HOW DO ELECTRICITY SHORTAGES AFFECT PRODUCTIVITY? EVIDENCE
FROM INDIA

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

How Do Electricity Shortages Affect Productivity? Evidence from India
Hunt Allcott, Allan Collard-Wexler, and Stephen D. O'Connell
NBER Working Paper No. 19977
March 2014, Revised February 2015
JEL No. D04,D24,L11,L94,O12,O13,Q41

ABSTRACT

We estimate the effects of electricity shortages on Indian manufacturers, instrumenting with supply shifts from hydroelectric power availability and power plant construction. We estimate that India's average reported level of shortages reduces the average plant's revenues and producer surplus by five to ten percent, but average productivity losses are significantly smaller because most inputs can be stored during outages. Shortages distort the plant size distribution, as there are significant economies of scale in generator costs and shortages more severely affect plants without generators. Simulations show that offering interruptible retail electricity contracts could substantially reduce the impacts of shortages.

Hunt Allcott
Department of Economics
New York University
19 W. 4th Street, 6th Floor
New York, NY 10012
and NBER
hunt.allcott@nyu.edu

Allan Collard-Wexler
Department of Economics
Duke University
233 Social Sciences
Durham, NC 27708
and NBER
allan.collard-wexler@duke.edu

Stephen D. O'Connell
City University of New York
Department of Economics
The Graduate Center
365 Fifth Ave
New York, NY 10016-4309
soconnell@gc.cuny.edu

1 Introduction

One of the potential contributors to the large productivity gap between developed and developing countries is low-quality infrastructure, and one of the most stark examples of infrastructure failures is electricity supply in India. In the summer of 2012, India suffered the largest power failure in history, a cascading blackout that plunged 600 million people into darkness at its peak (Yardley and Harris 2012). Even under normal circumstances, however, the Indian government estimates that shortages currently amount to about ten percent of demand at current prices. In the 2005 World Bank Enterprise Survey, one-third of Indian business managers named poor electricity supply as their biggest barrier to growth. According to these managers, blackouts are far more important than other barriers that economists frequently study, including taxes, corruption, credit, regulation, and low human capital.[1]

In this paper, we ask: *how do electricity shortages affect input choices, revenue, and productivity in the Indian manufacturing sector*? One potential prior is that because electricity is an essential input - most factories cannot produce anything without electricity for lights, motors, and machines - shortages could significantly reduce output. On the other hand, many firms might insure themselves against outages by purchasing generators or otherwise substituting away from grid electricity precisely because the potential losses are so large. The limited existing evidence could support either argument. Foster and Steinbuks (2009) and others argue that the cost of self-generation is relatively small, and Alam (2013) and Fisher-Vanden, Mansur, and Wang (2015) highlight ways in which plants substitute away from electricity when shortages worsen. In contrast, Hulten, Bennathan, and Srinivasan (2006) argue that growth of roads and electric generation capacity accounts for a remarkable 50 percent of productivity growth in Indian manufacturing between 1972 and 1992.

There are at least two reasons why this question is difficult to answer empirically. First, the necessary data on electricity shortages are typically not available: countries that have shortages are often the same types of countries that do not gather and disclose high-quality data on their infrastructure. Second, shortages are not exogenous to productivity or production. For example, rapid economic growth could cause an increase in electricity demand that leads to shortages, or poor institutions could lead to insufficient power supply and also reduce productivity. Either of these two mechanisms would bias estimates of the effects of shortages, albeit in opposite directions.

We begin by detailing an extensive array of data that we have gathered on the Indian electric power sector, including official state-specific electricity shortage estimates dating to 1992. To our knowledge, these are the only systematic accounts of electricity shortages available historically in any country suffering endemic blackouts. We document how electricity supply in India has continually lagged demand over the past 20 years, but shortage levels vary substantially within states over time.

We present a modified Cobb-Douglas production function model to predict how variation in

---

[1]For a tally of responses, see Online Appendix Table A1.

electricity shortages affects existing plants. For plants with generators, shortages act like a time-varying electricity input tax: during a grid power outage, the plants self-generate electricity at higher cost. Plants without generators shut down during outages, as if hit by an infinite input tax. This "input tax effect" causes all plants to contract, especially those without generators. The primary productivity loss is that plants without generators waste non-flexible inputs. For example, when textile plants shut down, their buildings and machines continue depreciating, but they leave thread on the looms without waste. Percent revenue losses must exceed percent productivity losses, because productivity feeds directly into revenue and plants' contraction due to the input tax effect further reduces revenue.

Drawing on the model, we then estimate how variation in shortages affects plants in India's official manufacturing survey, the Annual Survey of Industries (ASI). We instrument for shortages with shifts in electricity supply from new power plant construction and hydroelectric power availability, conditional on state-level rainfall and other controls. In support of the exclusion restriction, we show that these supply shifts are not conditionally associated with agricultural output, electricity prices, or official estimates of what demand would be in the absence of shortages.

Our instrumental variables estimates show that shortages have positive but economically small effect on variable energy input costs for plants with generators: a one percentage point increase in shortages increases their average fuel expenditures by 0.18 percent of revenues, and this is largely offset by the decrease in grid electricity purchases. Average revenues drop by 1.2 percent, but materials input drops by almost exactly that same amount. Since materials represent 70 percent of revenues on average, revenue productivity (TFPR) does not decrease nearly as much as revenues. The results are economically similar and statistically indistinguishable under a battery of alternative specifications.

As the instrumental variables estimates are identified by annual variation in supply shifters, they capture primarily "short-run" effects of shortages, i.e. holding constant decisions such as generator capital stock and plant entry and exit. To shed light on long-run effects, we briefly study the association between plant characteristics and the average shortages in the two years preceding plant entry. We find suggestive evidence that plants in electricity-intensive industries are less likely to enter when shortages worsen, implying that shortages may have deeper effects on the composition of Indian industry.

Finally, we apply our production function model to ASI plants to simulate the effects of shortages. Analogous to Todd and Wolpin (2006), we validate the structural model using the agreement of the model's prediction with the reduced form results. The simulated effects and IV estimates are statistically indistinguishable, which builds confidence that the estimates are reasonable and the model captures the first-order effects of shortages. The officially-assessed level of shortages is controversial because it is difficult to accurately assess demand in the absence of shortages. Subject to that caveat, we simulate that the assessed level of shortages reduced producer surplus by 9.5 percent, revenues by 5.6 percent, and productivity by 1.5 percent for the average plant in 2005.

Aside from these headline numbers, the simulations deliver two additional insights. First, shortages more severely affect plants that do not have generators, and generator costs have significant economies of scale. We simulate that as a result, variable profit losses average 2-3 times larger for small plants compared to large plants, which could distort India's plant size distribution in favor of large plants. Second, we simulate the effect of interruptible electricity contracts, which offer plants reduced retail prices in exchange for accepting more frequent power outages. These contracts efficiently allocate shortages to plants that can best deal with them, and our simulations show that if implemented nationwide, they could reduce producer surplus losses by more than an order of magnitude. While interruptible contracts do require additional physical infrastructure to implement, they may be a useful partial solution in the Indian context because political barriers have prevented other market reforms to reduce shortages.

The remainder of this section discusses related literature. Section 2 details the data. Section 3 provides background on the Indian electricity sector, the causes of electricity shortages, and manufacturers' responses to shortages. Section 4 presents the production function model and TFPR estimates. Sections 5 and 6 present the empirical strategy and results. Section 7 details the counterfactual simulations, and Section 8 concludes.

1.1 Related Literature

Our paper builds on an extensive literature that estimates the economic effects of investment in electricity, transportation, and other infrastructure. One early group of studies examines the effects of infrastructure investment on growth in panel data from U.S. states, including Aschauer (1989), Holtz-Eakin (1994), Fernald (1999), Garcia-Mila, McGuire, and Porter (1996); see Gramlich (1994) for a review. Easterly and Rebelo (1993), Esfahani and Ramirez (2002), and Roller and Waverman (2001) carry out analogous studies using cross-country panels.

The cross-state and cross-country literatures faced two basic problems. First, infrastructure spending is often endogenous to economic growth. Second, using aggregate infrastructure spending or quantity as the independent variable often hides important variation in effects between infrastructure of different types or quality levels. In the Indian context, for example, spending on power plants does not necessarily translate into electricity provision, because plants are frequently offline due to mechanical failure or fuel shortages.

Our paper is part of a recently-growing literature that evaluates the effects of infrastructure by combining microdata with within-country variation generated by natural experiments. This includes Banerjee, Duflo, and Qian (2012), Donaldson (2012), and Donaldson and Hornbeck (2013) on the effects of railroads in China, India, and the United States, Duflo and Pande (2007) on irrigation dams in India, Jensen (2007) on information technology, Baisa, Davis, Salant, and Wilcox (2008) on the benefits of reliable water provision in Mexico, and Baum-Snow (2007, 2013), Baum-Snow, Brandt, Henderson, Turner, and Zhang (2013), and Baum-Snow and Turner (2012) on urban transport expansions in China and the United States.

A subset of this literature focuses on electricity supply: Chakravorty, Pelli, and Marchand (2013), Dinkelman (2011), Lipscomb, Mobarak, and Barham (2013), and Rud (2012a) study the effects of electricity grid expansions, while Alby, Dethier, and Straub (2011), Foster and Steinbuks (2009), Steinbuks (2011), Steinbuks and Foster (2010), Reinikka and Svensson (2002), and Rud (2012b) study firms' generator investment decisions. Several recent papers focus specifically on Indian electricity supply: Ryan (2013) estimates the potential welfare gains from expanding transmission infrastructure, Cropper, Limonov, Malik, and Singh (2011) and Chan, Cropper, and Malik (2014) study the efficiency of Indian coal power plants, Abeberese (2012) tests how changes in electricity prices affect manufacturing productivity, and Alam (2013) studies how India's steel and rice milling industries respond differently to variable electricity supply. Fisher-Vanden, Mansur, and Wang (2015) is perhaps the most closely related paper to ours. They quantify the impacts of electricity shortages in the early 2000s on a sample of the largest Chinese manufacturing firms, finding that as shortages worsened, firms purchased more electricity-intensive inputs.

Our main contribution to the literature is to estimate the effects of electricity shortages across an entire country's manufacturing sector. Such an aggregate estimate is important because while Indian policymakers are well aware that shortages are a problem, India also has many other problems. Quantifying the losses from this and other distortions helps policymakers to allocate scarce time and political capital to the most "binding" constraints to growth, as suggested by the framework of Hausmann, Rodrik, and Velasco (2008). Aside from quantifying the magnitude of the problem, we also quantify a potential partial solution: interruptible contracts, which due to their technocratic nature could be more politically feasible than market liberalization. In addition to policy insights, we provide additional economic insights about industry in developing countries: we show how shortages might affect the plant size distribution and point out that while shortages might substantially affect manufacturing output, the short-run effects we estimate explain little of the manufacturing productivity gap between India and more developed countries.[2]

2 Data

We have collected comprehensive data from 1992 to 2010 on weather, the power sector, and manufacturing production.[3] All financial amounts are deflated to real 2004 Rupees (Rs).[4] Throughout the paper, we use the word "state" to refer to states, Union Territories, and the National Capital Region (New Delhi).

---

[2]See Tybout (2000) and Hsieh and Olken (2014) for a broader discussion of the firm size distribution in developing countries. See Banerjee and Duflo (2005), Hall and Jones (1999), Hsieh and Klenow (2009), and others for discussions of the manufacturing productivity gap.

[3]All data are originally reported in, or calculated to correspond to, the Indian fiscal year, which is April 1 through March 31. In this paper, "year" thus refers to the fiscal year, and for simplicity we refer to only the fiscal year's initial calendar year. (For example, "1998" always means "April 1998 through March 1999.")

[4]The exchange rate was approximately Rs 50 per U.S. dollar at that time.

Table 1: **State-by-Year Summary Statistics on Weather and Power Sector Data**

	Mean	SD	Min.	Max.	Obs.
Rainfall (meters)	1.33	0.75	0.26	5.02	536
Average Cooling Degrees (F)	12.2	3.33	2.67	18.3	536
Assessed Demand (TWh)	20.5	22.4	0.14	128	536
Energy Available (TWh)	18.6	19.9	0.12	107	536
Shortage	0.076	0.075	0	0.36	536
Peak Shortage	0.12	0.11	0	0.50	536
Total Electricity Sold (TWh)	14.0	15.0	0.08	87.5	536
Hydro Generation (TWh)	2.61	3.13	0.00	15.3	536
Hydro Capacity (MW)	840	969	0	3618	536
Total Capacity (MW)	2744	3099	0	16062	536
Reservoir Inflows (billion cubic meters)	8.78	16.4	0	116	536
Run-of-River Generation (TWh)	0.33	0.95	0	8.89	536
Capacity Added in Previous Year (MW)	117	250	-472	2070	536

Notes: See text for variable sources and definitions. TWh stands for terawatt-hours of electricity, and MW stands for megawatts of capacity.

2.1 Weather Data

Rainfall data are from the University of Delaware, which provides monthly rainfall for geographic gridpoints spaced at 1/2 degree intervals (Willmott and Matsuura 2012). We sum to total annual rainfall by gridpoint, then calculate state-by-year average rainfall by averaging across all gridpoints within each state. Temperature data are from the National Climate Centre, which provides daily temperatures for geographic gridpoints at one degree intervals (Srivastava, Rajeevan, and Kshirsagar 2009). For each day at each gridpoint, we construct cooling degrees in Fahrenheit: max{0,Day's Average Temperature - 65}. We then calculate state-by-year average cooling degrees by averaging across gridpoints within each state.[5] Table 1 summarizes the state-by-year observations of weather data, as well as the power sector data described below.

2.2 Power Sector Data

Power sector data are from India's Central Electricity Authority (CEA). The CEA collects many of the same data that the U.S. Energy Information Administration makes available online. Unfortunately, the CEA's website includes only a scattered set of recent information, and their on-site archive of hard copies is incomplete, so individual data series must be hand-collected from CEA staff. With the cooperation of CEA management and the help of several research assistants in New Delhi, we gathered, digitized, and cleaned extensive data on the Indian power sector dating back to 1992 or before.

---

[5]"Rainfall" is more precisely "precipitation," as it includes winter snowfall in the Himalayan states. University of Delaware and NCC both provide precipitation and temperatures. We use the University of Delaware rainfall because of the finer geographic scale, although the two data sources are extremely highly correlated. We need daily temperatures to construct cooling degrees, and the University Delaware only provides monthly average temperatures.

The primary measure of electricity shortages is the percent energy deficit reported in the Load Generation Balance Reports (CEA 1993-2011b). At the end of each year, analysts from the CEA and Regional Power Committees estimate the counterfactual quantity that would have been demanded in each state and month at current prices in the absence of shortages. We refer to this state-level annual figure as "Assessed Demand." "Energy Available" is the sum of electricity available at power plants and from net imports. The CEA measure of shortages (hereafter, "Shortage") is the percent of demand in state s in year t that is unmet:

$$S_{st} = \frac{Assessed\,Demand_{st} - Energy\,Available_{st}}{Assessed\,Demand_{st}} \tag{1}$$

Both Assessed Demand and Energy Available are growing rapidly due to economic growth: nationwide totals of both variables increased by a factor of 2.9 between 1992 and 2010. Thus, shortages can be thought of as the extent to which supply growth lags demand growth.

The CEA also estimates "Peak Shortage," an analogous measure of power shortage in peak demand periods. Peak Shortage and Shortage are highly correlated ($R^2 = 0.56$), and robustness checks will show that results are similar when we use Peak Shortage instead of Shortage.

The Shortage variable depends on an administrative assessment of counterfactual demand, so it is almost certainly measured with error. Potential attenuation bias is one reason why it is important to instrument for Shortage in our empirical analysis. However, correlations with independent data suggest that the CEA's estimates do contain meaningful information. Alam (2013) shows that Peak Shortage is correlated with her measure of blackouts, which is based on nighttime lights measured by satellites. We also present several additional correlations in Online Appendix Table A2. First, in the World Bank Enterprise Survey, plants in higher-Shortage states self-generate a larger share of electricity, report worse power quality, and are more likely to report that electricity is their primary obstacle to growth. Second, conditional on year indicators and plant fixed effects, coal power plant capacity factors are positively associated with Shortage, suggesting that coal plants are being run more to respond to a tighter supply-demand balance.[6] Furthermore, our main empirical results will show that Shortage is strongly correlated with hydroelectricity supply and with manufacturing outcomes such as self-generation.

From a publication called the General Review (CEA 1994-2012a), we observe total electricity sold to consumers, hydroelectric generation, total hydroelectric generation capacity, total generation capacity, and previous year capacity added, all at the state-by-year level. The General Review also includes state-specific allocations for power plants that are jointly owned by multiple states; see Online Appendix Table A3.

From the annual Review of Performance of Hydro Power Stations, we observe annual inflows into major reservoirs and electricity generation at plants behind each reservoir. Coverage grows over the sample, due to both new construction and entry of existing reservoirs into the sample

---

[6]This test is inspired by Fisher-Vanden, Mansur, and Wang (2015), whose empirical analysis directly uses thermal plant capacity factors as a proxy for shortages.

(CEA 1993-2011c). The 1992 data include 16 reservoirs, while the 2010 data include 31 reservoirs with plants comprising 50 percent of national hydroelectric generation capacity. To get inflow-predicted capacity factors separately for each reservoir-year, we regress generation on inflows and divide predicted generation by annual generation capacity.[7] We will use f_{dt}, the demeaned inflow-predicted capacity factor for reservoir d in year t, in our instrument.

The Review of Performance of Hydro Power Stations also includes capacity and annual generation for all hydro plants in India. We divide generation by annual generation capacity and de-mean within plant to construct demeaned capacity factors f_{dt}. We also have collected information on each plant's design, primarily from the Global Energy Observatory database (Gupta and Shankar 2014). About 18 percent of plants have run-of-river designs without reservoirs, meaning that they cannot adjust generation in response to electricity demand.

The CEA has lost its reservoir data for 2000 and its hydro plant generation data for 1992, so we impute data in those years using rainfall within the watershed.[8] Online Appendix B provides more information on the power sector data.

2.3 Annual Survey of Industries Data

We use India's Annual Survey of Industries (ASI) for establishment-level microdata. Registered factories with over 100 workers (the "census scheme") are surveyed every year, while smaller establishments (the "sample scheme") are typically surveyed every three to five years. The publicly available ASI includes establishment identifiers that are consistent across years beginning in 1998. While working in India, we also gained access to a version of the ASI with establishment identifiers before 1998, allowing us to construct a plant-level panel for the entire 1992-2010 sample.

The ASI is comparable to manufacturing surveys in the United States and other countries. Variables include revenues, value of fixed capital stock, total workers employed, total costs of labor, materials, fuel, and grid electricity purchased, and the physical quantity of grid electricity purchased, self-generated electricity, and electricity consumed. Industries are grouped using India's NIC (National Industrial Classification) codes, which are closely related to SIC (Standard Industrial Classification) codes. Online Appendix C gives more detail on the ASI data preparation and cleaning.

Table 2 presents sample-weighted summary statistics. There are 615,721 plant-by-year observations at 224,684 unique plants, although observation counts differ across variables due to

---

[7]By "annual generation capacity," we refer to potential generation at full capacity. Annual generation capacity in megawatt-hours is power production capacity in megawatts multiplied by 8760 hours per year.

[8]Specifically, we use GIS elevation maps and the latitude-longitude coordinate of each dam to determine each hydro plant's "watershed," i.e. all higher elevations that drain through the plant. We aggregate rainfall across gridpoints within the watershed to get annual within-watershed rainfall. For each plant (or reservoir), we run a regression of rainfall on generation (or reservoir inflows) using all years of data, then predict generation in 1992 (or reservoir inflows in 2000). Because there are not weather stations in every watershed, these within-watershed rainfall data are largely interpolated and/or simulated, so they are highly correlated with state-level rainfall and are not perfect predictors of reservoir inflows or generation by run-of-river plants.

Table 2: **Annual Survey of Industries Summary Statistics**

	Mean	Std. Dev.	Min.	Max.	Obs.
Plant Number of Observations	2.20	2.13	1	19	224,684
Revenues (million Rupees)	139	2156	0	788,868	613,930
Capital Stock (million Rupees)	51	1044	0	297,370	612,424
Number of Employees	79	431	0	52,148	576,901
Labor Cost (million Rupees)	6.39	70.5	0	16,074	602,124
Materials Purchased (million Rupees)	90	1562	0	636,095	607,522
Fuels Purchased (million Rupees)	5.07	102	0	39,360	596,036
Electricity Purchased (million Rupees)	3.81	48.1	0	9,935	561,284
Electricity Purchased (GWh)	0.95	19.2	0	6,545	594,925
Electricity Self-Generated (GWh)	0.44	20.8	0	7,147	553,515
Electricity Consumed (GWh)	1.38	30.0	0	7,357	596,010
1(Self-Generator)	0.44	0.50	0	1	615,721
Self-Generation Share	0.06	0.16	0	1	546,328
Fuel Revenue Share	0.05	0.13	0	5.48	596,036
Electric Intensity (kWh/Rs)	0.013	0.022	0	0.37	594,882
1(Census Scheme)	0.14	0.34	0	1	615,721

Notes: Plant number of observations is reported at the plant level; all other variables are reported at the plant-by-year level. Rupees are constant 2004 Rupees. Means and standard deviations are weighted by ASI sample weights. Observation counts differ due to non-response and due to variable-specific cleaning procedures described in Online Appendix C.

non-response and cleaning procedures described in Online Appendix C. 107,032 plants will be immediately dropped from our fixed effects and difference estimators because they are observed only once. For plants observed multiple times, 60 percent of intervals between observations are one year, while 91 percent are five years or less.

The mean (median) plant employs 79 (34) people and has gross revenues of 139 million (20 million) Rupees, or in U.S. dollars approximately $3 million ($400,000). 1(Self-Generator) is an indicator variable for whether a plant self-generates electricity.[9] Self-Generation Share is the ratio of electricity purchased to electricity consumed, and Fuel Revenue Share is the value of fuels purchased divided by revenues. Forty-one percent of unweighted observations are in the census scheme, although the table shows that 14 percent of registered factories are in the census scheme after applying sample weights.

---

[9]While the ASI does not explicitly ask plants whether they own a generator, 1(Self-Generator) is a very good proxy, as it would be unusual in India for a plant to own a generator but never use it. Comparison with the 2005 World Bank Enterprise Survey (WBES) provides supporting evidence: 81 percent of plants with 100 or more workers and 38 percent with fewer than 100 workers in the 2005 ASI ever self-generate electricity, while 83 percent of plants with 100 or more workers and 46 percent with fewer than 100 workers in the 2005 WBES report owning generators.

3 Background

3.1 Reasons for Systemic Shortages

At the end of our study period in 2010, India had 174 gigawatts of utility-scale power generation capacity, or about one-sixth the U.S. total.[10] Of this, 53 percent was coal, ten percent was natural gas, and 22 percent was hydroelectric. While power generation has been open to private investment since India's 1991 liberalization, 80 percent of electricity supply in 2010 remained government owned: 51 percent by states and 29 percent by the central government. Most retail distribution companies are also state-run, although some have been privatized.

The immediate reason for shortages is that the retail distribution companies cannot raise retail prices to clear the market. In fact, conditional on state and year effects, there is no correlation between the Shortage variable and the median electricity price paid by plants in the ASI. However, such a disconnect between retail price and market conditions is common to nearly all power systems around the world, including many that do not experience endemic shortages. There are several underlying reasons why shortages arise in India and many other developing countries.

The first reason is the "infrastructure quality and subsidy trap" (McRae 2015): distribution companies provide low-quality electricity, consumers tolerate this low quality because they pay low prices, government subsidies cover distribution companies' losses from the low prices, and politicians support the subsidies to avoid voter backlash. At least since the 1970s, state-run distribution companies have offered un-metered electricity at a monthly fixed fee and zero marginal cost to agricultural consumers (Bhargava and Subramaniam 2009). These low agricultural prices are cross-subsidized by industrial electricity prices that are nearly four times higher. Furthermore, 26 percent of electricity generated in India in 2010 was lost due to "technical and commercial losses," meaning poor transmission infrastructure and uncollected bills.[11] As a result of low prices and losses, distribution companies' revenues from power sales were 24 percent less than costs in 2010.

To cover the low agricultural prices, state governments promise large subsidies, which amounted to 20 percent of distribution companies' revenues in 2009 and 11 percent in 2010. Even after receiving subsidies, however, distribution companies lost a total of $61 billion (in real 2004 dollars) between 1992 and 2010. Resulting financial constraints constrain infrastructure investment and maintenance, and degraded infrastructure further increases the probability of power outages. The government bails out the state power utilities at irregular intervals.

A second underlying reason for shortages is underinvestment in new generation capacity: supply is not keeping up with rapid demand growth. After the 1991 liberalization, investors signed 200 memoranda of understanding with the government to build 50 gigawatts of generation capacity, but less than four gigawatts of this was actually built (Bhargava and Subramaniam 2009). Of the 71 gigawatts of capacity targeted to be built between 1997 and 2007, only half was actually

---

[10]Capacity and ownership statistics in this paragraph are from the General Review (CEA 2012a).

[11]Statistics on technical and commercial losses, average revenues and costs, subsidies, and losses by state distribution companies are from Power Finance Corporation (various years).

achieved (CEA 2013). Potential power plant investors face concerns over both output demand and input supply. Their main customers, the distribution companies, face serious financial problems. Meanwhile, the main coal supplier is Coal India, a government-run monopoly that is struggling to keep pace with demand growth (Chilkoti and Crabtree 2014).

Furthermore, existing capacity is systematically underutilized. Between 1994 and 2009, Indian coal power plants were offline about 28 percent of the time due to forced outages, planned maintenance, coal shortages, and other factors. When capacity is utilized, India's coal plants are also substantially less efficient than comparable plants in the United States (Chan, Cropper, and Malik 2014).

3.2   Variation in Shortages

Figure 1: **Average Shortages and Per Capita GDP by State**
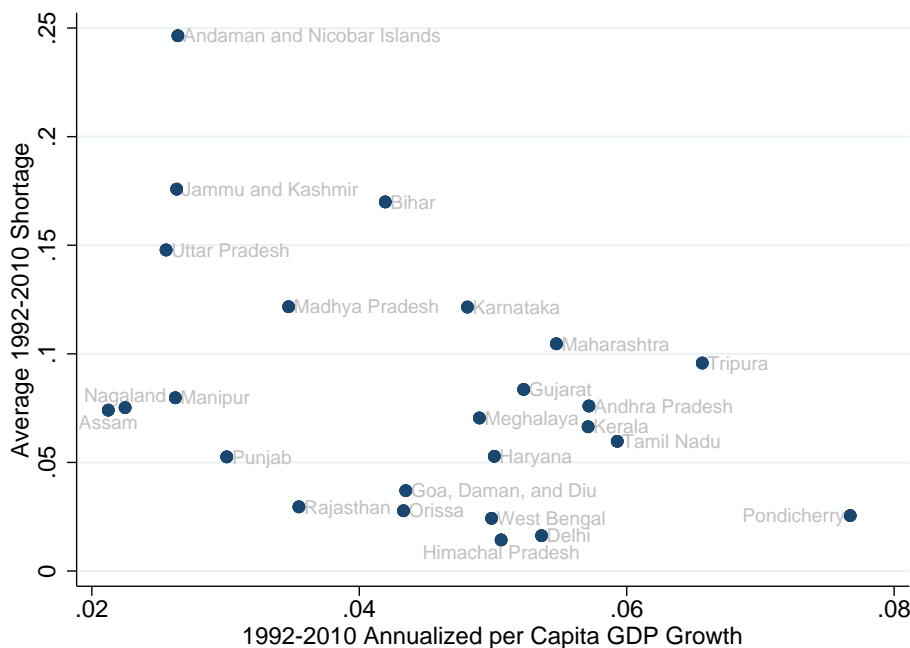
Notes: This figure plots each state's average shortage over the 1992-2010 study period vs. its per capita GDP growth rate over that period. Shortage data are estimated by the Central Electricity Authority.

Figure 1 plots the average shortage for each state over our 1992-2010 study period against its annualized per-capita GDP growth over the same period. The figure illustrates that shortages vary substantially across states. The cross-sectional correlation between shortages and growth is noisy but potentially negative, suggesting that poor institutions and other factors both cause slow growth and worsen shortages.[12] On the other hand, states such as Rajasthan and Punjab have both slow

---

[12]The best fit line slopes downward ($p = 0.032$), but the correlation is less strong ($p = 0.145$) when excluding Andaman and Nicobar Islands and Pondicherry, which are very small Union Territories.

growth and low shortages, partially because slow growth makes it easier for supply to keep up with demand. This discussion highlights the importance of instrumenting for shortages in the empirical analysis.

Shortages also vary substantially within states over time. Figure 2 illustrates this for five large states. Uttar Pradesh experiences relatively high and variable shortages, while West Bengal has had consistently low shortages for the past 20 years. Karnataka faced almost zero shortage in the mid-2000s but had significant shortages in the early to mid-1990s. Gujarat had relatively reliable power supply by the end of the study period, but shortages were more severe in the mid-2000s.

Figure 2: **Shortages Over Time in Five Large States**
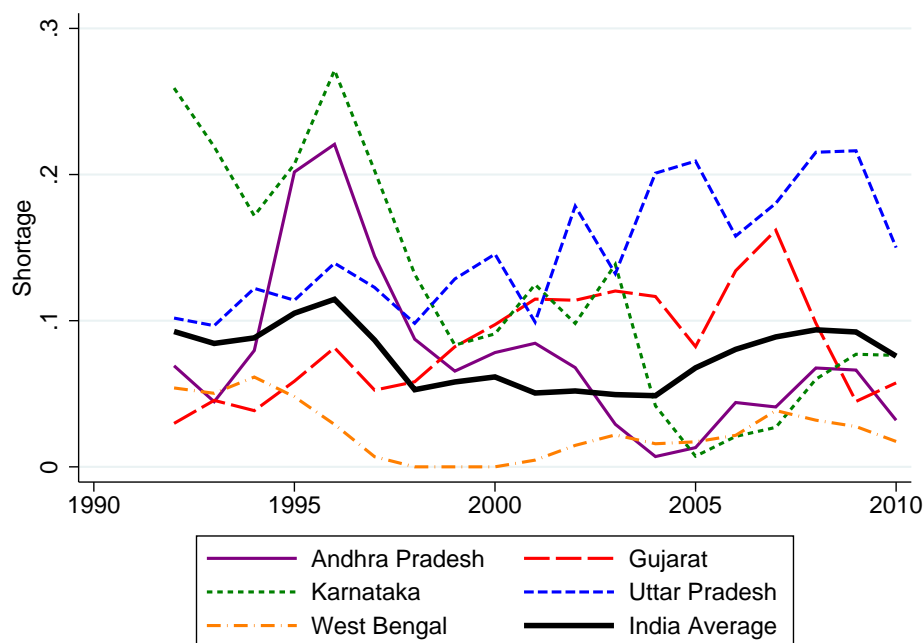
Note: Shortage data are estimated by the Central Electricity Authority.

Several factors drive this year-to-year variation in shortages. On the demand side, fast economic growth over a few years can increase shortages, and air conditioner use during an unusually hot summer can substantially increase electricity demand. On the supply side, power plant outages due to maintenance or coal shortages can worsen electricity shortages. For our instrument, we will focus on two supply shifters: new plants coming online, which can temporarily alleviate shortages, and variation in water availability for hydroelectricity production.

In 2009 and 2010, respectively, five and seven percent of electricity generated nationwide was exchanged across states.[13] Because distribution companies are able to procure power from other states, supply-demand imbalances do not vary as much as they would under autarky.

---

[13]These are our calculations based on data in the Load Generation Balance Reports.

3.3 Industrial Electricity Use in India

A natural response to shortages is to self-generate electricity. We calculate that manufacturing plants in the ASI generate 35 percent of their total electricity consumption, more than twice the 15.8 percent for US manufacturers reported in the Manufacturing Energy Consumption Survey (MECS) (U.S. DOE 2013). Figure 3 compares the ratio of electricity generated to electricity consumed at Indian plants (from the ASI) to US plants (from the MECS), with each dot reflecting a three-digit NIC industry code.[14]

Figure 3: **Manufacturing Electricity Generation in India vs. the U.S.**
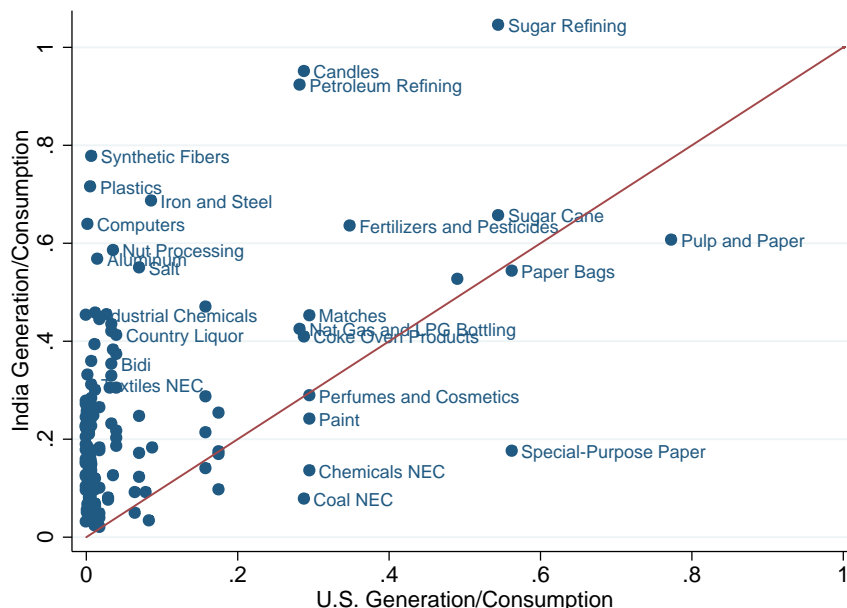
Notes: This figure presents the ratio of electricity generation to consumption by three-digit industry. Indian and U.S. data are from the Annual Survey of Industries and the Manufacturing Energy Consumption Survey, respectively.

The figure highlights two important facts. First, there is a strong correlation between the US and Indian data, suggesting that the ASI self-generation data are meaningful. In the United States, where power outages are relatively unimportant, many industries still produce a substantial fraction of electricity consumed. For example, in the sugar refining industry, byproducts from sugarcane processing can be burned to generate electricity, so there is a natural complementarity between manufacturing operations and electricity generation. Second, the mass of points along the vertical

---

[14]This ratio of generation to consumption differs slightly from Self-Generation Share because electricity generated also includes electricity sales by manufacturing plants to others. Several industries don't match well between the two datasets: chemicals and refining are not broken out into many different sub-industries in the public US data, so Indian sub-industries such as Explosives, Chemicals Not Elsewhere Classified (NEC), Matches, and Perfumes and Cosmetics are matched to "Chemicals," a broader industry where other establishments are more likely to have feedstock for self-generation, and thus a higher self-generation share. Similarly, Natural Gas and LPG Bottling, Coal NEC, and Coke Oven Products are matched to "Petroleum and Coal Products," another very broad category.

axis implies that many industries in India produce much more than their counterparts in the U.S. For instance, plastics manufacturers in the U.S. produce none of their power, while production by Indian plastics manufacturers equals 70 percent of their electricity consumption.

4 Model and Production Function Estimation

4.1 Setup

This section presents a model of how electricity shortages affect manufacturing plants. τ indexes points in time, which we refer to as "days." Every day, a plant uses capital K, labor L, electricity E, and materials M to produce output Q. $Q_{it\tau}$ denotes the output for plant i in year t on day τ, and $Q_{it} \equiv \int_{\tau} Q_{it\tau} d\tau$ is the annual aggregate. The measure of "days" in a year is normalized to one. We do not model the possibility for inter-day substitution.[15]

The daily production function is a Cobb-Douglas aggregate of capital, labor, electricity, and materials, with physical productivity A_{it}[16]:

$$Q_{it\tau} = A_{it} K_{it\tau}^{\alpha_K} L_{it\tau}^{\alpha_L} M_{it\tau}^{\alpha_M} E_{it\tau}^{\alpha_E}. \tag{2}$$

Since we observe revenues rather than physical quantities produced, we need to relate revenues to our production function in Equation (2). We assume that plants sell into a perfectly competitive output market with price p, and we define revenue productivity (TFPR) as $\Omega_{it} = A_{it}p$.[17] This yields the following daily revenue-generating production function:

$$R_{it\tau} = \Omega_{it} K_{it\tau}^{\alpha_K} L_{it\tau}^{\alpha_L} M_{it\tau}^{\alpha_M} E_{it\tau}^{\alpha_E}. \tag{3}$$

4.2 Power Outages, Inputs, and Timing

A power outage occurs on each day with probability δ, and δ is known at the beginning of the year. If there is no outage, plants can purchase electricity from the grid at price $p^{E,G}$. If there is an outage, plants with generators can self-generate electricity at price $p^{E,S} > p^{E,G}$. Plants without generators have zero electricity input during an outage, so they produce zero output. Notice that setting electricity input to zero for δ percent of the year is different from a δ percent reduction in

---

[15] In Online Appendix F, we present a model that does allow inter-day substitution.

[16] Our earlier working paper used a production function that is Leontief in electricity, thus ruling out substitution away from electricity during power outages: $Q_{it\tau} = \min\{AK_{it\tau}^{\alpha_K} L_{it\tau}^{\alpha_L} M_{it\tau}^{\alpha_M}, \frac{1}{\lambda} E_{it\tau}\}$, where λ indexes the plant's electricity intensity. We show in Online Appendix F that for our particular outcomes of interest, the Leontief model's main predictions are similar to those from the full Cobb-Douglas model.

[17] We could alternatively assume that plants sell into an imperfectly competitive output market with daily demand curve $Q^D = BP^\epsilon$. This yields identical results, except with revenue productivity redefined as $\Omega \equiv A^{1+\frac{1}{\epsilon}} B^{-\frac{1}{\epsilon}}$ and production function coefficients $\beta_X = \alpha_X(1+\frac{1}{\epsilon})$ for $X = \{K, L, M, E\}$. Using a demand curve that depends on annual output would introduce dynamics into this problem, as production on day τ would affect prices, and consequently input choices, on other days of the year.

electricity use, and the effect of power outages can be much larger than the electricity coefficient α_E.

There are three types of inputs:

1. *Fixed inputs* are chosen before δ and Ω_{it} are known. For the model and simulations, we assume that capital stock K_{it} is fixed.

2. *Yearly-flexible inputs* can be modified at the beginning of a year t after observing δ and Ω_{it}, but they cannot be modified from day to day. For the model and simulations, we treat labor as yearly-flexible, as plants cannot hire and fire workers from one moment to the next as blackouts occur. This gives $L_{it\tau} = L_{it}$.[18]

3. *Fully-flexible inputs* can be modified for each day τ after observing whether or not there is an outage. We treat materials and electricity as fully flexible.

4.3 The Plant's Problem

Plants have daily profit function

$$\Pi_{it\tau} = \Omega K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it\tau}^{\alpha_M} E_{it\tau}^{\alpha_E} - p^L L_{it} - p^M M_{it\tau} - p^E E_{it\tau}, \tag{4}$$

where p^L, p^M and p^E are labor, materials, and electricity prices, respectively. We exclude capital costs because they are sunk before the plant makes any production decisions.

For plants with generators, the first-order condition with respect to materials yields

$$\alpha_M \frac{R_{it\tau}}{M_{it\tau}} = p^M, \tag{5}$$

and the first-order condition for electricity yields

$$\alpha_E \frac{R_{it\tau}}{E_{it\tau}} = \begin{cases} p^{E,S} & \text{if outage} \\ p^{E,G} & \text{if grid power} \end{cases}. \tag{6}$$

Define M_{itG} and M_{itS} as materials input during grid power and outage periods, respectively, and define E_{itG} and E_{itS} as well as R_{itG} and R_{itS} analogously. For non-generators, the same first order conditions hold except that there is no production during outages, so $M_{itS} = E_{itS} = R_{itS} = 0$.

When setting labor, the plant begins with its yearly profit function, which is simply the weighted average of Equation (4) over grid power and outage periods. The first-order condition yields

$$\alpha_L \left[(1 - \delta) \frac{R_{itG}}{L_{it}} + \delta \frac{R_{itS}}{L_{it}} \right] = p^L. \tag{7}$$

---

[18]In some industries such as plastics, material inputs can be spoiled during a power outage. In these cases, it might be more plausible to assume that materials are also semi-flexible, and it is straightforward to change the model in this way.

For non-generators, this simplifies to $\alpha_L(1-\delta)\frac{R_{itG}}{L_{it}} = p^L$, because $R_{itS} = 0$.

These first-order conditions show that for plants with generators, power outages act exactly like a time-varying input tax on electricity. The electricity price increase causes plants to reduce electricity input, which reduces the marginal revenue products of materials and labor, so plants reduce those inputs as well. We call this input reduction the *input tax effect.* The effect is larger for non-generators, because power outages require them to shut down, as if facing an infinite input tax.

4.4 Effects of Shortages on Revenues and Measured Productivity

Annual revenues are the weighted average of revenues during grid power and outage periods:

$$R_{it} = \Omega_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}\left((1-\delta)M_{itG}^{\alpha_M}E_{itG}^{\alpha_E} + \delta M_{itS}^{\alpha_M}E_{itS}^{\alpha_E}\right) \tag{8}$$
$$= \Omega_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}M_{it}^{\alpha_M}E_{it}^{\alpha_E} \cdot V_{it},$$

where

$$V_{it} = \frac{(1-\delta)M_{itG}^{\alpha_M}E_{itG}^{\alpha_E} + \delta M_{itS}^{\alpha_M}E_{itS}^{\alpha_E}}{((1-\delta)M_{itG} + \delta M_{itS})^{\alpha_M}((1-\delta)E_{itG} + \delta E_{itS})^{\alpha_E}} = \frac{(1-\delta)M_{itG}^{\alpha_M}E_{itG}^{\alpha_E} + \delta M_{itS}^{\alpha_M}E_{itS}^{\alpha_E}}{M_{it}^{\alpha_M}E_{it}^{\alpha_E}}. \tag{9}$$

Using lower-case variables to denote natural logs, revenues are:

$$r_{it} = \alpha_K k_{it} + \alpha_L l_{it} + \alpha_M m_{it} + \alpha_E e_{it} + \omega_{it} + v_{it}. \tag{10}$$

Define "measured TFPR" $\hat{\omega}_{it}$ as the difference between logged revenues and input contributions:

$$\hat{\omega}_{it} = r_{it} - \alpha_K k_{it} - \alpha_L l_{it} - \alpha_M m_{it} - \alpha_E e_{it} = \omega_{it} + v_{it}. \tag{11}$$

Equation (10) shows that outages reduce revenue through two mechanisms. First, plants reduce inputs l_{it}, m_{it}, and e_{it} through the input tax effect. Second, measured TFPR $\hat{\omega}_{it}$ changes by v_{it}. The revenue loss must be larger than the measured TFPR loss because the input tax effect also causes plants to contract.

For plants with generators, Equation (9) shows that V_{it} is the effect of using different bundles of fully-flexible inputs in outage and grid power periods, relative to using the same weighted average bundle in all periods. In the standard case where $\alpha_M + \alpha_E < 1$, the daily production function is concave in $M_{it\tau}$ and $E_{it\tau}$, so $V_{it} < 1$. For small δ, V_{it} is decreasing in δ: more outages lead to more productivity loss. We call this the *input variation effect.*

For plants without generators, $M_{itS} = E_{itS} = 0$, so v_{it} simplifies to $v_{it} = (1-\alpha_M-\alpha_E)\ln(1-\delta)$. Given that $\ln(1-\delta) \approx -\delta$, we have $v_{it} \approx -\delta(1 - \alpha_M - \alpha_E)$. This illustrates that the measured TFPR loss is the waste of all inputs when outages force the plant to shut down, net of $\alpha_M + \alpha_E$, the share of inputs that can be "turned off" during the outage.

16

4.5 Production Function Estimation

4.5.1 Static Inputs

Under the assumption of profit maximization, we use the first-order conditions to recover production function coefficients α_L, α_M, and α_E from yearly ASI data.[19] Although the first-order conditions derived above depend on variables that change between outage and non-outage periods and are thus unobserved in the ASI's annual aggregates, each can be re-arranged to obtain the usual result that the production function coefficient equals the annual aggregate input revenue share.

The first-order condition for labor gives:

$$\alpha_L = \frac{p^L L_{it}}{(1-\delta)R_{itG} + \delta R_{itS}} = \frac{p^L L_{it}}{R_{it}}. \tag{12}$$

The first-order condition for materials shows that the materials to revenue ratio never varies, so

$$\alpha_M = \frac{p^M M_{it\tau}}{R_{it\tau}} = \frac{p^M M_{it}}{R_{it}}. \tag{13}$$

To derive α_E, we re-arrange the first-order condition and take a weighted average across outage and non-outage periods:

$$\alpha_E \left((1-\delta)R_{itG} + \delta R_{itS}\right) = (1-\delta)p^{E,G}E_{itG} + \delta p^{E,S}E_{itS}. \tag{14}$$

This gives

$$\alpha_E = \frac{(1-\delta)p^{E,G}E_{itG} + \delta p^{E,S}E_{itS}}{R_{it}}, \tag{15}$$

where the numerator is total expenditures on grid electricity plus fuel for self-generated electricity.[20]

We use median regressions to separately estimate each α parameter for each of 143 three-digit NIC industries, allowing separate linear time trends by two-digit industry. We prefer median regressions because they are highly robust to outliers. See Online Appendix C for additional details.

4.5.2 Estimating the Capital Coefficient

It is not realistic to assume a static first-order condition for capital analogous to those for other inputs, because capital has substantial adjustment costs and irreversibilities (Asker, Collard-Wexler, and De Loecker 2014). Instead, we estimate α_K using GMM.

We define $\tilde{r}_{it} \equiv r_{it} - \hat{\alpha}_L l_{it} - \hat{\alpha}_M m_{it} - \hat{\alpha}_E e_{it}$ as "transformed revenue," after netting out the fitted contribution from labor, materials, and electricity. We then regress transformed revenue on

---

[19]For additional discussion of the first-order condition approach to production function estimation, see De Loecker and Warzynski (2012) and Bartelsman, Haltiwanger and Scarpetta (2013).

[20]$p^{E,S}$ is unobserved, so we assume that $p^{E,S}$ is 7 Rs/kWh, reflecting the median price reported in the 2005 World Bank Enterprise Survey.

capital:

$$\tilde{r}_{it} = \alpha_K k_{it} + \omega_{it}. \tag{16}$$

If more productive plants invest in more capital and if productivity ω_{it} is serially correlated, $E[\omega_{it}k_{it}] \neq 0$ and estimation in OLS would yield an upward-biased $\hat{\alpha}_K$. We thus use a standard approach from control function estimators, exploiting the fact that capital investments take time to plan and implement. Specifically, we assume that capital requires a one year time-to-build, so $K_{it} = (1-\kappa)K_{it-1} + I(K_{it-1}, \omega_{it-1})$, where κ represents depreciation and $I(K_{it-1}, \omega_{it-1})$ represents the investment policy function, which depends on past productivity. We allow productivity to evolve according to a general first-order Markov process $\omega_{it} = g(\omega_{it-1}) + \xi_{it}$. Under the one-year time-to-build assumption, the productivity innovation ξ_{it} is uncorrelated with contemporaneous capital stock: $E[\xi_{it}k_{it}] = 0$. We use this moment condition in GMM to estimate α_K from Equation (16), with bootstrapped standard errors.[21]

4.5.3   Production Function Estimates

Table 3 presents summary statistics on the estimated production function parameters. There are 19 years of coefficients for 143 different industries, so we present the mean and 25th and 75th percentiles of the distribution of each α. The mean labor, materials, electricity, and capital coefficients are 0.078, 0.71, 0.019, and 0.16, respectively.

The mean returns to scale coefficient is 0.96. Slightly decreasing returns to scale are common in the production function literature; see, for instance, Collard-Wexler and De Loecker (2015). For comparison, we also include the α_K from OLS estimation of Equation (16). As predicted above, the OLS estimates are biased slightly upward.

5   Estimation Strategy

5.1   Translating the Model to an Estimating Equation

The model suggests a set of outcomes that might be affected by electricity shortages. Plants with generators will use them during outages, so shortages should be positively associated with Self-Generation Share and Fuel Revenue Share for self-generators. Both self-generators and non-generators will reduce electricity, materials, and labor inputs, and measured TFPR will also decrease. Revenues should decrease due to decreases in both inputs and measured TFPR.

We use i, j, s, and t, respectively, to index plants, two-digit NIC industries, states, and years. A simplified version of our estimation strategy is to regress outcomes Y_{ijst} on the CEA Shortage

---

[21] We could also estimate equation (16) in OLS with first differences if we made the stronger assumption that productivity is a random walk, i.e. $\omega_{it} = \omega_{it-1} + \xi_{it}$. Thus, the purpose of the GMM procedure is to allow a more flexible time-series process for productivity.

Table 3: **Production Function Parameter Estimates and TFPR**

	(1)	(2)	(3)
	Mean	25%	75%
Labor (α_L)	0.078	0.053	0.101
Materials (α_M)	0.71	0.66	0.76
Electricity (α_E)	0.019	0.016	0.022
Capital (α_K)	0.16	0.10	0.21
Capital (α_K) from OLS	0.19	0.12	0.24
Returns to Scale	0.96	0.94	1.00
Measured TFPR (ω_{it})	2.12	1.44	2.61

Notes: Distribution statistics for production function coefficients and returns to scale are based on 2,424 three-digit NIC industry-by-year observations. Distribution statistics for measured TFPR are based on 589,779 plant-year observations.

variable S_{st}, controlling for year indicators θ_t, industry-by-year indicators μ_{jt}, and plant indicators ϕ_i:

$$Y_{ijst} = \rho S_{st} + \theta_t + \mu_{jt} + \phi_i + \varepsilon_{ijst}. \tag{17}$$

This equation uses Shortage S_{st} as a proxy for the model's outage probability δ. In reality, the supply-demand imbalances captured by S_{st} are translated to plant-specific outage probabilities δ through potentially-endogenous decisions by electricity distribution companies. For example, larger or more productive plants may be able to secure preferential electricity access as shortages worsen. Using Shortage instead of δ captures the "reduced form" of this process. Of course, estimates of ρ from this equation are also "reduced form" in the sense that they are not constrained by the assumptions of the model.[22]

5.2 Instrumenting for Shortages

5.2.1 Overview

There are several reasons why shortages might be endogenous in Equation (17). For example, improvements in state-level economic conditions or institutions could increase productivity and revenue, and the resulting increase in electricity demand could cause shortages. Alternatively, shortages could be measured with error, causing attenuation bias.

---

[22]Several unmodeled effects of shortages might be relevant. First, if plants substitute production across days in response to outages, our estimates capture this by estimating net effects of shortages over a year. Second, if shortages reduce output quality, this will appear in revenues and/or TFPR. Third, effects on input demand are not constrained by our yearly-flexible vs. fully-flexible categorizations or by Cobb-Douglas substitution patterns. Fourth, if inputs or outputs are traded on local markets instead of national or international markets with exogenous prices, shortages could affect plants through changes in other input or output prices, not just through the price and availability of electricity.

A valid instrument for shortages must shift electricity supply but affect manufacturers only through shortages. Our instrument exploits two supply shifters. The first is hydroelectricity generation: because hydro plants have very low physical marginal cost, their annual output depends primarily on a water availability constraint determined by rainfall at higher elevations (or snowfall in the Himalayan states). The second is previous-year capacity addition: 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.

To construct the supply instrument we combine the two sources of variation after normalizing each into a share of state electricity demand. Denote predicted hydro generation as H_{st}, denote predicted generation from previous year capacity addition as N_{st}, and denote predicted electricity demand as Q_{st}. The supply instrument is:

$$Z_{st} = \frac{H_{st} + N_{st}}{Q_{st}}. \tag{18}$$

5.2.2 Details of the Instruments

Because shortages directly affect actual consumption, we predict state s electricity demand using the product of total electricity sold in all other states in year t and the sample average ratio of sales in state s to sales in all other states. Indexing other states by r and years by y, this is:

$$Q_{st} = \sum_{r \neq s} Q_{rt} \cdot \sum_{y=1992}^{2010} \frac{Q_{sy}}{\sum_{r \neq s} Q_{ry}}. \tag{19}$$

For H_{st}, one option is to use actual hydroelectricity generation. In theory, however, reservoirs can store water for future years if demand is low in the present year, which means that hydro generation might be simultaneously determined with manufacturing production. For our primary specification, we thus predict state-level hydro generation using reservoir inflows and generation from "run-of-river" hydro plants that have no reservoirs to store water.

More precisely, recall that Section 2 introduced f_{dt}, reservoir or plant d's demeaned capacity factor in year t. We get the state average capacity factor, weighting by annual generation capacity c_d and by $a_{ds} \in [0, 1]$, the share of d's output that is contractually allocated to state s. Multiplying by C_{st}, the state's predicted annual hydro generation capacity, translates the capacity factor to predicted hydro generation H_{st}. Denoting $\mathcal{D}_s$ as the set of reservoirs and run-of-river plants ever observed in state s, we have:

$$H_{st} = C_{st} \cdot \frac{\sum_{d \in \mathcal{D}_s} c_d a_{ds} f_{dt}}{\sum_{d \in \mathcal{D}_s} c_d a_{ds}}. \tag{20}$$

We predict hydro capacity C_{st} using Equation (19), except with C substituted for Q. We

use predicted instead of actual hydro capacity because the latter is determined endogenously by economic growth over the study period.

To get N_{st}, the predicted generation from plants that came online in the one previous year, we simply multiply the capacity added in the previous year by the national average thermal plant capacity factor in year t. We will later see that although the moderate increase in precision from adding N_{st} is helpful in the main estimates and robustness checks, nearly all of the identifying power comes from H_{st}, not N_{st}. For this reason, much of the discussion of the instrument below focuses on H_{st}.

5.2.3 Rainfall Controls

Of course, precipitation affects India's economy in ways other than just relaxing water availability constraints for hydroelectricity plants. To address this, we include rainfall controls in our primary specification. We get each state's sample average annual rainfall $\bar{R}_s$ and include indicator variables for five bins above and five bins below $\bar{R}_s$ at 60 millimeter intervals. (We also present alternative specifications with 50 millimeter bins, 100 millimeter bins, and a simple linear control.) Because rainfall is also correlated with temperature, which also can affect manufacturers, we control for average cooling degrees W_{st}.

There are three sources of variation in the hydro instrument conditional on rainfall. First, the cross-state contracts measured by allocations a_{ds} mean that in some cases, shortages in a given state will depend on water availability in another state. Second, rainfall in hydro plant watersheds differs from rainfall in the same state outside of the watersheds. Half of the geographic gridpoints in our rainfall data are not in or near a hydro plant watershed, and rainfall can vary substantially across even small geographic regions (Dell, Jones, and Olken 2014, Auffhammer, Hsiang, Schlenker, and Sobel 2013). Furthermore, dams are more likely to be placed in more steeply-sloped terrain (Duflo and Pande 2007), where local rainfall variation can be especially significant (Lipscomb, Mobarak, and Barham 2013).

Third, even if there were no cross-state contracts and every point in a state received the same rainfall, the slope of the relationship between rainfall and Z_{st} differs across states because states capture different shares of rainfall for hydroelectricity generation. Some states, such as West Bengal, receive relatively high rainfall but convert little into hydroelectricity. Other states, such as Karnataka, have lower-than-average rainfall but derive a relatively large share of electricity from hydro. The same variation in rainfall thus has a much larger effect on electricity supply in Karnataka than in West Bengal.[23]

---

[23]In fact, the first stage relationship between hydro production and shortages is clearly visible in the raw data in Karnataka. We illustrate this in Online Appendix Figure A4. Online Appendix Figure A5 plots each state's sample average rainfall against the average ratio of hydro generation to total consumption; the two quantities are uncorrelated.

5.3 Estimating Equation

Our empirical specification takes the intuition from Equation (17) and instruments for S_{st}, adding the vector of rainfall bins $\mathbf{R}_{st}$, cooling degree controls W_{st}, state-specific time trends $\psi_s t$, and also state split indicators Λ_{st} that control for three changes in state geographic definitions:[24]

$$Y_{ijst} = \rho S_{st} + \beta \mathbf{R}_{st} + \nu W_{st} + \Lambda_{st} + \psi_s t + \theta_t + \mu_{jt} + \phi_i + \varepsilon_{ijst}. \tag{21}$$

Observations are weighted by ASI sample weights. Our primary specifications use the fixed effects estimator, although we also present estimates using the difference estimator. Because S_{st} and Z_{st} vary by state and year, errors are clustered by state-year. Our primary specifications use two-way clustering by plant and state-year (Cameron, Gelbach and Miller 2011).[25]

5.4 Discussion

Several features of the empirical strategy merit discussion. First, we include state-specific trends $\psi_s t$ because they substantially improve first stage power. They are especially important in two robustness checks, one which uses ln(Energy Available) instead of Shortage S_{st} as the endogenous variable, and another that uses actual hydro generation for H_{st}. Both ln(Energy Available) and actual hydro generation have highly heterogeneous state-specific trends, the former because of differing economic growth rates and the latter because of differing trends in the importance of hydro relative to other generation sources.[26] Unless they are absorbed by the $\psi_s t$ controls, these heterogeneous trends substantially reduce precision.

Second, because the estimates include year indicators, ρ measures effects of shortages S_{st} on plants in state s relative to plants in other states in the same year. If there are cross-state spillover effects, for example if customers substitute to plants in other states as production in state s is slowed by outages, our estimates reflect some reallocation of output across states, not just a loss of aggregate output. Thus, the estimates are relevant for state-level policymakers who want to

---

[24]In November 2000, three new states (Chhattisgarh, Jharkhand, and Uttarakhand) were split from three existing states (Madhya Pradesh, Bihar, and Uttar Pradesh). Weather, hydro plants and reservoirs, and ASI plants can be assigned to post-2000 state definitions in all years of the sample. However, the CEA reports Shortages and other state-level variables only for the combined state areas before 2001 or 2002 (depending on the variable). Before those years, we assign data for the combined states to plants in either of the eventually-separate states. State split indicators Λ_{st} take different values when a state's geographic definition changes.

[25]We do not have the downward-biased standard errors described by Bertrand, Duflo, and Mullainathan (2004). They study traditional differences-in-differences estimators, in which the dependent variable of interest is serially correlated. By contrast, Online Appendix Table A7 shows that the instrument Z_{st} is not serially correlated. Notwithstanding, we also present robustness checks with standard errors clustered by state.

[26]Online Appendix Figures A6 and A7 illustrate this. For example, the economy in Andhra Pradesh grows rapidly over the study period, while Uttar Pradesh grows relatively slowly. Residual of year effects θ_t, ln(Energy Available) trends steeply downward in Uttar Pradesh and steeply upward in Andhra Pradesh. For hydro generation, states such as Andhra Pradesh and Karnataka have relatively steep downward trends in the share of consumption supplied by hydro, while states such as Gujarat and West Bengal have small and relatively flat hydro generation over the study period. Thus, residual of year effects θ_t, the instrument constructed with actual hydro generation slopes steeply downward in Karnataka and steeply upward in West Bengal.

know how improvements in electricity supply in their state affect outcomes for their manufacturers relative to manufacturers in other states. The estimates are informative about a national-level supply improvement only if cross-state spillovers are relatively small.

Third, the coefficients will largely be identified by states with more variation in hydro generation, which tend to also be the states where hydro represents a large share of total supply. While some hydro-heavy states are small mountainous areas such as Himachal Pradesh, Meghalaya, and Uttarakhand, other states such as Andhra Pradesh, Karnataka, Kerala, Orissa, and Punjab are both large and rely heavily on hydroelectricity.[27]

Fourth, the empirical estimates will primarily reflect short-run effects. Because of the plant fixed effects ϕ_i, the ρ parameter reflects impacts of shortages only on continuing plants. Entry, exit, and even generator adoption or resale are unlikely responses to variation in shortages S_{st} driven by year-to-year variation in hydro availability and new supply Z_{st}.

5.4.1 Evaluating the Instrument with State-Level Data

Table 4 evaluates the instrument. All regressions are analogues to the first stage, regressing a dependent variable on the hydro and new supply instrument Z_{st}, controlling for weather $\mathbf{R}_{st}$ and W_{st}, state trends $\psi_s t$, year indicators θ_t, and state split controls Λ_{st}. These regressions are estimated at the state-by-year level, with observations weighted by the number of ASI establishments. Column 1 presents the direct analogue to the first stage, showing that a one percentage point increase in instrumented supply decreases shortages by 0.177 percentage points. The fact that the coefficient is smaller than one (in absolute value) implies that states partially offset an increase in supply by decreasing generation from other sources and by importing less from or exporting more to other states.

For an instrument to be valid, it must shift supply without affecting manufacturing other than through shortages. Columns 2 and 3 of Table 4 present evidence that this is the case. Recall that the CEA separately reports the two components of shortages: Assessed Demand and Energy Available. Column 2 shows that Energy Available is associated with the instrument, but column 3 shows that Assessed Demand is not. It is difficult to conceive of a story under which the exclusion restriction is violated but the instrument is not associated with electricity demand. Column 4 provides further evidence on this, showing that agricultural output is not associated with the instrument (conditional on rainfall).

If supply shocks affected electricity prices instead of shortages, this would be equally interesting but should be interpreted differently. Column 5 regresses the median electricity price paid by ASI plants in each state and year on the instrument, finding no statistically significant correlation. This confirms that that when new plants or hydro production shift supply, distribution companies clear the market by adjusting outages, not by adjusting prices.

---

[27]Online Appendix Figure A5 and Table A5 present more detail.

Table 4: **Evaluating the Hydro and New Supply Instrument**

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Shortage	ln(Energy Available)	ln(Assessed Demand)	ln(Agri Output)	ln(Median Price)
Hydro & New Supply	-0.177	0.200	0.007	-0.242	-0.168
	(0.059)***	(0.072)***	(0.089)	(0.245)	(0.139)
Number of Obs.	536	536	536	504	509

Notes: Observations weighted by number of ASI establishments in the state-year cell. Robust standard errors. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

6 Empirical Results

6.1 First Stages

Table 5 presents first stage estimates. There are two panels of estimates, corresponding to two groups of outcomes: energy inputs and then materials, labor, revenue, and TFPR. Regressions with Self-Generation Share and ln(Fuel Revenue Share) include only the 54 percent of plants in the 2005 ASI that ever self-generate electricity, because Self-Generation Share is mechanically zero for non-generators, and our model similarly predicts zero impact of outages on Fuel Revenue Share for non-generators.[28] For this reason, and also because of missing data and outlier flags, the sample sizes differ across columns.

While the exact coefficient estimates differ slightly due to the different samples, the estimates across all columns are very similar to each other and to the state-level estimates in column 1 of Table 4. The instruments are powerful: the cluster and heteroskedasticity-robust Kleibergen-Paap F-statistics range from 17 to 20. For comparison, the Stock and Yogo (2005) critical values for one instrument and one endogenous regressor are 8.96 and 16.38 for maximum 15 and 10 percent bias, respectively.

6.2 Effects of Shortages

6.2.1 Energy Inputs

Table 6 presents estimates of Equation (21) for energy inputs. Panels A and B present OLS and instrumental variables results, respectively. The IV estimates show that a one percentage point increase in Shortage (for example, an increase from 0.1 to 0.11) causes manufacturers to increase the share of electricity self-generated by 0.463 percentage points. If manufacturing electricity demand were fully inelastic and shortages mapped one-for-one into increases in outages δ for manufacturers, then the coefficient on Shortage in column 1 would be one. However, plants with generators will

---

[28]Online Appendix Table A18 plus additional regressions confirm that shortages have no effect on fuel revenue shares for non-generators.

Table 5: **First Stages for Base IV Estimates**

Panel A: Energy Inputs

	(1)	(2)	(3)
Second Stage	Self-Gen	ln(Fuel	ln(Electric
Dependent Var:	Share	Rev Share)	Intensity)
Hydro & New Supply	-0.179	-0.184	-0.168
	(0.0404)***	(0.0417)***	(0.0400)***
Obs.	240,743	291,759	479,616
Clusters	47,575	55,939	111,819
Clusters (2)	535	535	536
1st Stage F-Stat	19.74	19.36	17.73

Panel B: Materials, Labor, Revenue, and TFPR

	(1)	(2)	(3)	(4)	(5)
Second Stage			ln(Earnings/		
Dependent Var:	ln(Materials)	ln(Workers)	Worker)	ln(Revenue)	ln(TFPR)
Hydro & New Supply	-0.165	-0.165	-0.173	-0.165	-0.168
	(0.0400)***	(0.0400)***	(0.0419)***	(0.0401)***	(0.0398)***
Obs.	495,043	502,724	456,443	501,130	479,313
Clusters	115,040	116,803	110,213	116,231	112,371
Clusters (2)	536	536	482	536	536
1st Stage F-Stat	17.05	16.99	17.09	16.95	17.75

Notes: This table presents the first stage estimates for the IV regressions in Tables 6 and 7. 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 plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

substitute away from electricity through the input tax effect, and electricity distribution companies may impose more or less of the marginal shortage on manufacturers instead of agricultural, residential, or other consumers.

The IV estimate in column 2 shows that a one percentage point increase in Shortage increases fuel revenue share by 3.318 percent. Using the fact that the mean fuel revenue share is 0.055, the point estimate suggests that a one percentage point increase in shortages increases fuel costs by $(3.318 \times 1\%) \times 0.055 \approx 0.18$ percent of revenues. Even a large shortage increase thus imposes only a relatively small fuel cost increase to plants with generators - and this is largely offset by a decrease in purchased electricity costs.

Can plants become less electricity-intensive, in the short-run, as electricity becomes scarce? Column 3 shows that on average, the answer is no: shortages do not statistically affect the ratio of physical electricity use to revenue. The standard errors rule out that the average plant responds to a one percentage point increase in shortages by reducing electric intensity by more than one percent.

The OLS point estimates differ from the IV estimates, and the directions are consistent with two

biases hypothesized earlier: simultaneity bias (economic growth causes shortages) and measurement error. Self-Generation Share is less subject to the former bias, as it is not clear that economic growth would have any direct effect on this variable. The fact that the OLS point estimates are smaller than the IV estimates suggests that instrumenting helps slightly to address attenuation bias. The variables in columns 2-3, however, are ratios to revenues, which increase with economic growth. The coefficients are thus strongly downward biased.

Table 6: **Effects of Shortages on Energy Inputs**

	(1)	(2)	(4)
	Self-Gen	ln(Fuel	ln(Electric
Dependent Var:	Share	Rev Share)	Intensity)
Panel A: OLS			
Shortage	0.282	0.917	-0.539
	(0.0337)***	(0.176)***	(0.122)***
Panel B: IV			
Shortage	0.463	3.318	0.173
	(0.142)***	(0.955)***	(0.698)
Obs.	240,743	291,759	479,616
Clusters	47,575	55,939	111,819
Clusters (2)	535	535	536
1st Stage F-Stat	19.74	19.36	17.73

Notes: This table presents estimates of Equation (21). Panel B instruments for Shortage using hydro availability and new supply. 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 plant and state-year. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

6.2.2 Non-Energy Inputs, Revenue, and TFPR

Table 7 presents effects of shortages on non-energy inputs, revenue, and revenue productivity. Column 1-3 measure impacts on materials and labor. Column 1 shows that a one percentage point increase in Shortage causes a 1.216 percent decrease in materials input. Column 2 shows that shortages do not statistically affect labor input, as measured by total persons ever employed during the year. This is not entirely surprising, since our instrument captures year-to-year variation in shortages that might not be observed by the plant at the time it hires labor. Even though plants might not be able to reduce their number of workers, they might reduce hours worked in response to shortages. However, column 3 confirms that plants are unable to reduce earnings per worker as shortages worsen. This suggests that materials inputs tend to be determined after plants recognize electricity supply shifts - perhaps on a day-to-day basis - while labor input is less flexible.

Columns 4 and 5 present effects on log revenue r and log TFPR ω, respectively. The IV estimates in column 4 show that a one percentage point increase in shortages causes a 1.182 percent decrease in revenues. In the special case distribution companies impose the marginal shortage proportionally

Table 7: **Effects of Shortages on Materials, Labor, Revenue, and TFPR**

	(1)	(2)	(3)	(4)	(5)
			ln(Earnings/		
Dependent Var:	ln(Materials)	ln(Workers)	Worker)	ln(Revenue)	ln(TFPR)
Panel A: OLS					
Shortage	-0.00711	-0.0138	0.161	0.116	0.0543
	(0.0631)	(0.0461)	(0.0421)***	(0.0631)*	(0.0387)
Panel B: IV					
Shortage	-1.216	-0.302	-0.225	-1.182	-0.297
	(0.473)**	(0.313)	(0.199)	(0.498)**	(0.236)
Obs.	495,043	502,724	456,443	501,130	479,313
Clusters	115,040	116,803	110,213	116,231	112,371
Clusters (2)	536	536	482	536	536
1st Stage F-Stat	17.05	16.99	17.09	16.95	17.75

Notes: This table presents estimates of Equation (21). Panel B instruments for Shortage using hydro availability and new supply. 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.

to manufacturers vs. other consumers, no plants self-generate, plants have no yearly-flexible inputs, and δ is small, this coefficient would be approximately one - for example, an increase in shortages from 0 to 0.01 would reduce revenue by one percent. In reality, some plants have generators, which attenuates the revenue effects, while all plants can reduce yearly-flexible inputs, which magnifies the revenue effects.

Column 5 shows that shortages have statistically zero effect on log TFPR, with a 90-percent confidence interval from -0.69 to 0.094 percent. Given the standard errors, a statistically zero effect is exactly what the model predicts. For plants without generators, Equation (11) predicts a ln(TFPR) loss of $v_{it} = (1 - \alpha_M - \alpha_E) \ln(1 - \delta)$. As shown in Table 3, the mean α_M and α_E are 0.71 and 0.019, respectively. Thus, a one percentage point increase in δ would reduce non-generators' TFPR by approximately $(1-0.71-0.019) \times 1\% \approx 0.27\%$, and the effect on self-generators' TFPR should be smaller. In Section 7, we show that our point estimate of -0.297 is very close to the simulated model prediction.

The model also predicts that the ln(TFPR) loss is weakly smaller than the ln(revenue) loss, because the revenue loss equals the TFPR loss plus losses from reductions in yearly-flexible and fully-flexible inputs. The point estimates in columns 4 and 5 bear this out. Statistically, the TFPR coefficient in column 5 is less negative than the revenue coefficient in column 4 with a p-value of 0.054 in a one-sided test and 0.11 in a two-sided test.

Comparing the OLS estimates in Panel A to the IV estimates in Panel B strongly suggests simultaneity bias. All input, output, and TFPR coefficients are positively biased relative to the IV estimates, suggesting that the improved economic conditions that increase shortages also increase wages, productivity, and revenue. Without instrumenting for shortages, one might falsely conclude

that shortages are good for manufacturers.

6.3 Robustness Checks

Tables 6 and 7 are highly robust to a battery of robustness checks and alternative specifications. Here we present the overview; see Online Appendix E for details.

Estimates are very similar and statistically indistinguishable when using the difference estimator, omitting industry-by-year controls μ_{jt}, eliminating or tightening the criteria for dropping outliers, or using the CEA's estimated Peak Shortage or supply (i.e. ln(Energy Available)) instead of Shortage as the endogenous right-hand-side variable. The fact that using supply instead of Shortage as the endogenous variable gives similar estimates is unsurprising, because as shown in Table 4, the instrument affects only supply, not demand. It is reassuring, however, because unlike Shortage, Energy Available does not rely on the CEA's assessment of counterfactual demand. The similarly in results suggests that the Shortage variable does not suffer from some non-classical measurement error that an instrumental variable cannot address.

Clustering by state instead of using two-way clustering only mildly changes the standard errors; no discrete significance levels change, and no first-stage F-statistic drops below 10. Estimates are economically similar and statistically indistinguishable when using linear rainfall or 100mm or 50mm rainfall bins instead of 60mm bins, or when using rainfall data from the National Climate Centre instead of University of Delaware.

Estimates are statistically indistinguishable when constructing the instrument with actual hydro generation for H_{st} instead the prediction from reservoirs and run-of-river plants, except that the Self-Generation Share coefficient is statistically larger and the ln(Electric Intensity) coefficient is statistically negative. This latter result suggests that plants may indeed become less electric intensive as electricity shortages worsen. Constructing the instrument with only hydro generation and not new supply N_{st} reduces the first stage F-statistics by about 15 percent and increases the IV standard errors slightly, but again the estimates are economically and statistically similar. Additional regressions show that using only new supply and not hydro generation to construct Z_{st} gives a highly underpowered first stage. Thus, the role of the new supply part of the instrument is purely to slightly increase precision.

Finally, Table A17 in the Online Appendix shows that results are qualitatively similar under six different approaches to calculating production functions and TFPR.

6.4 Associations with Longer-Term Shortage Variation

Our main IV results identify the effects of year-to-year shortage variation. How do shortages affect manufacturers in the long run? Rigorous causal evidence is challenging, because we do not have a plausible instrument for long-run shortage variation. This section addresses long-run effects using correlational evidence.

Specifically, we test how shortages preceding the year of entry are associated with entrant plants' technology choices. Define Y_{ijst} as an outcome for plant i which enters in year t, and define ζ_s as state-specific indicator variables. $\bar{S}_{st}$ is the average of Shortage in state s in year t and the one year before; this proxies for potential entrants' beliefs about future shortages. Averaging over two years both reduces measurement error and smoothes over idiosyncratic year-to-year variation that should not impact potential entrants' beliefs about future shortages. We run the following regression with plant-level data:

$$Y_{ijst} = \rho \bar{S}_{st} + \Lambda_{st} + \zeta_s + \theta_t + \mu_{jt} + \epsilon_{st} \tag{22}$$

Standard errors are clustered by state.

Table 8 presents results. Column 1 considers the industry average electric intensity, i.e. the average ratio of electricity use to revenues for all plants in the same three-digit industry as plant i. (This column excludes industry-by-year indicators μ_{jt}.) For readability, electric intensity is normalized to units of kWh per 100 Rupees of revenue; the mean is 1.35 kWh/100 Rupees. A negative coefficient would imply that plants in electricity-intensive industries are less likely to enter when shortages are higher. This would be predicted by an extension of our model in Section 4 in the direction of Melitz (2003), in which potential entrant plants decide whether to enter based on expectations of δ and their productivity shock Ω_{it}. For each expected δ and production technology α, there is a cutoff productivity $\underline{\Omega}(\delta, \alpha)$ above which plants will enter. An increase in δ reduces profitability more (and thus increases $\underline{\Omega}(\delta, \alpha)$ more) for self-generators in industries with high α_E, because more self-generation causes larger cost increases when α_E is large. Thus, an increase in expected shortages should reduce entry more for higher-α_E industries, reducing entrants' overall average electric intensity. Indeed, the point estimate shows that a one percentage point increase in shortages is associated with a 0.00463 kWh/Rs decrease in the industry-level electric intensity of entrant plants, or about 0.34 percent of the mean.

Higher shortages should also increase the return to generator ownership. Indeed, Column 2 shows that shortages in the entry year $\bar{S}_{st}$ are positively associated with higher generator ownership rates for entrants. This association is small, however: a one percentage point increase in shortages is associated with a 0.246 percentage point increase in generator ownership. Such a small coefficient is quite remarkable: generator ownership rates vary substantially across states, ranging from a low near 10 percent in some lower-income states to more than 80 percent in Uttar Pradesh, New Delhi, and Haryana. Column 2 suggests that little of this variation is driven by state average shortages.

Aside from the generator adoption and extensive margin choices, entrant plants can also choose production technologies. Columns 3-5 test whether shortages at the time of entry are associated with entrants' factor shares.[29] Theory gives no unambiguous prediction about whether shortages

---

[29]Specifically, these are the natural log of the ratio of average input to average revenue, where the averages are taken across all years a plant is observed. The logs help to reduce the influence of outliers, which is the same role that median regressions play when estimating the production function parameters α in Section 4.5.

Table 8: **Association Between Entering Plant Characteristics and Shortages**

Dependent Variable:	(1) Ind. Elec/ Rev	(2) 1(Self Gen)	(3) ln(Capital/ Rev)	(4) ln(Labor/ Rev)	(5) ln(Matls/ Rev)
Shortage at Entry	-0.463	0.246	-0.701	-0.078	-0.025
	(0.179)**	(0.074)***	(0.422)	(0.135)	(0.073)
Number of Obs.	100,496	100,496	99,306	100,246	99,529
Industry-by-year effects	No	Yes	Yes	Yes	Yes

Notes: This table presents estimates of Equation (22). "Shortage at Entry" is the mean of S_{st} in the plant's entry year and the one year before. Robust standard errors, clustered by state. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

should increase or decrease various factor shares. For example, higher shortages increase generator adoption, which mechanically increases capital stock. On the other hand, higher shortages also decrease expected capital utilization, especially for non-generators, and this could cause them to switch to a less capital-intensive technology. Column 3 shows that a one percentage point increase in shortages is associated with a marginally significant 0.7 percent decrease in capital intensity. As predicted, additional (unreported) regressions shows that the negative association is stronger ($\hat{\rho} = -0.96$) for non-generators. Columns 4 and 5 show that labor and materials shares are not associated with $\bar{S}_{st}$.

7 Simulations

7.1 Overview

In this section we apply the model in Section 4 to the plants and production function parameters in the Annual Survey of Industries data. There are several objectives. First, we simulate effects of shortages and test whether they line up with the estimates in Section 6. Second, we simulate heterogeneous effects of shortages, which we do not have the power to detect empirically. Third, we simulate a counterfactual policy, offering interruptible electricity contracts, which could substantially reduce the effects of shortages.

The simulations consider the population of plants that appear in the ASI in 2005.[30] Each plant is characterized by estimated production function parameters α and exogenous state variables K_{i2005}, Ω_{i2005}, and generator ownership. For the price ratio of grid-purchased to self-generated electricity, we use the World Bank Enterprise Survey medians: $p^{E,S}/p^{E,G} = 7/4.5 \approx 1.56$. Given

---

[30]We chose 2005 because it allows comparison to the 2005 World Bank Enterprise Survey and because the assessed shortages are close to average: over the 1992-2010 sample, the average nationwide shortage ranged from 6.4 to 11.1 percent, and the mean reported in Table 1 is 7.6 percent. Online Appendix Figure A8 presents predicted revenue losses when simulating with each year between 1992 and 2010, showing that 2005 is not unusual.

this information and an assumed δ, we simulate production decisions using the first-order conditions in Section 4. See Online Appendix F for more details on the simulation procedure.

For the next several subsections until Section 7.5, our reported "effects of shortages" reflect a change from $\delta = 0$ to $\delta = S_{s2005}$, the CEA estimated Shortage in the plant's state in 2005. The mean S_{s2005} across all plants observed in 2005 is 7.2 percent. Although we have shown that the variation in S_{st} is meaningful for empirical estimates, we caution that this official assessment is a highly imperfect proxy for the level of Indian manufacturers' true outage probability δ.

To more closely match our empirical estimates, we weight by each plant's sampling weight and report the weighted average simulated effect of the change in δ. This is different than reporting output-weighted effects, which might be informative about aggregate sectoral implications. As we discuss below, output-weighted effects on most outcomes would be smaller, because the largest plants are much more likely to own generators.

7.1.1 Modeling Generator Adoption

We model the decision to purchase a generator as a binary choice: either a plant owns a generator that is large enough to provide full backup during a power outage, or it does not own a generator. Because the ASI does not include data on generator capacity, we construct a back-of-the-envelope estimate of each plant's capacity requirement: we transform a plant's total electricity consumption E_{it} (in kilowatt-hours) into required generation capacity assuming that plants use a constant flow of power while operating an average of six hours per day.[31] Under this assumption, the median generator capacity requirement is 500 kilowatts.

We model that a plant will purchase a generator if the simulated variable profit increase exceeds the cost. The simulated variable profit increase is just the simulated profits with vs. without a generator, given the plant's α, K_{i2005}, Ω_{i2005}, and an outage probability δ. We assume that generator cost $C^{\mathcal{G}}(c_i)$ is log-linear in plant i's required capacity c_i, i.e. $\ln C^{\mathcal{G}}(c_i) = \sigma_0 + \sigma_1 \ln(c_i)$.

We estimate σ_0 and σ_1 in GMM using the 1(Self-Generator) variable from the ASI and assuming $\delta = S_{st}$, matching mean generator adoption and also the covariance between generator adoption and $\ln(c_i)$. Estimates reflect strong economies of scale in generator costs: the estimated σ_1 is 0.83 (se=0.01), and this is statistically and economically similar to direct quotes of generator purchase costs. The estimated intercept σ_0 is also comparable to direct price quotes, implying that fixed costs are large enough to explain why so many Indian plants do not have generators. See Online Appendix 7.1.1 for details.

---

[31]Changing this assumption about the number of operating hours per day would only rescale our generator cost estimates; the simulation predictions would otherwise be unchanged.

7.2   Effect of Shortages

7.2.1   Comparing Empirical Estimates to Short-Run Simulations

Panel A of Table 9 presents the effects of changing δ from zero to $S_{s,2005}$ on average revenue, TFPR, self-generation share, material and labor inputs. Column 1 presents "exogenous generator simulations," i.e. simulated short-run effects holding fixed the generator ownership observed in the ASI. This is appropriate for comparison to the IV estimates because the instrument uses annual electricity supply shifters, and plants are unlikely to have responded by adopting or selling generators. Column 2 adds predictions based on the instrumental variable estimates: the predicted effects of a 7.2 percent shortage given the IV point estimates in Tables 6 and 7. Column 3 presents the p-value of a test that the simulation result differs from the IV estimate.

There is remarkable agreement between the empirical estimates and simulation results. Average revenue losses are 5.6 percent in the simulations and 8.4 percent from the IV. Both the simulations and IV predict that TFPR losses are much smaller than revenue losses, with predictions of 1.5 percent and 2.1 percent from the simulations and IV, respectively. Given that the TFPR effect is relatively small, the decrease in revenues must be caused by decreased input use. This is indeed the case: the model predicts a 5.6 percent decrease in materials and labor, while the IV estimates show 8.6 and 2.1 percent decreases in materials and labor, respectively. The simulated and estimated increases in Self-Generation Share also line up closely. Column 3 shows that the simulation and IV estimates are highly statistically indistinguishable except for the effects on labor input, which has a p-value of 0.12. The model assumes that labor is set with knowledge of the current year's shortage; if workers are instead hired before the effects of new power plant capacity or hydro availability are known, this would explain the difference between the estimated and simulated effects.

Column 4 presents self-reported revenue losses from the 2005 World Bank Enterprise Survey; this number is between the simulation and IV estimates. Overall, the close correspondence between our model and estimates using different datasets and identification strategies gives us confidence in the empirical estimates as well as our understanding of the underlying mechanisms.

7.2.2   Effects for Generators and Non-Generators

Panel B of Table 9 presents effects of the same change in δ (from zero to S_{s2005}) on producer surplus separately for plants with vs. without generators. Producer surplus is relevant because under our modeling assumption that output markets are perfectly competitive and there are no other market failures, the change in producer surplus is equivalent to overall welfare effects. We decompose producer surplus loss into generator fixed costs and variable profit changes, including the component due to decreases in measured TFPR. We compute fixed costs of generators observed in the 2005 ASI by applying estimated generator costs $\hat{\sigma}_0$ and $\hat{\sigma}_1$ to required capacity c_i, and we assume for simplicity that no plants would own generators in the scenario with $\delta = 0$. We find that generator costs amount to 7.7 percent of profit losses for plants with generators, or 3.9 percent

when averaged across all plants.

The simulations predict very different effects on plants with vs. without generators. Plants with generators lose only 0.3 percent of variable profits. They lose no output to shutdown; all revenue losses come from the decrease in the marginal products of materials and labor due to higher-cost self-generated electricity. The decrease is small because this implicit tax is small: given that electricity costs increase by 55 percent when self-generating, electricity has a 5 percent revenue share, and S_{s2005} averages 7.2 percent, the weighted average input cost rises by only about $55\% \times 5\% \times 7.2\% \approx 0.2$ percent of revenues.

In contrast, variable profits fall by 10 percent for the average plant without a generator. This effect is larger than the 7.2 percent of the time when the plants shut down, because the input tax effect magnifies the effect of shortages: expected power outages reduce the expected marginal revenue product of the semi-flexible input (labor) by δ, causing plants to reduce labor input.

The bottom row of Panel B shows that predicted effects on measured TFPR are also very different for non-generators vs. self-generators: 2.6 percent vs. almost zero, respectively, with an average over all plants of 1.5 percent. Recall that for plants with generators, the TFPR loss is due to the input variation effect: with a concave daily production function, it would have been more efficient to produce with a constant input bundle instead of different input bundles during outage vs. non-outage periods. This input variation effect is so small because plants do not reduce production very much during outages given the small implicit input tax.

By contrast, the TFPR loss for non-generators is much larger. To understand the magnitude, recall from the model that measured TFPR loss $v_{it} \approx -\delta(1 - \alpha_M - \alpha_E)$. In the limiting case with no fully flexible inputs, shutting down 7.2 percent of the time would cause a 7.2 percent TFPR loss. In another extreme with $\alpha_M + \alpha_E = 1$, there is also no TFPR loss, because with constant return to scale in the daily production function the plant could costlessly shift production to a non-outage day. In the data, the average $\alpha_E + \alpha_M$ is 0.73, and the simulations predict a TFPR loss is 2.6 percent for the average plant without a generator.[32]

7.3 Counterfactual with Interruptible Contracts

Given that 44 percent of manufacturing plants have generators, this "distributed generation" provides production capacity that would optimally be exploited when power is scarce. Currently, there are plants that have generators but do not use them because they receive grid power, while other nearby plants without generators simultaneously experience outages. Interruptible electricity contracts offer manufacturers and other electricity consumers a rebate for accepting outages during times of scarcity. If distribution companies offer both uninterruptible and interruptible contracts

---

[32]Online Appendix E.6 presents empirical estimates of how generator ownership moderates effects of shortages; we do not have the power to detect differences between plants with vs. without generators. Notice that such empirical estimates would likely understate the average causal effects of generator ownership because generator adoption is endogenous, so plants without generators will tend to have unobservably smaller losses. For example, plants with unobservably better electricity supply are both less likely to adopt generators and potentially less affected by an increase in shortages.

Table 9: **Effect of Shortages**

Panel A: Effects of Shortages: Model and IV Estimates

	(1)	(2)	(3)	(4)
			p-Value for	World
		IV	Columns	Bank
	Simulation	Estimate	(1) vs. (2)	Survey
Self-Generation Share Increase	3.1%	3.3%	(0.85)	
Materials Reduction	5.6%	8.6%	(0.37)	
Labor Reduction	5.6%	2.1%	(0.12)	
Revenue Loss	5.6%	8.4%	(0.29)	7.8%
TFPR Loss	1.5%	2.1%	(0.47)	

Panel B: Producer Surplus Effects of Shortages

	All	With Generators	Non-Generators
Producer Surplus Loss:	9.5%	8.0%	10.0%
of which Δ Generator Costs:	3.9%	7.7%	N.A.
of which Δ Variable Profit:	5.6%	0.3%	10.0%
of which Δ TFPR:	1.5%	0.0%	2.6%

Panel C: Producer Surplus Effects of Shortages with Interruptible Contracts

	All	With Generators	Non-Generators
Share of plants opting into interruptible contracts	8%	13%	3%
Variable Profit Loss:	0.4%	0.1%	0.7%
of which Δ TFPR:	0.0%	0.0%	0.1%

Notes: This table presents the effects of changing δ from zero to the CEA's 2005 assessed shortage level. In Panel A, "Simulation" refers to the predictions of the model using the 2005 ASI and described in the text. "IV Estimate" refers to the estimates in Tables 6 and 7, extrapolated under a 7.2 percent shortage. "p-Value" is the p-value for the test of whether the model's prediction with exogenous generators is equal to the IV estimate, with standard errors calculated using the Delta method. "World Bank Survey" refers to self-reported data from the 2005 World Bank Enterprise Survey. Panels B and C are simulated effects, as described in the text.

and allow consumers to sort into their preferred contract, this provides a mechanism to allocate outages to plants that are least affected.[33]

We simulate the effects of allowing plants to select into one of two electricity supply contracts, an uninterruptible contract with no outages and an interruptible contract that will allow outages 14.4 percent of the time - twice the assessed national average outage rate. The market-clearing rebate for the interruptible contract is pinned down by the maximum profit loss among the plants comprising 50 percent of manufacturers' grid electricity consumption. When this 50 percent of consumption is interrupted 14.2 percent of the time, this allows the uninterruptible contracts to be fulfilled. Panel C of Table 9 presents the effects of shortages under this contract structure. Because larger plants are more likely to have generators and are thus willing to accept interruptible contracts at lower rebates, only eight percent of plants need to opt into the interruptible contract to clear the market. Under this counterfactual policy, shortages of $\delta = S_{s2005}$ instead of $\delta = 0$ reduce variable profits by 0.4 percent, more than an order of magnitude less than the 5.6 percent loss when outages are evenly distributed within states. With interruptible contracts, there is no noticeable decrease in TFPR.

7.4 Comparing Effects for Large vs. Small Plants

How do shortages differentially affect small vs. large plants? Panel (a) of Figure 4 shows that the share of plants that report self-generating electricity increases with plant size. The figure reflects the same economy of scale in generator ownership discussed above. Because outages more severely affect plants without generators, this economy of scale makes small plants more exposed to outages than large plants. The solid line in Panel (b) of Figure 4 illustrates this, showing the relationship between employment and the predicted variable profit loss from shortages in the exogenous generator simulation. While the average ten-employee plant is predicted to lose about six percent of profits, the average 100-employee plant loses only about two percent. These differences between large and small plants are driven entirely by differential generator ownership: the dashed line in Panel (b) shows that the predicted profit losses would be roughly constant in plant size in an alternative simulation in which no plants have generators.

7.5 Effects of Shortages with Endogenous Generator Adoption

Given that generators moderate the effects of shortages, increased shortages should increase generator adoption. How much does this moderate the producer surplus losses from shortages? Figure 5 plots the effects of raising all plants' δ from zero to a homogeneous δ up to 0.2, as listed on the x-axis. Panel (a) uses the "exogenous generator simulations," while Panel (b) uses "endogenous generator simulations," which endogenize generator adoption at the given δ using the model

---

[33]Interruptible contracts have been studied by Baldick, Kolos, and Tompaidis (2006) and are common in the U.S. They are more suitable for large electricity consumers such as manufacturing plants, because a fixed cost is required to install switches to turn off electricity to the specific establishment.

detailed above in Section 7.1.1.[34]

In the exogenous generator simulation in Panel (a), the variable profit losses increase approximately linearly from zero to 18 percent as shortages worsen from zero to 20 percent. Since generator capital stock is held fixed in this simulation, the fixed cost of generators is invariant to shortages at 2.9 percent.

Panel (b) illustrates a starkly different pattern when generator ownership is endogenously adjusted. As shortages worsen, the variable profit loss peaks at 1.7 percent with a three percent shortage, drops slightly, and approaches only 1.5 percent as shortages approach 20 percent. The reason for this non-monotonicity is that many plants adopt generators at low levels of δ, and for these low δ, generator adoption more than offsets the "direct effect" of an increase in δ. Once δ is large and most plants have generators, the effects of shortages rise relatively slowly with increases in δ. Generator investments are costly, however, and at a 20 percent shortage, generator fixed costs decrease profits by 4.2 percent. Thus, as shortages worsen, the cost of shortages increasingly consists of the cost of purchasing backup generators.

---

[34]Online Appendix Table A22 presents more detailed statistics used to construct these figures, and Online Appendix Figure A9 presents generator adoption rates as a function of δ.

Figure 4: **Generator Ownership and Shortage Effects by Plant Size**

kernel = epanechnikov, degree = 0, bandwidth = .12, pwidth = .18
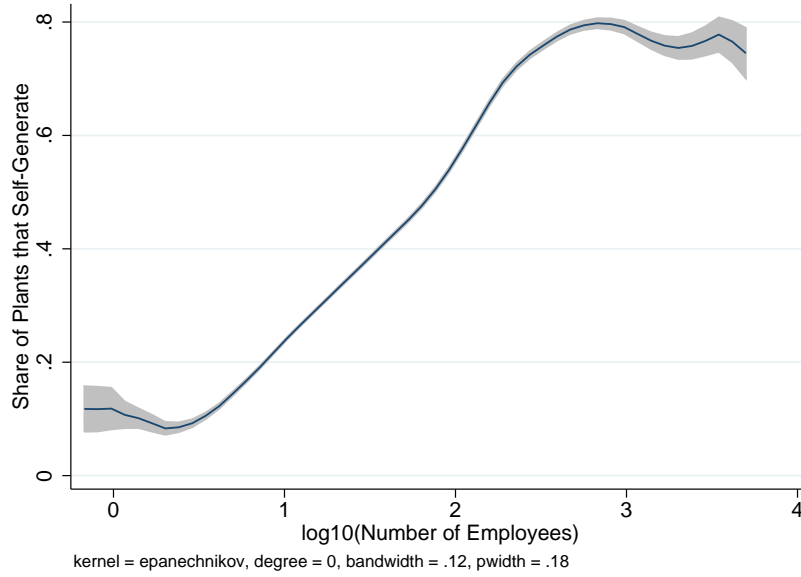
(a) **Generator Ownership and Plant Size**
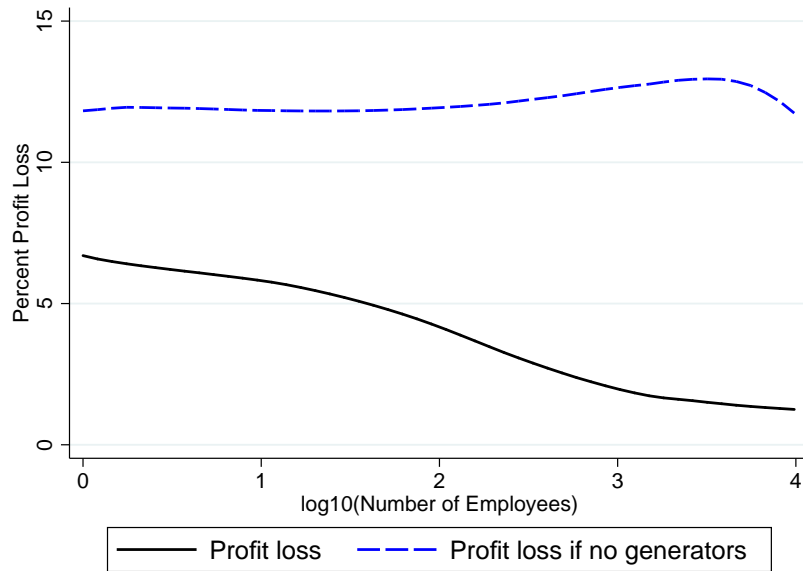
(b) **Simulated Profit Loss and Plant Size**

Notes: Panel (a) presents local mean-smoothed estimates of the share of plants in all years of the Annual Survey of Industries that ever self-generate electricity, as a function of number of employees. Panel (b) shows the simulated effect of shortages on profits as a function of the number of employees.

Figure 5: **Counterfactuals Under Varying Shortage Levels**
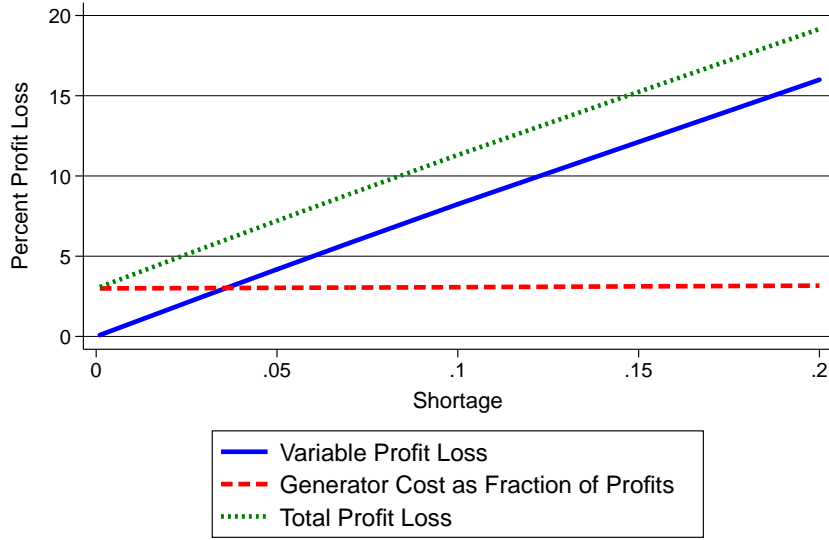
(a) **Profit Loss (Exogenous Generators)**
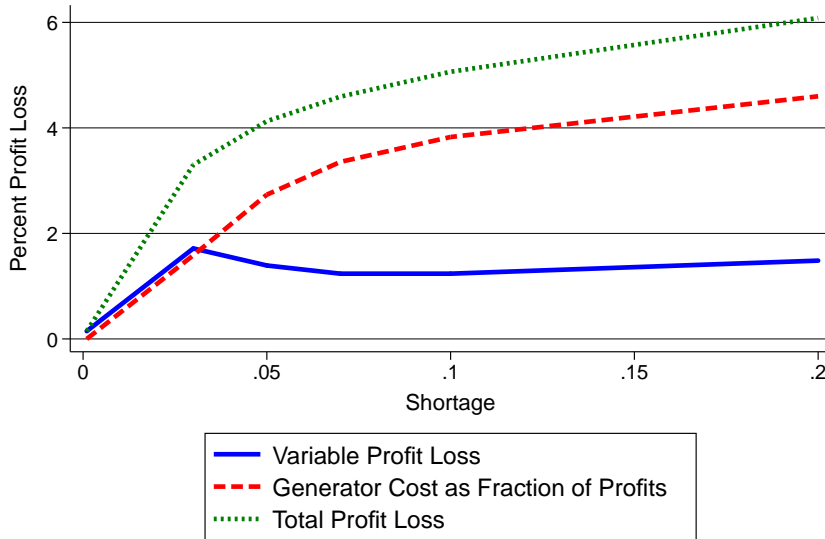
(b) **Profit Loss (Endogenous Generators)**

Note: These figures show the simulated effects of assigning the δ on the x-axis to all plants in the 2005 ASI, relative to $\delta = 0$. Panel (a) holds fixed the generator ownership observed in the ASI. Panel (b) uses the model in Section 7.1.1 to endogenously adjust generator ownership at the given δ.

8 Conclusion

India's lack of reliable electricity supply is a stark example of how poor infrastructure might affect economic growth. But while the problem is apparent, there had been no quantification of the resulting losses, making it difficult to prioritize relative to other economic distortions. We estimate the effects of shortages on manufacturers using archival data on shortages, instrumenting for shortages with supply shifts from hydro availability and newly-constructed power plants.

There are four main conclusions. First, we estimate that shortages are a substantial drag on Indian manufacturing, reducing revenue by 5.6 percent (in the simulations) to 8.6 percent (in the IV estimates) for the average plant in the short run. Producer surplus drops by 9.5 percent for the average plant, of which 3.9 percent is due to the capital costs of backup generators. Second, because plants also reduce inputs in response to shortages, shortages affect productivity much less than they affect revenue. Thus, at least in the short run, electricity shortages alone do not explain much of the productivity gap between firms in developing vs. developed countries discussed by Banerjee and Duflo (2005), Hall and Jones (1999), Hsieh and Klenow (2009), and others. Third, because shortages more strongly affect plants without generators and there are substantial economies of scale in generator costs, shortages more severely affect small plants. This adds another distortion to the firm size distribution in developing countries, related to the discussion of Hsieh and Olken (2014), Tybout (2000), and others. Fourth, policy can help: even though some potential solutions to India's power shortage problem may be politically infeasible, more technocratic solutions such as interruptible contracts can substantially reduce the costs of shortages.

References

[1] Abeberese, Ama Baafra (2012). "Electricity Cost and Firm Performance: Evidence from India." Working Paper, Columbia University (November).

[2] Alam, Muneeza (2013). "Coping with Blackouts: Power Outages and Firm Choices." Working Paper, Yale University (November).

[3] Alby, Philippe, Jean-Jacques Dethier, and Stephane Straub (2011). "Let There be Light! Firms Operating under Electricity Constraints in Developing Countries." Working Paper, Toulouse School of Economics (July).

[4] Aschauer, David (1989). "Is Public Expenditure Productive?" *Journal of Monetary Economics*, Vol. 23, No. 2 (March), pages 177-200.

[5] Asker, John, Allan Collard-Wexler, and Jan De Loecker (2014) "Dynamic Inputs and Resource (Mis)Allocation" *Journal of Political Economy* Vol. 122, No. 5, pages 1013-1063.

[6] Auffhammer, Maximilian, Solomon Hsiang, Wolfram Schlenker and Adam Sobel (2013). "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change." *Review of Environmental Economics and Policy*, Vol. 7, No. 2, pages 181-198.

[7] Baisa, Brian, Lucas Davis, Stephen Salant, and William Wilcox (2008). "The Welfare Costs of Unreliable Water Service." *Journal of Development Economics*, Vol. 92, No. 1 (May), pages 1-12.

[8] Baldick, Ross, Sergey Kolos, and Stathis Tompaidis (2006). "Interruptible electricity contracts from an electricity retailer's point of view: valuation and optimal interruption." *Operations Research,* Vol. 54, No. 4, pages 627-642.

[9] Banerjee, Abhijit, and Esther Duflo (2005). "Growth Theory through the Lens of Development Economics." In Aghion, Philippe, and Steven Durlauf, Eds., Handbook of Economic Growth.

[10] Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2012). "On the Road: Access to Transportation Infrastructure and Economic Growth in China." Working Paper, Yale University (February).

[11] Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta (2013). "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review*, Vol. 103, No. 1 (February), pages 305-334.

[12] Baum-Snow, Nathaniel (2007). "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, Vol. 122, No. 2, pages 775-805.

[13] Baum-Snow, Nathaniel (2013). "Urban Transport Expansions, Employment Decentralization, and the Spatial Scope of Agglomeration Economies." Working Paper, Brown University (October).

[14] Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew Turner, and Qinghua Zhang (2013). "Roads, Railroads, and the Decentralization of Chinese Cities." Working Paper, Brown University (October).

[15] Baum-Snow, Nathaniel, and Matthew Turner (2012). "Transportation and the Decentralization of Chinese Cities." Working Paper, Brown University (May).

[16] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004). "How Much Should We Trust Differences-In-Differences Estimates?" *Quarterly Journal of Economics*, Vol. 119, No. 1 (February), pages 249-275.

[17] Bhargava, Anjuli, and Kandula Subramaniam (2009). "Watts Not Happening." Business World, July 6th.

[18] Cameron, Colin, Jonah Gelbach, and Douglas Miller (2011). "Robust Inference with Multi-Way Clustering." *Journal of Business and Economic Statistics*, Vol. 29, No. 2, pages 238-249.

[19] CEA (Central Electricity Authority) (1994-2012a). "All India Electricity Statistics: General Review." New Delhi: Central Electricity Authority.

[20] CEA (Central Electricity Authority) (1993-2011b). "Load Generation Balance Report." New Delhi: Central Electricity Authority.

[21] CEA (Central Electricity Authority) (1993-2011c). "Review of Performance of Hydro Power Stations." New Delhi: Central Electricity Authority.

[22] CEA (Central Electricity Authority) (2013). "Planwise Capacity Addition." http://www.cea.nic.in/reports/planning/plan_addition.pdf

[23] Chakravorty, Ujjayant, Martino Pelli, and Beyza Ural Marchand (2013). "Does the Quality of Electricity Matter? Evidence from Rural India." Working Paper, Tufts University (April).

[24] Chan, Hei Sing (Ron), Maureen Cropper, and Kabir Malik (2014). "Are Power Plants in India Less Efficient than Power Plants in the United States?" *American Economic Review, Papers and Proceedings*, forthcoming.

[25] Chilkoti, Avantika, and James Crabtree (2014). "India Eyes End to State Coal Monopoly." *Financial Times* (October 23).

[26] Collard-Wexler, Allan and Jan De Loecker (2015). "Reallocation and Technology: Evidence from the US Steel Industry" *American Economic Review*, Vol. 105, No. 1, pp. 131-71.

[27] Cropper, Maureen, Alexander Limonov, Kabir Malik, and Anoop Singh (2011). "Estimating the Impact of Restructuring on Electricity Generation Efficiency: The Case of the Indian Thermal Power Sector." NBER Working Paper No. 17383 (September).

[28] De Loecker, Jan, and Frederic Warzynski (2012). "Markups and Firm-Level Export Status." *American Economic Review*, Vol. 102, No. 6 (October), pages 2437-2471.

[29] Dell, Melissa, Benjamin Jones, and Benjamin Olken (2014). "What Do We Learn from the Weather? The New Climate-Economy Literature." *Journal of Economic Literature*, Vol. 52, No. 3 (September), pages 740-798.

[30] Dinkelman, Taryn (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review*, Vol. 101, No. 7 (December), pages 3078-3108.

[31] Donaldson, David (2012). "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, forthcoming.

[32] Donaldson, David, and Richard Hornbeck (2013). "Railroads and American Economic Growth: A "Market Access" Approach." NBER Working Paper No. 19213 (July).

[33] Duflo, Esther, and Rohini Pande (2007). "Dams." *Quarterly Journal of Economics*, Vol. 122, No. 2, pages 601-646.

[34] Easterly, William, and Sergio Rebelo (1993). "Fiscal Policy and Economic Growth." *Journal of Monetary Economics*, Vol. 32, pages 417-458.

[35] Esfahani, Hadi Salehi, and Maria Teresa Ramirez (2002). "Institutions, Infrastructure, and Economic Growth." *Journal of Development Economics*, Vol. 70, pages 443-477.

[36] Fernald, John (1999). "Roads to Prosperity? Assessing the Link Between Public Capital and Productivity." *American Economic Review*, Vol. 89, No. 3 (June), pages 619-638.

[37] Fisher-Vanden, Karen, Erin Mansur, and Qiong (Juliana) Wang (2015). "Electricity Shortages and Firm Productivity: Evidence from China's Industrial Firms." *Journal of Development Economics*, Vol. 114, Pages 172-188.

[38] Foster, Vivien, and Jevgenijs Steinbuks (2009). "Paying the Price for Unreliable Power Supplies: In-House Generation of Electricity by Firms in Africa." World Bank Policy Research Paper 4913 (April).

[39] Garcia-Mila, T., McGuire, T. and Porter, R. H., (1996). "The effect of public capital in state-level production functions reconsidered." *The Review of Economics and Statistics*, Vol. 78, No. 1 (February), pages 177-180.

[40] Gramlich, Edward (1994). "Infrastructure Investment: A Review Essay." *Journal of Economic Literature,* Vol. 32 (September), pages 1176-1196.

[41] Gupta, Rajan and Harihar Shankar (2014). *Global Energy Observatory.* Available: http://globalenergyobservatory.org/. Accessed: September 20, 2014.

[42] Hall, Robert and Charles Jones (1999) "Why do Some Countries Produce So Much More Output Per Worker than Others?" *The Quarterly Journal of Economics* (1999), Vol. 114, No. 1, pages 83-116.

[43] Hausmann, Ricardo, Dani Rodrik, and Andrés Velasco (2008). "Growth diagnostics." in *The Washington consensus reconsidered: Towards a new global governance*, edited by Narcís Serra and Joseph Stiglitz, pages 324-355.

[44] Holtz-Eakin, Douglas (1994). "Public-Sector Capital and the Productivity Puzzle." *Review of Economics and Statistics*, Vol. 76, No. 1 (February), pages 12-21.

[45] Hsieh, Chang-Tai, and Peter Klenow (2014). "The Life Cycle of Plants in India and Mexico." *Quaterly Journal of Economics*, Vol. 129, No. 3, pages 1035-1084.

[46] Hsieh, Chang-Tai, and Peter Klenow (2009). "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, Vol. 124, No. 4 (November), pages 1403-1448.

[47] Hsieh, Chang-Tai, and Benjamin Olken (2014). "The Missing "Missing Middle."" *The Journal of Economic Perspectives*, Vol. 28, No. 3 (Summer), pp. 89-108.

[48] Hulten, Charles, Esra Bennathan, and Sylaja Srinivasan (2006). "Infrastructure, Externalities, and Economic Development: A Study of the Indian Manufacturing Industry." *World Bank Economic Review*, Vol. 20, No. 2, pages 291–308.

[49] Jensen, Robert (2007). "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *Quarterly Journal of Economics*, Vol. 122, No. 3 (August), pages 879-924.

[50] Lipscomb, Molly, Ahmed Mushfiq Mobarak, and Tania Barham (2013). "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics*, Vol. 5 No. 2, (April), pages 200–231.

[51] McRae, Shaun (2015). "Infrastructure Quality and the Subsidy Trap." *American Economic Review*, Vol. 105, No. 1, pages 35-66.

[52] Melitz, Marc (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, Vol. 71, No. 6 (November), pages 1695-1725.

[53] Power Finance Corporation (various years). "Report on The Performance of The State Power Utilities." New Delhi, India: Power Finance Corporation Limited.

[54] Reinikka, Ritva, and Jakob Svensson (2002). "Coping with Poor Public Capital." *Journal of Development Economics*, Vol. 69, pages 51-69.

[55] Roller, Lars-Hendrik, and Leonard Waverman (2001). "Telecommunications Infrastructure and Economic Development: A Simultaneous Approach." *American Economic Review*, Vol. 91, No. 4 (September), pages 909-923.

[56] Ryan, Nicholas (2013). "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market." Working Paper, MIT (September).

[57] Rud, Juan Pablo (2012a). "Electricity Provision and Industrial Development: Evidence from India." *Journal of Development Economics*, Vol. 97, pages 352-367.

[58] Rud, Juan Pablo (2012b). "Infrastructure Regulation and Reallocations within Industry: Theory and Evidence from Indian Firms." *Journal of Development Economics*, Vol. 99, pages 116-127.

[59] Srivastava, AK, M Rajeevan, and SR Kshirsagar (2009). "Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region." *Atmospheric Science Letters*, Vol. 10, No. 4 (October), pages 249-254.

[60] Steinbuks, Jevgenijs, and Vivien Foster (2010). "When Do Firms Generate? Evidence on In-House Electricity Supply in Africa." *Energy Economics*, Vol. 32, pages 505-514.

[61] Steinbuks, Jevgenijs (2011). "Financial Constraints and Firms' Investment: Results from a Natural Experiment Using Power Interruption." Working Paper, Purdue University (October).

[62] Stock, James, and Motohiro Yogo (2005). "Testing for Weak Instruments in Linear IV Regression." In Stock, James, and Donald Andrews, Eds., Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg. Cambridge, England: Cambridge University Press.

[63] Todd, Petra and Kenneth Wolpin (2006) "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility", *American Economic Review*, Vol. 96, No. 5, pages 1384-1417.

[64] Tybout, James (2000). "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature*, Vol. 38, No. 1 (March), pages 11-44.

[65] U.S. DOE (Department of Energy) (2013). "Manufacturing Energy Consumption Survey." http://www.eia.gov/consumption/manufacturing/index.cfm

[66] Willmott, Cort, and Kenji Matsuura (2012). "Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (1900 - 2010) V3.02." http://climate.geog.udel.edu/~climate/html_pages/download.html#P2011rev.

[67] Yardley, Jim, and Gardiner Harris (2012). "2nd Day of Power Failures Cripples Wide Swath of India." New York Times, July 31. http://www.nytimes.com/2012/08/01/world/asia/power-outages-hit-600-million-in-india.html

Online Appendix: Not for Publication

How Do Electricity Shortages Affect Industry? Evidence from India

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

A   2005 World Bank Enterprise Survey: Barriers to Growth

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?"

B   Power Sector Data Appendix

This appendix provides additional details on the power sector data in Section 2.2. Appendix Table A2 correlates the CEA Shortage variable with outside measures. Columns 1, 2, and 3 correlate outcomes Y_{ij} for firm i in industry j in state s from the 2005 World Bank Enterprise Survey with Shortage in 2005, controlling for industry indicators μ_j:

$$Y_{ijs} = \rho S_{s,2005} + \mu_j + \varepsilon_{ijs} \tag{23}$$

The dependent variable in column 1 is an indicator for whether the firm's manager reports that electricity is the primary barrier to growth. The dependent variable in column 2 is the manager's rating of the quality of power, from 1 (extremely bad) to 10 (excellent). The dependent variable in column 3 is Self-Generation Share; the sample in this column includes only plants that own generators. Standard errors are clustered by state. There are only 16 states in the WBES data, and this small number of clusters might cause standard errors to be biased. Reassuringly, however, the standard errors (and point estimates) are almost exactly identical in analogous state-level regressions with 16 data points, and the discrete significance levels are identical.

Column 4 correlates coal power plant capacity factors F with Shortage, conditional on year indicators, state split controls Λ_{st}, and plant fixed effects ϕ_d. Indexing power plants by d, we have:

$$F_{dst} = \rho S_{st} + \theta_t + \Lambda_{st} + \phi_d + \varepsilon_{dst} \tag{24}$$

Standard errors are clustered by state.

Appendix Table A3 presents the state allocations from jointly-owned power plants.

Appendix Table A4 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 A4 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 A5 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 A2: **Correlations with the Shortage Variable**

	(1)	(2)	(3)	(4)
	1(Largest Barrier)	Power Quality	Self-Gen Share	Capacity Factor
Shortage	1.107	-8.378	0.699	0.167
	(0.492)**	(3.383)**	(0.213)***	(0.062)**
N	2,280	2,265	1,124	1,286
Dependent Var. Mean	.31	6.38	5.9	.64
Sample	WBES Firms	WBES Firms	WBES Firms	Coal Plants

Notes: Sample for columns 1, 2, and 3 is the 2005 World Bank Enterprise Survey (WBES). Sample for column 4 is panel data on all Indian coal power plants from 1994-2009. Robust standard errors, clustered by state. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A3: **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 A4: **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 A5: **Electricity Statistics for the Ten Largest States**

State	1992-2010 Shortages			2010 Capacity (gigawatts)	1992-2010 Generation Share Hydro
	Mean	Min.	Max.		
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**
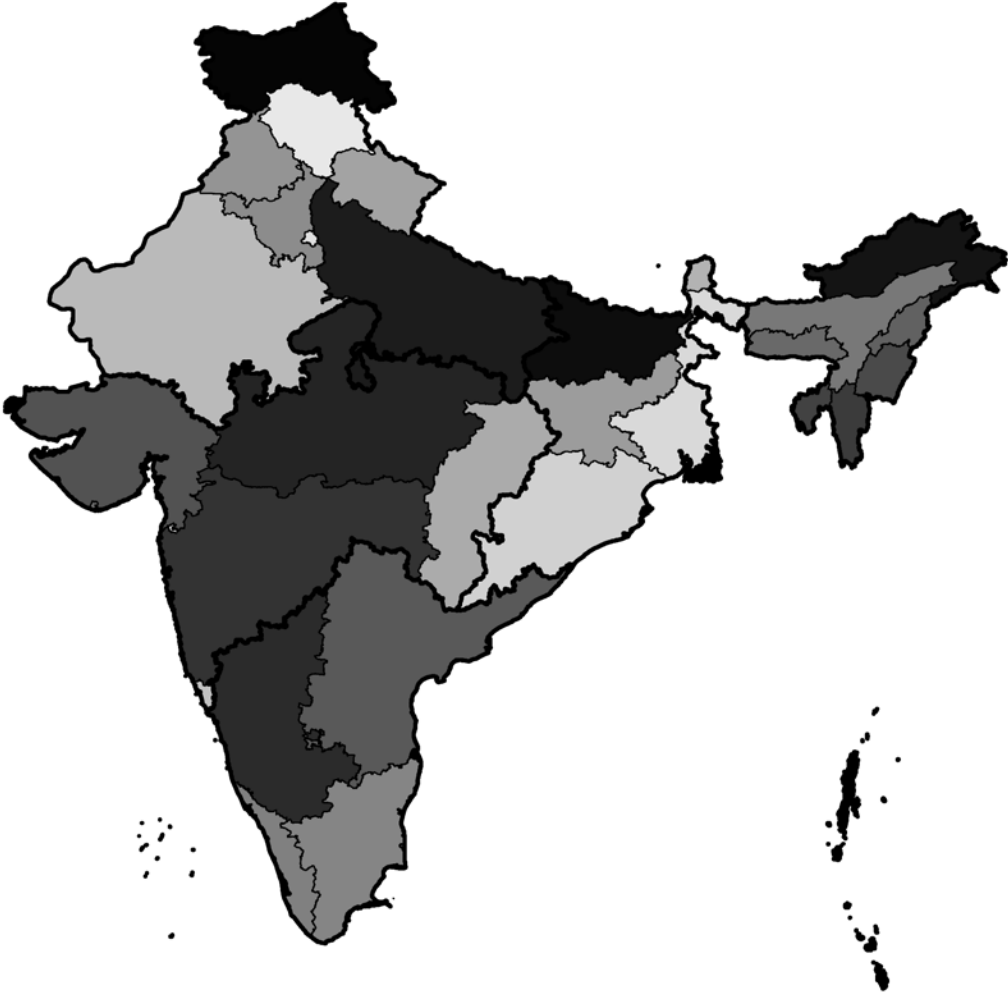
Notes: This figure presents each state's average Shortage assessed by the Central Electricity Authority over 1992-2010, with darker color illustrating higher Shortage.

Figure A2: **Hydro Power Plants, Rainfall Gridpoints, and Weather Stations**
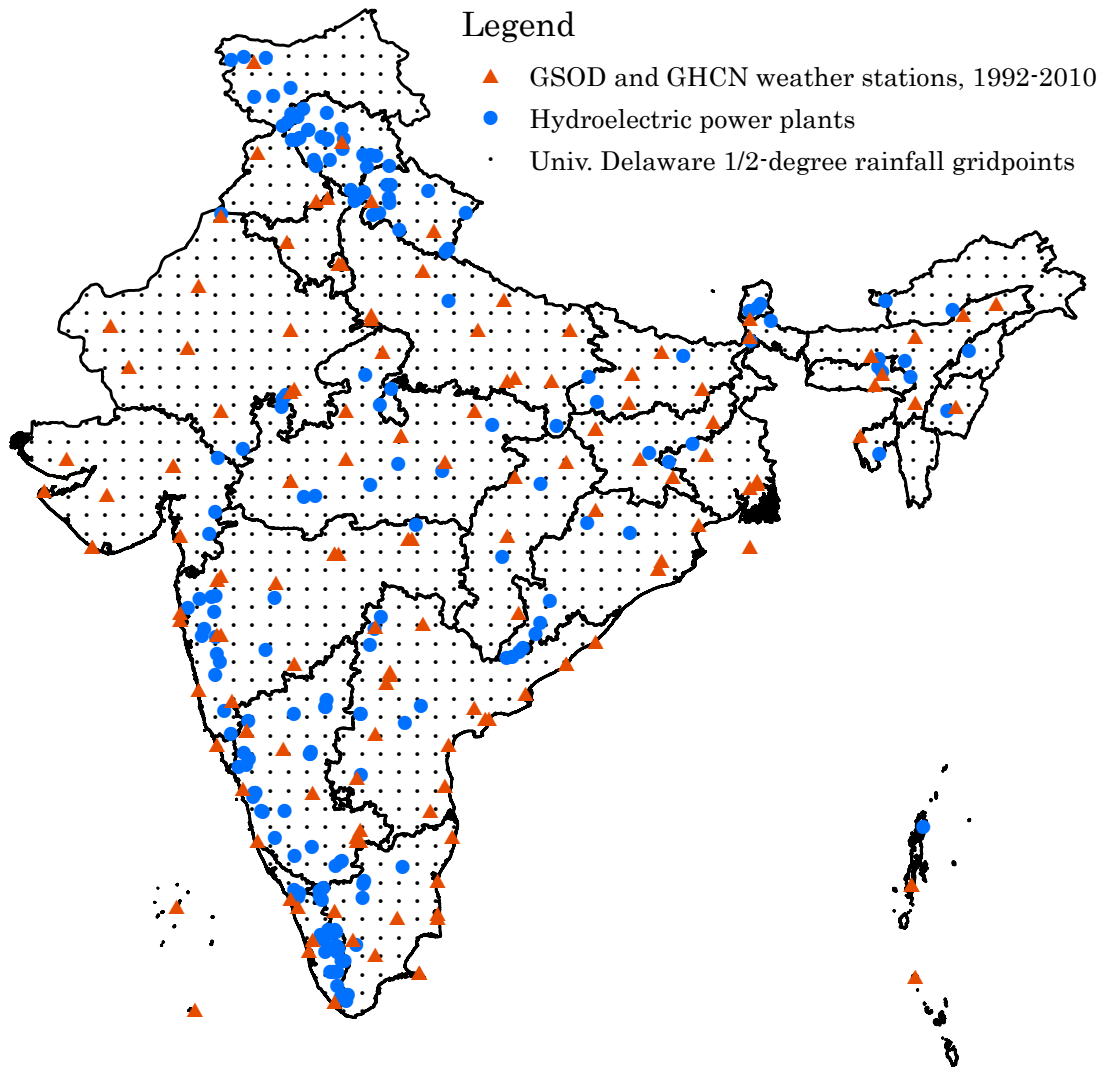
Notes: This figure plots the 1/2 degree gridpoints in the University of Delaware rainfall data, the weather stations whose measurements underlie the gridded data, and the locations of all hydroelectric power stations in India.
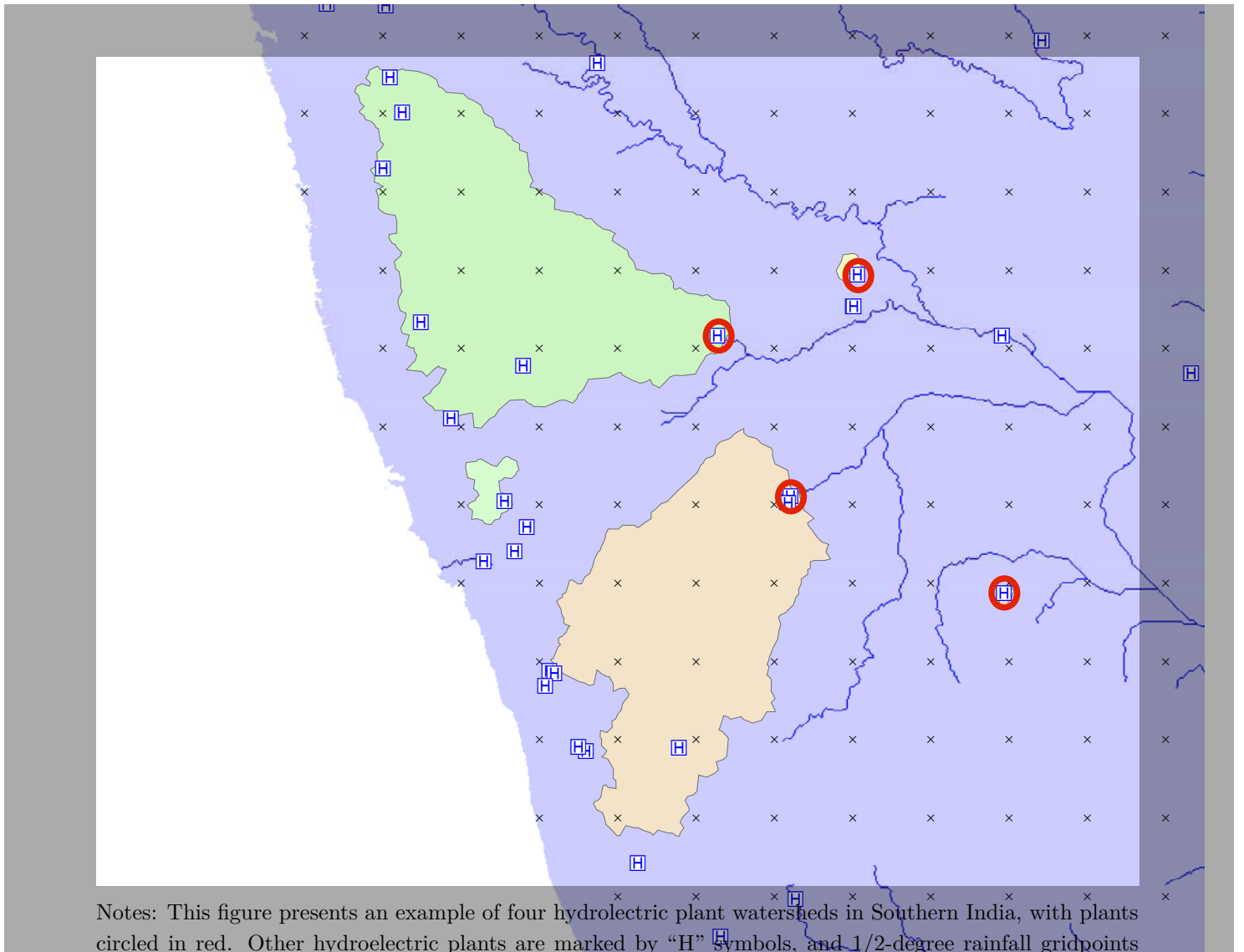
Notes: This figure presents an example of four hydrolectric plant watersheds in Southern India, with plants circled in red. Other hydroelectric plants are marked by "H" symbols, and 1/2-degree rainfall gridpoints are marked by "+" symbols. The river network is extrapolated based on the predicted flow of water accross India.

C   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.[35] 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.[36] 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.[37] 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 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

---

[35]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).

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

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

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.

C.1   Determination of Base Sample

Appendix Table A6 details how the sample in Table 2 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.

C.2   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:

- 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.[38] Because we also observe physical quantities for labor and

---

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

Table A6: **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 Table 2 is determined from the original set of observations in the Annual Survey of Industries.

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.

C.2.1 Cleaning Electricity Variables

We clean plant electricity measures in the following ways:

- 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.

C.2.2   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 A17. 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.

D   Empirical Strategy Appendix

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

Table A7: **Serial Correlation Tests for Hydro and New Supply Instrument**

	(1)	(2)
1st Lag Z	0.099	0.105
	(0.080)	(0.088)
2nd Lag Z		0.007
		(0.074)
3rd Lag Z		-0.123
		(0.066)*
4th Lag Z		-0.101
		(0.084)
5th Lag Z		-0.024
		(0.070)
Number of Obs.	506	386
F-Stat	1.52	1.67
R-Squared	0.01	0.06

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

Figure A4: **Shortages and Hydro Generation in Karnataka**
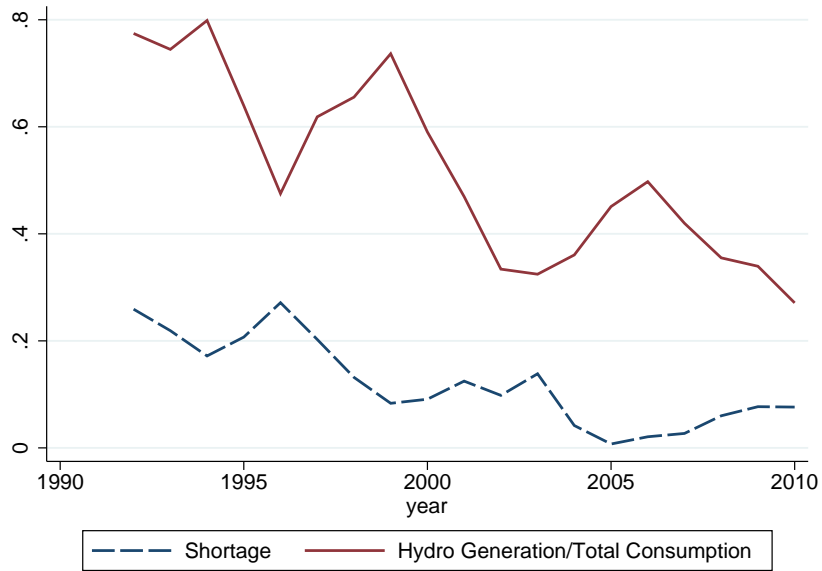
Notes: Shortage data are estimated by the Central Electricity Authority. Hydro Generation/Total Electricity Consumption is a simplified version of the hydro availability instrument. The figure gives a simple graphical example of the first stage of our IV estimator.

Figure A5: **Rainfall and Hydro Share of Electricity by State**
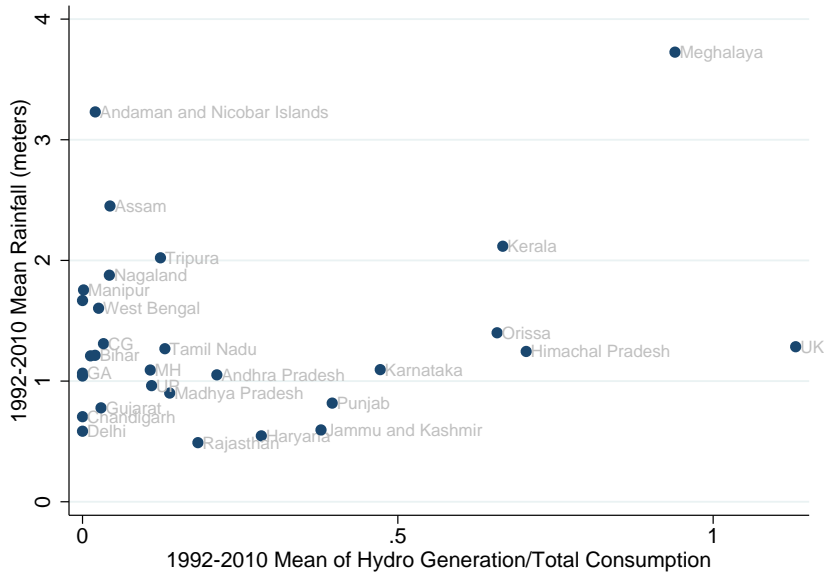
Notes: This figure plots sample average annual rainfall against the mean ratio of hydroelectricity generation to total electricity consumption. The figure emphasizes that there is substantial variation in hydro generation conditional on rainfall.

Figure A6: **Hydro Share of Predicted Consumption Over Time in Five Large States**
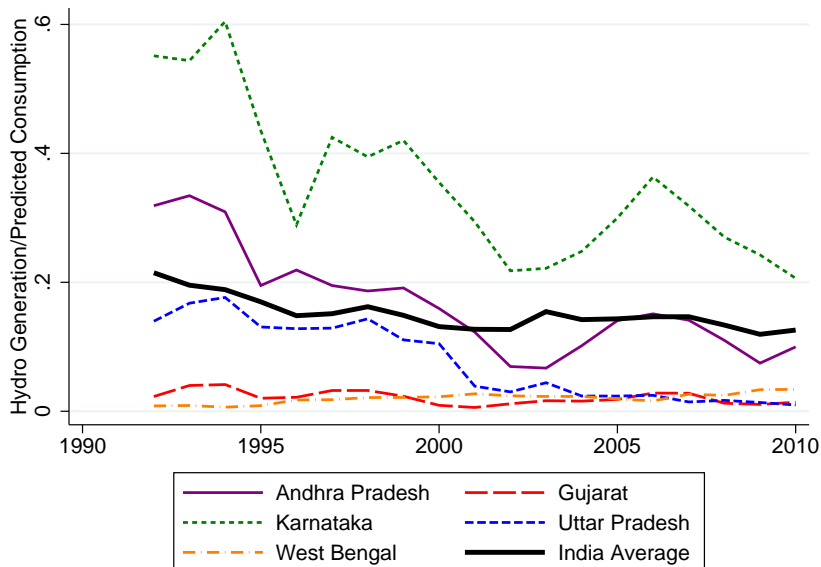
Notes: This figure presents the ratio of hydro generation to predicted consumption over 1992-2010 for five large states. Different states have different slopes, illustrating the importance of including state-specific time trends as control variables.

Figure A7: **ln(Energy Available) Over Time in Five Large States**
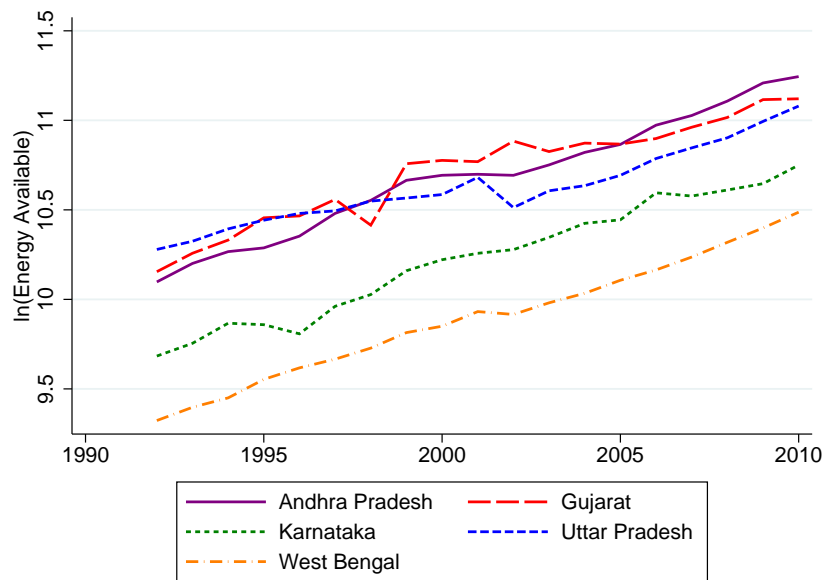
Notes: This figure presents the natural log of Energy Available over 1992-2010 for five large states. Different states have different slopes, illustrating the importance of including state-specific time trends as control variables.

E   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.

E.1   Estimates with Difference Estimator

Tables A8, A9, and A10 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 A8: **First Stages Using the Difference Estimator**

Panel A: Energy Inputs

	(1)	(2)	(3)
Second Stage	Self-Gen	ln(Fuel	ln(Electric
Dependent Var:	Share	Rev Share)	Intensity)
Hydro & New Supply	-0.179	-0.184	-0.168
	(0.0404)***	(0.0417)***	(0.0400)***
Obs.	240,743	291,759	479,616
Clusters	47575	55939	111819
Clusters (2)	535	535	536
1st Stage F-Stat	19.74	19.36	17.73

Panel B: Other Inputs and Outputs

	(1)	(2)	(3)	(4)	(5)
Second Stage			ln(Earnings/		
Dependent Var:	ln(Materials)	ln(Workers)	Worker)	ln(Revenue)	ln(TFPR)
Hydro & New Supply	-0.180	-0.180	-0.189	-0.180	-0.182
	(0.0402)***	(0.0402)***	(0.0432)***	(0.0403)***	(0.0403)***
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	20.02	19.97	19.03	19.96	20.55

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 A9: **Effects of Shortages on Energy Inputs Using the Difference Estimator**

	(1)	(2)	(4)
	Self-Gen	ln(Fuel	ln(Electric
Dependent Var:	Share	Rev Share)	Intensity)
Panel A: OLS			
Shortage	0.274	0.874	-0.595
	(0.0383)***	(0.202)***	(0.132)***
Panel B: IV			
Shortage	0.367	2.500	0.410
	(0.129)***	(0.677)***	(0.680)
Obs.	177,822	234,384	363,572
Clusters	504	504	506
Clusters (2)	505	505	506
1st Stage F-Stat	24.13	25.02	20.59

Notes: This table presents estimates of Equation (21) using the difference estimator. Panel B instruments for Shortage using hydro availability and new supply. 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 A10: **Effects of Shortages on Materials, Labor, Revenue, and TFPR Using the Difference Estimator**

	(1)	(2)	(3)	(4)	(5)
			ln(Earnings/		
Dependent Var:	ln(Materials)	ln(Workers)	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	-0.956	-0.375	-0.183	-0.816	-0.100
	(0.432)**	(0.293)	(0.207)	(0.459)*	(0.221)
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	20.02	19.97	19.03	19.96	20.55

Notes: This table presents estimates of Equation (21) using the difference estimator. Panel B instruments for Shortage using hydro availability and new supply. 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.

E.2 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.2, 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 A11: **Robustness Checks: Energy Inputs**

	(1)	(2)	(3)	(4)	(5)	(6)
	No Ind.-	Tighter	No	Use	Use	Cluster
	by-Year	Outlier	Outlier	ln(Energy	Peak	by
Change from Base Spec:	Effects μ_{jt}	Flags	Flags	Available)	Shortage	State
Self-Generation Share						
Shortage	0.484	0.412	0.473	0.461	0.429	0.463
	(0.146)***	(0.131)***	(0.146)***	(0.158)***	(0.161)***	(0.115)***
Number of Obs.	240,743	223,128	293,866	240,743	240,743	240,743
First Stage F-Stat	17.48	19.68	19.87	10.40	11.50	43.64
ln(Fuel Revenue Share)						
Shortage	3.725	2.756	3.105	3.231	3.170	3.318
	(1.084)***	(0.929)***	(1.206)**	(1.182)***	(1.203)***	(0.915)***
Number of Obs.	291,759	268,663	300,697	291,759	291,759	291,759
First Stage F-Stat	17.27	19.28	19.44	10.90	10.42	47.58
ln(Electric Intensity)						
Shortage	-0.0851	0.0155	0.135	0.157	0.138	0.173
	(0.680)	(0.652)	(0.698)	(0.620)	(0.554)	(0.618)
Number of Obs.	479,616	453,482	483,843	479,616	479,616	479,616
First Stage F-Stat	16.39	17.55	17.78	13.43	13.42	15.72

Notes: This table presents alternative estimates for Table 6, instrumenting for Shortage using hydro availability and new supply. 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 A12: **Robustness Checks: Materials, Labor, Revenue, and TFPR**

Change from Base Spec:	(1) No Ind.-by-Year Effects μ_{jt}	(2) Tighter Outlier Flags	(3) No Outlier Flags	(4) Use ln(Energy Available)	(5) Use Peak Shortage	(6) Cluster by State
	ln(Materials)					
Shortage	-1.123	-1.101	-1.272	-1.076	-0.993	-1.216
	(0.506)**	(0.442)**	(0.534)**	(0.408)***	(0.411)**	(0.513)**
Number of Obs.	495,043	478,152	498,464	495,043	495,043	495,043
First Stage F-Stat	15.77	16.99	17.06	13.78	12.29	14.94
	ln(Workers)					
Shortage	-0.277	-0.307	-0.297	-0.266	-0.247	-0.302
	(0.328)	(0.292)	(0.319)	(0.260)	(0.254)	(0.384)
Number of Obs.	502,724	496,474	503,217	502,724	502,724	502,724
First Stage F-Stat	15.71	16.98	17.01	13.97	12.26	14.82
	ln(Earnings/Worker)					
Shortage	-0.290	-0.348	-0.316	-0.207	-0.212	-0.225
	(0.219)	(0.200)*	(0.219)	(0.174)	(0.205)	(0.224)
Number of Obs.	456,443	440,524	461,131	456,443	456,443	456,443
First Stage F-Stat	15.70	17.21	16.89	12.84	9.310	10.99
	ln(Revenue)					
Shortage	-1.135	-1.097	-1.315	-1.042	-0.967	-1.182
	(0.513)**	(0.466)**	(0.576)**	(0.401)***	(0.436)**	(0.596)**
Number of Obs.	501,130	484,753	503,664	501,130	501,130	501,130
First Stage F-Stat	15.66	16.75	17.00	13.87	12.19	14.81
	ln(TFPR)					
Shortage	-0.0889	-0.248	-0.381	-0.264	-0.240	-0.297
	(0.232)	(0.209)	(0.256)	(0.204)	(0.199)	(0.314)
Number of Obs.	479,313	472,612	480,243	479,313	479,313	479,313
First Stage F-Stat	16.47	17.82	17.68	13.86	13.06	15.73

Notes: This table presents alternative estimates for Table 7, instrumenting for Shortage using hydro availability and new supply. 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.

E.3   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 A13: **Alternative Weather Controls: Energy Inputs**

	(1)	(2)	(3)	(4)
		100mm	50mm	NCC
	Linear	Rainfall	Rainfall	Rainfall
Change from Base Spec:	Rainfall	Bins	Bins	Data
Self-Generation Share				
Shortage	0.458	0.488	0.463	0.444
	(0.169)***	(0.150)***	(0.142)***	(0.145)***
Rainfall	0.00299			
	(0.00885)			
First Stage F-Stat	17.79	20.58	19.74	23.65
ln(Fuel Revenue Share)				
Shortage	3.008	3.204	3.318	2.092
	(1.025)***	(0.911)***	(0.955)***	(0.803)***
Rainfall	0.192			
	(0.0664)***			
First Stage F-Stat	17.17	19.89	19.36	23.16
ln(Electric Intensity)				
Shortage	0.0314	0.0809	0.173	0.143
	(0.733)	(0.666)	(0.698)	(0.627)
Rainfall	-0.0249			
	(0.0352)			
First Stage F-Stat	17.45	20.15	17.73	20.93

Notes: This table presents alternative estimates for Table 6, instrumenting for Shortage using hydro availability and new supply. 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 A14: **Alternative Weather Controls: Materials, Labor, Revenue, and TFPR**

	(1)	(2)	(3)	(4)
		100mm	50mm	NCC
	Linear	Rainfall	Rainfall	Rainfall
Change from Base Spec:	Rainfall	Bins	Bins	Data
ln(Materials)				
Shortage	-1.079	-1.092	-1.216	-1.013
	(0.456)**	(0.409)***	(0.473)**	(0.403)**
Rainfall	-0.0145			
	(0.0231)			
First Stage F-Stat	17.14	19.70	17.05	20.87
ln(Workers)				
Shortage	-0.287	-0.285	-0.302	-0.309
	(0.316)	(0.285)	(0.313)	(0.280)
Rainfall	0.00481			
	(0.0149)			
First Stage F-Stat	17.06	19.61	16.99	20.73
ln(Earnings/Worker)				
Shortage	-0.158	-0.186	-0.225	-0.170
	(0.195)	(0.181)	(0.199)	(0.181)
Rainfall	0.00255			
	(0.0114)			
First Stage F-Stat	17.92	19.76	17.09	22.69
ln(Revenue)				
Shortage	-1.034	-1.074	-1.182	-0.901
	(0.482)**	(0.436)**	(0.498)**	(0.412)**
Rainfall	-0.0292			
	(0.0237)			
First Stage F-Stat	16.98	19.53	16.95	20.64
ln(TFPR)				
Shortage	-0.302	-0.297	-0.297	-0.238
	(0.237)	(0.218)	(0.236)	(0.207)
Rainfall	-0.0142			
	(0.0114)			
First Stage F-Stat	17.73	20.27	17.75	21.32

Notes: This table presents alternative estimates for Table 7, instrumenting for Shortage using hydro availability and new supply. 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.

E.4   Alternative Instruments

This section presents estimates of Tables 6 and 7 with alternative instruments. Columns 1 and 2 isolate identification from hydro availability, constructing the instrument without new supply N_{st}. Column 1 uses actual hydro generation, while column 2 uses predicted hydro generation from reservoirs and run-of-river plants per Equation (20). Column 3 uses new supply and actual hydro generation instead of predicted hydro generation.

Table A15: **Alternative Instruments: Energy Inputs**

	(1)	(2)	(3)
	Without New Supply		With New Supply
	Actual	Predicted with	Actual
	Hydro	Run-of-River	Hydro
Instrument:	Generation	and Reservoirs	Generation
Self-Generation Share			
Shortage	0.794	0.442	0.788
	(0.176)***	(0.153)***	(0.167)***
Number of Obs.	240,743	240,743	240,743
First Stage F-Stat	17.61	17.00	19.44
ln(Fuel Revenue Share)			
Shortage	3.597	3.294	3.596
	(1.049)***	(1.032)***	(1.003)***
Number of Obs.	291,759	291,759	291,759
First Stage F-Stat	17.76	16.53	19.72
ln(Electric Intensity)			
Shortage	-1.392	0.0926	-1.217
	(0.718)*	(0.755)	(0.673)*
Number of Obs.	479,616	479,616	479,616
First Stage F-Stat	14.24	14.98	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 A16: **Alternative Instruments: Materials, Labor, Revenue, and TFPR**

	(1)	(2)	(3)
	Without New Supply		With New Supply
	Actual	Predicted with	Actual
	Hydro	Run-of-River	Hydro
Instrument:	Generation	and Reservoirs	Generation
ln(Materials)			
Shortage	-1.370	-1.137	-1.415
	(0.607)**	(0.511)**	(0.576)**
Number of Obs.	495,043	495,043	495,043
First Stage F-Stat	13.88	14.23	15.89
ln(Workers)			
Shortage	-0.232	-0.243	-0.280
	(0.356)	(0.339)	(0.341)
Number of Obs.	502,724	502,724	502,724
First Stage F-Stat	13.82	14.19	15.82
ln(Earnings/Worker)			
Shortage	-0.542	-0.267	-0.487
	(0.270)**	(0.218)	(0.247)**
Number of Obs.	456,443	456,443	456,443
First Stage F-Stat	13.25	14.63	15.07
ln(Revenue)			
Shortage	-1.019	-1.091	-1.097
	(0.586)*	(0.536)**	(0.560)*
Number of Obs.	501,130	501,130	501,130
First Stage F-Stat	13.84	14.17	15.83
ln(TFPR)			
Shortage	0.158	-0.304	0.128
	(0.274)	(0.259)	(0.257)
Number of Obs.	479,313	479,313	479,313
First Stage F-Stat	14.21	14.90	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.

E.5   Alternative TFPR Measures

Table A17 presents estimates of Equation (21), using alternative measures of TFPR described in Appendix C.2.2.

Table A17: **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	-0.283	-0.142	-0.097	-0.107	-0.115	-0.199
	(0.219)	(0.225)	(0.192)	(0.190)	(0.201)	(0.242)
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	17.71	17.72	17.69	17.18	17.7	17.73

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.

E.6 Heterogeneous Effects of Shortages

Table A18 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 A18 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 A18: **Heterogeneous Effects of Shortages**

	(1)	(2)	(3)	(4)
	Self-Gen	ln(Fuel		
Dependent Variable:	Share	Rev Share)	ln(Revenue)	ln(TFPR)
Shortage	-0.026	-0.937	-0.479	-0.119
	(0.054)	(1.556)	(0.602)	(0.380)
Shortage x Elec Intensive	0.021	0.206	-0.882	-0.010
	(0.108)	(1.797)	(0.996)	(0.445)
Shortage x Self-Generator	0.493	4.173	-0.535	-0.400
	(0.145)***	(1.798)**	(0.660)	(0.353)
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.

F   Simulation Appendix

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

F.1   Simulation Inputs

Table A19 presents the sources of the parameters used in the simulations.

Table A19: **Simulation Inputs**

Parameter	Source	Level
$\alpha_K, \alpha_L, \alpha_M, \alpha_E$	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.

F.2   Exogenous Generators: Cobb-Douglas

This section presents full details on how the simulations in Section 7 determine optimal input and output bundles conditional on exogenous generator ownership. Here we present the Cobb-Douglas model in Section 4; 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. <u>Plants without generators</u>

      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

      $$\alpha_M \Omega K^{\alpha_K}[M^G]^{\alpha_M-1}L^{\alpha_L}[E^G]^{\alpha_E} = p^M. \tag{25}$$

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

$$\alpha_E \Omega K^{\alpha_K}[M^G]^{\alpha_M} L^{\alpha_L}[E^G]^{\alpha_E - 1} = p^{E,G}. \tag{26}$$

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

$$\alpha_L(1-\delta)\Omega K^{\alpha_K}[M^G]^{\alpha_M} L^{\alpha_L - 1}[E^G]^{\alpha_E} = p^L. \tag{27}$$

Rearranging these three equations, we obtain:

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

$$L^* = (1-\delta)\frac{p^M \alpha_L}{p^L \alpha_M} M^{*G} \tag{28}$$

$$E^{*G} = \frac{p^M \alpha_E}{p^E \alpha_M} M^{*G}$$

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, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}^S} = 0$, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}^G} = 0$, $\frac{\partial \pi_{it\tau}}{\partial E_{it\tau}^S} = 0$, $\frac{\partial \pi_{it\tau}}{\partial E_{it\tau}^G} = 0$, and $\frac{\partial \pi_{it}}{\partial L_{it}} = 0$. These yield

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

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

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

$$\alpha_M \Omega K^{\alpha_K}[M^G]^{\alpha_M} L^{\alpha_L}[E^G]^{\alpha_E - 1} = p^{E,G} \tag{29}$$

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

$$+\alpha_L \delta p \Omega K^{\alpha_K}[M^S]^{\alpha_M} L^{\alpha_L - 1}[E^S]^{\alpha_E} = p^L$$

The set of equations in system (29) 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 (28), the no-shortage values.

Annual revenue is:

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

F.3   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:

$$Q = \min\{AK^{\alpha_K}L^{\alpha_L}M^{\alpha_M}, \frac{1}{\lambda}E\} \tag{30}$$

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 (30). 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 $\Pi_{it\tau}$:

$$\begin{aligned}
\Pi_{it\tau} =& p \min\{A_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}M_{it\tau}^{\alpha_M}, \frac{1}{\lambda}E_{it\tau}\} \\
&- p^L L_{it} - p^M M_{it\tau} - p^E E_{it\tau},
\end{aligned} \tag{31}$$

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:

$$A_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}M_{it\tau}^{\alpha_M} = \frac{1}{\lambda}E_{it\tau}, \tag{32}$$

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

$$\Pi_{it\tau} = (1 - \frac{\lambda p^E}{p})\Omega_{it}A_{it}K_{it}^{\alpha_K}L_{it}^{\alpha_L}M_{it\tau}^{\alpha_M} - p^L L_{it} - p^M M_{it\tau} \tag{33}$$

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.[39]

The first-order condition with respect to materials yields:

$$\alpha_M(1 - \gamma)\frac{R_{it\tau}}{M_{it\tau}} - p^M = 0. \tag{34}$$

The marginal revenue product of materials is:

$$MRPM = \begin{cases} \alpha_M(1 - \gamma)\frac{R_{it\tau}}{M_{it\tau}} & \text{if grid power} \\ \mathcal{T}\alpha_M(1 - \gamma)\frac{R_{it\tau}}{M_{it\tau}} & \text{if power outage} \end{cases} \tag{35}$$

When setting labor, the firm begins with its yearly profit function, which is simply the weighted average of Equation (33) 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:

$$MRPL = \alpha_L(1 - \gamma)\left[(1 - \delta)\frac{R_{it}^G}{L_{it}} + \delta\mathcal{T}\frac{R_{it}^S}{L_{it}}\right] = p^L, \tag{36}$$

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

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. <u>If in Case 3:</u>

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

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

   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

---

[39]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.

FOC for materials during grid power periods, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}} = 0$, yields

$$\alpha_M (1 - \gamma^G) \Omega K^{\alpha_K} [M^G]^{\alpha_M - 1} L^{\alpha_L} = p^M. \tag{37}$$

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

$$\alpha_L (1 - \gamma^G)(1 - \delta) \Omega K^{\alpha_K} [M^G]^{\alpha_M} L^{\alpha_L - 1} = p^L. \tag{38}$$

Rearranging these first-order equation, we obtain

$$M^{*G} = \left[(1 - \gamma) \Omega K^{\alpha_K} \left(\frac{\alpha_M}{p_M} \right)^{1 - \alpha_L} \left(\frac{(1 - \delta)\alpha_L}{p_L} \right)^{\alpha_L} \right]^{\frac{1}{1 - \alpha_L - \alpha_M}}.$$
$$L^* = (1 - \delta) \frac{p^M \alpha_L}{p^L \alpha_M} M^{*G} \tag{39}$$

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, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}^S} = 0$, $\frac{\partial \pi_{it\tau}}{\partial M_{it\tau}^G} = 0$, and $\frac{\partial \pi_{it}}{\partial L_{it}} = 0$.

$$\alpha_M (1 - \gamma^G) p \Omega K^{\alpha_K} [M^G]^{\alpha_M - 1} L^{\alpha_L} = p^M$$
$$\alpha_M (1 - \gamma^S) p \Omega K^{\alpha_K} [M^S]^{\alpha_M - 1} L^{\alpha_L} = p^M$$
$$\alpha_L (1 - \gamma^G)(1 - \delta) p \Omega K^{\alpha_K} [M^G]^{\alpha_M} L^{\alpha_L - 1}$$
$$+ \alpha_L \delta (1 - \gamma^S) p \Omega K^{\alpha_K} [M^S]^{\alpha_M} L^{\alpha_L - 1} = p^L \tag{40}$$

The system of equations in (40) is solved numerically in MATLAB using the fsolve routine.

Finally, electricity usage is

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

Annual revenue is:

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

F.4 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.

F.4.1 First-Order Conditions for Non-Generators

For firms that do not have generators, revenue is:

$$
\begin{aligned}
R_{it} &= \left[\int_\tau (R_{it\tau})^\sigma d\tau\right]^{\frac{1}{\sigma}} \\
&= [(1-\delta)M_{it\tau}^{\sigma\alpha_M} E_{it\tau}^{\sigma\alpha_E} L_{it}^{\sigma\alpha_L} K_{it}^{\sigma\alpha_K}]^{\frac{1}{\sigma}} \\
&= (1-\delta)^{\frac{1}{\sigma}} M_{it\tau}^{\alpha_M} L_{it}^{\alpha_L} E_{it\tau}^{\alpha_E} K_{it}^{\alpha_K} \\
&= (1-\delta)^{\frac{1}{\sigma}} \left(\frac{M_{it}}{(1-\delta)}\right)^{\alpha_M} \left(\frac{E_{it}}{(1-\delta)}\right)^{\alpha_E} L_{it}^{\alpha_L} K_{it}^{\alpha_K} \\
&= (1-\delta)^{\frac{1}{\sigma}-\alpha_M-\alpha_E} M_{it}^{\alpha_M} E_{it}^{\alpha_E} L_{it}^{\alpha_L} K_{it}^{\alpha_K}
\end{aligned}
\tag{41}
$$

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 $(1-\delta)^{1-\alpha_M-\alpha_E}$.

F.4.2   First-Order Conditions for Generators

The profit function is:

$$\Pi_{it} = R_{it} - p^M \int_\tau M_{it\tau}d\tau - p^E \int_\tau E_{it\tau}d\tau - p^L \int_\tau L_{it\tau}d\tau \tag{42}$$

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

$$\frac{1}{\sigma}R_{it}^{\frac{\frac{1}{\sigma}-1}{\frac{1}{\sigma}}}\sigma\alpha_M\frac{R_{it\tau}^\sigma}{M_{it\tau}} = p^M. \tag{43}$$

The electricity FOC is similar:

$$\frac{1}{\sigma}R_{it}^{\frac{\frac{1}{\sigma}-1}{\frac{1}{\sigma}}}\alpha_E\frac{R_{it\tau}^\sigma}{E_{it\tau}} = p^E. \tag{44}$$

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

$$\frac{1}{\sigma}R_{it}^{\frac{\frac{1}{\sigma}-1}{\frac{1}{\sigma}}}\sigma\alpha_L\left[(1-\delta)\frac{\left(R_{it}^G\right)^\sigma}{L_{it\tau}} + \delta\frac{\left(R_{it}^S\right)^\sigma}{L_{it\tau}^S}\right] = p^L, \tag{45}$$

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 (43,44,45) 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.

F.5   Comparing Predictions from Cobb-Douglas, Leontief, and CES Models

Table A20 presents the simulated effects of the 2005 assessed Shortage levels S_{s2005} relative to zero shortage. Indeed, Table A20 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.[40] 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.

---

[40]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.

Table A20: **Predictions from Different Production Function Models**

	Cobb-Douglas	Leontief	CES $\sigma = 0.9$	CES $\sigma = 0.5$
	(1)	(2)	(3)	(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.

F.6   Endogenous Generators

For the model with endogenous generators, the equations given in Online Appendix F.2 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

$$\Pi_{it} = R_{it} - p^L L^* - p^M \left(\delta M^{*G} + (1-\delta)M^{*S} \right) - \delta p^{E,S} E^{*S} - (1-\delta)p^{E,G} E^{*G}, \qquad (46)$$

where optimal inputs are according to the equations in Online Appendix F.2 above.

F.6.1   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 A21 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.[41] This gives a 1.6:1 ratio between generator costs and rental

---

[41] This 30 percent discount rate is high by U.S. standards, but as Banerjee and Duflo (2014) discuss, Indian firms

Table A21: **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

Note: 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.

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 A21. 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.

---

pay far higher interest rates - on the order of 30 to 60 percent, if they have access to capital at all.

F.7   Additional Simulation Figures and Table

Table A22: **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.

Figure A8: **Predicted Average Revenue Loss by Simulation Year**
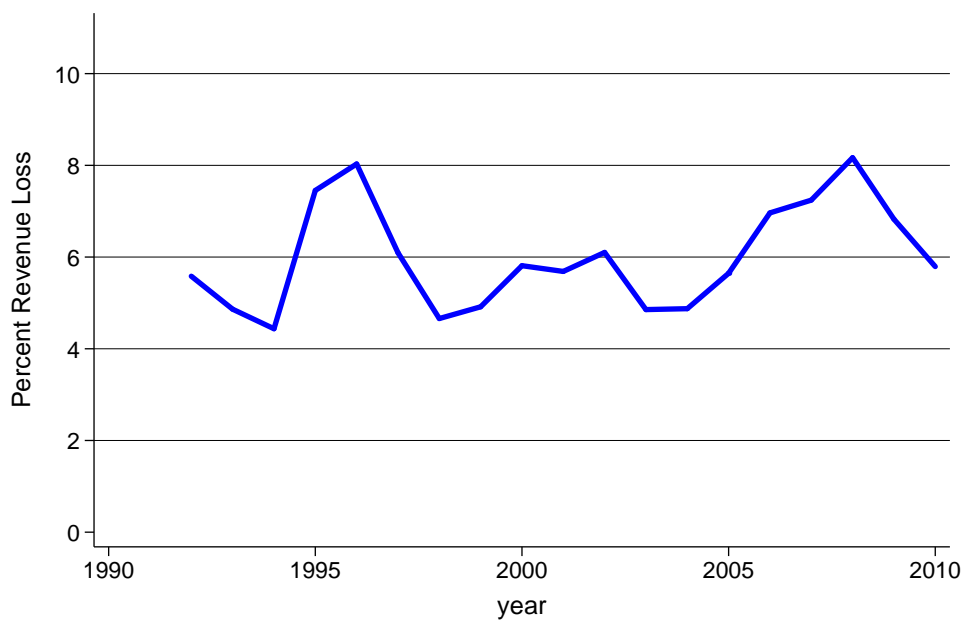
Notes: In the body of the paper, we simulate effects of moving from $\delta = 0$ to $\delta = S_{s2005}$ for all plants in the data in 2005. This figure presents revenue effects from the same simulations for each year of the 1992-2010 sample, i.e. taking the sample of plants in year t and changing δ from $\delta = 0$ to $\delta = S_{st}$.

Figure A9: **Generator Adoption Under Varying Shortage Levels**
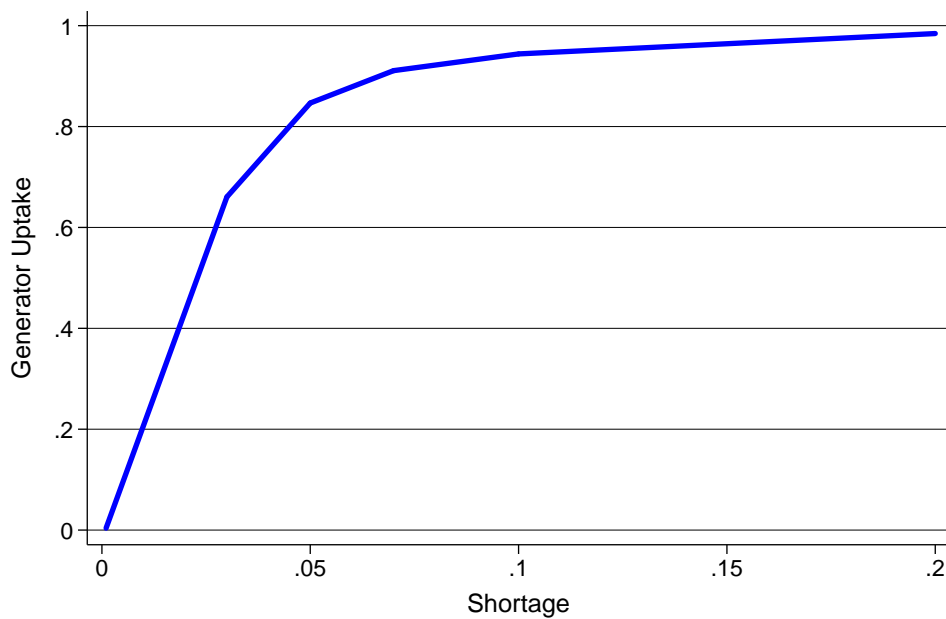
Note: These figures show the simulated generator adoption rate when the δ on the x-axis is assigned to all plants in the 2005 ASI, using the generator adoption model in Section 7.1.1. Note that generator takeup exceeds 90 percent at a seven percent δ, which may seem puzzling given that the generator 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.

References

[1] 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.

[2] 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.

[3] Nagaraj, R. (2002). "How to Improve India's Industrial Statistics," *Economic and Political Weekly*, Vol. 37, No. 10 (Mar. 9-15), pages 966-970.