker, Michael Greenstone, Peter Klenow, Kabir Malik, Rohini Pande, Nick Ryan, Jagadeesh Sivadasan, Anant Sudarshan, and seminar participants at Brown, Drexel, Duke, the Federal Trade Commission, Harvard, the 2013 NBER Summer Institute, the 2014 NBER Winter IO/EEE meetings, Society for Economic Dynamics, Stanford, KU Leuven, Toulouse, Universidad de los Andes, University of Chicago, University of Cologne, and the World Bank for helpful comments. We are particularly grateful to Nick Bloom, Troy Smith, and Shaleen Chavda for insight into the textile industry. We thank Deepak Choudhary, Anuradha Bhatta, Sherry Wu, and Mark Thomas for helpful research assistance and the Stern Center for Global Economy and Business for financial support. We have benefited from helpful conversations with Jayant Deo of India Energy Exchange, Gajendra Haldea of the Planning Commission, Partha Mukhopadhyay of the Centre for Policy Research, and Kirit Parikh of IGIDR. We also thank A. S. Bakshi and Hemant Jain of the Central Electricity Authority for help in collecting archival data. Of course, all analyses are the responsibility of the authors, and no other parties are accountable for our conclusions. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

Table A1—: **Biggest Obstacle for Growth**

Problem	**Percent**
Electricity	33
High Taxes	16
Corruption	10
Tax Administration	8
Cost of and Access to Financing	6
Labor Regulations and Business Licensing	5
Skills and Education of Available Workers	4
Access to Land	3
Customs and Trade Regulations	2
Other	12

Notes: These data are from the 2005 World Bank Enterprise Survey in India. The table presents responses to the question, "Which of the elements of the business environment included in the list, if any, currently represents the biggest obstacle faced by this establishment?"

Power Sector Data Appendix

This appendix provides additional details on the power sector data in Section I.I.B.

Appendix Table A1 presents the state allocations from jointly-owned power plants.

Appendix Table A2 presents summary statistics for the reservoir and hydro plant microdata. Panel A presents the reservoir microdata; 31 reservoirs ever appear. A reservoir scheme may include multiple hydro plants, so generation and generation capacity are for all plants within the reservoir scheme. All reservoir data are missing for the year 2000, so inflows are imputed using rainfall at gridpoints within the reservoir watershed. For each reservoir, we then run a regression of generation on inflows; the fitted values are then divided by generation capacity and transformed into a predicted capacity factor.

Panel B of Appendix Table A2 presents the hydro plant microdata; 181 plants ever appear, of which 18 percent (32 plants) are known to be run-of-river plants. All plant-level data are missing for the year 1992, and generation data are occasionally missing in other years. Just less than six percent of generation observations are imputed using rainfall at gridpoints within the plant's watershed.

Appendix Table A3 presents summary statistics on electricity supply for the ten largest states.

Figures A1, A2, A3 present maps of shortage severity, hydro power plants and weather stations, and four example hydro plant watersheds, respectively.

Table A1—: **State Allocations from Jointly-Owned Hydro Plants**

Power Station	State	Share (Percent)	Total Capacity (MW)
Bhakra Nangal Complex	Haryana	33.91	1479.5
	Punjab	50.87	
	Rajasthan	15.22	
Dehar	Haryana	32	990
	Punjab	48	
	Rajasthan	20	
Pong	Haryana	16.6	396
	Punjab	24.9	
	Rajasthan	58.5	
Gandhi Sagar	Madhya Pradesh	50	115
	Rajasthan	50	
Jawahar Sagar	Madhya Pradesh	50	99
	Rajasthan	50	
Rana Pratap Sagar	Madhya Pradesh	50	172
	Rajasthan	50	
Machkund	Andhra Pradesh	70	114.75
	Orissa	30	
Tungabhadra/Hampi	Andhra Pradesh	80	72
	Karnataka	20	
Pench	Madhya Pradesh	66.67	160
	Maharashtra	33.33	
Sardar Sarovar	Gujarat	16	1450
	Madhya Pradesh	57	
	Maharashtra	27	
Rajghat	Madhya Pradesh	50	45
	Uttar Pradesh	50	
Ranjit Sagar	Punjab	75.4	600
	Jammu & Kashmir	20	
	Himachal Pradesh	4.6	

Source: Central Electricity Authority, General Review 2012.

Table A2—: **Reservoir and Hydro Plant Microdata Summary Statistics**

Panel A: Reservoir Microdata

	Mean	Std. Dev.	Min.	Max.	Obs.
Reservoir Years Observed	12.4	7.0	4	19	31
Reservoir Inflows (billion cubic meters)	9.0	11.1	0.14	77.4	362
Reservoir-Level Generation (GWh)	1926	1669	27	8016	367
Reservoir-Level Generation Capacity (MW)	676	547	75	1956	383
Capacity Factor Predicted by Inflows	0.36	0.14	0.09	0.75	383

Panel B: Hydro Plant Generation Microdata

	Mean	Std. Dev.	Min.	Max.	Obs.
Plant Years Observed	14.0	6.1	1	19	181
Run-of-River Plant	0.19	0.4	0.00	1.0	177
Generation (GWh)	654	1042	0	13211	2395
Capacity (MW)	207	304	0	1956	2505
Capacity Factor	0.37	0.19	0	1.57	2387

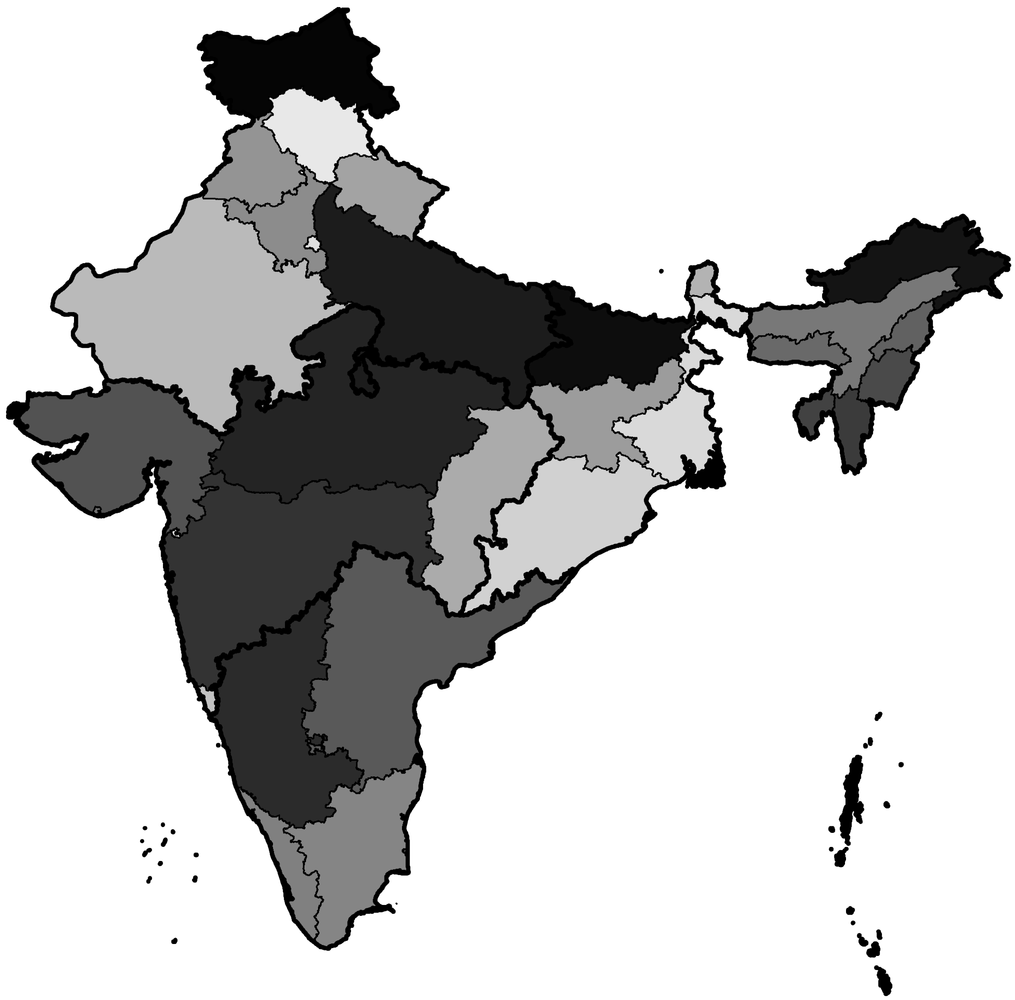
Notes: Reservoir Years Observed is at the reservoir level; all other variables in Panel A are at the reservoir-by-year level. Plant Years Observed and run-of-river categorization are at the plant level; all other variables in Panel B are at the plant-by-year level.

Table A3—: **Electricity Statistics for the Ten Largest States**

	1992-2010 Shortages			2010	1992-2010
				Capacity	Generation
State	Mean	Min.	Max.	(gigawatts)	Share Hydro
Andhra Pradesh	0.08	0.01	0.22	10.8	0.21
Gujarat	0.08	0.03	0.16	11.4	0.03
Karnataka	0.12	0.01	0.27	9.1	0.47
Madhya Pradesh	0.12	0.05	0.20	4.9	0.14
Maharashtra	0.10	0.02	0.21	16.1	0.11
Punjab	0.05	0.01	0.14	5.1	0.40
Rajasthan	0.03	0.00	0.07	5.7	0.18
Tamil Nadu	0.06	0.00	0.14	11.6	0.13
Uttar Pradesh	0.15	0.10	0.22	5.5	0.11
West Bengal	0.02	0.00	0.06	7.4	0.03

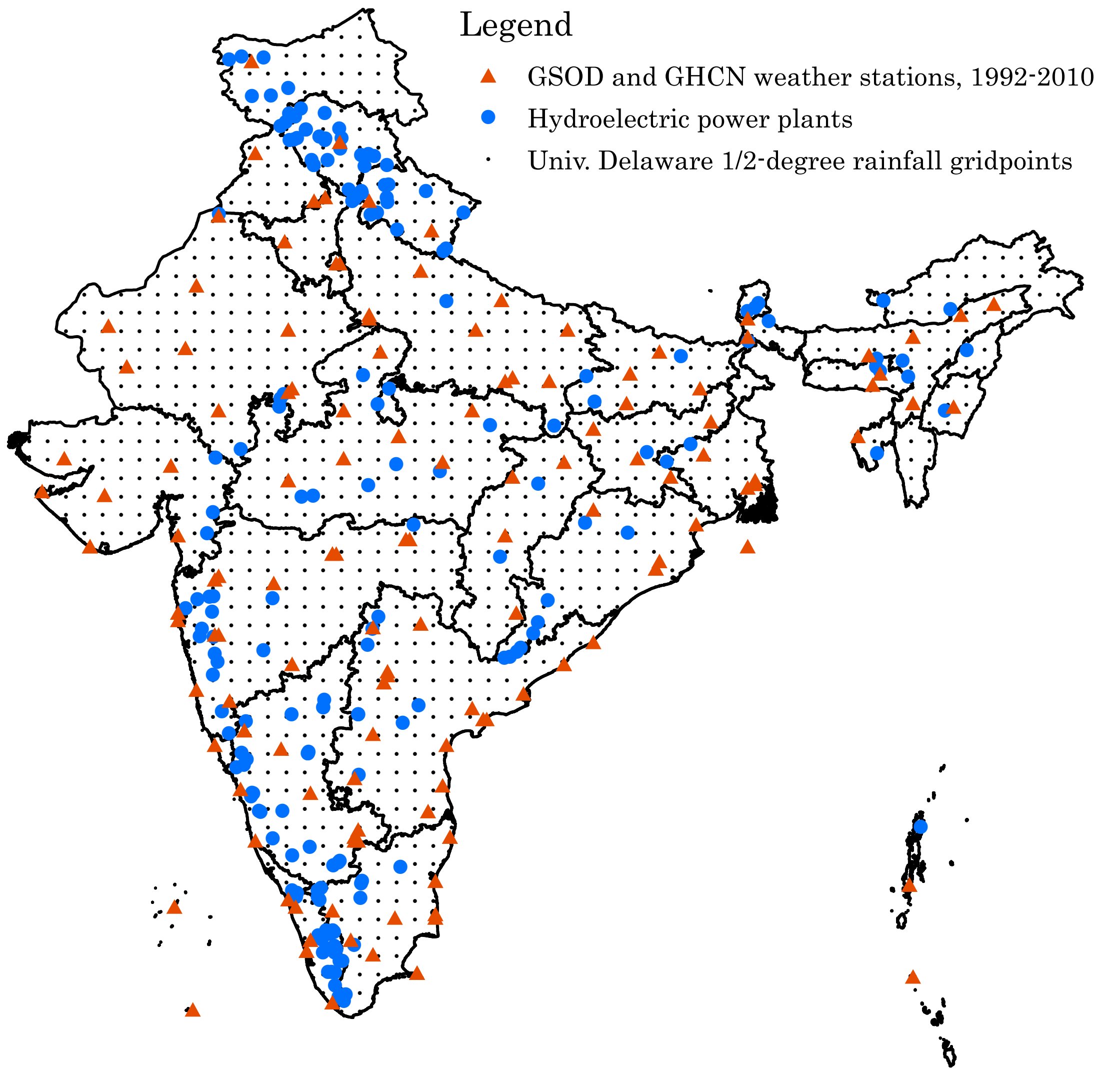
Notes: Shortage data are estimated by the Central Electricity Authority. 2010 Generation Capacity is reported in the CEA General Review 2012, and Generation Share Hydro is is the ratio of hydroelectricity to total generation, both of which are reported in the General Review.

Figure A1. : **Map of Average Shortages by State**



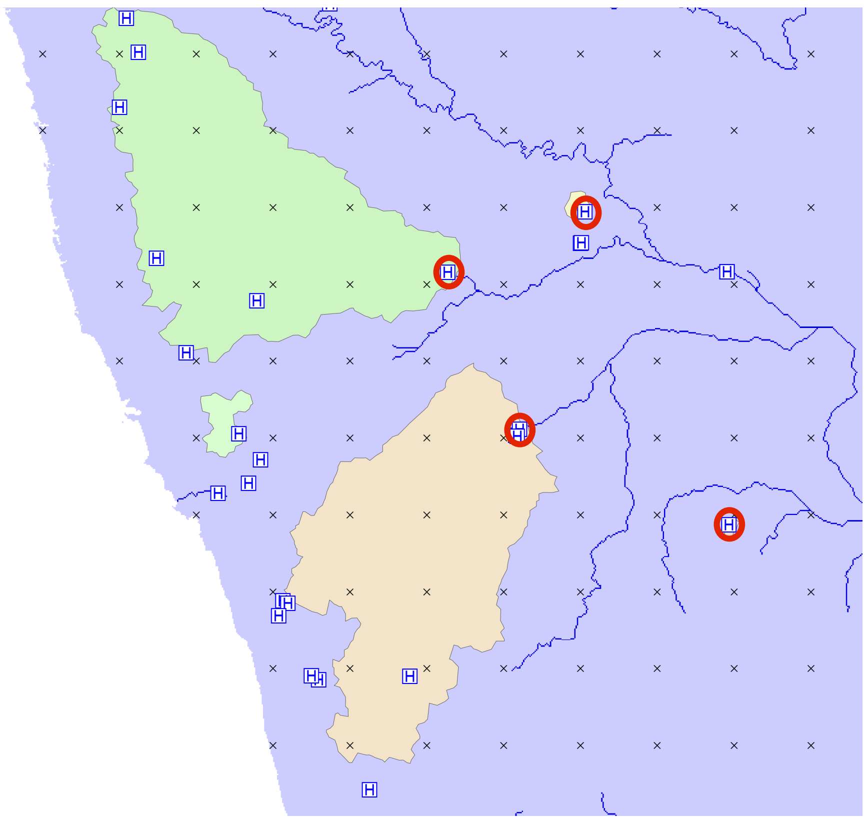
Notes: This figure presents each state's average Shortage assessed by the Central Electricity Authority over 1992-2010, with darker color illustrating higher Shortage.

Figure A2. : **Hydro Power Plants, Rainfall Gridpoints, and Weather Stations**



Notes: This figure plots the 1/2 degree gridpoints in the University of Delaware rainfall data, the weather stations whose measurements underlie the gridded data, and the locations of all hydroelectric power stations in India.

Figure A3. : **Example of Hydroelectric Plant Watersheds**



Notes: This figure presents an example of four hydrolectric plant watersheds in Southern India, with plants circled in red. Other hydroelectric plants are marked by "H" symbols, and 1/2-degree rainfall gridpoints are marked by "+" symbols. The river network is extrapolated based on the predicted flow of water accross India.

Annual Survey of Industries Data Appendix

This appendix presents additional information on the Annual Survey of Industries data.

We extract a subset of variables from the raw data separately for each year and then stack all years of data together to apply the following cleaning processes.[1] We correct observations in 1993-94 to 1997-98 whose values have been supplied in "pre-multiplied" format from the India's Ministry of Statistics and Programme Implementation (MOSPI). We create a separate consistently-defined state variable which takes into account the creation of Jharkhand, Chhattisgarh and Uttaranchal (now Uttarakhand) in 2001 from Bihar, Madhya Pradesh and Uttar Pradesh, respectively. We assign establishments to the last observed state, which correctly places establishments despite state splits, as long as the establishment is surveyed after 2001.

India classifies manufacturing establishments with its National Industrial Classification (NIC), which resembles industrial classifications commonly used in other countries. The classifications were revised in 1987, 1998, 2004, and 2008. We convert all industry classifications to the NIC-1987 scheme using concordances provided by MOSPI.

All financial amounts are deflated to constant 2004-05 Rupees. Revenue (gross sales) is deflated by a three-digit commodity price deflators as available in the commodity-based table "Index Numbers Of Wholesale Prices In India – By Groups And Sub-Groups (Yearly Averages)" produced by the Office of the Economic Adviser-Ministry of Commerce & Industry.[2] Each three-digit NIC-1987 code is assigned to a commodity listed in this table. The corresponding commodity deflator is used to deflate revenues. To deflate material inputs, we construct the average output deflator of a given industry's supplier industries based on India's 1993-94 input-output table, available from the Central Statistical Organization. Fuel and electricity costs are deflated by the price index for "Fuel, Power, Light, and Lubricants." Capital is deflated by an implied national deflator calculated from "Table 13: Sector-wise Gross Capital Formation" from the Reserve Bank of India's Handbook of Statistics on the Indian Economy.[3] Electricity costs are deflated using a national GDP deflator.

The sampling rules have changed somewhat over time. The census scheme included factories with 100 or more workers in all years except 1997-2003, when it included only factories with 200 or more workers. The sample scheme included one-third of factories until 2004 and one-fifth since then (MOSPI 2014).

The ASI data have at least two well-known shortcomings. First, while the data

[1] We thank Jagadeesh Sivadasan for helpful discussions and for providing Stata code that facilitated the read-in of 1992-1997 ASI data, and Olivier Dupriez for similarly helpful discussions and pointing us to read-in programs for ASI data from 1998 to 2007 available at the International Household Survey Network (http://catalog.ihsn.org/index.php/catalog/central).

[2] Available from http://www.eaindustry.nic.in/

[3] Available from http://www.rbi.org.in

are representative of small registered factories and a 100 percent sample of large registered factories, not all factories are actually registered under the Factories Act. Nagaraj (2002) shows that only 48 percent and 43 percent of the number of manufacturing establishments in the 1980 and 1990 economic censes appear in the ASI data for those years. Although it is not clear how our results might differ for unregistered plants, the plants that are observed in the ASI are still a significant share of plants in India. Second, value added may be under-reported, perhaps associated with tax evasion, by using accounting loopholes to overstate input costs or under-state revenues (Nagaraj 2002). As long as changes in this under-reporting are not correlated with electricity shortages, this will not affect our results.

C1. Determination of Base Sample

Appendix Table A1 details how the sample in Panel B of Table 1 is determined from the original set of observations in the ASI. The 1992-2010 ASI dataset begins with 949,992 plant-year observations. Plants may still appear in the data even if they are closed or did not provide a survey response. We drop 172,697 plants reported as closed or non-responsive. We drop a trivial number of observations missing state identifiers and observations in Sikkim, which has only been included in the ASI sampling frame in the most recent years. We drop 45,664 observations reporting non-manufacturing NIC codes. We remove a small number of observations (primarily in the early years of our sample) which are exact duplicates in all fields, assuming these are erroneous multiple entries made from the same questionnaire form. Due to the importance of revenue and productivity results, we remove the 102,036 observations with missing revenues. We also drop the 9,095 observations with two or more input revenue share flags, from the flagging process described below.

With this intermediate sample, we use median regression to estimate revenue productivity (TFPR) under a full Cobb-Douglas model in capital, labor, materials, and energy and assuming constant returns to scale. This full Cobb-Douglas revenue productivity term is used only for the final sample restriction, which is to drop 4,521 plant-years which have log-TFPR greater than 3.5 in absolute value from the sample median. Such outlying TFPR values strongly suggest misreported inputs or revenues. The final sample includes 615,721 plant-year observations, of which 362,151 are from the sample scheme and 253,570 are from the census scheme.

C2. Variable-Specific Sample Restrictions

After the final sample is determined, there may still be observations which have correct data for most variables but misreported data for some individual variable. When analyzing specific variables (such as self-generation share, energy revenue share, or output in Table 6), we therefore additionally restrict the sample using the following criteria:

Table A1—: **Determination of Base Sample**

Step	Dropped Observations	Resulting Sample Size
Original ASI dataset		949,992
Closed plants	172,697	777,295
Missing state codes or in Sikkim	99	777,196
Non-manufacturing NIC codes	45,664	731,532
Exact duplicates	311	731,221
Missing revenues	102,036	629,185
Multiple input revenue share outliers	8,943	620,242
Productivity outliers	4,521	615,721
Total observations		615,721

Notes: This table details how the sample in Panel B of Table 1 is determined from the original set of observations in the Annual Survey of Industries.

- We generate "input revenue share flags" for labor and materials if input cost is more than two times revenues, and we generate input revenue share flags for electricity and fuels if input cost is greater than revenues.[4] Because we also observe physical quantities for labor and electricity, we generate analogous input revenue share flags by multiplying physical quantities by prices, resulting in an implied revenue share based on these physical quantities. For electricity, we use the median real price (in Rs/kWh) of purchased electricity in any given state and year. For labor, we assume a very conservative 1,000 Rs per person per annum wage rate. When using either of these inputs as an outcome, we omit observations with an input revenue share flag for that input.
- There are a trivial number of observations which report unrealistic count of workers (greater than 200,000 persons engaged), which we make missing in those cases.
- We generate "within-plant outlier" flags for observations with unrealistically large year-to-year fluctuations in revenue, TFPR, or any input. We flag observations if the change in logged value is more than 3.5 (or 1.5 in a robustness check) from both adjacent observations. For a plants' first or last year, an observation is flagged if the change is more than 3.5 (or 1.5) from the subsequent or previous observation.

Cleaning Electricity Variables

We clean plant electricity measures in the following ways:

[4]The flags would be slightly different if applied to deflated inputs and revenues, but this will have minimal implications for the results.

- We make electricity consumption missing for all observations (other than brick kilns) that report zero electricity consumption.

- We make all electricity variables missing if the plant reports consuming more than 110 percent or less than 90 percent of the total amount of electricity they report purchasing and generating.

- We make missing the values of electricity purchased and sold if the implied price per kilowatt-hour is less than 2 percent or more than 5,000 percent of the median grid electricity price calculated across plants in the same state and year. We also make missing the reported quantities of electricity purchased and sold if the respective price flag is triggered.

Production Function and Productivity Estimation

We recover production function coefficients given by Equations (12), (13), and (15) for each of the 143 three-digit industries in the dataset. (To ensure sufficient sample size in each three-digit industry, we adjust industry definitions slightly to ensure each three-digit industry has at least 100 plant-year observations.) We use separate median regression for each two-digit industry, allowing for a linear time trend and separate intercepts for each underlying three-digit industry. After calculating production function coefficients, we compute TFPR using Equation (11).

For our main TFPR estimates, we define materials to be the plant's original reported materials plus fuels not used for self-generation. This latter variable is: Total Fuel Cost - (7 Rs/kWh)×(kWh Self-Generated), where 7 Rs/kWh is the median price reported in the 2005 World Bank Enterprise Survey. This allows us to account for the plant's full input costs when calculating production function parameters and TFPR. (In regressions where we use materials as the outcome variable, we use the original reported materials without adding any fuel costs.)

We use several alternative methods for calculating production function coefficients and TFPR for robustness checks, seen in Appendix Table A10. In the order of that table, these are:

- Including or excluding all fuel costs from the materials variable

- Removing the linear time trend when estimating production coefficients, which amounts to taking the unconditional median revenue share by industry

- Relaxing the assumption that factor shares are constant by plant size, allowing all production function coefficients to vary by plant median ln(Revenue). To implement, we add ln(Revenue) as a term in the median regressions for α_L, α_M, and α_E and then segment plants into five size classes when estimating α_K in GMM.

- Backing out the capital coefficient α_K under an assumption of constant returns to scale
- Assuming production is Leontief in electricity and calculating the capital coefficient under an assumption of constant returns to scale. (This is the approach in our original working paper.)

None of these changes affects the estimated coefficient by more than about half of the standard error.

Empirical Strategy Appendix

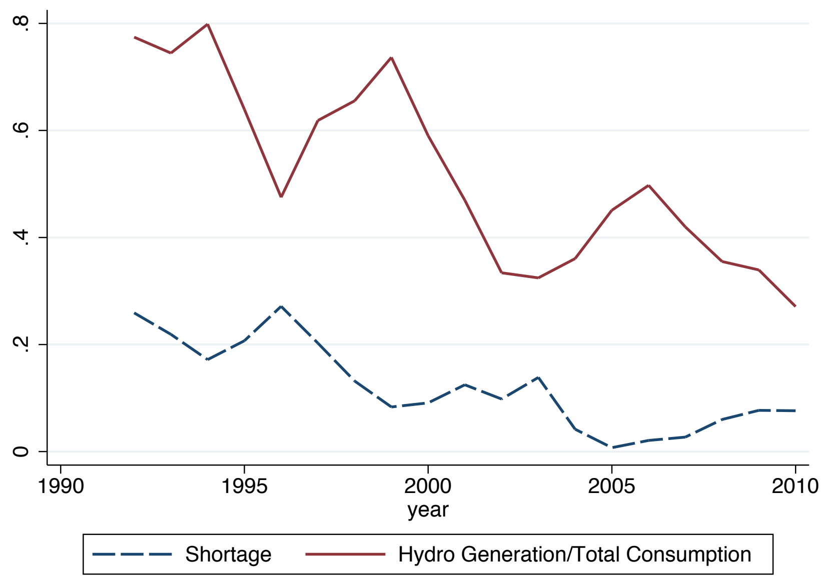
This appendix presents a table and figures that support the empirical strategy section.

Table A1—: **Serial Correlation Tests for the Hydro Instrument**

	(1)	(2)
1st Lag Z	0.107	0.111
	(0.081)	(0.088)
2nd Lag Z		0.008
		(0.075)
3rd Lag Z		-0.130
		(0.066)**
4th Lag Z		-0.104
		(0.083)
5th Lag Z		-0.025
		(0.068)
Number of Obs.	540	420
F-Stat	1.72	1.90
R-Squared	0.01	0.07

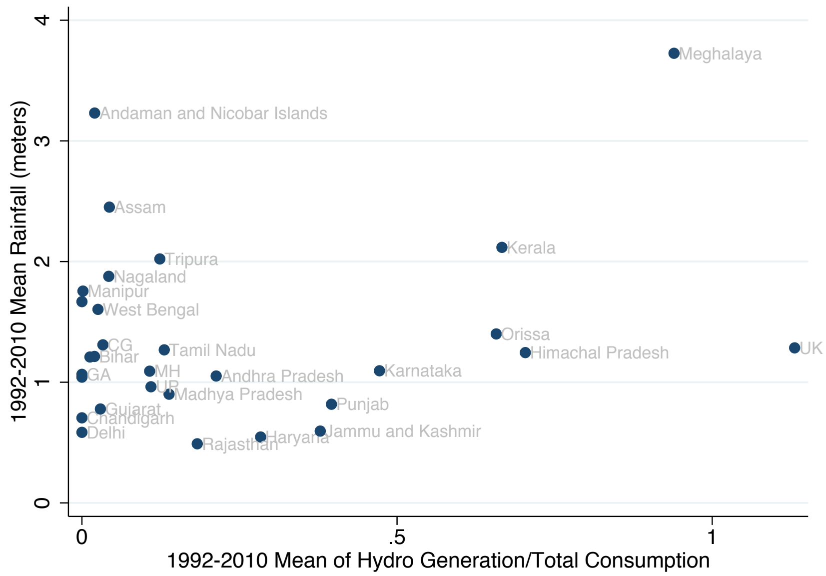
Notes: This table presents regressions of the hydro instrument Z_{st} on its lags. Robust standard errors. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Figure A1. : **Shortages and Hydro Generation in Karnataka**



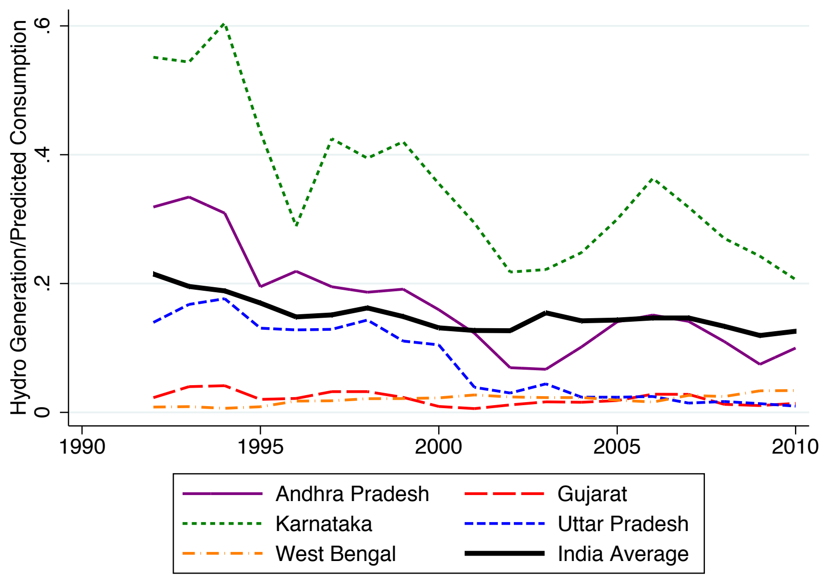
Notes: Shortage data are estimated by the Central Electricity Authority. Hydro Generation/Total Electricity Consumption is a simplified version of the hydro availability instrument. The figure gives a simple graphical example of the first stage of our IV estimator.

Figure A2. : **Rainfall and Hydro Share of Electricity by State**



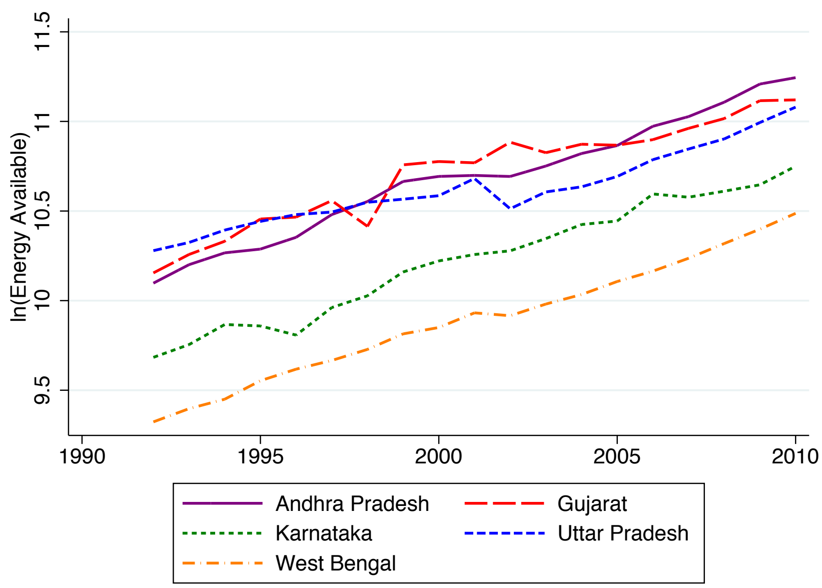
Notes: This figure plots sample average annual rainfall against the mean ratio of hydroelectricity generation to total electricity consumption. The figure emphasizes that there is substantial variation in hydro generation conditional on rainfall.

Figure A3. : **Hydro Share of Predicted Consumption Over Time in Five Large States**



Notes: This figure presents the ratio of hydro generation to predicted consumption over 1992-2010 for five large states. Different states have different slopes, illustrating the importance of including state-specific time trends as control variables.

Figure A4. : **ln(Energy Available) Over Time in Five Large States**



Notes: This figure presents the natural log of Energy Available over 1992-2010 for five large states. Different states have different slopes, illustrating the importance of including state-specific time trends as control variables.

Robustness Checks for Tables 6 and 7

This appendix presents robustness checks for Tables 6 and 7. We first present estimates from the difference estimator. We then present a series of robustness checks for the fixed effects estimator, including alternative weather controls, alternative instruments, and alternative constructions of TFPR.

E1. Estimates with Difference Estimator

Tables A1, A2, and A3 are analogous to Tables 5, 6, and 7, except that they use the difference estimator. Mechanically, we difference each variable within-plant and run OLS regressions with the differenced observations; the sample sizes thus differ from the fixed effects estimates. Note that the initial and final years of a differenced observation are not necessarily one year apart due to the irregularly-spaced ASI sample.

Table A1—: **First Stages Using the Difference Estimator**

Panel A: Energy Inputs

	(1)	(2)	(3)
Second Stage Dependent Var:	Self-Gen Share	ln(Fuel Rev Share)	ln(Electric Intensity)
Hydro	-0.189	-0.193	-0.174
	(0.0401)***	(0.0405)***	(0.0405)***
Obs.	177,822	234,384	363,572
Clusters	504	504	506
Clusters (2)	505	505	506
1st Stage F-Stat	22.15	22.68	18.41

Panel B: Other Inputs and Outputs

	(1)	(2)	(3)	(4)	(5)
Second Stage Dependent Var:	ln(Materials)	ln(Workers)	ln(Earnings/ Worker)	ln(Revenue)	ln(TFPR)
Hydro	-0.170	-0.170	-0.178	-0.170	-0.173
	(0.0404)***	(0.0403)***	(0.0432)***	(0.0405)***	(0.0404)***
Obs.	378,256	385,716	332,324	384,713	360,996
Clusters	506	506	451	506	506
Clusters (2)	506	506	451	506	506
1st Stage F-Stat	17.69	17.66	16.98	17.67	18.24

Notes: This table presents the first stage estimates for the IV regressions, estimated using the difference estimator. The dependent variable for these first stage regressions is Shortage S_{st}. Samples for columns 1 and 2 in Panel A are limited to plants that ever self-generate electricity.F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by state-by-initial year and state-by-final year of the differenced observation. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A2—: **Effects of Shortages on Energy Inputs Using the Difference Estimator**

	(1)	(2)	(4)
Dependent Var:	Self-Gen Share	ln(Fuel Rev Share)	ln(Electric Intensity)
Panel A: OLS			
Shortage	0.274	0.874	-0.595
	(0.0383)***	(0.202)***	(0.132)***
Panel B: IV			
Shortage	0.349	2.419	0.339
	(0.135)***	(0.713)***	(0.724)
Obs.	177,822	234,384	363,572
Clusters	504	504	506
Clusters (2)	505	505	506
1st Stage F-Stat	22.15	22.68	18.41

Notes: This table presents estimates of Equation (21) using the difference estimator. Panel B instruments for Shortage using hydro availability. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by state-by-initial year and state-by-final year of the differenced observation. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A3—: **Effects of Shortages on Materials, Labor, Revenue, and TFPR Using the Difference Estimator**

	(1)	(2)	(3)	(4)	(5)
Dependent Var:	ln(Materials)	ln(Workers)	ln(Earnings/ Worker)	ln(Revenue)	ln(TFPR)
Panel A: OLS					
Shortage	0.0203	0.0254	0.201	0.158	0.0724
	(0.0841)	(0.0526)	(0.0508)***	(0.0762)**	(0.0394)*
Panel B: IV					
Shortage	-0.959	-0.397	-0.243	-0.828	-0.106
	(0.460)**	(0.315)	(0.224)	(0.491)*	(0.238)
Obs.	378,256	385,716	332,324	384,713	360,996
Clusters	506	506	451	506	506
Clusters (2)	506	506	451	506	506
1st Stage F-Stat	17.69	17.66	16.98	17.67	18.24

Notes: This table presents estimates of Equation (21) using the difference estimator. Panel B instruments for Shortage using hydro availability. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by state-by-initial year and state-by-final year of the differenced observation. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

E2. Robustness Checks

This section presents robustness checks for Tables 6 and 7. Tables are organized separately for each of the main outcomes for ease of comparison. Column 1 excludes industry-by-year effects μ_{jt}. Column 2 uses a tolerance of 1.5 natural logs in the outlier flagging process described in Online Appendix C.C2, while column 3 does not exclude any flagged outliers. Columns 4 and 5 use ln(Energy Available) and Peak Shortage, respectively, instead of Shortage. Column 6 clusters standard errors by state.

We make two explanatory comments. First, the first stage F-statistics for ln(Energy Available) in column 4 are smaller than when using Shortage as the endogenous variable in the main estimates, which is unsurprising: unlike Shortage, ln(Energy Available) grows monotonically within states over the sample, and the state-specific linear time trends $\psi_s t$ do not control very well for different states' actual growth rates. Second, the first stage F-statistics increase in two specifications when clustering by state in column 6, and this may be a small sample bias from having only 30 state-level clusters.

Table A4—: **Robustness Checks: Energy Inputs**

	(1)	(2)	(3)	(4)	(5)	(6)
Change from Base Spec:	No Ind.-by-Year Effects μ_{jt}	Tighter Outlier Flags	No Outlier Flags	Use ln(Energy Available)	Use Peak Shortage	Cluster by State
Self-Generation Share						
Shortage	0.455	0.394	0.450	0.433	0.404	0.442
	(0.156)***	(0.141)***	(0.158)***	(0.163)***	(0.169)**	(0.129)***
Number of Obs.	240,743	223,128	293,866	240,743	240,743	240,743
First Stage F-Stat	15.08	16.97	17.02	9.285	9.705	35.12
ln(Fuel Revenue Share)						
Shortage	3.675	2.700	3.022	3.133	3.107	3.294
	(1.158)***	(1.001)***	(1.298)**	(1.215)***	(1.274)**	(0.961)***
Number of Obs.	291,759	268,663	300,697	291,759	291,759	291,759
First Stage F-Stat	14.79	16.49	16.56	9.829	8.773	37.85
ln(Electric Intensity)						
Shortage	-0.177	-0.0320	0.0247	0.0764	0.0726	0.0926
	(0.735)	(0.708)	(0.753)	(0.616)	(0.591)	(0.694)
Number of Obs.	479,616	453,482	483,843	479,616	479,616	479,616
First Stage F-Stat	13.89	14.86	14.99	13.71	11.38	13.41

Notes: This table presents alternative estimates for Table 6, instrumenting for Shortage using hydro availability. Samples for the first two panels are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A5—: **Robustness Checks: Materials, Labor, Revenue, and TFPR**

	(1)	(2)	(3)	(4)	(5)	(6)
Change from Base Spec:	No Ind.-by-Year Effects μ_{jt}	Tighter Outlier Flags	No Outlier Flags	Use ln(Energy Available)	Use Peak Shortage	Cluster by State
ln(Materials)						
Shortage	-1.048	-1.017	-1.237	-0.917	-0.915	-1.137
	(0.548)*	(0.472)**	(0.595)**	(0.392)**	(0.433)**	(0.561)**
Number of Obs.	495,043	478,152	498,464	495,043	495,043	495,043
First Stage F-Stat	13.20	14.20	14.22	13.91	10.29	12.52
ln(Workers)						
Shortage	-0.230	-0.253	-0.248	-0.195	-0.196	-0.243
	(0.358)	(0.311)	(0.349)	(0.260)	(0.271)	(0.391)
Number of Obs.	502,724	496,474	503,217	502,724	502,724	502,724
First Stage F-Stat	13.15	14.21	14.20	14.12	10.27	12.45
ln(Earnings/Worker)						
Shortage	-0.321	-0.384	-0.367	-0.234	-0.260	-0.267
	(0.239)	(0.220)*	(0.244)	(0.180)	(0.241)	(0.243)
Number of Obs.	456,443	440,524	461,131	456,443	456,443	456,443
First Stage F-Stat	13.45	14.76	14.45	12.49	7.354	9.508
ln(Revenue)						
Shortage	-1.050	-0.993	-1.255	-0.877	-0.880	-1.091
	(0.555)*	(0.494)**	(0.638)**	(0.385)**	(0.458)*	(0.646)*
Number of Obs.	501,130	484,753	503,664	501,130	501,130	501,130
First Stage F-Stat	13.13	14.02	14.19	14.04	10.22	12.45
ln(TFPR)						
Shortage	-0.0733	-0.246	-0.408	-0.247	-0.242	-0.304
	(0.252)	(0.231)	(0.283)	(0.203)	(0.216)	(0.348)
Number of Obs.	479,313	472,612	480,243	479,313	479,313	479,313
First Stage F-Stat	13.86	14.98	14.84	14.05	10.99	13.28

Notes: This table presents alternative estimates for Table 7, instrumenting for Shortage using hydro availability. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

E3. Alternative Weather Controls

This section presents estimates of Tables 6 and 7 with alternative weather controls. Column 1 controls linearly for rainfall instead of including rainfall bins. Columns 2 and 3 use 100mm and 50 mm rainfall bins, respectively, instead of 60mm bins. Column 4 uses rainfall data from the National Climate Centre instead of the University of Delaware.

Table A6—: **Alternative Weather Controls: Energy Inputs**

	(1)	(2)	(3)	(4)
Change from Base Spec:	Linear Rainfall	100mm Rainfall Bins	50mm Rainfall Bins	NCC Rainfall Data
Self-Generation Share				
Shortage	0.412	0.458	0.442	0.406
	(0.175)**	(0.155)***	(0.153)***	(0.150)***
Rainfall	0.00130			
	(0.00892)			
First Stage F-Stat	16.12	19.05	17.00	21.63
ln(Fuel Revenue Share)				
Shortage	2.797	3.049	3.294	1.901
	(1.052)***	(0.934)***	(1.032)***	(0.826)**
Rainfall	0.185			
	(0.0663)***			
First Stage F-Stat	15.41	18.20	16.53	20.98
ln(Electric Intensity)				
Shortage	-0.0294	0.0583	0.0926	0.0894
	(0.775)	(0.696)	(0.755)	(0.665)
Rainfall	-0.0264			
	(0.0358)			
First Stage F-Stat	15.37	18.12	14.98	18.57

Notes: This table presents alternative estimates for Table 6, instrumenting for Shortage using hydro availability. Samples for the first two panels are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A7—: **Alternative Weather Controls: Materials, Labor, Revenue, and TFPR**

	(1)	(2)	(3)	(4)
Change from Base Spec:	Linear Rainfall	100mm Rainfall Bins	50mm Rainfall Bins	NCC Rainfall Data
ln(Materials)				
Shortage	-0.969	-1.014	-1.137	-0.915
	(0.481)**	(0.431)**	(0.511)**	(0.426)**
Rainfall	-0.0118			
	(0.0228)			
First Stage F-Stat	14.99	17.56	14.23	18.42
ln(Workers)				
Shortage	-0.219	-0.228	-0.243	-0.249
	(0.337)	(0.301)	(0.339)	(0.297)
Rainfall	0.00649			
	(0.0152)			
First Stage F-Stat	14.93	17.49	14.19	18.30
ln(Earnings/Worker)				
Shortage	-0.181	-0.214	-0.267	-0.189
	(0.206)	(0.191)	(0.218)	(0.190)
Rainfall	0.00188			
	(0.0116)			
First Stage F-Stat	16.14	18.24	14.63	20.50
ln(Revenue)				
Shortage	-0.913	-0.988	-1.091	-0.792
	(0.504)*	(0.456)**	(0.536)**	(0.433)*
Rainfall	-0.0262			
	(0.0233)			
First Stage F-Stat	14.87	17.44	14.17	18.25
ln(TFPR)				
Shortage	-0.299	-0.294	-0.304	-0.235
	(0.254)	(0.232)	(0.259)	(0.221)
Rainfall	-0.0142			
	(0.0116)			
First Stage F-Stat	15.55	18.13	14.90	18.87

Notes: This table presents alternative estimates for Table 7, instrumenting for Shortage using hydro availability. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

E4. Alternative Instruments

This section presents estimates of Tables 6 and 7 with alternative instruments. Column 1 replicates the base estimates except using actual hydro generation instead of generation predicted from reservoirs and run-of-river plants. Columns 2 and 3 add N_{st}, the predicted generation from plants that came online in the previous year, as an additional supply shifter to increase power. Because Indian states are still not large compared to generation from a single plant, new plants generate lumpy reductions in shortages the year they come online. Power plants have a long and potentially unpredictable time-to-build, so we assume that the year that a plant comes online is exogenous conditional on state trends. The instrument in columns 2 and 3 is:

$$Z_{st} = \frac{H_{st} + N_{st}}{\tilde{Q}_{st}}. \tag{E1}$$

To get N_{st}, we simply multiply the capacity added in the previous year by the national average thermal plant capacity factor in year t. Column 2 uses H_{st} from reservoirs and run-of-river plants (as in the base estimates), while column 3 instead uses actual hydro generation (as in column 1).

The results below show that adding N_{st} provides a moderate increase in precision but does not otherwise change the results. We used this in an earlier working paper version, although it does not appear in the body of the published version due to concerns about the exogeneity of N_{st}.

Table A8—: **Alternative Instruments: Energy Inputs**

	(1)	(2)	(3)
	Base	With New Supply	
Instrument:	Actual Hydro Generation	Predicted with Run-of-River and Reservoirs	Actual Hydro Generation
		Self-Generation Share	
Shortage	0.794	0.463	0.788
	(0.176)***	(0.142)***	(0.167)***
Number of Obs.	240,743	240,743	240,743
First Stage F-Stat	17.61	19.74	19.44
		ln(Fuel Revenue Share)	
Shortage	3.597	3.318	3.596
	(1.049)***	(0.955)***	(1.003)***
Number of Obs.	291,759	291,759	291,759
First Stage F-Stat	17.76	19.36	19.72
		ln(Electric Intensity)	
Shortage	-1.392	0.173	-1.217
	(0.718)*	(0.698)	(0.673)*
Number of Obs.	479,616	479,616	479,616
First Stage F-Stat	14.24	17.73	16.18

Notes: This table presents estimates of Table 6 with alternative instruments. Samples for the first two panels are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A9—: **Alternative Instruments: Materials, Labor, Revenue, and TFPR**

	(1)	(2)	(3)
	Base	With New Supply	
	Actual Hydro Generation	Predicted with Run-of-River and Reservoirs	Actual Hydro Generation
Instrument:			
ln(Materials)			
Shortage	-1.370	-1.216	-1.415
	(0.607)**	(0.473)**	(0.576)**
Number of Obs.	495,043	495,043	495,043
First Stage F-Stat	13.88	17.05	15.89
ln(Workers)			
Shortage	-0.232	-0.302	-0.280
	(0.356)	(0.313)	(0.341)
Number of Obs.	502,724	502,724	502,724
First Stage F-Stat	13.82	16.99	15.82
ln(Earnings/Worker)			
Shortage	-0.542	-0.225	-0.487
	(0.270)**	(0.199)	(0.247)**
Number of Obs.	456,443	456,443	456,443
First Stage F-Stat	13.25	17.09	15.07
ln(Revenue)			
Shortage	-1.019	-1.182	-1.097
	(0.586)*	(0.498)**	(0.560)*
Number of Obs.	501,130	501,130	501,130
First Stage F-Stat	13.84	16.95	15.83
ln(TFPR)			
Shortage	0.158	-0.297	0.128
	(0.274)	(0.236)	(0.257)
Number of Obs.	479,313	479,313	479,313
First Stage F-Stat	14.21	17.75	16.22

Notes: This table presents estimates of Table 6 with alternative instruments. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

E5. Alternative TFPR Measures

Table A10 presents estimates of Equation (21), using alternative measures of TFPR described in Appendix C.C2.

Table A10—: **Robustness Check: Estimates with Alternative TFPR Measures**

	(1)	(2)	(3)	(4)	(5)	(6)
	Include All Fuels	Include No Fuels	No Time Trend	α Varies by Size	CRS	Leontief CRS
Shortage	-0.285	-0.150	-0.097	-0.112	-0.110	-0.211
	(0.240)	(0.248)	(0.211)	(0.212)	(0.221)	(0.266)
Number of Obs.	479,609	479,484	480,100	494,210	479,755	477,720
Number of Clusters	112,405	112,330	112,472	115,015	112,397	112,014
Number of Clusters (2)	536	536	536	536	536	536
First Stage F-Stat	14.87	14.88	14.84	14.32	14.85	14.89

Notes: F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

E6. Heterogeneous Effects of Shortages

Table A11 presents estimates of heterogeneous effects of shortages for plants with generators and for plants in industries with above-median electric intensity. Denote $\mathbf{M}_i$ as a 3-by-1 vector of these two moderators and a constant. The estimating equation is identical to Equation 21 except with $\mathbf{M}_i$ interacted with all right-hand-side variables other than μ_{jt}.

Table A11 presents the estimated interactions with the Shortage variable. As expected, column 2 shows that self-generators increase fuel use more when shortages worsen, while non-generators do not. However, we do not have the power to detect heterogeneous effects on revenues or TFPR. As a benchmark, in the World Bank Enterprise Survey, generators and non-generators report 7.3 and 8.4 percent losses from power cuts, respectively - a ratio of $8.4/7.3 \approx 1.15$. Our revenue estimates are statistically indistinguishable from this benchmark ratio.

These empirical results are not interpretable as the average causal effects of generator ownership, because endogenous generator adoption decisions could imply that the plants without generators have unobservably smaller losses. For example, plants without generators might have unobservably better electricity supply, reducing their losses from not adopting generators and also reducing the effects of an increase in shortages.

Table A11—: **Heterogeneous Effects of Shortages**

	(1)	(2)	(3)	(4)
Dependent Variable:	Self-Gen Share	ln(Fuel Rev Share)	ln(Revenue)	ln(TFPR)
Shortage	-0.027	-0.867	-0.478	-0.201
	(0.065)	(1.748)	(0.695)	(0.445)
Shortage x Elec Intensive	0.022	0.189	-0.936	0.130
	(0.131)	(2.181)	(1.212)	(0.551)
Shortage x Self-Generator	0.470	4.050	-0.384	-0.413
	(0.155)***	(1.956)**	(0.716)	(0.386)
Number of Obs.	428,969	477,005	501,130	479,313
Number of Clusters	102,995	109,715	116,231	112,371
Number of Clusters (2)	536	536	536	536

Notes: Robust standard errors, with two-way clustering by plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Simulation Appendix

This appendix presents full detail on the simulations, as well as additional robustness checks using different assumed production functions.

F1. Simulation Inputs

Table A1 presents the sources of the parameters used in the simulations.

Table A1—: **Simulation Inputs**

Parameter	Source	Level
α_K, α_L, α_M, α_E	Production function estimates from ASI	Industry Level
δ	Shortage S_{st} from CEA or other assumed value	State-Year
Generator ownership	Inferred from non-zero electricity generation in ASI	Plant
K	Capital stock in ASI	Plant-Year
Ω	Estimated revenue productivity	Plant-Year
p^M, p^L, p	Normalized to 1	Constant
$p^{E,G}$=4.5 Rs/kWh	Median grid electricity price from WBES	Constant
$p^{E,S}$=7 Rs/kWh	Median self-generated electricity cost from WBES	Constant

Notes: WBES refers to the 2005 World Bank Enterprise Survey.

F2. Exogenous Generators: Cobb-Douglas

This section presents full details on how the simulations in Section VI determine optimal input and output bundles conditional on exogenous generator ownership. Here we present the Cobb-Douglas model in Section III; subsequent sections present alternative models (Leontief and Constant Elasticity of Substitution).

The procedure takes production function parameters $\{\alpha_E, \alpha_L, \alpha_M, \alpha_K\}$ and exogenous state variables capital K_{it}, productivity Ω_{it}, and shortages δ_{it}. The optimal input choices of labor and materials are solved using profit maximization conditions, and L^*, M^{*G}, M^{*S}, E^{*G}, E^{*S}, and R^* are determined (where the superscripts S and G refer to shortage and non-shortage – grid – respectively) . This procedure is repeated for each plant i observed in the ASI data.

1) Plants without generators

   The optimal input bundles can be found analytically, using the first-order conditions. Non-generators shut down during outages, so $M^S = E^S = 0$. The first-order condition for materials during non-outage periods, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}} = 0$, yields

$$\alpha_M \Omega K^{\alpha_K} [M^G]^{\alpha_M - 1} L^{\alpha_L} [E^G]^{\alpha_E} = p^M. \tag{F1}$$

Analogously, the FOC for electricity, $\frac{\partial \pi_{it\tau}}{\partial E_{it\tau}} = 0$, yields

$$\alpha_E \Omega K^{\alpha_K} [M^G]^{\alpha_M} L^{\alpha_L} [E^G]^{\alpha_E - 1} = p^{E,G}. \tag{F2}$$

Finally, the FOC for labor, $\frac{\partial \pi_{it}}{\partial L_{it}} = 0$ yields

$$\alpha_L (1-\delta) \Omega K^{\alpha_K} [M^G]^{\alpha_M} L^{\alpha_L - 1} [E^G]^{\alpha_E} = p^L. \tag{F3}$$

Rearranging these three equations, we obtain:

(F4)

$$\begin{aligned} M^{*G} &= \left[\Omega K^{\alpha_K} \left(\frac{\alpha_M}{p_M} \right)^{1-\alpha_L-\alpha_E} \left(\frac{(1-\delta)\alpha_L}{p_L} \right)^{\alpha_L} \left(\frac{\alpha_E}{p^{E,G}} \right)^{\alpha_E} \right]^{\frac{1}{1-\alpha_L-\alpha_M-\alpha_E}} \\ L^* &= (1-\delta) \frac{p^M \alpha_L}{p^L \alpha_M} M^{*G} \\ E^{*G} &= \frac{p^M \alpha_E}{p^E \alpha_M} M^{*G} \end{aligned}$$

Annual revenue is thus:

$$R = (1-\delta) \Omega K^{\alpha_K} [M^{*G}]^{\alpha_M} [L^*]^{\alpha_L} [E^{*G}]^{\alpha_E}$$

Notice that if there are no shortages; $\delta = 0$, the same equations can also be used to determine optimal input bundles for all plants (assuming that $p^{E,G} < p^{E,S}$).

2) Plants with generators

There are five first-order conditions, $\frac{\partial \pi_{it\tau}}{\partial M^S_{it\tau}} = 0$, $\frac{\partial \pi_{it\tau}}{\partial M^G_{it\tau}} = 0$, $\frac{\partial \pi_{it\tau}}{\partial E^S_{it\tau}} = 0$, $\frac{\partial \pi_{it\tau}}{\partial E^G_{it\tau}} = 0$, and $\frac{\partial \pi_{it}}{\partial L_{it}} = 0$. These yield

$$\begin{aligned} \alpha_M \Omega K^{\alpha_K} [M^G]^{\alpha_M - 1} L^{\alpha_L} [E^G]^{\alpha_E} &= p^M \\ \alpha_M \Omega K^{\alpha_K} [M^S]^{\alpha_M - 1} L^{\alpha_L} [E^S]^{\alpha_E} &= p^M \\ \alpha_M \Omega K^{\alpha_K} [M^S]^{\alpha_M} L^{\alpha_L} [E^S]^{\alpha_E - 1} &= p^{E,S} \\ \alpha_M \Omega K^{\alpha_K} [M^G]^{\alpha_M} L^{\alpha_L} [E^G]^{\alpha_E - 1} &= p^{E,G} \\ \alpha_L (1-\delta) \Omega K^{\alpha_K} [M^G]^{\alpha_M} L^{\alpha_L - 1} [E^G]^{\alpha_E} & \\ + \alpha_L \delta p \Omega K^{\alpha_K} [M^S]^{\alpha_M} L^{\alpha_L - 1} [E^S]^{\alpha_E} &= p^L \end{aligned} \tag{F5}$$

The set of equations in system (F5) are solved numerically in MATLAB using the fsolve routine. Rather than solving for L^*, E^{*G}, $E^{*,S}$, M^{*G}, and

M^{*S} in levels, we solve in logarithms, since these values can differ by several orders of magnitude for different plants in the data. The starting values for L^*, E^{*G}, E^{*S}, M^{*G}, and M^{*S} are given by the analytic values from equation (F4), the no-shortage values.

Annual revenue is:

$$\begin{aligned} R =&(1-\delta)\Omega K^{\alpha_K}[M^{*G}]^{\alpha_M}[L^*]^{\alpha_L}[E^{*G}]^{\alpha_E} \\ &+ \delta\Omega K^{\alpha_K}[M^{*S}]^{\alpha_M}[L^*]^{\alpha_L}[E^{*S}]^{\alpha_E} \end{aligned}$$

F3. *Exogenous Generators: Leontief in Electricity*

Our original working paper used production functions that were Leontief in electricity and a Cobb-Douglas aggregate of capital, labor, and materials. For comparison, we include this below.

Denoting physical productivity as A, the physical production function is:

$$Q = \min\{AK^{\alpha_K}L^{\alpha_L}M^{\alpha_M}, \frac{1}{\lambda}E\} \tag{F6}$$

The Leontief production function dictates that electricity is used in constant proportion $\frac{1}{\lambda}$ with output. Electricity intensity λ varies across industries. Having A inside the Cobb-Douglas aggregator ensures that electricity is used in fixed proportion to output instead of to the bundle of other inputs.

Since we will observe total revenues rather than physical quantities produced, we need to relate revenues to our production function in equation (F6). We assume that plants sell into a perfectly competitive output market with price p, and denote $\Omega \equiv pA$.

Firms have the following daily profit function $\Pi_{it\tau}$:

$$\begin{aligned} \Pi_{it\tau} =&p \min\{A_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}M_{it\tau}^{\alpha_M}, \frac{1}{\lambda}E_{it\tau}\} \\ &- p^L L_{it} - p^M M_{it\tau} - p^E E_{it\tau}, \end{aligned} \tag{F7}$$

where p^L p^M are the prices of labor and materials, respectively. Capital is excluded, as it is sunk before the plant makes any production decisions.

Given the Leontief-in-electricity structure of production, cost minimization implies that for any desired level of output Q, the firm produces at a "corner" of the isoquant where:

$$A_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}M_{it\tau}^{\alpha_M} = \frac{1}{\lambda}E_{it\tau}, \tag{F8}$$

Given this, one can rewrite the profit function, substituting in Ω_{it} and the optimized value of electricity:

$$\Pi_{it\tau} = (1 - \frac{\lambda p^E}{p})\Omega_{it} A_{it} K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it\tau}^{\alpha_M} - p^L L_{it} - p^M M_{it\tau} \tag{F9}$$

Let $\gamma \equiv \frac{\lambda p^{E,G}}{p} = \frac{p^{E,G} E_{it\tau}}{pQ_{it\tau}} = \frac{p^{E,G} E_{it}}{R_{it}}$, the electricity revenue share if a firm only uses electricity purchased from the grid. Notice that if $(1-\gamma) < 0$, then the firm will choose not to produce.

There are three cases that can occur, depending on electricity intensity and the relative price of electricity:

1) If $p > \lambda p^{E,S}$, the plant always produces, regardless of power outages.
2) If $\lambda p^{E,S} > p > \lambda p^{E,G}$, the plant does not produce during power outages, but does produce otherwise.
3) If $p < \lambda p^{E,G}$, the plant never produces.

We ignore case (3): if plants never produce, they never appear in the data. Plants without generators effectively have $p^{E,S} = \infty$, so case (1) cannot arise. Of the plants with generators, those with higher λ will be in case (2). In other words, higher-electricity intensity plants will be more likely to shut down during grid power outages.[5]

The first-order condition with respect to materials yields:

$$\alpha_M (1-\gamma) \frac{R_{it\tau}}{M_{it\tau}} - p^M = 0. \tag{F10}$$

The marginal revenue product of materials is:

$$MRPM = \begin{cases} \alpha_M (1-\gamma) \frac{R_{it\tau}}{M_{it\tau}} & \text{if grid power} \\ \mathcal{T} \alpha_M (1-\gamma) \frac{R_{it\tau}}{M_{it\tau}} & \text{if power outage} \end{cases} \tag{F11}$$

When setting labor, the firm begins with its yearly profit function, which is simply the weighted average of Equation (F9) over grid power and outage periods. If a plant is in case (1), meaning that it self-generates during power outages, then the first-order condition is given by:

$$MRPL = \alpha_L (1-\gamma) \left[(1-\delta) \frac{R_{it}^G}{L_{it}} + \delta \mathcal{T} \frac{R_{it}^S}{L_{it}} \right] = p^L, \tag{F12}$$

where R_{it}^S and R_{it}^G indicate revenues during outage and grid power periods, respectively.

[5] While a firm would not invest in a generator if it expected to be in case (2), unexpected changes in p, $p^{E,S}$, or $p^{E,G}$ could cause firms with generators to not use them.

We now solve for profit-maximizing inputs and output. Define $\gamma^G \equiv \frac{\lambda p^{E,G}}{p}$ and $\gamma^S \equiv \frac{\lambda p^{E,S}}{p}$.

1) If in Case 3:

   The plant never produces. Thus $L^* = M^{*G} = M^{*S} = E^{*G} = E^{*S} = R^* = 0$.

2) If in Case 2 (including non-generators with $\gamma^G < 1$):

   The plant operates when there is grid power, but not during an outage. The optimal input bundles can be found analytically, using the first-order conditions. Clearly, $M^{*S} = 0$. The FOC for materials during grid power periods, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}} = 0$, yields

$$\alpha_M(1-\gamma^G)\Omega K^{\alpha_K}[M^G]^{\alpha_M-1}L^{\alpha_L} = p^M. \tag{F13}$$

   The FOC for labor, $\frac{\partial \pi_{it}}{\partial L_{it}} = 0$ yields

$$\alpha_L(1-\gamma^G)(1-\delta)\Omega K^{\alpha_K}[M^G]^{\alpha_M}L^{\alpha_L-1} = p^L. \tag{F14}$$

   Rearranging these first-order equation, we obtain

$$\begin{aligned} M^{*G} &= \left[(1-\gamma)\Omega K^{\alpha_K}\left(\frac{\alpha_M}{p_M}\right)^{1-\alpha_L}\left(\frac{(1-\delta)\alpha_L}{p_L}\right)^{\alpha_L}\right]^{\frac{1}{1-\alpha_L-\alpha_M}} \\ L^* &= (1-\delta)\frac{p^M\alpha_L}{p^L\alpha_M}M^{*G} \end{aligned} \tag{F15}$$

   Given labor and material choices, it is straightforward to compute revenue:

$$R = \Omega K^{\alpha_K}[M^{*G}]^{\alpha_M}[L^*]^{\alpha_L}.$$

   Electricity consumption is:

$$E^{*G} = \gamma^S R$$

   Notice that if there are no shortages; $\delta = 0$, the same equations can also be used to get optimal input bundles and revenue for all plants.

3) If in Case 1 (plants with generators only):

   The plant always operates, running its generator during outages. There are three first-order conditions, $\frac{\partial \pi_{it\tau}}{\partial M^S_{it\tau}} = 0$, $\frac{\partial \pi_{it\tau}}{\partial M^G_{it\tau}} = 0$, and $\frac{\partial \pi_{it}}{\partial L_{it}} = 0$.

$$
\begin{aligned}
&\alpha_M(1-\gamma^G)p\Omega K^{\alpha_K}[M^G]^{\alpha_M-1}L^{\alpha_L} = p^M \\
&\alpha_M(1-\gamma^S)p\Omega K^{\alpha_K}[M^S]^{\alpha_M-1}L^{\alpha_L} = p^M \\
&\alpha_L(1-\gamma^G)(1-\delta)p\Omega K^{\alpha_K}[M^G]^{\alpha_M}L^{\alpha_L-1} \\
&\quad +\alpha_L\delta(1-\gamma^S)p\Omega K^{\alpha_K}[M^S]^{\alpha_M}L^{\alpha_L-1} = p^L
\end{aligned} \tag{F16}
$$

The system of equations in (F16) is solved numerically in MATLAB using the fsolve routine.

Finally, electricity usage is

$$
\begin{aligned}
E^{*G} &= (1-\delta)\gamma^G\Omega K^{\alpha_K}[M^{*G}]^{\alpha_M}[L^*]^{\alpha_L} \\
E^{*S} &= \delta\gamma^S\Omega K^{\alpha_K}[M^{*G}]^{\alpha_M}[L^*]^{\alpha_L}.
\end{aligned}
$$

Annual revenue is:

$$
\begin{aligned}
R =&(1-\delta)\Omega K^{\alpha_K}[M^{*G}]^{\alpha_M}[L^*]^{\alpha_L} \\
&+\delta\Omega K^{\alpha_K}[M^{*S}]^{\alpha_M}[L^*]^{\alpha_L}
\end{aligned}
$$

F4. Exogenous Generators: Constant Elasticity of Substitution

One issue with the production functions that we have considered is that there is no direct intertemporal substitution in production. Suppose instead that we consider a CES aggregator with constant elasticity of substitution σ between days given by:

$$
R_{it} = \left[\int_\tau (R_{it\tau})^\sigma d\tau\right]^{\frac{1}{\sigma}}
$$

Notice that this is a CES type aggregator, so there is symmetric substitution across all days of the year. If $\sigma = 1$, we have the process considered in the paper. If $\sigma < 1$, outputs are interday complements, and if $\sigma > 1$, then there is inter-day substitution.

Given that in the daily production function, only materials and electricity can be varied, we can think of the daily production function as being written as:

$$
R_{it} = \Omega_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}\left[\int_\tau {E_{it\tau}}^{\sigma\alpha_E}{M_{it\tau}}^{\sigma\alpha_M}d\tau\right]^{\frac{1}{\sigma}}
$$

Notice that the daily returns to scale in the production function will be given by $\sigma(\alpha_M + \alpha_E)$, an issue we return to below.

First-Order Conditions for Non-Generators

For firms that do not have generators, revenue is:

$$\begin{aligned}
R_{it} &= \left[\int_\tau (R_{it\tau})^\sigma d\tau\right]^{\frac{1}{\sigma}} \\
&= [(1-\delta) M_{it\tau}^{\sigma\alpha_M} E_{it\tau}^{\sigma\alpha_E} L_{it}^{\sigma\alpha_L} K_{it}^{\sigma\alpha_K}]^{\frac{1}{\sigma}} \\
&= (1-\delta)^{\frac{1}{\sigma}} M_{it\tau}^{\alpha_M} L_{it}^{\alpha_L} E_{it\tau}^{\alpha_E} K_{it}^{\alpha_K} \\
&= (1-\delta)^{\frac{1}{\sigma}} \left(\frac{M_{it}}{(1-\delta)}\right)^{\alpha_M} \left(\frac{E_{it}}{(1-\delta)}\right)^{\alpha_E} L_{it}^{\alpha_L} K_{it}^{\alpha_K} \\
&= (1-\delta)^{\frac{1}{\sigma}-\alpha_M-\alpha_E} M_{it}^{\alpha_M} E_{it}^{\alpha_E} L_{it}^{\alpha_L} K_{it}^{\alpha_K}
\end{aligned} \tag{F17}$$

where we have assumed that the same input choice will be optimal across days to go from the first to the second line of this equation. Notice that shortages can cause anything between a zero and infinite decrease in revenues by changing σ, holding inputs fixed. Thus, our predictions are not robust to a range of σ. Moreover, the only difference between this setup and the setup without intertemporal substitution is that the plant's revenue decreases by $(1-\delta)^{\frac{1}{\sigma}-\alpha_M-\alpha_E}$ instead of $(1-\delta)^{1-\alpha_M-\alpha_E}$.

First-Order Conditions for Generators

The profit function is:

$$\Pi_{it} = R_{it} - p^M \int_\tau M_{it\tau} d\tau - p^E \int_\tau E_{it\tau} d\tau - p^L \int_\tau L_{it\tau} d\tau \tag{F18}$$

For plants with generators, the materials first-order condition, $\frac{\partial \Pi_{it}}{\partial M_{it\tau}} = 0$, yields

$$\frac{1}{\sigma} R_{it}^{\frac{\frac{1}{\sigma}-1}{\frac{1}{\sigma}}} \sigma \alpha_M \frac{R_{it\tau}^\sigma}{M_{it\tau}} = p^M. \tag{F19}$$

The electricity FOC is similar:

$$\frac{1}{\sigma} R_{it}^{\frac{\frac{1}{\sigma}-1}{\frac{1}{\sigma}}} \alpha_E \frac{R_{it\tau}^\sigma}{E_{it\tau}} = p^E. \tag{F20}$$

The labor FOC, $\frac{\partial \Pi_{it}}{\partial L_{it}} = 0$, yields:

$$\text{(F21)} \qquad \frac{1}{\sigma} R_{it}^{\frac{\frac{1}{\sigma}-1}{\frac{1}{\sigma}}} \sigma \alpha_L \left[(1-\delta) \frac{\left(R_{it}^G\right)^{\sigma}}{L_{it\tau}} + \delta \frac{\left(R_{it}^S\right)^{\sigma}}{L_{it\tau}^S} \right] = p^L,$$

where $R_{it}^G = \Omega L^{\alpha_L} (M^G)^{\alpha_M} (E^G)^{\alpha_E} K^{\alpha_K}$ and $R_{it}^S = \Omega L^{\alpha_L} (M^S)^{\alpha_M} (E^S)^{\alpha_E} K^{\alpha_K}$. The set of the equations (F19,F20,F21) are solved numerically in MATLAB using the fsolve routine. Rather than solving for L^*, E^{*G}, $E^{*,S}$, M^{*G}, and M^{*S} in levels, we solve in logarithms.

F5. Comparing Predictions from Cobb-Douglas, Leontief, and CES Models

Table A2 presents the simulated effects of the 2005 assessed Shortage levels S_{s2005} relative to zero shortage. Indeed, Table A2 replicates Panel A of Table 9. Column 1 shows the Cobb-Douglas production function used in the body of the paper. Column 2 presents the Leontief model, while columns 3 and 4 present results of the CES model with $\sigma = 0.9$ and $\sigma = 0.5$, respectively.[6] The predictions from the Cobb-Douglas and Leontief model are virtually identical, despite the different functional forms and different approaches to production function estimation. In the CES model, simulated losses are almost identical for plants with generators, but much larger for non-generators.

Table A2—: **Predictions from Different Production Function Models**

	Cobb-Douglas (1)	Leontief (2)	CES $\sigma = 0.9$ (3)	CES $\sigma = 0.5$ (3)
Revenue Loss: Average	5.6%	5.7%	7.5%	18%
Revenue Loss: Non-Generators	10.0%	9.8%	13%	32%
Revenue Loss: Generators	0.4%	0.6%	0.4%	0.4%
TFPR Loss: Average	1.5%	1.3%	1.9%	5.8%
TFPR Loss: Non-Generators	2.6%	2.4%	3.5%	10.6%
TFPR Loss: Generators	0.0%	0.0%	0.0%	0.0%

Notes: This table presents the effects of the 2005 assessed Shortage levels relative to zero Shortage. Cobb-Douglas, Leontief, and CES, refer to the production functions used for estimation and prediction, and are described in text.

[6] Using a higher value of σ such as $\sigma = 1.5$ yields the implication that a firm will have increasing returns at the daily level, since the returns to scale in the daily production function are $\sigma(\alpha_M + \alpha_E)$, so $\sigma(\alpha_M + \alpha_E) > 1$ means that it is optimal to produce all output on a single day of the year. For the CES simulation, we use production function coefficients α_E, α_L, α_M, α_K, and Ω, estimated from the Cobb-Douglas model. However, for non-generators, the CES α coefficients can be estimated using the same equations as in the Cobb-Douglas model, and recall that there is little difference in the estimates for generators and non-generators.

F6. Endogenous Generators

For the model with endogenous generators, the equations given in Online Appendix F.F2 above are used to obtain the optimal input bundle, *conditional* on the presence of a generator at the plant. In this model, however, we also endogenously solve for generator adoption.

Plant i purchases a generator if and only if $\Pi_{it}^{\mathcal{G}} - C_{it}^{\mathcal{G}} > \Pi^{\mathcal{NG}}$, where $\Pi^{\mathcal{G}}$ and $\Pi^{\mathcal{NG}}$ are profitability with and without generators, respectively, and $C_{it}^{\mathcal{G}}$ is the annualized generator cost.

Profits $\Pi^{\mathcal{G}}$ and $\Pi^{\mathcal{NG}}$ are both

$$\Pi_{it} = R_{it} - p^L L^* - p^M \left(\delta M^{*G} + (1-\delta) M^{*S}\right) - \delta p^{E,S} E^{*S} - (1-\delta) p^{E,G} E^{*G}, \tag{F22}$$

where optimal inputs are according to the equations in Online Appendix F.F2 above.

Estimating Generator Costs

We parametrize the generator cost as $\ln C_{it}^{\mathcal{G}} = \sigma_0 + \sigma_1 \ln(c_{it})$, where σ_1 is the economy of scale parameter, and c_{it} is the generator capacity in kilowatts. We estimate σ_0 and σ_1 using GMM, matching the mean generator adoption rate and the covariance between generator adoption and log generator capacity. We use the identity matrix as a weighting matrix, since the two moments that we match are of comparable scales. Column 1 of Table A3 presents GMM estimates of the generator cost function.

For comparison, we collected generator purchase price data from the United States. To compare to the estimated $C^{\mathcal{G}}$, we must first convert the purchase prices into yearly rental prices. First, we convert generators rated in KVA into KW using a 0.8 conversion ratio. Second, we convert US dollars into Rupees using a 50 to 1 exchange rate. Finally, we convert the purchase price of a generator into an annual rental price assuming a 30 percent discount rate, a ten percent depreciation rate, and a ten-year generator lifespan.[7] This gives a 1.6:1 ratio between generator costs and rental rates. Column 2 presents a regression of the natural logs of these observed prices on natural log of capacity.

The estimates of σ_1 are close to 0.8 in both columns of Table A3. The estimates of σ_0 are also comparable (10.67 vs. 11.14), although the point estimate in column 1 is smaller. This gives us some confidence that the estimated generator costs are approximately reasonable and that generator costs can explain the fact that many manufacturing plants in the ASI do not have generators.

[7] This 30 percent discount rate is high by U.S. standards, but as Banerjee and Duflo (2014) discuss, Indian firms pay far higher interest rates - on the order of 30 to 60 percent, if they have access to capital at all.

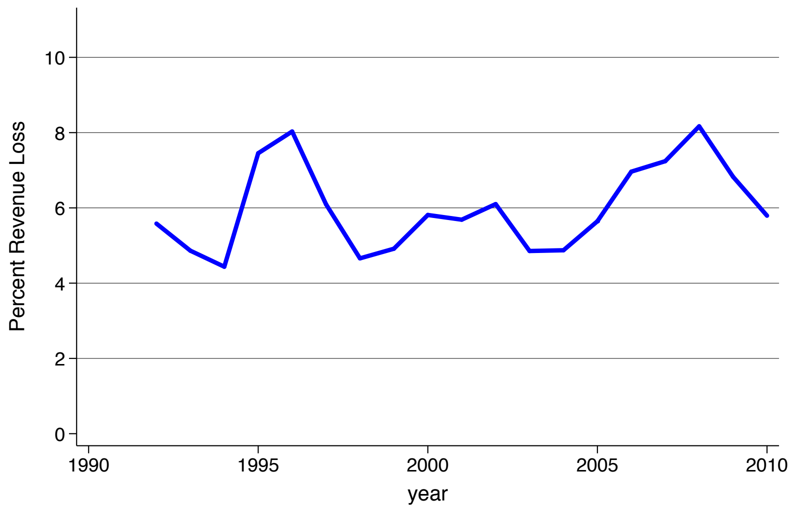
Table A3—: **Generator Cost Estimates**

	(1)	(2)
	GMM Estimates	Observed Prices
σ_0	10.67	11.14
	(0.18)	(0.25)
σ_1	0.83	0.79
(economy of scale)	(0.01)	(0.05)
Observations	33,871	223
Predicted Generator Takeup Rate	0.53	0.47
Covariance: Generator Takeup and Generator Size	0.63	0.72

Notes: Column 1 shows estimates of generator cost using generator adoption decisions via GMM as described in text. Column 2 shows a regression of log generator rental rate on log generator capacity.

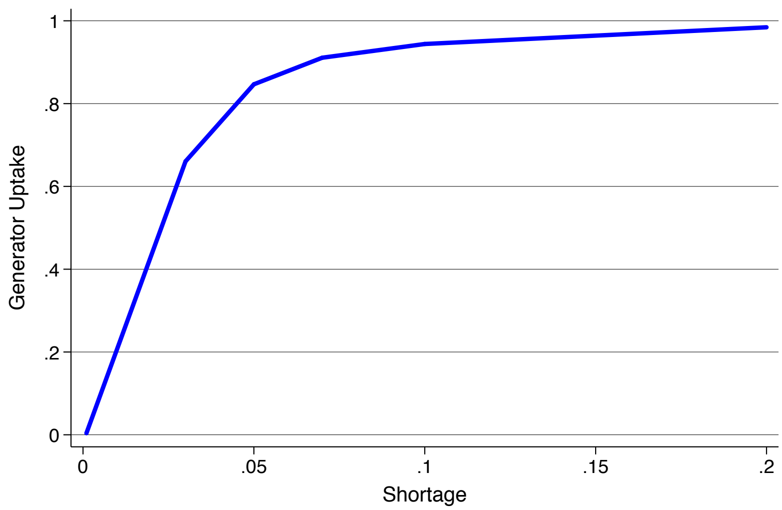
F7. Additional Simulation Figures and Tables

Figure A1. : **Predicted Average Revenue Loss by Simulation Year**



Notes: In the body of the paper, we simulate effects of moving from $\delta = 0$ to $\delta = S_{s2005}$ for all plants in the data in 2005. This figure presents revenue effects from the same simulations for each year of the 1992-2010 sample, i.e. taking the sample of plants in year t and changing δ from $\delta = 0$ to $\delta = S_{st}$.

Figure A2. : **Generator Adoption Under Varying Shortage Levels**



Notes: These figures show the simulated generator adoption rate when the δ on the x-axis is assigned to all plants in the 2005 ASI, using the generator adoption model in Section VI.VI.A. Note that generator takeup exceeds 90 percent at a seven percent δ, which may seem puzzling given that the generator cost estimates are based on a 44 percent takeup rate in the ASI at a 7.2 percent mean shortage. The reason is that the distribution of S_{s2005} across plants is right skewed; the median of S_{s2005} across plants is only 3.5 percent.

Table A4—: **Effect of Shortages: Semi-elasticities from Model and IV Estimates**

	(1)	(2)	(3)
	Simulation	IV Estimate	p-Value for Columns (1) vs. (2)
Self-Generation Share Increase	0.29%	0.44%	(0.33)
Materials Reduction	0.91%	1.14%	(0.66)
Labor Reduction	0.91%	0.24%	(0.05)
Revenue Loss	0.91%	1.09%	(0.74)
TFPR Loss	0.19%	0.30%	(0.66)

Notes: This table parallels columns 1 and 2 of Panel A of Table 9, except that it presents semi-elasticities, i.e. the effect of a one percentage point increase in shortages on percent changes in the dependent variable.

Table A5—: **Counterfactuals Under Varying Shortage Levels**

	(1)	(2)	(3)	(4)	(5)
Shortage Percent (δ):	3%	5%	7%	10%	20%
Exogenous Generators					
Revenue Loss: Average	2.5%	4.2%	5.8%	8.3%	16%
Revenue Loss: Generators	0.2%	0.3%	0.5%	0.6%	1.3%
Revenue Loss: Non-Generators	4.5%	7.4%	10%	15%	28%
TFPR Loss: Average	0.5%	0.9%	1.3%	1.9%	3.9%
Input Cost Increase: Generators	0.2%	0.3%	0.4%	0.5%	1.0%
Variable Profit Loss: Average	2.5%	4.2%	5.8%	8.2%	16%
Generator Cost (Percent of Profits)	3.0%	3.0%	3.0%	3.1%	3.2%
Total Profit Loss: Average	5.5%	7.2%	8.8%	12%	19%
Endogenous Generators					
Generator Take-up	66%	85%	91%	94%	98%
Revenue Loss: Average	1.7%	1.4%	1.3%	1.3%	1.7%
Revenue Loss: Generators	0.2%	0.3%	0.4%	0.6%	1.2%
Revenue Loss: Non-Generators	4.2%	7.1%	10%	15%	29%
TFPR Loss: Average	0.4%	0.2%	0.2%	0.2%	0.1%
Input Cost Increase: Generators	0.1%	0.2%	0.3%	0.5%	1.0%
Variable Profit Loss: Average	1.7%	1.4%	1.2%	1.2%	1.5%
Generator Cost (Percent of Profits)	1.6%	2.7%	3.4%	3.8%	4.6%
Total Profit Loss: Average	3.3%	4.1%	4.6%	5.1%	6.1%

Notes: This table presents predictions of the simulation model described in the text. The simulations with “exogenous” generators hold fixed the generator adoption decision observed in the ASI, while the simulations with “endogenous” generators use the model’s prediction of which plants will purchase generators at the different shortage levels. Input Cost Increase is reported as a share of revenues. In this table, the electricity shortage is uniform across all plants in all states.

*

REFERENCES

Banerjee, Abhijit and Esther Duflo (2014). "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." *Review of Economic Studies*, Vol. 81, pp. 572–607.

MOSPI (Ministry of Statistics and Programme Implementation (2014). "Annual Survey of Industries." Available at: http://mospi.nic.in/Mospi_New/upload/asi/ASI_main.htm?status=1&menu_id=88. Accessed 17 December 2014.