American Economic Review 2016, 106(3): 587–624
http://dx.doi.org/10.1257/aer.20140389

How Do Electricity Shortages Affect Industry? Evidence from India†

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

We estimate the effects of electricity shortages on Indian manufacturers, instrumenting with supply shifts from hydroelectric power availability. We estimate that India's average reported level of shortages reduces the average plant's revenues and producer surplus by 5 to 10 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. (*JEL* D24, L60, L94, O13, O14, Q41)

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

*Allcott: Department of Economics, New York University, 19 W. 4th Street, New York, NY 10012, NBER, and Poverty Action Lab (e-mail: hunt.allcott@nyu.edu); Collard-Wexler: Duke University, 230 Social Sciences Building, Durham, NC 27708, and NBER (e-mail: collardwexler@gmail.com); O'Connell: Department of Economics, City University of New York, Graduate Center, 365 5th Avenue, Room 5313, New York, NY 10016 (e-mail: soconnell@gradcenter.cuny.edu). We thank Maureen Cropper, Jan De Loecker, Michael Greenstone, Peter Klenow, Kabir Malik, Rohini Pande, Nick Ryan, Jagadeesh Sivadasan, Anant Sudarshan, and seminar participants at Brown, Drexel, Duke, the Federal Trade Commission, Harvard, the 2013 NBER Summer Institute, the 2014 NBER Winter IO/EEE meetings, Society for Economic Dynamics, Stanford, KU Leuven, Toulouse, Universidad de los Andes, University of Chicago, University of Cologne, and the World Bank for helpful comments. We are particularly grateful to Nick Bloom, Troy Smith, and Shaleen Chavda for insight into the textile industry. We thank Deepak Choudhary, Anuradha Bhatta, Sherry Wu, and Mark Thomas for helpful research assistance and the Stern Center for Global Economy and Business for financial support. We have benefited from helpful conversations with Jayant Deo of India Energy Exchange, Gajendra Haldea of the Planning Commission, Partha Mukhopadhyay of the Centre for Policy Research, and Kirit Parikh of IGIDR. We also thank A. S. Bakshi and Hemant Jain of the Central Electricity Authority for help in collecting archival data. Of course, all analyses are the responsibility of the authors, and no other parties are accountable for our conclusions. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.
† Go to http://dx.doi.org/10.1257/aer.20140389 to visit the article page for additional materials and author disclosure statement(s).

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. We have made these data publicly available as the India Energy Data Repository (Allcott, Collard-Wexler, and O'Connell 2015). 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 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 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 a positive but economically small effect on variable energy input costs for plants with generators: a 1 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 1.1 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 hydro availability, 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 two to three 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 because political barriers have prevented reforms to India's significantly distorted retail electricity prices.

The remainder of this section discusses related literature. Section I details the data. Section II provides background on the Indian electricity sector, the causes of electricity shortages, and manufacturers' responses to shortages. Section III presents the production function model and TFPR estimates. Sections IV and V present the empirical strategy and results. Section VI details the counterfactual simulations, and Section VII concludes.

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 US states, including Aschauer (1989); Holtz-Eakin (1994); Fernald (1999); Garcia-Mila, McGuire, and Porter (1996); and see Gramlich (1994) for a review. Easterly and Rebelo (1993); Esfahani and Ramirez (2003); 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 (forthcoming), 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 et al. (2010) on the benefits of reliable water provision in Mexico; and Baum-Snow (2007, 2013), Baum-Snow et al. (2015), 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 (2014); Dinkelman (2011); Lipscomb, Mobarak, and Barham (2013); and Rud (2012a) study the effects of electricity grid expansions, while Alby, Dethier, and Straub (2013); 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 et al. (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.[1]

[1] 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.

I. Data

We have collected comprehensive data from 1992 to 2010 on weather, the power sector, and manufacturing production.[2] All financial amounts are deflated to real 2004 rupees (Rs).[3] Throughout the paper, we use the word "state" to refer to states, Union Territories, and the National Capital Region (New Delhi).

A. *Weather Data*

Rainfall data are from the University of Delaware, which provides monthly rainfall for geographic grid points spaced at 1/2 degree intervals (Matsuura and Willmott 2012). We sum to total annual rainfall by grid point, then calculate state-by-year average rainfall by averaging across all grid points within each state. Temperature data are from the National Climate Centre (NCC), which provides daily temperatures for geographic grid points 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 grid points within each state.[4] Panel A of Table 1 summarizes the state-by-year observations of weather data, as well as the power sector data described below.

B. *Power Sector Data*

Power sector data are from India's Central Electricity Authority (CEA). The CEA collects many of the same data that the US 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. The cleaned and digitized data are now available as the India Energy Data Repository (www.indiaenergydata.info).

The primary measure of electricity shortages is the percent energy deficit reported in the Load Generation Balance Reports (CEA 1993–2011a). 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

[2] 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.")

[3] The exchange rate was approximately Rs 50 per US dollar at that time.

[4] "Rainfall" is more precisely "precipitation," as it includes winter snowfall in the Himalayan states. The 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 of Delaware only provides monthly average temperatures.

Table 1—Summary Statistics

	Mean	SD	Min.	Max.	Obs.
Panel A. Weather and power sector data (state-by-year)					
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	3,618	536
Total capacity (MW)	2,744	3,099	0	16,062	536
Reservoir inflows (billion cubic meters)	8.78	16.4	0	116	536
Run-of-river generation (TW)	0.33	0.95	0	8.89	536
Capacity added in previous year (MW)	117	250	−472	2,070	536
Panel B. ASI Data (plant-by-year)					
Plant number of observations	2.20	2.13	1	19	224,684
Revenues (million Rs)	139	2,156	0	788,868	613,930
Capital stock (million Rs)	51	1,044	0	297,370	612,424
Number of employees	79	431	0	52,148	576,901
Labor cost (million Rs)	6.39	70.5	0	16,074	602,124
Materials purchased (million Rs)	90	1,562	0	636,095	607,522
Fuels purchased (million Rs)	5.07	102	0	39,360	596,036
Electricity purchased (million Rs)	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: See text for variable sources and definitions. TWh stands for terawatt-hours of electricity, and MW stands for megawatts of capacity. 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 Rs. Means and standard deviations are weighted by ASI sample weights. Observation counts differ due to nonresponse and due to variable-specific cleaning procedures described in online Appendix C.

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

Table 2—Correlations with the Shortage Variable

	1(Largest barrier) (1)	Power quality (2)	Self-gen share (3)	Capacity factor (4)
Shortage	1.107	−8.378	0.699	0.167
	(0.492)**	(3.383)**	(0.213)***	(0.062)**
Observations	2,280	2,265	1,124	1,286
Dependent variable mean	0.31	6.38	5.9	0.64
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). Dependent variables in columns 1–3 are, respectively, an indicator for whether the firm's manager reports that electricity is the primary barrier to growth, the manager's rating of grid power quality (from 1 (extremely bad) to 10 (excellent)), and Self-Generation Share (omitting plants that do not own generators). Sample for column 4 is panel data on all Indian coal power plants from 1994–2009. Columns 1–3 condition on industry indicators, and column 4 conditions on year indicators and plant fixed effects. Robust standard errors, clustered by state.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

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. Columns 1–3 of Table 2 show that in the World Bank Enterprise Survey (WBES), 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. Column 4 shows that 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.[5] Furthermore, Alam (2013) shows that Peak Shortage is correlated with her measure of blackouts, which is based on nighttime lights measured by satellites, and 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 1994–2012), 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 A1.

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 (CEA 1993–2011b). 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

[5] 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.

and divide predicted generation by annual generation capacity.[6] 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.[7] Online Appendix B provides more information on the power sector data.

C. *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. We use the ASI sample weights to produce estimates valid for the population of registered factories in India.[8] 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 National Industrial Classification (NIC) codes, which are closely related to Standard Industrial Classification (SIC) codes. Online Appendix C gives more detail on the ASI data preparation and cleaning.

[6] 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 8,760 hours per year.

[7] 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.

[8] While the weighted ASI sample is representative of factories registered under the Factories Act, not all manufacturing establishments are registered. Small factories with fewer than ten workers (or those with fewer than 20 workers without electricity) are not required to register under the Factories Act and are thus excluded from the sampling frame. Nagaraj (2002) points out that there may also be some registration avoidance, especially for smaller plants near the registration threshold. While unregistered plants comprise around 80 percent of employment in the manufacturing sector (Hsieh and Klenow 2014), they contribute a smaller share (around one-third) to total manufacturing output (Sincavage, Haub, and Sharma 2010). If these smaller unregistered plants are less likely to have generators or high-quality grid power, the effects of shortages on the full manufacturing sector could be larger than for our ASI sample.

Panel B of Table 1 presents sample-weighted summary statistics for the ASI. There are 615,721 plant-by-year observations at 224,684 unique plants, although observation counts differ across variables due to nonresponse 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 US 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.

II. Background

A. *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 US total.[10] Of this, 53 percent was coal, 10 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. These low agricultural prices are cross-subsidized by industrial

[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 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.

[10] Capacity and ownership statistics in this paragraph are from the *General Review* (CEA 1994–2012).

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 4 gigawatts of this was actually built. Of the 71 gigawatts of capacity targeted to be built between 1997 and 2007, only half was actually 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).

B. *Variation in Shortages*

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 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.

[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).

[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.

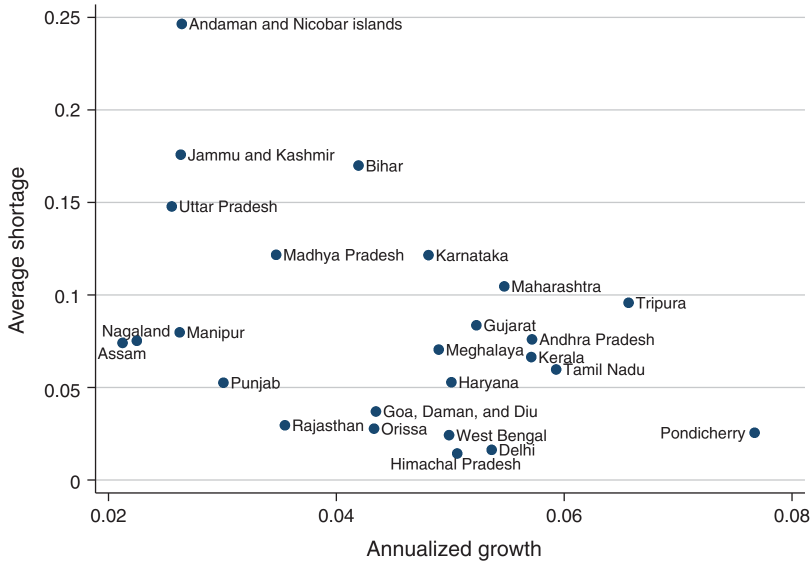

Figure 1. Average Shortages and Per Capita GDP by State, 1992–2010

Notes: This figure plots each state's average shortage over the 1992–2010 study period versus its per capita GDP growth rate over that period. Shortage data are estimated by the Central Electricity Authority.

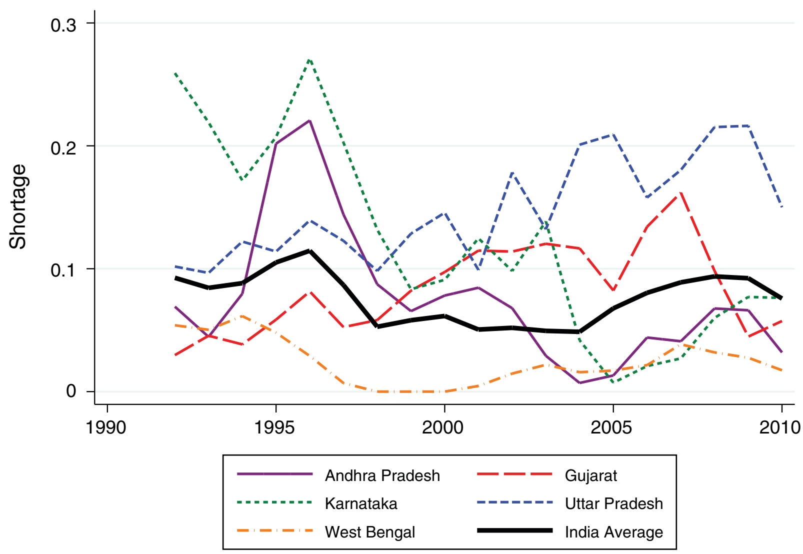

Figure 2. Shortages Over Time in Five Large States

Note: Shortage data are estimated by the Central Electricity Authority.

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.

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

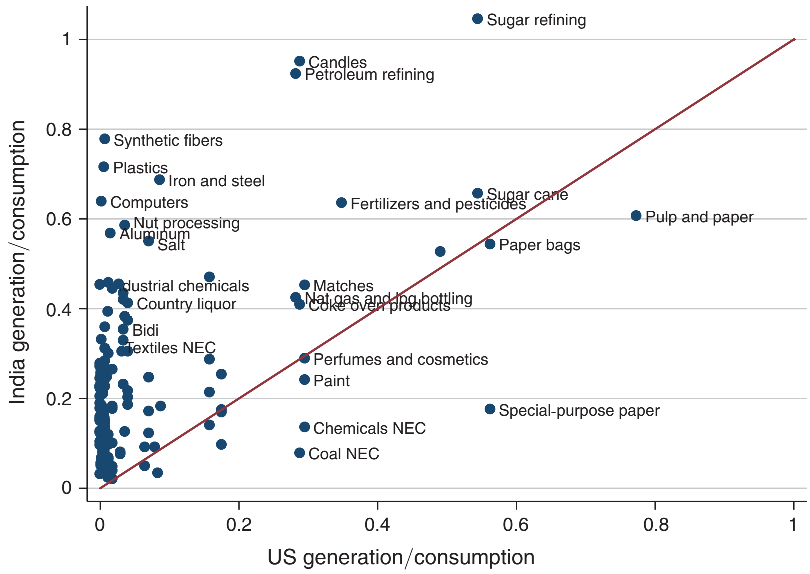

Figure 3. Manufacturing Electricity Generation in India versus the United States

Notes: This figure presents the ratio of electricity generation to consumption by three-digit industry. Indian and US data are from the Annual Survey of Industries and the Manufacturing Energy Consumption Survey, respectively.

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, 5 and 7 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.

C. *Industrial Electricity Use in India*

A natural response to shortages is to self-generate electricity. We calculate that total electricity generation by ASI manufacturing plants equals 35 percent of their total consumption, more than twice the 15.8 percent for US manufacturers reported in the Manufacturing Energy Consumption Survey (MECS) (US 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]

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

[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

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

III. Model and Production Function Estimation

A. *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:

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

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.

[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: $Q_{it\tau} = \min\left\{A K_{it\tau}^{\alpha_K} L_{it\tau}^{\alpha_L} M_{it\tau}^{\alpha_M}, \frac{1}{\lambda} E_{it\tau}\right\}$, 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 $\beta_X = \alpha_X\left(1 + \frac{1}{\epsilon}\right)$ 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.

B. *Power Outages, Inputs, and Timing*

On each day, there are two states of the world: a power outage state and a no-outage state. The outage occurs 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 electricity use, and the effect of power outages can be much larger than the electricity coefficient α_E.

There are three types of inputs:

(i) *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.

(ii) *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]

(iii) *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.

C. *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

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

[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.

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

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

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.

D. *Effects of Shortages on Revenues and Measured Productivity*

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

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

where

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

Using lowercase 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.

E. *Production Function Estimation*

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 rearranged 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 rearrange the first-order condition and take a weighted average across outage and non-outage periods

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

This gives

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

[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).

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.

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 2015). 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 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]

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 twenty-fifth and seventy-fifth 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.

[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 WBES.

[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

	Mean (1)	25th percentile (2)	75th percentile (3)
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.

IV. Estimation Strategy

A. *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 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]

[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 versus 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.

B. *Instrumenting for Shortages*

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.

A valid instrument for shortages must shift electricity supply but affect manufacturers only through shortages. Our instrument exploits supply shifts from 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 instrument Z_{st} is predicted hydro generation H_{st} as a share of predicted electricity demand $\tilde{Q}_{st}$,

$$(18) \qquad Z_{st} = \frac{H_{st}}{\tilde{Q}_{st}}.$$

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

$$(19) \qquad \tilde{Q}_{st} = \sum_{r \neq s} Q_{rt} \cdot \sum_{y=1992}^{2010} \frac{Q_{sy}}{\sum_{r \neq s} Q_{ry}}.$$

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 I 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 $\tilde{C}_{st}$, 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

$$(20) \qquad H_{st} = \tilde{C}_{st} \cdot \frac{\sum_{d \in \mathcal{D}_s} c_d a_{ds} f_{dt}}{\sum_{d \in \mathcal{D}_s} c_d a_{ds}}.$$

We predict hydro capacity $\tilde{C}_{st}$ 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.

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 $\overline{R}_s$ and include indicator variables for five bins above and five bins below $\overline{R}_s$ 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 et al. 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]

C. *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]

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

Observations are weighted by ASI sample weights. Our primary specifications use the fixed effects estimator, although we also present estimates using the

[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 A1. Online Appendix Figure A2 plots each state's sample average rainfall against the average ratio of hydro generation to total consumption; the two quantities are uncorrelated.

[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.

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]

D. *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 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 Z_{st}.

[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 A1 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 A3 and A4 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.

[27] Online Appendix Figure A2 and Table A3 present more detail.

Table 4—Evaluating the Hydro Instrument

	Shortage (1)	ln(Energy available) (2)	ln(Assessed demand) (3)	ln(Agri output) (4)	ln(Median price) (5)
Hydro	−0.163	0.206	0.029	−0.274	−0.149
	(0.059)***	(0.074)***	(0.092)	(0.246)	(0.140)
Observations	536	536	536	518	540

Notes: Observations weighted by number of ASI establishments in the state-year cell. Robust standard errors.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

E. *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 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 1 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 when hydro production shifts supply, distribution companies clear the market by adjusting outages, not by adjusting prices.

V. Empirical Results

A. *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

Table 5—First Stages for Base IV Estimates

Second stage dependent variable:	Self-gen share (1)	ln(Fuel rev share) (2)	ln(Electric intensity) (3)
Panel A. Energy inputs			
Hydro	−0.168	−0.171	−0.156
	(0.0407)***	(0.0422)***	(0.0403)***
Observations	240,743	291,759	479,616
Clusters	47,575	55,939	111,819
Clusters (2)	535	535	536
First stage F-statistic	17.00	16.53	14.98

Second stage dependent variable:	ln(Materials) (1)	ln(Workers) (2)	ln(Earnings/worker) (3)	ln(Revenue) (4)	ln(TFPR) (5)
Panel B. Materials, labor, revenue, and TFPR					
Hydro	−0.152	−0.152	−0.161	−0.152	−0.155
	(0.0403)***	(0.0403)***	(0.0422)***	(0.0404)***	(0.0400)***
Observations	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
First stage F-statistic	14.23	14.19	14.63	14.17	14.90

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.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

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 14 to 17. 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.

B. *Effects of Shortages*

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 1 percentage point increase in Shortage (for example, an

[28] Online Appendix Table A11 plus additional regressions confirm that shortages have no effect on fuel revenue shares for non-generators.

Table 6—Effects of Shortages on Energy Inputs

	Self-Gen share (1)	ln(Fuel rev share) (2)	ln(Electric intensity) (4)
Panel A. OLS			
Shortage	0.282	0.917	−0.539
	(0.0337)***	(0.176)***	(0.122)***
Panel B. IV			
Shortage	0.442	3.294	0.0926
	(0.153)***	(1.032)***	(0.755)
Observations	240,743	291,759	479,616
Clusters	47,575	55,939	111,819
Clusters (2)	535	535	536
First stage *F*-statistic	17.00	16.53	14.98

Notes: This table presents estimates of equation (21). Panel B instruments for Shortage using hydro availability. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. *F*-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

increase from 0.1 to 0.11) causes manufacturers to increase the share of electricity self-generated by 0.442 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 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 1 percentage point increase in Shortage increases fuel revenue share by 3.294 percent. Using the fact that the mean fuel revenue share is 0.055, the point estimate suggests that a 1 percentage point increase in shortages increases fuel costs by $(3.294 \times 1\%) \times 0.055 \approx 0.18$ 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 1 percentage point increase in shortages by reducing electric intensity by more than about 1.5 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 and 3, however, are ratios to revenues, which increase with economic growth. The coefficients are thus strongly downward biased.

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

	ln(Materials) (1)	ln(Workers) (2)	ln(Earnings/worker) (3)	ln(Revenue) (4)	ln(TFPR) (5)
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	−1.137	−0.243	−0.267	−1.091	−0.304
	(0.511)**	(0.339)	(0.218)	(0.536)**	(0.259)
Observations	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
First stage *F*-statistic	14.23	14.19	14.63	14.17	14.90

Notes: This table presents estimates of equation (21). Panel B instruments for Shortage using hydro availability. *F*-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. Robust standard errors, with two-way clustering by plant and state-year.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Nonenergy Inputs, *Revenue*, *and TFPR*.—Table 7 presents effects of shortages on nonenergy inputs, revenue, and revenue productivity. Columns 1–3 measure impacts on materials and labor. Column 1 shows that a 1 percentage point increase in Shortage causes a 1.137 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 1 percentage point increase in shortages causes a 1.091 percent decrease in revenues. In the special case distribution companies impose the marginal shortage proportionally to manufacturers versus 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 1 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 0.085 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 1 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 VI, we show that our point estimate of -0.304 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 0.077 in a one-sided test and 0.15 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.

C. *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 similarity in results suggests that the Shortage variable does not suffer from some nonclassical 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 100 mm or 50 mm rainfall bins instead of 60 mm bins, or when using rainfall data from the National Climate Centre instead of the 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. Finally, results are qualitatively similar under six different approaches to calculating production functions and TFPR.

D. *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. $\overline{S}_{st}$ 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 smooths over idiosyncratic year-to-year variation that should not impact potential entrants' beliefs about future shortages. We estimate run the following regression with plant-level data:

$$Y_{ijst} = \rho \overline{S}_{st} + \Lambda_{st} + \zeta_s + \theta_t + \mu_{jt} + \varepsilon_{st}. \tag{22}$$

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 Rs of revenue; the mean is 1.35 kWh/100 Rs. 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 III 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 1 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 $\overline{S}_{st}$ are positively associated with higher generator ownership rates for entrants. This association is small, however: a 1 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

Table 8—Association between Entering Plant Characteristics and Shortages

Dependent variable	Ind. elec/ rev (1)	1(Self gen) (2)	ln(Capital/ rev) (3)	ln(Labor/ rev) (4)	ln(Matls/ rev) (5)
Shortage at entry	−0.463	0.246	−0.701	−0.078	−0.025
	(0.179)**	(0.074)***	(0.422)	(0.135)	(0.073)
Observations	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.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

the time of entry are associated with entrants' factor shares.[29] Theory gives no unambiguous prediction about whether shortages 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 1 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 $\overline{S}_{st}$.

VI. Simulations

A. *Overview*

In this section we apply the model in Section III 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 V. 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

[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 IIIE.

[30] We chose 2005 because it allows comparison to the 2005 WBES 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 A1 presents predicted revenue losses when simulating with each year between 1992 and 2010, showing that 2005 is not unusual.

ratio of grid-purchased to self-generated electricity, we use the WBES medians (in Rs per kWh): $p^{E,S}/p^{E,G} = 7/4.5 \approx 1.56$. Given this information and an assumed δ, we simulate production decisions using the first-order conditions in Section III. See online Appendix F for more details on the simulation procedure.

For the next several subsections until Section VIE, 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.

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 versus 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 VIA for details.

B. *Effect of Shortages*

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,

[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.

TABLE 9—EFFECT OF SHORTAGES

	Simulation percent (1)	IV estimate percent (2)	p-value for columns (1) versus (2) (3)	World Bank Survey percent (4)
Panel A. Effects of shortages: Model and IV estimates				
Self-generation share increase	3.1	3.1	(0.98)	
Materials reduction	5.6	8.1	(0.49)	
Labor reduction	5.6	1.7	(0.11)	
Revenue loss	5.6	7.7	(0.57)	7.8
TFPR loss	1.5	2.2	(0.72)	
		All percent	With generators percent	Non-generators percent
Panel B. Producer surplus effects of shortages				
Producer surplus loss		9.5	8.0	10.0
of which Δ generator costs		3.9	7.7	NA
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				
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 models 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.

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.[32] 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 7.7 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 2.2 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, respectively, while the IV estimates show 8.1 and 1.7 percent decreases. The simulated and estimated

[32] Because the empirical estimates are identified off of relatively modest annual variation, column 2 is not a realistic estimate of the long-run effects of eliminating shortages. Online Appendix Table A4 shows very similar agreement between simulated and estimated results when comparing in terms of semi-elasticities, i.e., the effects of a 1 percentage point change in Shortage.

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 0.11. 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 WBES; this number is closely comparable to 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.

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 versus 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 versus 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 $55\% \times 5\% \times 7.2\% \approx 0.2\%$ 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 versus self-generators: 2.6 percent versus 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 versus 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.[33]

C. *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 and allow consumers to sort into their preferred contract, this provides a mechanism to allocate outages to plants that are least affected.[34]

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 8 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.

There are two important caveats to this analysis. First, it holds constant the manufacturing sector's total power allocation. In practice, the grid electricity conserved through interruptible contracts with manufacturers could be allocated to homes, agriculture, or other sectors, where the marginal social value of electricity could be different. Second, even holding constant the manufacturing power allocation, we

[33] Online Appendix E6 presents empirical estimates of how generator ownership moderates effects of shortages; we do not have the power to detect differences between plants with versus 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.

[34] Interruptible contracts have been studied by Baldick, Kolos, and Tompaidis (2006) and are common in the United States. 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.

do not observe which specific plants are subject to shortages, so the losses without interruptible contracts could be larger or smaller than we simulate here.

D. *Comparing Effects for Large versus Small Plants*

How do shortages differentially affect small versus 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 6 percent of profits, the average 100-employee plant loses only about 2 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.

E. *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 detailed above in Section VIA.[35]

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 3 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.

[35] Online Appendix Table A5 presents more detailed statistics used to construct these figures, and online Appendix Figure A2 presents generator adoption rates as a function of δ.

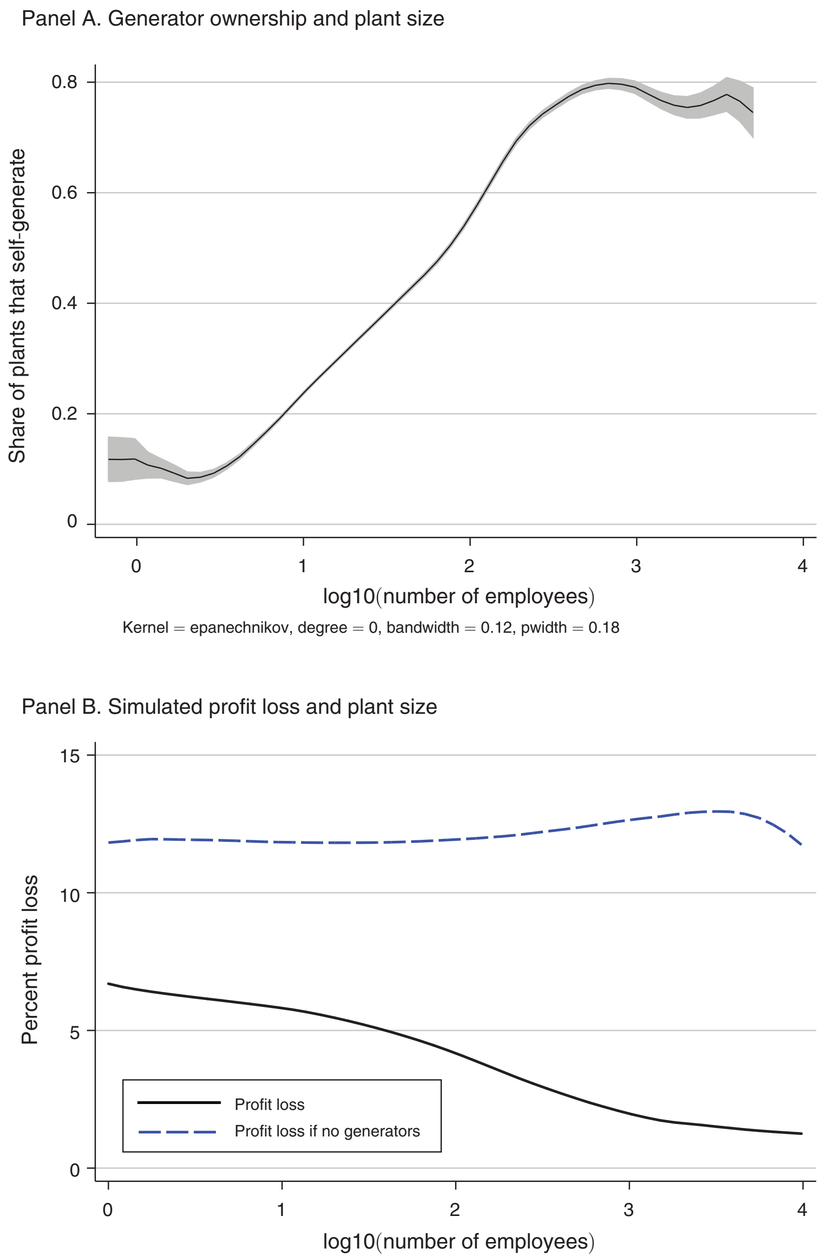

Figure 4. Generator Ownership and Shortage Effects by Plant Size

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.

VII. 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

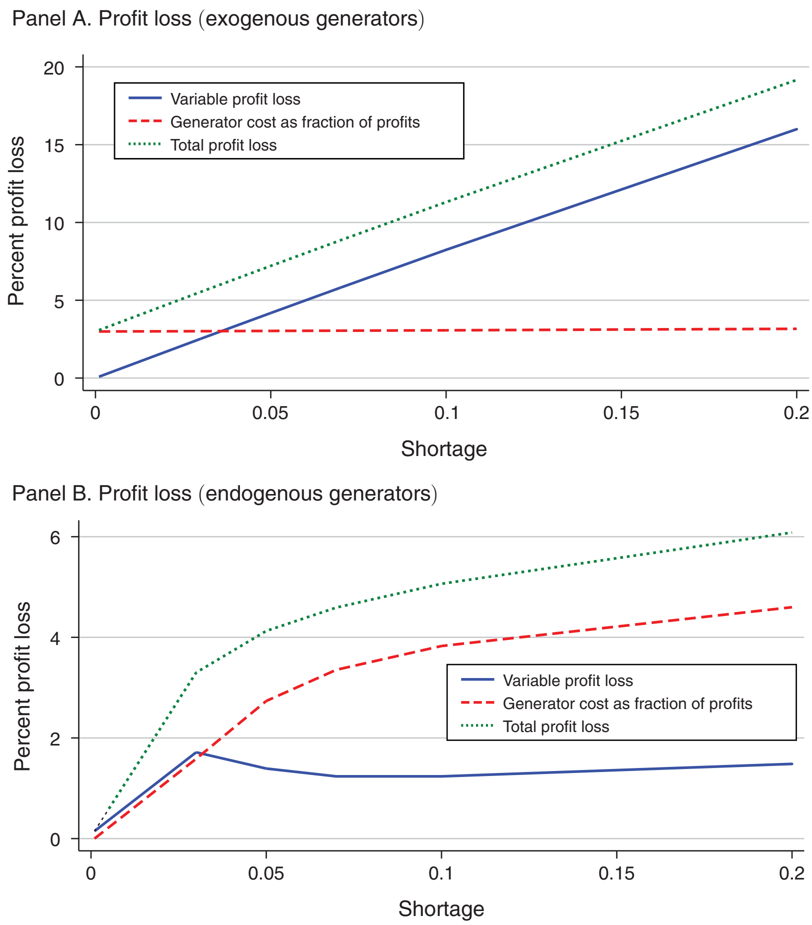

Figure 5. Counterfactuals Under Varying Shortage Levels

Notes: 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 VIA to endogenously adjust generator ownership at the given δ.

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.

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 7.7 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, the short-run effects of electricity shortages do not explain much of the productivity gap between firms in developing versus developed countries discussed by Banerjee and Duflo (2005); Hall and Jones (1999); Hsieh and Klenow (2009); and others, although the long-run effects could be different. 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: interruptible contracts or similar mechanisms can substantially reduce the costs of shortages. Because such contracts can be offered at a discount on an opt-in basis, they may be more politically feasible than raising prices or changing allocations.

REFERENCES

Abeberese, Ama Baafra. 2012. "Electricity Cost and Firm Performance: Evidence from India." https://sites.google.com/a/wellesley.edu/baafra/research (accessed January 15, 2016).

Alam, Muneeza M. 2013. "Coping with Blackouts: Power Outages and Firm Choices." http://economics.ucr.edu/seminars_colloquia/2013-14/econometrics/Alam%20paper%20for%202%203%2014%20seminar.pdf (accessed January 15, 2016).

Alby, Philippe, Jean-Jacques Dethier, and Stéphane Straub. 2013. "Firms Operating under Electricity Constraints in Developing Countries?" *World Bank Economic Review* 27 (1): 109–32.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell. 2015. "The India Energy Data Repository." http://www.indiaenergydata.info.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell. 2016. "How Do Electricity Shortages Affect Industry? Evidence from India: Dataset." *American Economic Review*. http://dx.doi.org/10.1257/aer.20140389.

Aschauer, David Alan. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics* 23 (2): 177–200.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)allocation." *Journal of Political Economy* 122 (5): 1013–63.

Auffhammer, Maximilian, Solomon M. 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* 7 (2): 181–98.

Baisa, Brian, Lucas W. Davis, Stephen W. Salant, and William Wilcox. 2010. "The Welfare Costs of Unreliable Water Service." *Journal of Development Economics* 92 (1): 1–12.

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* 54 (4): 627–42.

Banerjee, Abhijit, and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." In *Handbook of Economic Growth*, Vol. 1A, edited by Philippe Aghion and Steven N. Durlauf, 473–552. Amsterdam: Elsevier.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." http://economics.mit.edu/files/7782 (accessed January 15, 2016).

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103 (1): 305–34.

Baum-Snow, Nathaniel. 2007. "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics* 122 (2): 775–805.

Baum-Snow, Nathaniel. 2013. "Urban Transport Expansions, Employment Decentralization, and the Spatial Scope of Agglomeration Economies." http://www.parisschoolofeconomics.eu/IMG/pdf/baumsnow-pse-rues-oct2013.pdf (accessed January 15, 2016).

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2013. "Roads, Railroads and Decentralization of Chinese Cities." http://www.lse.ac.uk/geographyAndEnvironment/whosWho/profiles/henderson/RoadsRailroadsDecentralization.pdf (accessed January 15, 2016).

Baum-Snow, Nathaniel, and Matthew A. Turner. 2012. "Transportation and the Decentralization of Chinese Cities." http://www.brown.edu/Departments/Economics/Faculty/Matthew_Turner/papers/unpublished/china_baumsnow_turner_lincoln_all.pdf.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–75.

Cameron, A. Colin, Jonah B. Gelbach, Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business and Economic Statistics* 29 (2): 238–49.

Central Electricity Authority (CEA). 1993–2011a. *Load Generation Balance Report*. New Delhi: Central Electricity Authority.

Central Electricity Authority (CEA). 1993–2011b. *Review of Performance of Hydro Power Stations*. New Delhi: Central Electricity Authority.

Central Electricity Authority (CEA). 1994–2012. *All India Electricity Statistics: General Review*. New Delhi: Central Electricity Authority.

Central Electricity Authority (CEA). 2013. *Planwise Capacity Addition*. https://web.archive.org/web/20150427122546/http://cea.nic.in/reports/planning/plan_addition.pdf (accessed January 15, 2016).

Chakravorty, Ujjayant, Martino Pelli, and Beyza Ural Marchand. 2014. "Does the Quality of Electricity Matter? Evidence from Rural India." *Journal of Economic Behavior and Organization* 107: 228–47.

Chan, Hei Sing, Maureen L. Cropper, and Kabir Malik. 2014. "Why Are Power Plants in India Less Efficient Than Power Plants in the United States?" *American Economic Review* 104 (5): 586–90.

Chilkoti, Avantika, and James Crabtree. 2014. "India eyes end to state coal monopoly." *Financial Times*, October 23.

Collard-Wexler, Allan, and Jan De Loecker. 2015. "Reallocation and Technology: Evidence from the US Steel Industry." *American Economic Review* 105 (1): 131–71.

Cropper, Maureen L., 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." National Bureau of Economic Research Working Paper 17383.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review* 102 (6): 2437–71.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What Do We Learn from the Weather? The New Climate-Economy Literature." *Journal of Economic Literature* 52 (3): 740–98.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101 (7): 3078–3108.

Donaldson, Dave. Forthcoming. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*.

Donaldson, Dave, and Richard Hornbeck. 2013. "Railroads and American Economic Growth: A 'Market Access' Approach." National Bureau of Economic Research Working Paper 19213.

Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics* 122 (2): 601–46.

Easterly, William, and Sergio Rebelo. 1993. "Fiscal Policy and Economic Growth: An Empirical Investigation." *Journal of Monetary Economics* 32 (3): 417–58.

Esfahani, Hadi Salehi, and Maria Teresa Ramirez. 2003. "Institutions, Infrastructure, and Economic Growth." *Journal of Development Economics* 70 (2): 443–77.

Fernald, John G. 1999. "Roads to Prosperity? Assessing the Link between Public Capital and Productivity." *American Economic Review* 89 (3): 619–38.

Fisher-Vanden, Karen, Erin T. Mansur, and Qiong Wang. 2015. "Electricity Shortages and Firm Productivity: Evidence from China's Industrial Firms." *Journal of Development Economics* 114: 172–88.

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 Working Paper 4913.

Garcia-Mila, Teresa, Therese J. McGuire, and Robert H. Porter. 1996. "The Effect of Public Capital in State-Level Production Functions Reconsidered." *Review of Economics and Statistics* 78 (1): 177–80.

Gramlich, Edward M. 1994. "Infrastructure Investment: A Review Essay." *Journal of Economic Literature* 32 (3): 1176–96.

Gupta, Rajan, and Harihar Shankar. 2014. *Global Energy Observatory*. Database. http://globalenergyobservatory.org/ (accessed September 20, 2014).

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114 (1): 83–116.

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 E. Stiglitz, 324–55. Oxford: Oxford University Press.

Holtz-Eakin, Douglas. 1994. "Public-Sector Capital and the Productivity Puzzle." *Review of Economics and Statistics* 76 (1): 12–21.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–48.

Hsieh, Chang-Tai, and Peter J. Klenow. 2014. "The Life Cycle of Plants in India and Mexico." *Quarterly Journal of Economics* 129 (3): 1035–84.

Hsieh, Chang-Tai, and Benjamin A. Olken. 2014. "The Missing 'Missing Middle.'" *Journal of Economic Perspectives* 28 (3): 89–108.
Hulten, Charles R., Esra Bennathan, and Sylaja Srinivasan. 2006. "Infrastructure, Externalities, and Economic Development: A Study of the Indian Manufacturing Industry." *World Bank Economic Review* 20 (2): 291–308.
Jensen, Robert. 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *Quarterly Journal of Economics* 122 (3): 879–924.
Lipscomb, Molly, A. 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* 5 (2): 200–31.
Matsuura, Kenji, and Cort J. Willmott. 2012. "Terrestrial Precipitation: 1900–2010 Gridded Monthly Time Series (1900–2010) V3.02." http://climate.geog.udel.edu/~climate/html_pages/download.html#P2011rev (accessed January 15, 2016).
McRae, Shaun. 2015. "Infrastructure Quality and the Subsidy Trap." *American Economic Review* 105 (1): 35–66.
Melitz, Marc J. 2003. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.
Nagaraj, R. 2002. "How to Improve India's Industrial Statistics." *Economic and Political Weekly* 37 (10): 966–70.
Power Finance Corporation. Various years. *Report on the Performance of State Power Utilities*. New Delhi, India: Power Finance Corporation Ltd.
Reinikka, Ritva, and Jakob Svensson. 2002. "Coping with Poor Public Capital." *Journal of Development Economics* 69 (1): 51–69.
Roller, Lars-Hendrik, and Leonard Waverman. 2001. "Telecommunications Infrastructure and Economic Development: A Simultaneous Approach." *American Economic Review* 91 (4): 909–23.
Rud, Juan Pablo. 2012a. "Electricity Provision and Industrial Development: Evidence from India." *Journal of Development Economics* 97 (2): 352–67.
Rud, Juan Pablo. 2012b. "Infrastructure Regulation and Reallocations within Industry: Theory and Evidence from Indian Firms." *Journal of Development Economics* 99 (1): 116–27.
Ryan, Nicholas. 2013. "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market." http://economics.mit.edu/files/9302 (accessed January 15, 2016).
Sincavage, Jessica R., Carl Haub, and O. P. Sharma. 2010. "Labor Costs in India's Organized Manufacturing Sector." *Monthly Labor Review* 133 (5): 3–22.
Srivastava, A. K., M. Rajeevan, and S. R. Kshirsagar. 2009. "Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region." *Atmospheric Science Letters* 10 (4): 249–54.
Steinbuks, Jevgenijs. 2011. "Financial Constraints and Firms' Investment: Results of a Natural Experiment Using Power Interruption." http://www.steinbuks.info/constraints.pdf (accessed January 15, 2016).
Steinbuks, J., and V. Foster. 2010. "When Do Firms Generate? Evidence on In-House Electricity Supply in Africa." *Energy Economics* 32 (3): 505–14.
Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by Donald W. K. Andrews and James H. Stock, 80–108. Cambridge, UK: Cambridge University Press.
Todd, Petra E., and Kenneth I. 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* 96 (5): 1384–1417.
Tybout, James R. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature* 38 (1): 11–44.
United States Department of Energy (DOE). 2013. "Manufacturing Energy Consumption Survey." http://www.eia.gov/consumption/manufacturing/index.cfm (accessed January 15, 2016).