How Do Electricity Shortages Affect Productivity? Evidence from India

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

January 14, 2014

Abstract

We develop a hybrid Leontief/Cobb-Douglas production function model that characterizes how input shortages affect firms. As a case study, we analyze how "power holidays" affect daily production at large Indian textile plants, using data from Bloom *et al.* (2013). We then study the short-run effects of electricity shortages on all Indian manufacturing plants between 1992 and 2010, using archival data on shortages, previously-unavailable panel data, and an instrument for shortages based on variation in hydro reservoir inflows. We estimate that electricity shortages are a substantial drag on Indian manufacturing, reducing output by about five percent. However, productivity effects are smaller: because electricity is a small share of costs, higher-cost self-generation increases energy costs by only about 0.15 to 0.5 percent of revenues, and because most inputs can be stored during outages, the productivity loss is only a fraction of the output loss. We also show that because of economies of scale in self-generation, shortages impose much greater losses on small plants, suggesting an additional distortion to the firm size distribution in developing economies.

JEL Codes: D04, D24, L10, L51, L94, Q41, Q47.

Keywords: Manufacturing productivity, India, electricity shortages.

We thank Nick Ryan, Jagadeesh Sivadasan, and seminar participants at the 2013 NBER Summer Institute and 2014 NBER Winter IO/EEE meetings for helpful comments. We are particularly grateful to Nick Bloom and coauthors for sharing the data from "Does Management Matter?" We thank Troy Smith for help understanding those data, as well as Shaleen Chavda for insight into the textile industry. We thank Deepak Choudhary for helpful research assistance. 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. Data and code for replication are available from Hunt Allcott's website.

*Allcott: New York University and NBER. hunt.allcott@nyu.edu. Collard-Wexler: New York University and NBER. wexler@stern.nyu.edu. O'Connell: City University of New York. soconnell@gc.cuny.edu.

1 Introduction

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

In this paper, we estimate the short-run effects of electricity shortages on manufacturing plants in India. One 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, precisely because the potential losses would be so large, many firms might insure themselves against outages by purchasing generators or otherwise substituting away from grid electricity. The limited existing evidence could support either argument. Foster and Steinbuks (2009), Zuberi (2012), and others argue that the cost of self-generation is relatively small, and Alam (2013), Fisher-Vanden, Mansur, and Wang (2013), and others highlight ways in which plants substitute away from electricity when shortages worsen. By 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 three reasons why this question is difficult. First, the standard production function model needs to be adapted for the case of input shortages, when firms cannot procure electricity for a part of the year. Second, the necessary data are difficult to acquire: some industrial surveys do not have useful questions on electricity use, and more detailed firm-level datasets are often unrepresentative. Meanwhile, countries that have electricity shortages are often the same types of countries that do not record or disclose high-quality data on the performance of public infrastructure. Third, shortages are not exogenous to productivity. 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 causal estimates of shortages, albeit in opposite directions.

We begin by providing background on electricity shortages and industrial electricity use in India. First, there is significant variation in shortages within states over time, driven by weather, coal shortages, fluctuating hydroelectric production, and other factors. Second, Indian manufacturers self-generate approximately 35 percent of their electricity, more than twice the share in the United States. Third, self-generation is sharply increasing in plant size: while only 10-20 percent of plants

[1] (For a tally of responses, see Appendix Table A18.)

with fewer than 10 employees self-generate, about 75 percent of plants with more than 500 employees do so.

We then present a production function model in which output is Leontief in electricity and a Cobb-Douglas aggregate of materials, capital, and labor. Shortages have very different effects on firms with vs. without generators. Firms that use generators face an increase in electricity costs (the *input cost effect*). This enters the profit function like an output tax and also reduces demand for other inputs (the *output tax effect*). Even if these firms never stop production during shortages, total factor productivity (TFP) is lower due to the *input variation effect*: using different bundles of fully flexible inputs during outage vs. non-outage periods is less efficient than having a constant flow of production. Firms without generators shut down during shortages, which reduces output and causes waste of non-storable inputs (the *shutdown effect*). The waste reduces demand for non-storable inputs when firms foresee periods of higher shortages (the *shutdown tax effect*).

The empirical analysis begins with a case study of large textile manufacturers in Gujarat and Maharashtra, using data shared by Bloom *et al.* (2013). These plants face pre-scheduled "power holidays" once each week, and they respond either by self-generation or by shutting down, depending on the week. While these data include only 22 plants, all of which have generators, they give very clean estimates of the effects of shortages. Despite the fact that grid power is unavailable approximately 1/7th of the time, the effects are quite small: energy costs rise by 0.24 percent of revenues, and while physical output drops by 1.1 percent, TFP only decreases by 0.05 percent because 95 percent of inputs (including both labor and materials) can be flexibly adjusted on power holidays.

We then broaden our scope to all Indian manufacturing plants using data from the Annual Survey of Industries (ASI). We use a difference estimator, exploiting changes in shortages within states over time. To address the potential endogeneity of shortages - for example, economic growth both increases manufacturing output and worsens shortages - we instrument with changes in electricity production from dams, which are driven by changes in the amount of water flowing intro reservoirs. While working in India, we acquired a version of the ASI with consistent plant identifiers dating to 1992, which allows an unusually long panel of Indian plants. To complement this longer panel, we gathered archival data from India's Central Electricity Authority on shortages, reservoir inflows, generation by hydro and other plants, and other aspects of the Indian power sector.

Our instrumental variables estimates show that for plants that own generators, a one percentage point increase in shortages increases the share of self-generated electricity by 0.57 percentage points, which raises total input costs by 0.02 to 0.07 percent of revenues. Across all plants, a one percentage point increase in shortages decreases revenues by 0.68 percent. The accompanying TFP loss, however, is much smaller - the confidence interval bounds it at no more than 0.29 percent.

The effects of shortages vary in ways predicted by the model. Only plants that self-generate experience an increase in total energy costs, while non-generators experience much larger revenue losses. Firms in industries with higher electric intensity are more exposed to shortages, experiencing a larger increase in energy revenue share and also a larger decrease in output. The results are essentially identical under a battery of alternative specifications, including using fixed effects instead of differences, controlling for rainfall, using an alternative measure of shortages, constructing TFP in different ways, omitting various controls, and trimming outliers with different tolerances.

We then use simulations calibrated to ASI plants and production functions to calculate the nationwide effects of the average level of shortages over our sample (7.1 percent), holding capital stock constant. Across all plants, revenue and TFP are 7.1 and 1.9 percent lower, respectively. The simulated effects on output and TFP are economically similar and statistically indistinguishable from the empirical estimates, which builds confidence that the estimates are reasonable and that the model captures the first-order issues.

As with the empirical estimates, however, simulated effects differ starkly for plants with vs. without generators: those with generators see revenue and TFP drop by 0.7 and 0.1 percent, while those without generators experiences losses of 10.3 and 2.9 percent. For self-generators and non-generators, the reasons why output losses are larger than the percent of time shut down are the output tax and shutdown tax effects: shortages act like taxes that cause firms to reduce other inputs. These input reductions are also one reason why TFP losses are much smaller than output losses; the other important reason is that when non-generators shut down, they lose output but only waste non-storable inputs. Thus, while electricity shortages are a large drag on manufacturing output, they do not in isolation explain much of the difference in TFP between India and more developed economies.[2]

We also use the simulations to explore how electricity shortages might affect the firm size distribution. Tybout (2000) discusses several potential causes for why the firm size distribution in developing countries tends to have a "missing middle," and Hsieh and Klenow (2012) suggest that electricity shortages combined with differential access to grid electricity could be an important factor favoring large plants. We build on this idea, focusing on a different channel: economies of scale in generator ownership. Simulations show that effects of outages on revenues and TFP are 50 percent larger for plants with fewer than 100 employees compared to larger plants.

The remainder of this section discusses related literature. Section 2 provides background on the Indian electricity sector, the causes of electricity shortages, and manufacturers' responses to shortages. Section 3 details the production function model. Section 4 is the case study of textile manufacturing in western India, using data from Bloom *et al.* (2013). Sections 5 and 6 present the ASI data and empirical results. Section 7 details the counterfactual simulations, and Section 8 concludes.

[2] See Banerjee and Duflo (2005), Hsieh and Klenow (2009), and others for discussions.

1.1 Related Literature

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

This literature has faced two basic problems. First, infrastructure spending is econometrically endogenous to economic growth. There could be reverse causality: fast growth increases tax revenues, which allow more infrastructure spending. There is also economic endogeneity: infrastructure may be specifically allocated to places that are growing more quickly or slowly. Second, using aggregate infrastructure spending or quantity as the independent variable often hides important variation in effects between infrastructure of different types or quality levels. In the Indian context, for example, spending on power plants does not necessarily translate into electricity provision, because plants are frequently offline due to mechanical failure or fuel shortages.

Our paper is part of a recently-growing literature that evaluates the effects of infrastructure by combining microdata with within-country variation generated by natural experiments. This includes Banerjee, Duflo, and Qian (2012), Donaldson (2012), and Donaldson and Hornbeck (2013) on the effects of railroads in China, India, and the United States, Duflo and Pande (2007) on irrigation dams in India, Jensen (2007) on information technology, Baisa, Davis, Salant, and Wilcox (2008) on the benefits of reliable water provision in Mexico, and Baum-Snow (2007, 2013), Baum-Snow, Brandt, Henderson, Turner, and Zhang (2013), and Baum-Snow and Tuner (2012) on urban transport expansions in China and the United States. A subset of this literature focuses on electricity supply: Chakravorty, Pelli, and Marchand (2013), Dinkelman (2011), Lipscomb, Mobarak, and Barham (2013), Rud (2012a), and Shapiro (2013) study the effects of electricity grid expansions, while Alby, Dethier, and Straub (2011), Foster and Steinbuks (2009), Steinbuks (2011), Steinbuks and Foster (2010), Reinikka and Svensson (2002), and Rud (2012b) study firms' generator investment decisions. Several recent papers focus specifically on Indian electricity supply: Ryan (2013) estimates the potential welfare gains from expanding transmission infrastructure, Cropper, Limonov, Malik, and Singh (2011) and Chan, Cropper, and Malik (2014) study the efficiency of Indian coal power plants, and Abeberese (2012) tests how changes in electricity prices affect manufacturing productivity.

Three recent papers study the effects of blackouts on manufacturers. Fisher-Vanden, Mansur, and Wang (2013) show that when shortages become more severe, Chinese firms purchase more energy-intensive inputs, but they do not self-generate more electricity. Zuberi (2012) estimates a dynamic model of manufacturing production using data from Pakistan, showing how firms reallocate production to non-shortage periods. Alam (2013) studies how India's steel vs. rice milling industries respond differently to blackouts. Relative to these important papers, our study benefits from particularly clean data and identification: we have a clear case study using the high-quality textile plant data from Bloom *et al.* (2013), a 19-year national panel using previously-unavailable ASI data, newly-gathered archival data on the severity of shortages across Indian states, and an instrument that addresses the endogeneity of blackouts with respect to growth. Our paper also benefits from the way that we integrate theory and empirics: our model formalizes the major channels through which shortages affect production, and the close correspondence between simulation and empirical results builds confidence in the estimates.

2 Background

2.1 Power Sector Data

Our power sector data are from India's Central Electricity Authority (CEA). Many of the same types of data available online from the U.S. Energy Information Administration are also collected by the CEA. Unfortunately, however, the online data are incomplete, and the hard copies of some printed materials have been misplaced, so data have to be hand-collected from CEA staff. With the cooperation of CEA management and the help of research assistants in New Delhi, we were able to compile, digitize, and clean about 25 years of data for this and related projects. Table 1 details these power sector variables and other state-level data.[3]

The primary measure of electricity shortages is the percent energy deficit reported in the Load Generation Balance Report. Analysts at CEA and Regional Power Committees estimate the quantity that would be demanded for each state and month at current prices in the absence of shortages. The state-by-year sum is our "Assessed Demand" variable. "Shortage" is the percent difference between this counterfactual quantity demanded and the actual quantity supplied. In the 2011-2012 fiscal year, nationwide shortage was 8.5 percent, and shortages average 7.2 percent over the sample period. The CEA also estimates "Peak Shortage," an analogous measure of power shortage in peak demand periods. While (total kilowatt-hour) Shortage is more appropriate for our analysis, Peak Shortage and Shortage are highly correlated, with an R^2 of 0.5, and robustness checks show that results are similar when we use Peak Shortage instead of Shortage.[4]

From an annual report called the Review of the Performance of Hydro Power Stations, we observe inflows into reservoirs behind 22 major dams covering about 40 percent of national hydroelectric capacity. From the CEA's General Review, we observe each state's total annual electricity

[3] Throughout the paper, we use the word "state" to refer to both states and Union Territories.

[4] Although it is likely that shortages are measured with error, correlations with independent data suggest that the CEA's estimates contain meaningful information. Alam (2013) shows that Peak Shortage is correlated with her measure of blackouts based on variation in nighttime lights measured by satellites; she does not report a correlation with Shortage. In the World Bank Enterprise Survey, plants in higher-Shortage states report a larger share of self-generated electricity and are more likely to report that electricity is their primary obstacle to growth. Furthermore, our empirical results show that Shortages are positively correlated with hydroelectric supply and correlated in theoretically-predicted ways with self-generation and other outcomes in the ASI.

generation by fuel type, including hydroelectric plants. From the General Review, we also collected total quantity of electricity sold by utilities to end users for each state and year.

Aside from these electricity market variables, our empirical analysis also uses weather and temperature data from the Meteorological Department of the National Climate Centre of India. These data provide daily average temperatures and rainfall at one-degree gridded intervals across India. Using state border coordinates, we associate the grid points with particular states to arrive at annual state-level measures. Cooling degrees is a commonly-used correlate of electricity demand; it is the difference between the day's average temperature and 65 degrees Fahrenheit, or zero if the day's average temperature is below 65.

2.2 Reasons for Systemic Shortages

As of February 2013, India had 214 gigawatts of utility-scale power generation capacity, or about one-fifth the US total (CEA 2013). Of this, 58 percent was coal, nine percent was natural gas, and 18 percent was hydro-electric. While power generation has been open to private investment since 1991, 70 percent of electricity supply remains government owned: 40 percent is owned by state governments, and 30 percent is owned by central government entities. Although some retail distribution companies have been privatized, most of distribution is managed by state-run companies, which are often called State Electricity Boards (SEBs).

The proximate reason for shortages is that distribution companies do not raise retail prices in order to clear the market. In fact, conditional on state and year effects, there is no correlation between shortages and the median electricity price paid by ASI plants. Aside from being stark evidence on how prices do not adjust to supply and demand conditions, this also means that the effects we estimate are caused by input shortages, not by input price changes.

There are several underlying systemic reasons for shortages. The first is the "infrastructure quality and subsidy trap" (McRae 2013): distribution companies provide low-quality electricity to consumers, who tolerate poor service because they pay very low prices, distribution companies' losses from low prices are covered by government subsidies, and politicians support the subsidies to avoid voter backlash. At least since the 1970s, State Electricity Boards have offered un-metered electricity at a monthly fixed fee and zero marginal cost to agricultural consumers, largely to run well pumps (Bhargava and Subramaniam 2009). In 2010, the national average retail electricity cost paid by agricultural consumers was 1.23 Rupees per kilowatt-hour (Rs/kWh), against Rs 4.78 for industrial consumers and 3.57 Rs/kWh for all consumers. (The exchange rate is about 50 Rupees per dollar, and the average electricity price across all consumers in the United States is about 10 cents/kWh.) Although determining optimal electricity prices would be complex due to the variety of distortions, agricultural electricity is almost certainly under-priced.

Distortions in pricing are relevant only for consumers who actually pay for electricity. Twenty-six percent of electricity generated in India in 2010-2011 was lost due to "technical and commercial losses," meaning theft or poor transmission infrastructure. This is down from 34 percent in 2004-2005. Distribution companies thus have no ability to charge any price, let alone raise prices, on a significant share of electricity.

Agricultural subsidies and technical and commercial losses have led to mounting losses. The SEBs receive large annual payments from state governments to cover these losses, and in particular to fund the subsidies for agricultural consumers, but these payments and the cross-subsidy from industrial customers are not sufficient to cover the SEBs' costs. Between 1992 and 2009, the SEBs lost $54 billion dollars (again, in real 2004 dollars). These mounting losses caused the SEBs to reduce infrastructure investment, and degraded infrastructure further increases the probability of blackouts. The SEBs are bailed out at irregular intervals by the government.

A second systemic reason for shortages is underinvestment in new generation capacity. For example, after the 1991 liberalization, 200 Memoranda of Understanding were signed between the government and investors to build 50 gigawatts of generation capacity, but less than four gigawatts of this was actually built (Bhargava and Subramaniam 2009). Of the 71 gigawatts of capacity targeted to be built between 1997 and 2007, only half was actually achieved (CEA 2013a). Potential power plant investors faced concerns over both output demand and input supply. Their main customers, the State Electricity Boards, faced serious financial problems, and it was not clear that they would be able to honor contracts. Meanwhile, the main supplier of coal is Coal India, a government-owned monopoly that is struggling to keep pace with demand growth.

In addition, the 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, or other factors such as equipment malfunction, coal shortages, or poor coal quality.[5] Furthermore, when capacity is utilized, it is substantially less efficient than comparable plants in the United States (Chan, Cropper, and Malik 2014).

One potential solution to problems with retail distribution companies is "open access": allowing consumers to contract directly with generators. The 2003 Electricity Act mandated open access, but in practice direct power sales to bulk consumers have not materialized (GOI 2009, 2012), partially because states have imposed additional charges on open access consumers and have also banned export of power to open access consumers in other states.

2.3 Variation in Shortages

These systemic factors differ across states, generating differences in shortages. A substantial part of these differences persist across years. Figure 1 shows a map of average shortages by state from 1992-2011, with higher-shortages states colored darker. Jammu and Kashmir was the highest-shortage state at 23 percent, followed by Bihar, Arunachal Pradesh, Uttar Pradesh, and Madhya Pradesh, respectively. Lakshadweep had zero shortages, and the next four lowest-shortage states

[5] We thank Cropper et al. (2013) for sharing their coal plant outage data with us. Our original intent was to use these outages as an instrument for shortages, but the first stage is not sufficiently powerful.

were Chandigarh, Himachal Pradesh, Delhi, and West Bengal, respectively. About 7.5 percent of electricity consumed in 2011-2012 was generated in another state. Because distribution companies are able to procure power from other states, supply-demand imbalances do not vary as much as they would under autarky.

Figure 2 shows that there is a negative association between 2010 per capita GDP and average shortage over the sample. This suggests that low levels of development are associated with poor institutions, bad management, and other factors that worsen provision of public infrastructure. However, there is substantial residual variation. Rajasthan, Jharkhand, and Sikkim have low GDP and low shortages, partially because their slow GDP growth makes it easier for supply to keep up with demand. Because end-of-sample GDP is highly correlated with GDP growth, this implies that shortages could be correlated with factors that also affect manufacturing growth and productivity. This highlights the importance of instrumenting for shortages in our empirical analysis.

There is also substantial variation in shortages within states over time. Figure 3 shows the time path of annual average shortage over our sample for five large states in different parts of the country. West Bengal has had consistently low shortages for the past 20 years. Maharashtra, which is now one of the highest-shortage states, had only small shortages in the early 1990s. Karnataka, which faced almost zero shortage in the mid-2000s, had significant shortages in the early to mid 1990s. Gujarat has reliable power supply now, but in the mid-2000s was experiencing shortages.

Several factors drive year-to-year fluctuations in shortages. On the demand side, fast or slow economic growth over a few years can increase or decrease shortages. In addition, low rainfall in a given year can increase farmers' utilization of groundwater pumps, which can markedly increase electricity demand. An unusually hot summer can also increase electricity demand to cool buildings. On the supply side, the electricity market is still small enough that individual plants can affect the aggregate supply-demand balance. Power plant outages for maintenance or due to fuel shortages can cause electricity shortages, and new plants coming online can temporarily reduce shortages. Later in the paper, we will discuss one other factor that satisfies the first stage and exclusion requirements to be used as an instrument for shortages: variation in hydroelectric production due to low rainfall in the south and low snowpack in the north.

2.4 Industrial Electricity Use in India

A natural response to outages is to self-generate electricity. Manufacturers in India generate 35 percent of manufacturing electricity consumption, more than twice the 15.8 percent for U.S. manufacturers reported in the Manufacturing Energy Consumption Survey (MECS) (U.S. DOE 2013). Figure 4 compares manufacturing electricity generation in India to the United States. Each dot reflects a three-digit industry code from India's National Industrial Classification (NIC), comparing Indian data from the Annual Survey of Industries to U.S. data from the MECS.[6]

[6]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

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. Many industries in the United States - where power outages are relatively unimportant - produce a large share of their power. For instance, in the sugar refining industry, leftovers 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 y-axis implies that many industries in India produce much more than their counterparts in the U.S. For instance, plastics manufacturers in the United States produce none of their power (U.S. DOE 2013), while in India, the plastics industry produces 70 percent of its electricity consumption.

2.5 Self-Generation and Plant Size

The reason why electricity is typically generated in large power plants instead of by individual manufacturing plants is that there are strong economies of scale in generation. Even within the range of manufacturing plant sizes, generator costs rise meaningfully per kilowatt of capacity. The result is that larger plants are much more likely to self-generate power, as shown in Figure 5. This economy of scale has important implications for how electricity shortages affect the plant size distribution, an issue which we return to later in the paper.

Although smaller plants may be more affected, large plants also report significant losses from shortages. The World Bank Enterprise Survey (WBES) for 2005 surveyed 2286 Indian manufacturing firms in 50 cities in 16 states about their inputs, outputs, and business environments. It was in this survey that managers reported electricity supply as their single most important barrier to growth. Table 2 compares the WBES data for "Small" plants (<100 workers) and "Large" plants. Large plants experience fewer outages. Consistent with the ASI, large plants are more likely to own generators, and conditional on owning a generator, they source a larger share of electricity internally. While small plants report that they lose eight percent of revenues to electricity input disruptions, large plants report losing five percent. Furthermore, 26 percent of large plants report the electricity is the biggest obstacle to growth.

3 Model

In this section, we develop a model of how electricity shortages affect manufacturers. τ indexes points in time, which we refer to as "days." Every day, a producer 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

datasets: chemicals and refining are not broken out into many different sub-industries in the public US data, so Indian sub-industries such as Explosives, Chemicals Not Elsewhere Classified (NEC), Matches, and Perfumes and Cosmetics are matched to "Chemicals," a broader industry where other establishments are more likely to have feedstock for self-generation, and thus a higher self-generation share. Similarly, Natural Gas and LPG Bottling, Coal NEC, and Coke Oven Products are matched to "Petroleum and Coal Products," another very broad category.

$Q_{it} \equiv \int_\tau Q_{it\tau} d\tau$ is the annual aggregate. We do not model the possibility for inter-day substitution of production; this is covered nicely by Alam (2013) and Zuberi (2012). To the extent that firms can adjust in this way, this reduces the losses from shortages relative to what we simulate in Section 7.

The daily production function is Leontief in electricity and a Cobb-Douglas aggregate of capital, labor, and materials, with physical productivity A:

$$Q = \min\{A^{1-\frac{1}{\epsilon}} K^{\alpha_K} L^{\alpha_L} M^{\alpha_M}, \frac{1}{\lambda} E\} \tag{1}$$

The Leontief production function dictates that electricity is used in constant proportion $\frac{1}{\lambda}$ with output. Electricity intensity λ varies across industries. As is common, we assume that the Cobb-Douglas aggregate, $A^{1-\frac{1}{\epsilon}} K^{\alpha_K} L^{\alpha_L} M^{\alpha_M}$, has constant returns to scale, so $\alpha_K + \alpha_L + \alpha_M = 1$. Having A inside the Cobb-Douglas aggregator ensures that electricity is used in fixed proportion to output instead of to the bundle of other inputs.[7]

Since we will observe total revenues rather than physical quantities produced, we need to relate revenues to our production function in equation (1). As in Foster, Haltiwanger, Syverson (2008), Bloom (2009), and Asker, Collard-Wexler, and De Loecker (2013), we consider a firm facing a constant elasticity demand curve (CES) given by $Q_{it} = B_{it} p_{it}^{-\epsilon}$, where p is the output price. Combining the production function and the demand curve, we obtain an expression for the revenue-generating production function $R_{it} = \min\{\Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it}^{\beta_M}, B_{it}^{\frac{1}{\epsilon}} \frac{1}{\lambda} E_{it}\}$ where $\Omega_{it} \equiv A_{it}^{1-\frac{1}{\epsilon}} B_{it}^{\frac{1}{\epsilon}}$, and $\beta_X \equiv \alpha_X (1 - \frac{1}{\epsilon})$, for $X \in \{K, L, M\}$. We will use an elasticity of demand of $\epsilon = -10$, but we will also verify our results with other elasticities such as $\epsilon = -4$, the value used by Bloom (2009), and Asker, Collard-Wexler, and De Loecker (2013).

3.1 Decision Variables

We assume that inputs fall into three categories: fixed, semi-flexible at the yearly level, and fully flexible at the daily level.

1. *Fixed Inputs* are chosen before the current year and are exogenous in this analysis. We assume that capital stock K is fixed.

2. *Semi-Flexible Inputs* can be modified at the beginning of a year t, but they cannot be modified from day to day. For the model and simulations, we treat labor as semi-flexible, since firms

[7] We have also considered a production function which is Cobb-Douglas in K, L, M, and also E. There are two main differences in this model's predictions. First, plants that own generators will always self-generate at least a small amount of electricity no matter how high the cost, because an input's marginal revenue product approaches infinity as quantity input limits to zero. By contrast, plants such as the textile factories in Section 4 sometimes choose to shut down completely during outages even if they have generators. Second, higher costs of self-generated electricity act like an input tax on electricity, while they act like an output tax in the Leontief-in-electricity model.

Quantitatively, the effects of blackouts are the same in the two models for plants that do not have generators. For plants that have and use generators, the Cobb-Douglas model would find a smaller effect of shortages on output and productivity than our Leontief-in-electricity model, since there is scope for substituting electricity with other inputs.

cannot hire and fire workers from one moment to the next as blackouts occur. This gives $L_{it\tau} = L_{it}$. An alternative interpretation is that these are non-storable inputs, which cannot be stockpiled and used another day.[8]

3. *Fully Flexible Inputs* can be modified for each day τ. We treat materials and electricity as fully flexible.

3.2 Power Outages

Power outages occur on each day with probability δ, and firms observe whether there is a power outage before setting their fully flexible inputs. If there is not an outage, firms can purchase electricity from the grid at price $p^{E,G}$. If there is an outage, firms with generators can self-produce electricity at price $p^{E,S} > p^{E,G}$. Firms without generators will have zero electricity use during an outage, and thus zero output.

3.3 Firm's Problem

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

$$\begin{aligned}\Pi_{it\tau} =& p \min\{A_{it}{}^{1-\frac{1}{\epsilon}} K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it\tau}^{\alpha_M}, \frac{1}{\lambda} E_{it\tau}\} \\ & - p^L L_{it} - p^M M_{it\tau} - p^E E_{it\tau},\end{aligned} \tag{2}$$

where p, p^L, p^M are the prices of output, labor, and materials, respectively. Note that capital is excluded, since it is fixed, and thus a sunk cost in the yearly decision problem.

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

$$A_{it}^{1-\frac{1}{\epsilon}} K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it\tau}^{\alpha_M} = \frac{1}{\lambda} E_{it\tau}, \tag{3}$$

Given this, one can rewrite the profit function, substituting in for the optimized value of electricity:

$$\begin{aligned}\Pi_{it\tau} &= (1 - \frac{\lambda p^E}{p}) p A_{it}^{1-\frac{1}{\epsilon}} K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it\tau}^{\alpha_M} - p^L L_{it} - p^M M_{it\tau} \\ &= (1 - \frac{\lambda p^E}{p}) \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it\tau}^{\beta_M} - p^L L_{it} - p^M M_{it\tau}.\end{aligned} \tag{4}$$

[8]Some of the high self-generation in Indian industries, such as in plastics, is presumably due to inputs being spoiled during a power outage. In these industries, it might be more plausible to assume that materials are also semi-flexible inputs.

Let $\gamma \equiv \frac{\lambda p^E}{p} = \frac{p^E E}{pQ}$, the electricity revenue share. Notice that if $(1-\gamma) < 0$, then the firm will choose not to produce.

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

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

We ignore case (3): if firms never produce, they never appear in the data. Firms without generators effectively have $p^{E,S} = \infty$, so case (1) cannot arise. Of the firms with generators, those with higher λ will be in case (2). In other words, higher-electricity intensity firms will be more likely to shut down during grid power outages.[9]

The first-order condition with respect to materials yields:

$$\beta_M(1-\gamma)\frac{R_{it\tau}}{M_{it\tau}} - p^M = 0. \tag{5}$$

This is the usual Cobb-Douglas first-order condition for materials, except that the marginal revenue product of materials is reduced by γ. Since λ is constant, γ is higher when a firm self-generates and pays a price for power of $p^{E,S}$, rather than purchasing from the grid and paying $p^{E,G}$. One can thus interpret $\mathcal{T} \equiv \lambda\frac{p^{E,S}-p^{E,G}}{p}$ as an "output tax" on output due to self-generation, and this tax is higher for industries which are more electricity intensive.

The marginal revenue product of materials is:

$$MRPM = \begin{cases} \beta_M(1-\gamma)\frac{R_{it\tau}}{M_{it\tau}} & \text{if no power outage} \\ \mathcal{T}\beta_M(1-\gamma)\frac{R_{it\tau}}{M_{it\tau}} & \text{if power outage} \end{cases} \tag{6}$$

The first-order condition for labor is more complicated, since a firm must integrate over outage and non-outage days when setting semi-flexible inputs. If a plant is in case (1), meaning that it self-generates during power outages, then the first-order condition is given by:

$$MRPL_1 = \beta_L(1-\lambda\frac{p^{E,G}}{p})\left[(1-\delta)\frac{R_{itG}}{L_{it}} + \delta\mathcal{T}\frac{R_{itS}}{L_{it}}\right] = p^L, \tag{7}$$

where R_{itS} indicates revenue during a shortage period. This expression also includes an "output tax"$\mathcal{T}$ during shortage periods. We call the reduction in the marginal revenue products of materials and labor for self-generating plants the *output tax effect*.

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

For firms in case (2), i.e. firms without generators or firms that will not produce during outages since $\lambda p^{E,S} < p$, the marginal revenue product of labor is:

$$MRPL_2 = (1-\delta)(1-\gamma)\beta_L \frac{R_{itG}}{L_{it}} = p^L, \tag{8}$$

where R_{itG} indicates revenue during a non-shortage period.

This is the usual Cobb-Douglas first-order condition, except that the marginal revenue product is reduced by $(1-\gamma)$ *and* $(1-\delta)$, the fraction of the time the plant will be down due to power shortages. We call this reduction in marginal product of labor for non-generators the *shutdown tax effect*.

3.4 Profits

The above first-order conditions imply optimal input bundles L^*_{it}, M^*_{itS}, and M^*_{itG}, where the latter two represent optimal materials input during outage and non-outage periods, respectively. The firm's annual profit can then be expressed as:

$$\begin{aligned}
\Pi_{i,t} =&(1-\delta)(1-\frac{\lambda p^{E,G}}{p})\left(\Omega_{it}K_{it}^{\beta_K}(L^*_{it})^{\beta_L}(M^*_{itG})^{\beta_M} - p^M M^*_{itG}\right) \\
&+\delta(1-\frac{\lambda p^{E,S}}{p})\left(\Omega_{it}K_{it}^{\beta_K}(L^*_{it})^{\beta_L}(M^*_{itS})^{\beta_M} - p^M M^*_{itS}\right) \\
&- p^L L^*_{it},
\end{aligned} \tag{9}$$

where the second line will be zero for firms which do not self-generate.

3.5 Productivity

3.5.1 Production Function Estimation

We use the first-order condition approach to production function estimation[10] to recover production function coefficients β_L, β_M, β_K and γ using yearly data from the Annual Survey of Industries. In our context, however, the first-order conditions depend on variables that vary between outage and non-outage periods and are thus unobserved in the yearly data. Below, however, we see that for plants that do not self-generate, the first-order conditions simplify to functions only of yearly aggregates.

For non-self-generators, γ equals the revenue share of electricity over the year:

$$\gamma = \frac{p^{E,G}E_{it\tau}}{R_{it\tau}} = \frac{p^{E,G}E_{it}}{R_{it}} \tag{10}$$

[10] For additional discussion, see De Loecker and Warzynski (2012) and Haltiwanger, Bartelsman and Scarpetta (2013).

The latter equality holds because $(1-\delta)E_{it\tau} = E_{it}$ and $(1-\delta)R_{it\tau} = R_{it}$: non-self-generators use zero electricity and earn zero revenues during outages.

Similarly, the first-order condition for labor gives:

$$\beta_L = \frac{p^L L_{it}}{(1-\delta)R_{it\tau}} \frac{1}{1-\gamma} = \frac{p^L L_{it}}{R_{it}} \frac{1}{1 - \frac{p^{E,G}E_{it}}{R_{it}}}. \tag{11}$$

We thus have the usual Cobb-Douglas equality of β_L with the revenue share of labor, except adjusted by the inverse of one minus the electricity revenue share.[11]

Likewise, the first-order condition for materials yields:

$$\beta_M = \frac{p^M M_{it\tau}}{R_{it\tau}} \frac{1}{1-\gamma} = \frac{p^M M_{it}}{R_{it}} \frac{1}{1 - \frac{p^{E,G}E_{it}}{R_{it}}}. \tag{12}$$

Again, the latter equality holds because $M_{it\tau} = (1-\delta)M_{it}$, $(1-\delta)E_{it\tau} = E_{it}$, and $R_{it\tau} = (1-\delta)R_{it}$ for non-self-generators.

Finally, using the constant returns to scale assumption that $\alpha_K + \alpha_L + \alpha_M = 1$, the capital coefficient is given by $\beta_K = (1 - \frac{1}{\epsilon}) - \hat{\beta}_L - \hat{\beta}_M$. We use median regression to estimate these coefficients by three-digit industry, using only plants in the ASI that never self-generate. See Appendix A for additional details.

3.5.2 Productivity Effect of Shortages

With no power shortages, productivity is:

$$\omega_{it} = r_{it} - \beta_K k_{it} - \beta_M m_{it} - \beta_L l_{it} \tag{13}$$

where lower case variables denote the logarithms of variables in upper case; i.e., $x_{it} = \log(X_{it})$.

For plants that do not have a generator or have one but choose not to self-generate, revenue is:

$$\begin{aligned} R_{it} &= \int_\tau \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it\tau}^{\beta_M} d\tau \\ &= (1-\delta)^{1-\frac{1}{\epsilon}-\beta_M} \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it}^{\beta_M} \end{aligned} \tag{14}$$

Taking logs, this yields:

$$r_{it} = \beta_K k_{it} + \beta_M m_{it} + \beta_L l_{it} + \underbrace{\omega_{it} + (1 - \frac{1}{\epsilon} - \beta_M)\log(1-\delta)}_{\hat{\omega}_{it}} \tag{15}$$

and since $1-\delta \leq 1$, $\log(1-\delta) < 0$, so shortages reduce measured productivity $\hat{\omega}_{it}$ relative to ω_{it}.

[11] In a production function that is Cobb-Douglas in electricity, a similar equation would hold with the absence of the $\frac{1}{1-\gamma}$ adjustment.

If plants have generators and use them during outages, then revenue is given by:

$$\begin{aligned} R_{it} &= \int_{\tau} \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it\tau}^{\beta_M} d\tau \\ &= \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} \left((1-\delta) M_{itG}^{\beta_M} + \delta M_{itS}^{\beta_M} \right), \end{aligned} \tag{16}$$

where M_{itS} is the bundle of materials chosen during shortages, and M_{itG} is the bundle of materials chosen otherwise.

Define C as the loss in output from using different bundles of materials in shortage and non-shortage periods, relative to using the same bundle in both periods:

$$C_{it} = \frac{(1-\delta) M_{itG}^{\beta_M} + \delta M_{itS}^{\beta_M}}{((1-\delta) M_{itG} + \delta M_{itS})^{\beta_M}}. \tag{17}$$

By Jensen's inequality, $C < 1$, since $\left((1-\delta) M_{itG}^{\beta_M} + \delta M_{itS}^{\beta_M} \right) < ((1-\delta) M_{itG} + \delta M_{itS})^{\beta_M}$. C_{it} is increasing in β_M, as $\beta_M < 1$ and C would be one for a $\beta_M = 1$ and $\epsilon = -\infty$, the linear production case. For small δ, C_{it} is decreasing in δ.

This gives:

$$\begin{aligned} R_{it} &= \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} \left((1-\delta) M_{itG}^{\beta_M} + \delta M_{itS}^{\beta_M} \right) \\ &= C_{it} \Omega_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it}^{\beta_M} \end{aligned} \tag{18}$$

Taking logs, this yields:

$$r_{it} = \beta_K k_{it} + \beta_M m_{it} + \beta_L l_{it} + \underbrace{\omega_{it} + c_{it}}_{\hat{\omega}_{it}}. \tag{19}$$

Since $C_{it} < 1$, $c_{it} < 0$, so outages again decrease measured productivity $\hat{\omega}_{it}$.

Collecting our results, we have the effects of shortages on measured TFP:

$$\hat{\omega}_{i,t} - \omega_{i,t} = \begin{cases} (1 - \frac{1}{\epsilon} - \beta_M) \log(1-\delta) & \text{If no self-generation} \\ c_{it} & \text{If self-generation} \end{cases} \tag{20}$$

We call the first line the *shutdown effect.* The $(1 - \frac{1}{\epsilon} - \beta_M) = \beta_L + \beta_K$ term represents the share of inputs that are not fully flexible - in this case, capital and labor. Thus, the shutdown effect on TFP is just the share of the year shut down multiplied by the share of inputs that are wasted during outages. We call the second line the *input variation effect.*

4 Case Study: Large Textile Manufacturers

Bloom *et al.* (2013) study how improved management practices increased productivity at textile plants in Gujarat, Maharashtra, and Dadra and Nagar Haveli. In the industrial areas where these plants operate, there are scheduled "power holidays" on a given day of the week, during which there is typically no grid electricity. As a case study to illustrate and calibrate the model, this section uses data shared by Bloom *et al.* (2013) to estimate how power holidays change inputs and production.

4.1 Overview and Data

Bloom *et al.* selected an initial random sample of 66 firms from the set of textile firms that had between 100 and 1,000 employees based in two towns near Mumbai. These 66 firms were contacted and offered free consulting services, and 17 agreed to participate in the consulting program. On average, the firms have 270 employees and revenues of $7.5 million per year. The data include 22 plants owned by the 17 firms.

The electric distribution companies spread power holidays throughout the week in order to reduce load on all days. Fourteen of the plants are in areas with power holidays on Fridays, while the remainder have holidays spread throughout the week. Appendix Table A1 presents information on plant locations and scheduled power holidays, while Appendix Table A2 summarizes the data.

The plants typically operate continuously: 24 hours a day, every day. Before each power holiday, however, plant managers can choose to reduce output or fully shut down for all or part of the day. As they do this in advance, they can inform some or all workers that they need not come to work. Production workers are on 12 hour shifts, and they are paid if and only if they are called in. In the context of our model, labor is thus fully flexible. Similarly, materials such as yarn are fully flexible: they can be stored if the plant does not operate.

Physical output Q is measured in "picks," where one pick is a single rotation of the weaving shuttle. Figure 6 presents the distributions of output at an example plant for each of the seven days of the week. The dashed line is output on Fridays, when the plant has power holidays, while the solid lines represent output on each of the other six days. The distributions are very similar for the six non-holidays, with a mode of about 12,000 picks per day. On most power holidays, output does not appear to differ. On some power holidays, however, output is roughly half, as the plant cancels one 12-hour shift. Output drops to zero on a small share of power holidays.

4.2 Empirical Estimates

4.2.1 Differences in Output on Power Holidays

We now estimate the reduction in output on power holidays. We observe physical output $Q_{i\tau}$ for each plant i on each day τ. $\widetilde{Q}_{i\tau}$ is output normalized by plant i's sample average production:

$\widetilde{Q}_{i\tau} = Q_{i\tau}/\overline{Q}_i$.[12] ϕ_i is a vector of 22 plant indicators, while θ_τ represents 1339 day-of-sample indicators. The estimating equation is:

$$\widetilde{Q}_{i\tau} = \rho \cdot 1(\text{Power Holiday}_{i\tau}) + \phi_i + \theta_\tau + \varepsilon_{i\tau} \tag{21}$$

Table 3 presents the results of this regression, with standard errors clustered by plant. Column 1 presents the above specification, except without the day-of-sample controls θ_τ. Column 2 is the exact specification above. These columns show that average production is 7.4 to 9.7 percent lower on power holidays.

According to the consulting team that worked with the factories, there is typically no grid electricity available for the 24 hours of the scheduled power holiday. However, there are both "type I" and "type II" classification errors. During the winter months when electricity demand is lower, there may be unscheduled power availability during scheduled power holidays. During all months, especially the summer months when electricity demand is higher, there can also be unscheduled power cuts on any day.

Column 3 measures this by estimating how production on power holidays and non-holidays varies with the monthly CEA shortage estimate for the state in which each plant is located. On non-power holidays, output is not statistically significantly associated with shortages. This is consistent with the fact that when power cuts occur on non-power holidays, plants typically self-generate instead of shutting down, as labor input for the day has already been fixed.

The coefficient on 1(Power Holiday), which in Column 3 represents the intercept in months when the CEA estimates zero shortages, is statistically zero. Output on power holidays decreases by 0.6 percentage points as shortages increase by 1 percentage point. Column 4 includes power holiday-by-month controls, to control for any time-varying factors such as demand or diesel prices. This does not change the results relative to column 3. These results suggest that the managers have some ability to predict when there will be more electricity on a scheduled power holiday, and when they expect more electricity they call in more workers and produce more. This highlights that the effects of scheduled power holidays on production depends on the severity of the underlying shortage that the holidays are designed to address.[13]

4.2.2 Input Cost Effect

We now estimate the input cost effect: the increase in electricity costs when power holidays force a switch from grid electricity to self-generated electricity. We observe total energy costs - electricity

[12] We normalize because production varies substantially within and between the different plants, and we do not want the coefficient estimates to be driven by outliers. This normalization is preferable to using $\ln(Q_{i\tau})$ or $\ln(Q_{i\tau}+1)$ as the dependent variable because $Q_{i\tau}=0$ on some days, and this large variation makes it incorrect to interpret the coefficients as approximately reflecting percent changes in Q.

[13] Our model does not capture potential effects of electricity shortages on output quality, and so we would understate productivity losses if shortages affect output quality along with quantity. However, we have tested this using two measures of output quality, finding no statistical difference on power holidays vs. non-holidays.

plus generator fuel - at the monthly level, not for each day. After conditioning on total monthly production, the relationship between energy costs and the share of production on power holidays tells us the incremental marginal cost of self-generated electricity.

Let G_{im} represent the share of output produced on power holidays at plant i during month m. We denote $\widetilde{F}_{im}$ and $\widetilde{Q}_{im}$ as plant i's energy costs and output for month m, normalized by the plant's average monthly values. Analogous to above, ϕ_i is a plant fixed effect, and θ_m is a full set of month-of-sample dummies. δ_{im} denotes the CEA's estimated shortage in plant i's state in month m. The regression is:

$$\widetilde{F}_{im} = \eta_1 G_{im} + \eta_2 \widetilde{Q}_{im} + \eta_3 \delta_{im} + \phi_i + \theta_m + \varepsilon_{im} \tag{22}$$

Table 4 presents the results, again with standard errors clustered by plant. Column 1 shows the unconditional correlation between energy costs and power holiday output share, while Columns 2-4 progressively add controls for month-of-sample, normalized output, and shortages. The estimates imply that shifting 100 percent of production from non-power holidays to power holidays would increase monthly energy costs by 61 to 81 percent. This is closely consistent with the World Bank Enterprise Survey data in Table 2, which suggest that grid electricity costs an average of Rs 4.5 per kilowatt-hour, while generator electricity costs Rs 7 per kilowatt-hour, or 56 percent more.

4.3 Estimating Losses from Power Holidays

Table 5 uses the empirical estimates to calculate the input cost effect, output loss, and TFP reduction from power holidays at this set of plants. The top panel calculates the input cost effect. This is the proportion of electricity that is self-generated, which we assume to be equal to the share of production on power holidays (G), multiplied by the percentage increase in energy costs when self-generating ($\widehat{\eta}_1$) and by the sample median[14] energy revenue share. Power holidays increase input costs by 0.32 percent of revenues. This is effect is small, both because the electricity input cost share is small and because only one-ninth of production is on power holidays.

Estimating output losses requires the additional assumption from Section 3 that production is not substitutable across "days." If plants operating at less than full capacity can substitute production across days, they should produce when lower-cost grid electricity is available. In this case, the reductions in output associated with power holidays would not reflect a reduction in total output caused by power holidays - instead, they would represent inter-day shifting of the same amount of production. If plants do substitute production across days, estimated output losses assuming static production thus overstate the true output losses. In additional regressions, we see little evidence that intertemporal substitution causes the static model to overstate output losses.[15]

[14] We use median instead of mean to avoid bias due to potential recording error for revenues.

[15] More specifically, we test for inter-day substitution in two ways. First, we find that days of week just before power holidays do not have higher output, and the day immediately after a power holiday actually tends to have slightly but statistically significantly *lower* output, which suggests delays in restarting plants. Second, we exploit

Under the static production assumption, the middle panel estimates that power holidays reduce output by 1.1 percent. This is the product of 1/7 days that are power holidays and an estimated 7.4 percent output reduction on those days. While this is meaningful, it is small relative to average output losses estimated later for all Indian plants, because this sample of textile plants all own generators and thus often do not shut down during outages. To the extent that there is any undetected inter-day substitution, this only strengthens the qualitative conclusion that the output losses are small.

The bottom panel presents measured TFP losses under the assumption that at a given time on a power holiday, a plant has either shut down completely or is operating at the typical production rate for a non-power holiday. Under this assumption, there is no input variation effect, and measured TFP losses accrue through the shutdown effect. Under constant returns to scale and using the assumption that labor and materials are fully storable, Equation (20) for the measured TFP losses from the shutdown effect can be modified to obtain the measured TFP loss from power holidays:

$$\hat{\omega}_{i,t} - \omega_{i,t} = (\beta_K) \log(1 - \widetilde{\delta}) \tag{23}$$

In this equation, $\widetilde{\delta}$ is the percent of time shut down, which under our assumptions equals the 1.1 percent output loss. We take β_{K} from the ASI for textile plants (NIC 1987 code 230), which is slightly less than five percent. (Variable profits are relatively low in this industry.) The table shows that power holidays reduce measured TFP by 0.05 percent. Intuitively, this effect is so small because the plants rarely shut down, and when they do they have the flexibility to reduce the vast majority of their inputs.

While these plants provide a clear case study of the model, the effects of blackouts might be smaller here compared to the average plant in India, because labor and materials are both storable during planned power holidays, these plants all can self-generate instead of needing to always shut down, and textile plants are only moderately electricity intensive. To learn more about the broader Indian manufacturing sector, we now to turn to data from the Annual Survey of Industries.

5 Data and Empirical Strategy

5.1 Data

We use India's Annual Survey of Industries (ASI) for 1992 through 2010. The survey is split into two schemes: the census scheme and the sample scheme. In each year, the census scheme surveys all manufacturing establishments with over 100 workers, while the sample scheme surveys a rotating

the fact that it is more difficult to substitute production away from power holidays when already producing closer to capacity. Interday substitution would thus generate more output reduction during periods when plants are producing less. By contrast, we find that output reductions are larger when plants are producing *more*.

sample of one-third of establishments below that size.[16] Since 1998, the publicly-available version of the ASI includes establishment identifiers that are consistent across years. We also have an earlier version of ASI data that was acquired in India and is not publicly available, which contains consistent establishment identifiers for years before 1998. Combining these two datasets gives us an establishment-level panel for our entire sample.

The ASI is comparable to manufacturing surveys in the United States and other countries. It contains several modules, covering value of fixed capital stock and inventories, loans and cash flow information, cost and quantities of labor, materials, fuels, and other inputs, value of output, and other occasional questions. The reporting period is the Indian fiscal year, which is April 1 through March 31; when we refer to an individual survey year, we refer to the calendar year when the fiscal year begins. All financial amounts are deflated to constant 2004 Rupees. Appendix A gives more detail on the ASI data preparation and cleaning.

Table 6 gives descriptive statistics for the full ASI dataset. There are 616,129 plant-by-year observations. The median plant employs 34 people and has gross revenues of 20 million Rupees, or slightly less than $500,000. One of the benefits of the ASI over other manufacturing datasets from India and other countries is that we observe the physical quantity of each plant's total electricity purchases and self-generation in each year. The sum of these two variables, minus reported sales of electricity, yields Electricity Consumed. Self-Generation Share equals Electricity Self-Generated divided by Electricity Consumed. Energy Revenue Share is the value of electricity and fuels purchased divided by revenues.

The mean plant uses 0.013 kWh per Rupee of revenues, which equals 5.7 percent of revenues at typical electricity prices of 4.5 Rupees/kWh. 1(Self-Generator) is an indicator variable for whether a plant self-generates electricity in at least one year. Approximately 54 percent of plants ever self-generate: 42 percent of "Small" plants with fewer than 100 workers and 77 percent of "Large" plants with 100 or more workers. This is closely consistent with the World Bank Enterprise Survey data in Table 2, in which 46 percent of Small plants and 83 percent of Large plants report owning or sharing a generator. We combine the ASI with the state-by-year electricity market and weather data summarized in Table 1.

5.2 Estimation Strategy

Define Y_{ijst} as an outcome at plant i in industry j in state s in year t. Our primary specifications focus on four outcomes: self-generation share, energy revenue share, revenues, and productivity.

[16]The sampling rules have changed somewhat over time. The census sector, from which 100 percent of factories are sampled, was factories with 50 or more workers (100 or more if without electric power) until 1986-1987, 100 or more workers between that year and 1996-1997, 200 or more workers from then until 2003-2004, and then 100 or more workers since then. One-fifth of smaller factories were sampled since 2004-2005. The selection was done as a rotating panel such that plants are surveyed approximately once every third (or fifth) year, subject to constraints that sufficient numbers of factories had to be sampled to assure representativeness at the state and industry level (MOSPI 2009).

We use a difference estimator for our primary specifications, although we present robustness checks using the fixed effects estimator. The difference estimator is conceptually preferable because it identifies coefficients using shorter-term variation, consistent with our focus on identifying the short-run effects of shortages.

Because of non-response and the ASI's irregular sampling procedure, the data form an unbalanced panel with plants observed at irregular intervals. Sixty percent of intervals are one-year, while 91 percent are five years or less. Let Δ_i denote the difference operator, and define δ_{st} as electricity shortage in state s in year t, ranging from zero to one. The variable $\Delta_i \delta_{st}$ is the difference between the shortage in year t and the shortage in the year of plant i's previous observation. We include indicators θ_{it} for each "year difference," by which we mean the initial and final year combination for each differenced observation. For example, there is a θ_{it} indicator that takes value 1 for all plants observed in 2008 whose preceding observation was in 2005, and another θ_{it} indicator for all plants observed in 2008 whose preceding observation was 2004, etc. The variables μ_{jt} and ψ_s are vectors of indicators for two-digit industry-by-year and state, respectively.

The estimating equation is:

$$\Delta_i \ln Y_{ijst} = \rho \Delta_i \delta_{st} + \mu_{jt} + \theta_{it} + \psi_s + \varepsilon_{ijst} \tag{24}$$

All observations are weighted equally in the empirical estimates. This increases power, and because the difference estimator drops plants observed only once, the estimates are already unrepresentative. The simulations in Section 7 do use the ASI sampling weights to construct estimates that are representative of registered plants nationwide. Standard errors are robust and clustered at the level of variation in the year difference $\Delta_i \delta_{st}$. [17]

In the model, electricity shortages affect firms only through input availability: demand is unaffected by shortages. This would reflect the case in which manufacturers sell into national or international markets. In reality, many manufacturers sell to customers in the same state whose demand might also be affected by shortages. Thus, our empirical estimates capture effects of shortages through both input availability and demand.

These empirical estimates using annual data can capture additional effects not contemplated in the model in Sections 3 and 7. For example, if plants substitute production across days in response to outages, our empirical estimates capture the net effect of outages on output and other variables

[17] Sample sizes will differ from the observation counts in Table 6 for precisely four reasons. First, the difference estimator drops the approximately 107,000 plants observed only once. Second, we exclude "within-plant outliers": observations of logged inputs or outputs that are flagged because they differ from both preceding and subsequent observations by more than 3.5. A one-time annual jump of 3.5 natural logs is almost certainly a reporting error, although robustness checks in Appendix B.3 show that the estimates are not sensitive to either tightening or eliminating this restriction. Third, the states of Jharkhand, Chhattisgarh and Uttaranchal (now Uttarakhand) were established in 2001 from parts of Bihar, Madhya Pradesh and Uttar Pradesh, respectively. State-level measures of shortages and hydroelectric generation thus do not represent consistent areas before vs. after the splits, so we drop observations that are differences of years that span this split. Fourth, when examining self-generation share or energy revenue share as the outcome in our basic specifications, we exclude the 46 percent of plants that never self-generate electricity.

over the year. The estimates reflect the causal impact of annual variation in blackouts except in the unlikely event that plants substitute production across years. As another example, the empirical estimates let the data tell us whether materials and labor are semi-flexible or fully flexible.

5.3 Instruments and First Stages

Shortages could be econometrically endogenous to some outcomes, in particular output and productivity. For example, improvements in economic conditions within a state could increase productivity and output in manufacturing and other sectors, and the resulting increase in electricity demand could cause shortages. Furthermore, shortages could also be measured with error, causing attenuation bias.

To address this, we need an instrument that causes shortages to vary but is otherwise unrelated to manufacturing. We instrument using hydroelectricity generation, which varies from year to year due to the availability of water in reservoirs. H_{st} is the predicted share of state consumption that can be fulfilled by hydroelectric generation in state s in year t. As there are positive shocks to reservoir inflows, and thus to hydroelectricity generation, a larger share of state consumption can be fulfilled, and shortages should decrease. Intuitively, the denominator should be total electricity consumption in state s in year t, but this is determined partially by the extent of shortages. Thus, the denominator is national electricity consumption multiplied by the average ratio of state to national consumption for all years of the sample:

$$H_{st} = \frac{\text{Hydro Electricity Production}_{st}\text{ (GWh)}}{\begin{bmatrix}\text{National Electricity}\\ \text{Consumption}_t\text{ (GWh)}\end{bmatrix} \cdot \begin{bmatrix}\text{Average Ratio of State}\\ \text{to National Consumption}_s\end{bmatrix}} \tag{25}$$

Figure 7 illustrates the cross-state variation in the importance of hydroelectricity. While some hydro-heavy states are small mountainous regions such as Himachal Pradesh, Meghalaya, and Uttaranchal, other states such as Andhra Pradesh, Karnataka, Kerala, Orissa, and Punjab are both large and rely heavily on hydroelectricity. Figure 8 hydroelectricity generation over the study period for these states. In essence, the instrument is the product of the cross-section variation in Figure 7 with the time series variation in Figure 8.

Karnataka, a large state in southwest India, is the country's largest producer of hydroelectricity. Figure 9 plots the hydro instrument and shortages for Karnataka over the study period. The two variables are highly negatively correlated: more hydro generation reduces shortages.

Column 1 of Table 7 presents an analogue to the first stage using only first-differenced state-level data. Specifically, we regress the change in shortage on the change in the hydro instrument, controlling for state and year fixed effects. A one percentage point increase in hydro production relative to predicted demand decreases shortages by 0.048 percentage points. If a state's own hydro plants were its only source of electricity, this coefficient should be one. In practice, states offset the loss of hydro production by increasing generation from coal and other sources and by importing from other states.

The exclusion restriction is that changes in hydroelectricity production are associated with changes in manufacturing outcomes only through their effects on electricity shortages. In theory, there are at least two reasons why this identifying assumption might be violated. First, there could be reverse causality: hydroelectricity generation could respond in equilibrium to changes in electricity demand associated with manufacturing outcomes. However, the marginal cost of hydroelectricity production is relatively low, and annual production is constrained by the amount of water available behind reservoirs. By contrast, the exclusion restriction would be violated for production technologies such as coal power plants that have higher marginal costs, because their output is determined in equilibrium with demand.

To substantiate this, we gathered data from the Central Electricity Authority on inflows into 22 large reservoirs. For each state with a reservoir, we regressed annual hydroelectricity generation on inflows and construct the fitted values. Figure 10 plots predicted and actual hydro generation; theR^2 is 0.86. While the R^2 should not be 1 because the data include reservoirs that supply only 40 percent of India's hydroelectric generating capacity, the very high R^2 indicates that inflows are the primary determinants of hydroelectric production. Note that it is not possible to directly use inflows as our instrument because only 2/5 of states that have positive hydro generation have reservoirs in the inflows data.

The second reason why the identifying assumption might be violated would be if rainfall or some other third variable influences both hydroelectricity generation and manufacturing productivity or input or output prices. To address this, we can simply control for rainfall in our regressions, along with cooling degrees, which are correlated with rainfall and may affect agriculture. Although rainfall is associated with the hydro instrument, Column 2 of Table 7 shows that conditioning on rainfall and cooling degrees has very little impact on the state-level estimates aside from increasing the standard error. By contrast, Column 3 shows that rainfall is associated with agricultural output, while there is a positive but not statistically significant association between the instrument and agricultural output.

Columns 4 and 5 of Table 7 present a placebo test that provides even more direct support for the exclusion restriction. For an instrument to be valid, it needs to affect electricity supply but should not be associated with demand. To test this, we exploit the fact that the CEA reports the two components of shortages: assessed quantity demanded at current prices as well as the actual quantity supplied. Column 4 shows that the instrument is associated with quantity supplied, but column 5 shows that it is not associated with assessed demand. It is difficult to conceive of a story under which the exclusion restriction is violated but the instrument is not associated with electricity demand.

6 Empirical Results

6.1 First Stages

Table A3 in Appendix B.2 presents first stage estimates using microdata. In theory, the coefficient estimates might differ from the state-level results in Table 7 because the microdata weights states with more plants more heavily and because the microdata includes one-year and multi-year differences instead of only one-year differences. In practice, the microdata coefficients are slightly larger in absolute value but roughly comparable, ranging from -0.100 to -0.139. The instruments are powerful: the cluster and heteroskedasticity-robust Angrist-Pischke F-statistics range from 39 to 53.[18] 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.

Appendix Tables A4 and A5 present first stages for the alternative specifications in the upcoming section that potentially have the least power. These two tables respectively consider the sample when self-generation share is the outcome variable, which is the smallest sample, and when log output is the outcome variable, which has the smallest F-stat in Appendix Table Table A3. When conditioning on rainfall and cooling degrees, including only one-year differences, or testing interactions with shortages, the smallest F-statistic is 15.52. When clustering by state instead of by state-by-year difference, the F-statistics are 12.52 and 7.51 for self-generation share and log output, respectively.

6.2 Regression Results

Table 8presents results of Equation (24) for four different outcomes: self-generation share, natural log of energy revenue share, natural log of revenues, and natural log of productivity. Panels A and B present OLS and instrumental variables results, respectively. The IV estimates are very reasonable. Columns 1 and 2 test for impacts on energy input within the 54 percent of plants that ever self-generate. Column 1 shows that a one percentage point increase in shortages, which would increase the shortage variable from (for example) 0.1 to 0.11, causes a 0.57 percentage point increase in the share of self-generated electricity. If shortages affected manufacturers and all other consumers equally and manufacturing electricity demand were fully inelastic, this coefficient should be 1. In reality, state electricity boards may impose more or less of the marginal shortage on manufacturers instead of residential and agricultural consumers, and when manufacturers are faced with shortages, they do not make up for them one-for-one with self-generation. Column 2 shows that a one percentage point increase in shortages causes a 0.64 percentage point increase in energy revenue share.

[18]The Angrist-Pischke F-statistics are identical to the Kleibergen-Paap F-statistics when there is one endogenous variable. The Angrist-Pischke F-statistics are more appropriate in the parts of Appendix Tables A4 and A5 that test for weak identification of individual endogenous regressors in regressions with multiple endogenous regressors.

Either of columns 1 and 2 can be used to derive an estimate of the input cost effect for plants that self-generate. If $p^{E,S} - p^{E,G}$=2.5 Rs/kWh (from the World Bank Enterprise Survey) and the mean electric intensity is 0.013 kWh/Rupee (from Table 6), a one percentage point increase in shortages translates to a 1%×0.57×2.5×0.013≈0.018 percent unit cost increase. In other words, a one percentage point increase in shortages increases self-generation by 0.57 percentage points, which increases average electricity costs by 0.57%×2.5 Rs/kWh≈0.0142 Rs/kWh, which increases total unit costs by 0.0142 Rs/kWh*0.013 kWh/Rs≈0.018 percent of revenues. Similarly, using the fact that the mean energy revenue share is 0.11, the point estimate in Column 2 suggest that a one percentage point increase in shortages increases energy input costs by 0.64%*0.11≈0.07 percent of revenues. While these two estimates differ slightly, both imply that the input cost increase imposed on plants with generators is relatively small.

The IV estimates in column 3 show that a one percentage point increase in shortages causes a 0.68 percent decrease in revenues. Hypothetically, if no plants self-generate and there were no shutdown tax effect (in which firms reduce semi-flexible inputs in response to shortages), this coefficient would be one. In reality, self-generation reduces the revenue loss for plants with generators. If firms can foresee and respond to changes in shortages driven by hydro generation, this would be offset by the fact that both self-generators and non-generators reduce semi-flexible inputs through the output tax effect and shutdown tax effect, respectively.

Column 4 shows that shortages have statistically zero effect on TFP. The 90 percent confidence interval bounds the TFP losses from a one percentage point increase in shortages at no more than a 0.29 percent decrease in TFP. In our model, TFP losses should be much smaller than output losses because the primary cause of TFP loss is waste of inputs that are not fully flexible, and most inputs are fully flexible - the average input cost share for materials across all plants is 70 percent. Thus, the fact that TFP losses are too small to detect is fully consistent with our model.

The OLS estimates are statistically and economically different from the IV estimates, and the direction suggest two forms of bias. With self-generation share in Column 1, we expect less omitted variables bias in OLS. The fact that the IV estimates are substantially larger than OLS suggest that the instrument corrects measurement error in the shortage variable. By contrast, with output and TFP in Columns 3 and 4, we expect potential upward bias in OLS, because economic growth can cause shortages. Indeed, the OLS coefficients are biased upwards from the IV coefficients, and TFP actually appears to be positively associated with shortages. This shows the importance of using instrumental variables: without the IV, the econometrician might erroneously conclude that shortages cause TFP to *increase*.

6.2.1 Robustness Checks and Fixed Effects Estimates

Appendix B.3 shows that the estimates in Table 8 are remarkably robust. None of the estimates differs statistically or loses statistical significance when omitting the industry-by-year controls μ_{jy}, eliminating or tightening the flags for within-plant outliers, or controlling for rainfall and cooling degrees. When using only one-year differences, this focuses estimates on larger census scheme plants whose output is less affected by shortages and also reduces the sample size. This slightly reduces the point estimate of effect on output and increases the standard error; the resulting coefficient is statistically indistinguishable from both the base case estimate and from zero. Clustering at the state level instead of state-by-year difference increases the standard errors slightly but does not affect statistical significance. Appendix Table A13 shows that results are qualitatively similar under five different approaches to calculating production functions and TFP.

Appendix B.4 estimates an analogue to Equation (24) using fixed effects instead of differences, including state-specific linear trends and clustering by state to address potential serial correlation in errors. The results are remarkably similar to Table 8, and none of the IV estimates differ statistically. However, the standard errors are slightly wider, and the first-stage F-statistics are smaller. Furthermore, although excluding state-specific linear trends does not affect the non-IV estimates, the IV first stages have no power when excluding the state-specific linear trends. The reason for this is suggested in Figure 9: the share of hydro in total electricity production has decreased over time in Karnataka and other states, so while annual changes in the hydro instrument are negatively associated with changes in shortages, levels of the instrument are not. Because the level of the hydro generation share decrease is mechanically larger in states with more hydro production, the year indicators do not properly control for this in the fixed effects estimator. These results give a practical reason why the difference estimator is slightly preferred to fixed effects.

6.3 Moderators of Shortage Effects

The model in Section 3 generates predictions for how shortages should differentially affect different types of plants. Electricity-intensive industries should be more likely to shut down instead of self-generate during shortages, meaning that revenues and TFP should drop more. Furthermore, shortages should have much smaller effects on revenues and TFP for plants that self-generate. Table 9 interacts the change in shortages with indicators for self-generation and whether the plant's industry is above-median electricity intensity; the regressions also include lower-order interactions with θ_{iy} and ψ_s.

We fail to reject that more electric-intensive plants change self-generation the same amount in response to shortages, although their energy revenue share increases more. Shortages reduce output more for electric intensive plants and reduce output less for self-generators. However, the standard errors are wider, and coefficient magnitudes should be interpreted with caution. Precision would be further reduced if we cut the data more finely or studied individual industries in isolation.

Table 10 tests for effects on other outcomes. The point estimate in column 1 suggests that plants reduce labor input in response to shortages, but the effect is not close to being statistically significant. Column 2 shows that a ten percentage point increase in shortages is associated with an 8.53 percent decrease in materials input. Column 3 shows that shortages decrease the labor to materials ratio, consistent with columns 1 and 2. These estimates provide support for our modeling assumption in Section 7 that materials are fully flexible, while labor is not. Column 4 tests for effects on fuel revenue share, where fuels equal total energy net of electricity. The effect should be and is statistically larger than the effect on energy revenue share, because the latter includes an increase in fuel input costs but a decrease in electricity purchases. Column 5 shows that shortages do not statistically affect electricity intensity. In reality, there should be some small effect, consistent with the results from Fisher-Vanden, Mansur, and Wang (2012) for Chinese manufacturers. Our standard errors rule out that a ten percent increase in shortages causes more than a 0.0005 kWh/Rupee decrease in electricity intensity. This is about 38 percent of the median, which is 0.013 kWh per Rupee. Although this could be economically meaningful, it provides some statistical support for our model's simplifying assumption that λ is exogenous.

7 Simulations

In this section we quantify the welfare loss from missing power supply using the structure provided by our model described in Section 3. We need additional information from the ASI to identify parameters from the model. Thus, one can think of our simulation exercise as working out the quantitative implications of our model, using the distribution of parameters estimated from the ASI.

7.1 Calibrated Parameters

Table A20 presents the parameters that we use to calibrate the model. We use the ASI data, presented in Section 5.1, as the universe of plants that we make a prediction for. We apply the approach described in Section 3.5 to recover plant-level production function coefficient β_l, β_k, and β_m, from the input cost shares of non-self generators according to equations (11) and (12). We estimate γ using the electricity cost share for plants that do not generate any power themselves (non-generators). The data on the fraction of the year in which there are shortages δ, comes from our own collection efforts described in Section 2.1.

Finally, we need to make an assumption on the relative cost of electric power bought from the electric grid ($p^{E,G}$), versus power that is self-generated ($p^{E,S}$). We use the responses to the World Bank Enterprise Survey to the question on the cost of grid power versus self-generated power presented previously in Table 2. Only the ratio between these two price matters, and we assume that self-generated power is 55% more expensive than grid power.

7.2 Predictions

To get a handle on the effect of power shortages on welfare, we compute the loss in output and measured TFP from these shortages. Specifically, we compute the predicted output under the observed shortages, and compare these to the predicted output without any shortages; i.e., $\delta = 0$.

Table 11 presents results from this exercise for 2005. However, these prediction are fairly similar across any year from 1992 to 2010, as the average shortage across all of India has not moved very much over times, and ranges between 6.4 and 11.1 percent from 1992 to 2010. In Column I, we present results using the assumptions outlined in the paper, while Columns II, III, and IV, we show results using alternative assumptions to investigate how robust our results are. Specially, in Column II and III, we use an elasticity of demand of $\epsilon = -4$ and $\epsilon = -20$, respectively, which makes firms both less and more responsive to changes in their costs.[19] In Column IV, we investigate whether a significantly higher price of self-generated power, twice as high as our estimate from the World Bank, changes how responsive firms are to power shortages.

In 2005, there are, on average, shortages 7.1 percent of the time. We predict an output loss of 7.1 percent due to these shortages. However, this loss in output is starkly different for plants that have a generator, versus those that do not. Plants with a generator essentially have no loss in output due to power outages, losing only 0.7%. To understand this small effect, which for generating plants we previous called the input cost effect, remember that electricity purchases are only 5 percent of revenues. Thus, for plants with a generator, a power outage is equivalent to a 55 percent increase in the price of power, or a 2.8 percent increase in costs, the 7.1 percent of the year shortages occur. This an input cost effect of 0.2 percent, assuming that plants do not reduce their output during blackouts. This tiny increase in costs during power outages rationalize the slight effect of shortages on output, that we found in Section 4 for large textile producers, all of whom have backup generators.

However, for plants that do not have a generators, output would fall by 10.3 percent due to the shutdown effect. Notice that the shutdown effect is larger than the 7.1 percent of the time where the plant cannot operate due to shortages. This is because the effect of shortages is amplified because of the non-storability of labor. Anticipating a power outage, firms hire less labor than they would otherwise, and this enhances the effect of shortages. Given the large losses for plants without generators, this begs the question of why these plants do not purchase generators in the first place, which we turn to in Section 7.3.

Next, we show the predicted losses in measured TFP due to shortages. Echoing the stark contrast between generators and non-generators we found for output, the loss in TFP is almost zero for plants that self-generate, versus a 2.9 percent TFP loss for Non-Generators. Thus, the predicted average TFP loss is only 1.9 percent. For plants that have generators, the loss in TFP is driven by the input variation effect: with a concave production function, it is less efficient to produce

[19] One can think of the elasticity of demand as a tuning parameter that alters the concavity of the profit function. Thus, higher elasticity implies a smaller response to changes in prices and productivity, and a higher markup

with a variable bundle of inputs. Since plants do not reduce their production very much during shortages if they have generators, the input variation effect is necessarily quite small. However, the shutdown effect on TFP, the loss in non-storable inputs during power outages such as capital and labor, has larger effects on productivity. If materials were not storable, then a 7.1 shortage would directly translate into a 7.1 loss in TFP. However, firms that shut down during power outages do not lose their material inputs, only their labor and capital. Given that labor and capital have a 30 percent share of input costs, the shutdown effect on TFP is only about a third of 7.1 percent of the time the plant is closed due to power outages.

7.3 Decision to Purchase a Generator

Most of the loss in output is at plants who do not have a generator. This begs the question of why some plants choose not to buy them. To get at this issue, we look a the plant's decision to purchase generators, and compute the cost of a generator at which a plant would break even from purchasing it.

First, we need to impute the type of generator that a plant would buy, since we do not observe generator capacities in any our data. To do this, we need to transform a firm's total power usage E in kWh, into the generator it would need to purchase in order to be fully backed-up during outages. Assuming that a plant uses power continuously 6 hours a day, 365 days a year, the median plant would require a generator with about 500 KW of capacity.

Second, we compute how much a firm would increase its profits if the plant purchases a generator. These profits are computed using equation (4) along with the optimized values for labor and materials, and hence, are based on our model. We find that profits would increase by an average 9.1 percent upon the purchase of a generator – excluding the cost of the generator itself. Again, notice that the firm's increase in profits from purchasing a generator are greater than the 7.1 percent outage frequency.

Third, we compute the cost of a generator at which a plant would break even from purchasing it. For the median plant in the ASI, it would break even at a cost of 134 Rupees per KW of capacity. However, there is considerable heterogeneity of this break even cost, given that plants have different endowment of capital, different productivity levels, and different electricity intensities in their production process. The standard deviation of the break even cost is almost 4,000 Rupees per KW of capacity.

Fourth, we compute the share of plants that would choose to purchase a generator. To do this, we need information on the cost of a generator, and for how many years a generator will last. Through conversations with sellers of generators in the United States, we reckon that a generator lasts roughly ten years. Our research efforts contacting sellers of generators in India give us a range of prices for generators, with a 25th and 75th percentile of 4,700 and 7,500 Rupees per KW of capacity, and a median of of 5,500 Rupees per KW. There are economies of scale in generation of electricity. Indeed, without these economies of scale, it would make sense for all electricity to be produced at the household or plant-level. A small generator of 60 KW costs around 13,000 Rupees per KW, while a larger generator of 1000 KW costs around 8,000 Rupees per KW. This makes purchasing a generator cheaper for larger plants.

Table 11 shows that if a generator costs 5,500 Rupees per KW, then 37 percent of plants would purchase generators. This is very close to the 37 percent of plants that report producing power in the ASI. Thus, we can rationalize the uptake of generators in the ASI data: even if there are large losses in output for plants that choose not purchase a generator, these generators are expensive enough to explain this decision. Note that the fraction of plants that purchase generators is fairly robust to the assumptions that we make on generator costs. For instance, at a cost of 4,700 R per KW, 42 percent buy a generator, while at a cost of 7,500 R per KW, 34 percent do. This is because of the large heterogeneity in the return on purchasing a generator, and the elasticity of generator adoption with respect to the price of a generator is -0.14 .[20]

7.4 Firm Size Distribution

An important question is the extent to which the shortages affect firms differently. Hsieh and Klenow (2012) propose that electricity shortages combined with differential access to grid electricity could be an important factor benefiting large establishments. Here we focus on a different mechanism for differential impacts by firm size: economies of scale in self-generation.

Table 12 presents similar statistics to Table 11, but broken down by whether a plant is large or small, as measured by having more or less than 50 employees, and whether a plant is in an electricity intensive industry, as measured by having an electricity revenue share that is either above or below the median for the ASI. For the size comparison, we also give a small plant a cost of a generator of 13,000 R per KW, while for a large plant, we assign these a generator cost of 8,000 R. per KW of capacity, given the scale economies of generation.

First, we find that the output loss for a small plant would be 7.8 percent, while it is 5.4 percent for a large plant. The reason for these differences is that in the ASI, in 2005, 23 percent of small plants have generators, while 55 percent of large plant have them. Since the effects of power outages are more pronounced for plants without generators, smaller plants will be more affected.

Second, using the assumed economies of scale of power generation that we previously discussed, large plants will be more likely to purchase generators, at a rate of 37 percent, versus small plants that will only purchase generators at rate of 26 percent.

Third, turning to the role of electric intensity, we would predict that plants that are not electricity intensive would be more likely to purchase a generator, at a rate of 46 percent for plants in industries below the median electricity revenue share, versus 19 percent for industries above the median electricity revenue share. To understand this effect, remember that our model assumed that

[20] For a point of reference on the cost of a generator, a quick search on Amazon.com yields, for instance, a 4KW diesel generator for $ 336. DuroStar DS4000S 4,000 Watt 7.0 HP OHV 4-Cycle Gas Powered Portable Generator. Price: $336.92, accessed December 23, 2013.

electricity is an essential input: it is impossible to produce without it. Therefore is much cheaper for an industry for which electricity is only 1 percent of revenues to buy backup generators, versus an electricity intensive industry such as steel, for which electricity represents 10 percent of revenue.

7.5 Robustness

A natural question has to do with the robustness to alternate assumptions in our model, and the plausibility of the model's predictions. This is particularly important given the paucity of papers estimating the effect of power outages.

We assumed that firms face a CES demand curve for their products, with an elasticity of demand of $\epsilon = -10$, which corresponds to firms setting a 10 percent markup for their products. This elasticity is critical for our model, since the curvature of the firms profit function affects both how much a firm will respond to cost shocks, and also how profitable it is. Suppose instead that we had assumed an elasticity of $\epsilon = -4$, such as is done in Bloom (2009) and Asker, Collard-Wexler, and De Loecker (2013). We would obtain an output loss of 5.0 percent instead of 7.1 percent, mainly because non-generators reduce their output by 8.2 percent rather than 10.3 percent. It makes sense that if we endow firms with a flatter profit function, this will amplify their response to shortages. However, the share of plants that would purchase a generator increases from 28 to 36 percent from Column I to II. A less elastic demand curve raises the markup that a firm earns: it makes producing more profitable. Thus the output loss is more costly with a higher markup, and firms are more likely to purchase a generator. Likewise, if we assumed a more elastic demand curve, with an elasticity of $\epsilon = -20$, thus a 5 percent markup, we would find a larger output response, at 9.4 percent versus 7.1 percent, and fewer plants purchasing generators. [21]

Second, we used data from the World Bank to motivate the assumption that self-generated power is 55 percent more expensive than grid generated power. Suppose instead that self-generated power is *four* times more expensive as grid supplied power. We find an output loss of 7.9 percent instead of 7.1 percent, in large part because plants who generate reduce their output by a little more than 3.2 percent, while plants that do not self-generate, will clearly not respond to changes in the price of generated power.[22]

[21] There are few papers that look at the elasticity of output with respect to power shortages. Most prominently, Davis, Grim and Haltiwanger (2008) investigate the elasticity of output with respect to the price of electricity in the United States. Since a power outage is – for generators – equivalent to an increase in the price of electricity, we can compare the elasticity of output with respect to price of power in our model to the one estimated in Davis, Grim and Haltiwanger (2008). We compute an elasticity of output with respect to the price of electricity of -0.2 when we assume $\epsilon = -10$, and -0.4 when we assume $\epsilon = -20$, while Davis, Grim and Haltiwanger (2008) find an elasticity of -0.6. So our model, if anything, under predicts the response of output to shortages for generators, but not by a large amount. On the other hand, Davis, Grim and Haltiwanger (2008)'s estimate is quite large, and somewhat difficult to explain given how small the share of electricity is as a share to total costs.

[22] We should stress that the results are sensitive to the measurement of generators versus non-generators. We find that almost half of plants in the ASI report producing no power whatsoever, and even for plants with over 500 employees, approximately 30 percent report no generation. If we overstate the fraction of plants that are non-generators, our welfare effects will also be overstated. However, the World Bank Enterprise Survey explicitly asks

7.6 Evaluating the Fit of the Model

A final test of our model and our instrumental variable strategy is to see whether our instrumental variable estimates and our model's predictions line up. Panel A of Table 13 presents in Column I the results in Table 11, and in Column II the predicted effects of a 7.1 percent shortage given the elasticities estimated in the second panel of Table 8. We find remarkably agreement between the estimates and our model. For instance, the IV estimates predict that shortages would reduce output by 4.8 percent, while our model has a prediction of 7.1 percent. These predictions are statistically indistinguishable with a 33 percent probability. Likewise, our estimates predict that generators and non-generators would reduce their output by -2.7 and 14.9 percent, while the model predicts 0.7 and 10.3 percent respectively. Again, these predictions are statistically indistinguishable with over a 19 percent confidence. Finally, our estimates show a 0.2 percent TFP loss, while the model predicts a 1.9 percent loss, but these estimates are also not statistically distinguishable at the 10 percent level.

Panel B of Table 13 goes a step further and adds our results in section 4 from our case study of textile plants, and from the World Bank Survey of firms discussed in section 2. For large textile plants, we estimated between a 0.9 and 0.7 percent reduction in output due to shortages. For a plant with a generator, our model predicts a drop of 0.7 percent. As well, for the World Bank Survey, plant owners reckoned that power cuts made them lose 7.8 percent output, which is fairly close to our prediction of 7.1 percent.

Overall, the close fit between our model and the estimates using various datasets and identification strategies gives us greater confidence that we have correctly identified the effects of shortages, but we also understand the mechanisms involved.

8 Conclusion

India's lack of reliable electricity supply provides a stark example of how poor infrastructure affects economic growth. We study the effects of shortages on manufacturing using archival data on shortages, previously-unavailable panel data from the Annual Survey of Industries, and a new instrument for shortages based on variation in hydro reservoir inflows. We augment this with a detailed case study of how textile plants in Bloom *et al.* (2013) respond to planned power holidays. We use a hybrid Leontief/Cobb-Douglas production function model to clarify the different ways in which input shortages affect firms and use simulations to confirm and extend the empirical results.

There are three main conclusions. First, electricity shortages are a large drag on Indian manufacturing, on the order of five percent of output. Second, however, electricity shortages affect productivity much less than output, and shortages alone do not explain a large share of the TFP gap between firms in developing vs. developed countries. Third, shortages have heterogeneous

plants if they own a generator, and we find a similar fraction of firms are flagged as owning generators in the ASI and in the World Bank data.

effects across plants with vs. without generators and with high vs. low electric intensity. Relatedly, because of economies of scale in self-generation, small plants are less likely to own generators, meaning that shortages have much stronger negative effects on small plants.

One way in which future work could extend this study is to include various kinds of dynamics, such as intertemporal substitution of production, investment decisions in generators, and entry and exit. Our study uses a static model, focusing on the effects of annual variation in shortages with fixed capital stock. However, because most of the policies available to address shortages would be unlikely to fully eliminate shortages for many years, this type of annual variation may identify the most policy-relevant effects.

References

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

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

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

[4] Asker, John, Allan Collard-Wexler, and Jan De Loecker (2013). "Productivity Volatility and the Misallocation of Resources in Developing Economies." Working Paper, New York University (July).

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

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

[7] Banerjee, Abhijit, and Esther Duflo (2010). "Giving Credit Where It Is Due." *Journal of Economic Perspectives*, Vol. 24, No. 3 (Summer), pages 61-80.

[8] Banerjee, Abhijit, and Esther Duflo (2012). "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." Working Paper, MIT.

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

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

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

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

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

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

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

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

[17] Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013). "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, Vol. 128, No. 1 (February), pages 1-51.

[18] Central Electricity Authority (2010). All India Electricity Statistics: General Review 2010. New Delhi, India: Central Electricity Authority.

[19] Central Electricity Authority (2012). "Load Generation Balance Report." http://www.cea.nic.in/reports/yearly/lgbr_report.pdf.

[20] Central Electricity Authority (2013). “Monthly Executive Summary Report.” http://www.cea.nic.in/reports/monthly/executive_rep/feb13/feb13.pdf

[21] Central Electricity Authority (2013a). “Planwise Capacity Addition.” http://www.cea.nic.in/reports/planning/plan_addition.pdf

[22] Central Statistical Organisation (2012). “Input-Output Transactions Table, 1993-94.”

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

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

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

[26] Steven Davis, Cheryl Grim, and John Haltiwanger (2008). "Productivity Dispersion and Input Prices: The Case of Electricity," Working Papers 08-33, Center for Economic Studies, U.S. Census Bureau.

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

[28] de Mel, Suresh, David McKenzie, and Christopher Woodruff (2008). "Returns to Capital in Microenterprises: Evidence from a Field Experiment.” *Quarterly Journal of Economics*, Vol. 123, No. 4, pages 1329-1372.

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

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

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

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

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

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

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

[36] Fisher-Vanden, Karen, Erin Mansur, and Qiong (Juliana) Wang (2012). “Costly Blackouts? Measuring Productivity and Environmental Effects of Electricity Shortages.” NBER Working Paper No. 17741 (January).

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

[38] Government of India (GOI) (2009, 2012). Measures for Operationalising Open Access in the Power Sector. New Delhi, India: Secretariat for the Committee on Infrastructure.

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

[40] Hsieh, Chang-Tai, and Peter Klenow (2012). "The Life Cycle of Plants in India and Mexico." Working Paper, University of Chicago (May).

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

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

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

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

[45] McRae, Shaun (2013). "Infrastructure Quality and the Subsidy Trap." *American Economic Review*, forthcoming.

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

[47] Office of the Economic Advisor (2013). "Wholesale Prices Index (WPI) Data." http://www.eaindustry.nic.in/

[48] Rauch, James (1995). "Bureaucracy, Infrastructure, and Economic Growth: Evidence from U.S. Cities During the Progressive Era." *American Economic Review*, Vol. 85, No. 4 (September), pages 968-979.

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

[50] Reserve Bank Of India (2012). "Handbook of Statistics on the Indian Economy." Tables 4, 5, 8 and 9. http://www.rbi.org.in/scripts/AnnualPublications.aspx?head=Handbook%20of%20Statistics%20on%20Indian%20Econ

[51] Restuccia, Diego, and Richard Rogerson (2008). "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics*, Vol. 11, pages 707-720.

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

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

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

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

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

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

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

[59] Straub, Stephane (2008). "Infrastructure and Development: A Critical Appraisal of the Macro-Level Literature." World Bank Policy Research Working Paper No. 4590 (April).

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

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

[62] Zuberi, James (2012). "Estimating the Cost of Power Outages for Large Scale Manufacturing Firms." Working Paper, University of California at Berkeley (November).

Tables

Table 1: **State-Level Data Summary Statistics**

Variable	Mean	Std. Dev.	Min.	Max.	N
Assessed Demand (TWh)	20.01	22.74	0	128.3	509
Quantity Supplied (TWh)	18.27	20.17	0	107.02	509
Shortage	0.07	0.07	0	0.36	507
Peak Shortage	0.11	0.1	0	0.5	507
Reservoir Inflows (Billion Cubic Meters)	5.19	13.73	0	115.98	570
Hydro Generation (TWh)	2.46	3.1	0	15.27	570
Total Electricity Sold (TWh)	12.86	15.14	0.05	87.53	543
Average Cooling Degrees (F), Base 65	12.32	3.3	2	18.94	543
Annual Rainfall (meters)	1.36	0.63	0.27	5.01	551

Notes: This table presents descriptive statistics for data that vary at the state-by-year level. The first six variables are from the Central Electricity Authority, while the temperature and rainfall data are from the National Climate Centre. Cooling degrees for day τ=max(0,Average Temperature$_\tau$(F)-65).

Table 2: **Power Cuts and Plant Size in the 2005 World Bank Enterprise Survey**

Plant Descriptions	**Small Plants**	**Large Plants**
Number of Plants Surveyed	1719	306
Number of Workers (Mean)	23	494
Gross sales, in million Rupees (Median)	5.7	172
Electricity Shortage Questions		
How [many times in 2005] did your establishment experience power outages or surges? (Mean)	132	99
Does your establishment own or share a generator? (Percent)	46	83
What percent of your electricity comes from the generator? (Mean)	10	17
What is the average cost [in Rs/kWh] for generator electricity? (Median)	7	7
What is the average cost [in Rs/kWh] for public grid electricity? (Median)	4.5	4.5
What were your percentage losses from power outages or surges? (Mean)	8.0	4.9
Electricity is the "biggest obstacle for operation/growth" (Percent)	34	26

Notes: “Small Plants” have less than 100 workers, while “Large Plants” have 100 or more workers.

Notes: Rupees are constant 2004 Rupees. 1(Census Scheme) takes value 1 for plants with more than 100 workers which are surveyed each year, and value 0 for Sample Scheme for smaller plants in the rotating panel.

Table 3: **Textile Output on Power Holidays**

Dependent Variable: Output	(1)	(2)	(3)	(4)
1(Power Holiday)	-0.097	-0.074	-0.006	0.025
	(0.025)***	(0.017)***	(0.022)	(0.027)
1(Power Holiday) x Shortage			-0.006	-0.006
			(0.003)**	(0.003)**
Shortage			0.001	0.001
			(0.004)	(0.004)
Number of Obs.	26,114	26,114	26,114	26,114
Number of Clusters	22	22	22	22
Day-of-Sample Controls	No	Yes	Yes	Yes
Power Holiday x Month Controls	No	No	No	Yes

Notes: This table presents estimates of Equation (21). The dependent variable for columns 1-4 is $\widetilde{Q}_{i\tau}$, plant i's production on day τ, normalized by plant i's average daily production. Robust standard errors, clustered by plant. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table 4: **Textile Monthly Energy Cost Regressions**

Dependent Variable: Energy Cost	(1)	(2)	(3)	(4)
Power Holiday Output Share	0.610	0.791	0.821	0.811
	(0.362)	(0.314)**	(0.335)**	(0.343)**
Normalized Output			0.237	0.232
			(0.167)	(0.167)
Shortage				-0.003
				(0.005)
N	307	307	307	307
Month-by-Year Controls	No	Yes	Yes	Yes

Notes: This table presents estimates of Equation (22). The dependent variable for columns 1-5 is $\widetilde{F}_{im}$, plant i's total energy cost in month m, normalized by plant i's average monthly energy cost. Robust standard errors, clustered by plant. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table 5: **Losses on Planned Power Holidays**

Input Cost Effect	
Mean share of output on power holidays (G)	0.11
Increase in energy cost share ($\widehat{\eta}_1$)	0.81
Median energy revenue share	0.026
Input cost increase (share of revenues)	0.0024

Output Loss	
Share of days that are power holidays	1/7
Output loss on power holidays ($\widehat{\rho}$)	0.074
Share of output lost	0.011

Shutdown Effect on Measured TFP	
Share of fixed inputs (capital) (β_K)	0.05
ln(TFP) change: $\beta_K ln(1 - Outputloss)$	-0.00053

Notes: This table presents estimates of textile plants' losses on planned power holidays, using empirical estimates from Tables 21 and 22.

Table 6: **Annual Survey of Industries Summary Statistics**

Variable	Mean	Std. Dev.	Min.	Max.	N
Revenues (million Rupees)	323.43	3693.23	0	788867.5	616129
Capital Stock (million Rupees)	126.87	1751.17	0	297370.25	612832
Total Persons Engaged	164.56	740.19	0	121007	577669
Materials Purchased (million Rupees)	210.12	2706.43	0	636136.94	609957
Fuels Purchased (million Rupees)	13.04	175.24	0	39359.95	576762
Electricity Purchased (million Rupees)	8.69	82.87	0	9935.30	561464
Electricity Consumed (GWh)	3.4	51.81	0	7356.86	599116
Electricity Purchased (GWh)	2.17	32.99	0	6544.51	599717
Electricity Self-Generated (GWh)	1.23	35.01	0	7147	598619
1(Self-Generator)	0.54	0.5	0	1	616129
Electric Intensity (kWh/Rs)	0.01	0.02	0	0.39	599116
Self-Generation Share	0.08	0.19	0	1	592914
Energy Revenue Share	0.11	0.16	0	3.23	598932
1(Census Scheme)	0.41	0.49	0	1	616129
Plant Number of Observations	5.86	4.97	1	19	616129

Table 7: **Assessing the Hydro Instrument**

	(1)	(2)	(3)	(4)	(5)
Outcome Variable:	Shortage	Shortage	ln(Agri Output)	ln(Quantity Supplied)	ln(Assessed Demand)
Δ Hydro	-0.048	-0.046	0.149	0.063	0.014
	(0.017)***	(0.025)*	(0.117)	(0.032)**	(0.026)
Δ ln(Rainfall)		-0.009	0.156		
		(0.009)	(0.059)**		
Δ Cooling Degrees		-0.002	-0.027		
		(0.003)	(0.020)		
R^2	0.14	0.15	0.27	0.17	0.22
N	469	454	398	469	469

Robust standard errors. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table 8: **Base Specifications**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.227	-0.035	0.020	0.095
	(0.023)***	(0.055)	(0.040)	(0.030)***
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	2,781	2,936	3,263	3,261

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.568	0.639	-0.682	-0.034
	(0.105)***	(0.232)***	(0.327)**	(0.155)
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	2,781	2,936	3,263	3,261

Notes: This table presents estimates of Equation (24). Panel B instruments for Shortage using the hydroelectric generation instrument. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity.

Table 9: **Instrumental Variables Estimates with Moderators**

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	-0.060	1.018	-1.979	0.075
	(0.043)	(0.822)	(0.976)**	(0.427)
Δ Shortage x Elec Intensive	0.124	1.108	-1.323	-0.107
	(0.084)	(0.450)**	(0.484)***	(0.304)
Δ Shortage x Self-Generator	0.571	-0.883	2.343	-0.121
	(0.107)***	(0.791)	(0.974)**	(0.417)
Number of Obs.	301,390	343,696	374,168	366,319
Number of Clusters	3,187	3,213	3,263	3,261

Notes: This table presents estimates of Equation (24), instrumenting for Shortage using the hydroelectric generation instrument. Electric Intensive is an indicator variable for being in an industry with above median electricity use per unit revenues. Regressions also include lower-order interactions of Electric Intensive and Self-Generator with year difference indicators θ_{iy} and state indicators ψ_s. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table 10: **Instrumental Variables Estimates for Additional Outcomes**

	(1)	(2)	(3)	(4)	(5)
Outcome Variable:	ln(Labor)	ln(Materials)	ln(Labor/ Materials)	ln(Fuel Rev Share)	Electric Intensity λ
Δ Shortage	-0.172	-0.853	-0.741	1.773	0.005
	(0.218)	(0.325)***	(0.259)***	(0.468)***	(0.006)
Number of Obs.	375,106	367,504	366,838	212,554	356,690
Number of Clusters	3,272	3,254	3,253	2,677	3,231

Notes: This table presents estimates of Equation (24) for additional outcomes, instrumenting for Shortage using the hydroelectric generation instrument. Electric Intensity for plant i in year t is the ratio of kWh of electricity consumed to revenues. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity.

Table 11: Predicted Losses from Electricity Shortages

	Baseline $\epsilon = -10$ (1)	Lower Elasticity $\epsilon = -4$ (2)	High Elasticity $\epsilon = -20$ (3)	Self-Generated Power 4 times more expensive (4)
Shortage	7.1%	7.1%	7.1%	7.1%
Output Loss				
Loss in Output	7.1%	5.0%	9.4%	7.9%
Loss in Output for Non-Generators	10.3%	8.2%	13.5%	10.3%
Loss in Output for Generators	0.7%	0.3%	1.4%	3.2%
TFP Loss				
TFP Loss	1.9%	2.2%	1.5%	2.0%
TFP Loss for Non-Generators	2.9%	4.6%	2.2%	2.9%
TFP Loss for Generators	0.1%	0.1%	0.1%	0.5%
Input Cost Effect	0.2%	0.2%	0.2%	1.1%
Generator Decision				
Profit Loss from No Generator	9.1%	6.0%	14%	10.1%
Median Break Even Cost of a Generator in Rupees per KW of Capacity	134	349	101	91
Share of Plants predicted to purchase a generator	27%	36%	26%	24%
Elasticity of Output with respect to the price of electricity	-0.20	-0.07	-0.39	-0.20

Note: Prediction from the ASI for 2005 using the model described in text. Weighted average refers to the average weighted by plant output. TFP defined as the residual of the sales generating production function using the approach described in section 3. Elasticity refers to the elasticity of the CES Demand Curve. 4 times more expensive self-generated power, sets the price of self-generated power, $p^{E,S}$ to 18 rupees per kWh, instead of 7 rupees per kWh. The share of plants purchasing a generator assumes a cost for a generator of 5,500 Rupees per KW of capacity, and this generator lasts for ten years.

Table 12: Differential Effects of Shortages

	Plant Size		Electricity Intensity	
	Large	Small	High	Low
Output Loss	5.4%	7.8%	6.8%	7.5%
TFP Loss	1.4%	2.1%	1.7%	2.2%
Profit Loss from No Generator	7.0%	9.9%	8.5%	9.8%
Generator Take-Up Rate	37%	26%	19%	46%

Note: Large Plants refer to plants with more than 100 employees, while small plants refers to plants with less than 100 employees. Electric Intensive and Not Electric Intensive refers to plants belonging to industries that are above are below the median electricity input share.

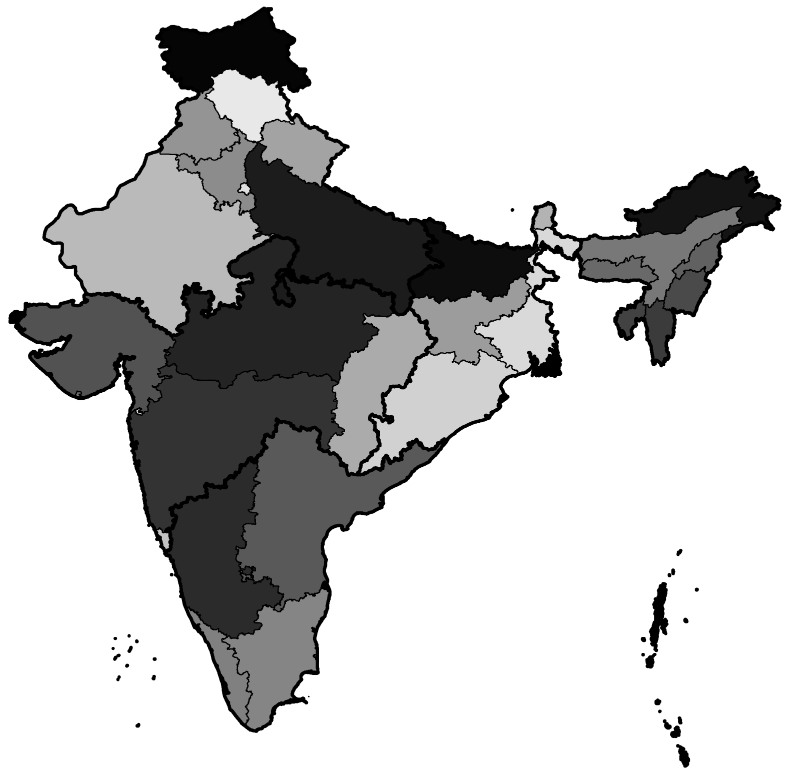
Table 13: Simulation versus Estimates

	Model	IV Estimates	P-Value	Textile	World Bank Survey
Output Loss	7.1%	4.8%	33%		7.8%
Output Loss for Generators	0.7%	-1.7%	24%	0.7%	
Output Loss for Non-Generators	10.3%	20.0%	19%		
TFP Loss	1.9%	0.2%	13%		
TFP Loss for Generators	0.1%	0.6%		0.1%	
Input Cost	0.2%	0.2%		0.2%	

Note: Model corresponds to the predictions in Column I of Table 11. IV Estimates corresponds to the estimates in Table 8 and Table 9, extrapolated under a 7.1 percent shortage. "P-Value" is the p-value for the test of whether the model's prediction is equal to the empirical estimate.

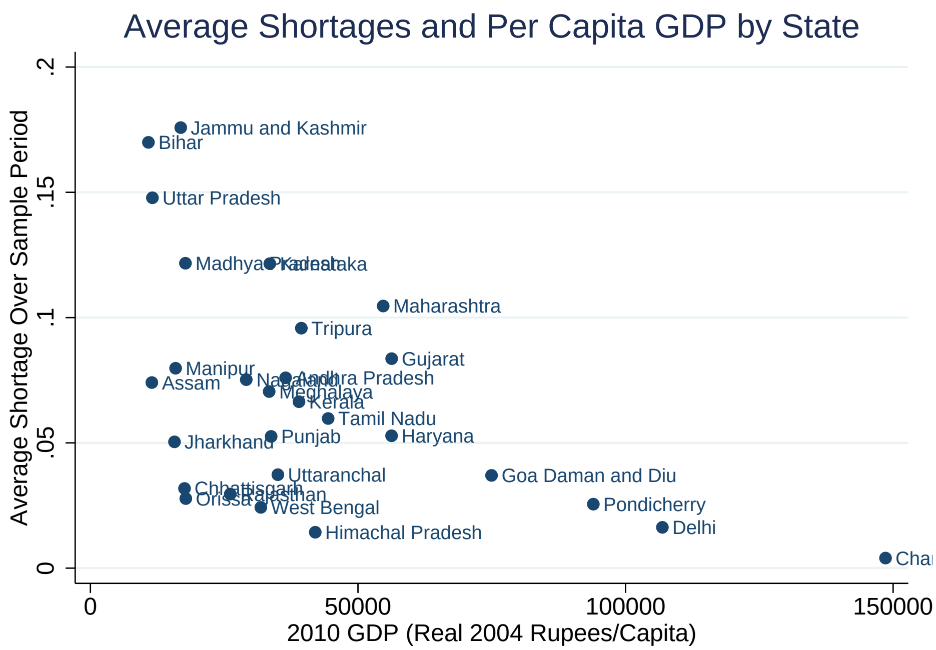
Figures

Figure 1: **Average Shortages and Per Capita GDP by State**



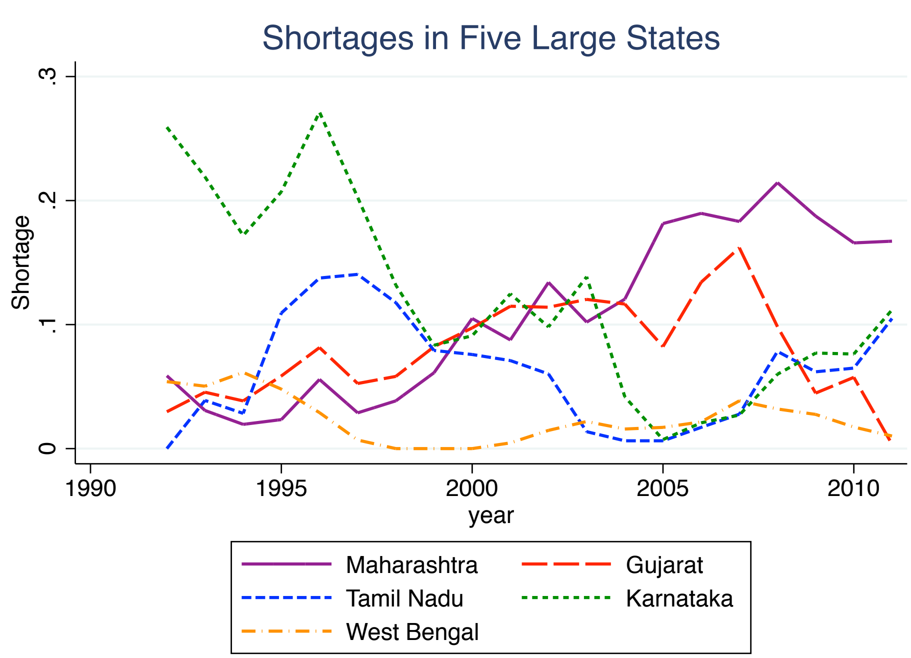
Notes: This figure compares the average of shortages estimated by the Central Electricity Authority to the 2010 per capita GDP, for all states and Union Territories.

Figure 2: Average Shortages and Per Capita GDP by State



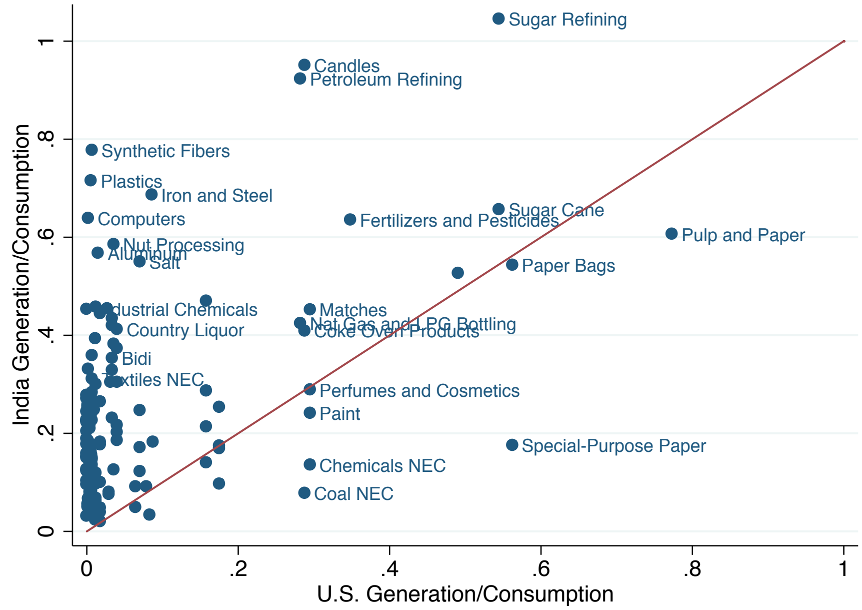
Notes: This figure compares the average of shortages estimated by the Central Electricity Authority to the 2010 per capita GDP, for all states and Union Territories.

Figure 3: **Variation in Shortages Over Time**



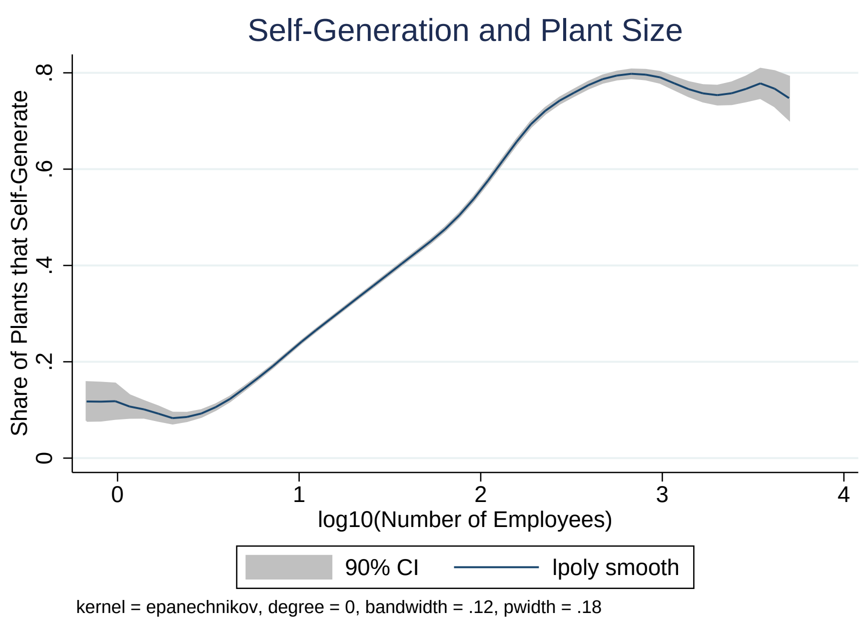
Notes: This figure presents shortages over the study period for five large states, as estimated by the Central Electricity Authority.

Figure 4: **Manufacturing Electricity Generation in India vs. the U.S.**



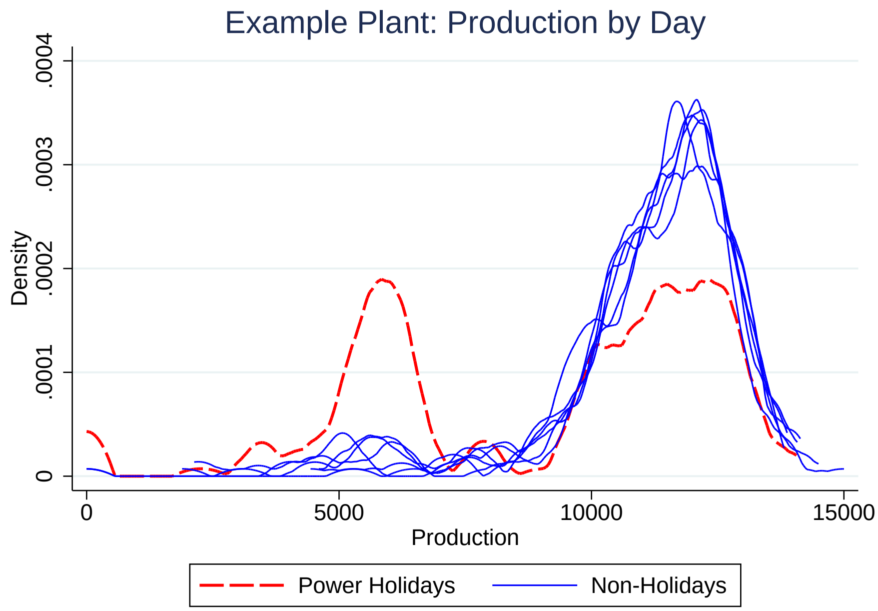
Notes: This figure presents the ratio of electricity generation to consumption by three-digit industry. Indian and U.S. data are from the Annual Survey of Industries and the Manufacturing Energy Consumption Survey, respectively.

Figure 5: **Generator Ownership and Plant Size**



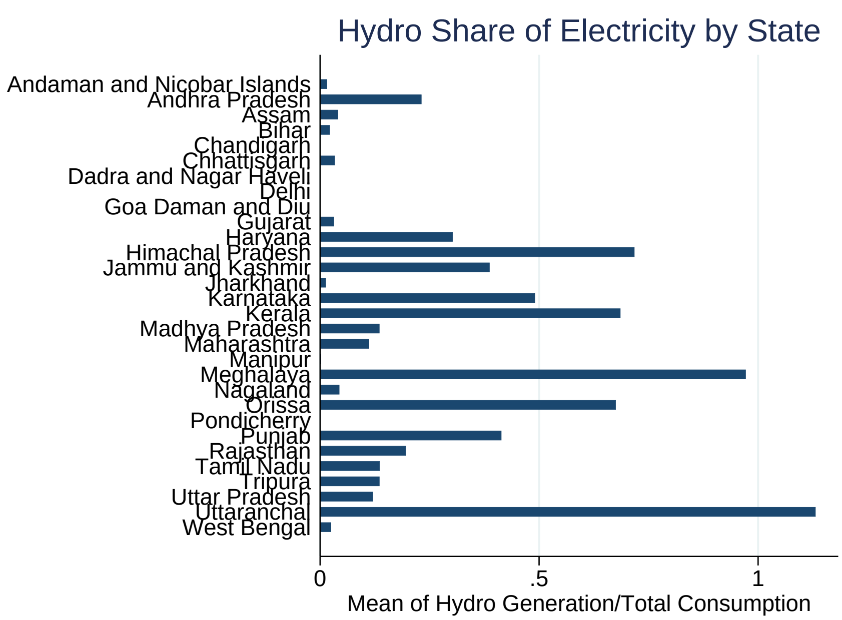
Notes: This figure presents local mean-smoothed estimates of the share of plants in all years of the Annual Survey of Industries sample that ever report self-generation, as a function of number of employees.

Figure 6: **Output on Power Holidays and Non-Holidays**



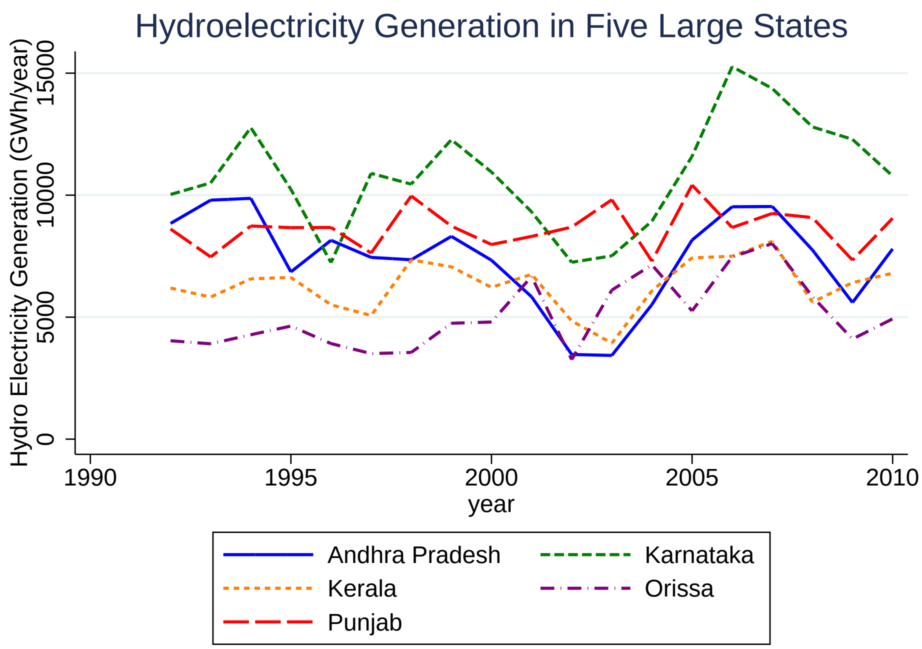
Notes: This figure presents the distribution of production by day of week for an example plant, using an Epanechnikov kernel with bandwidth 250. For this plant, every Friday is a power holiday.

Figure 7: **Hydro Share of Electricity by State**



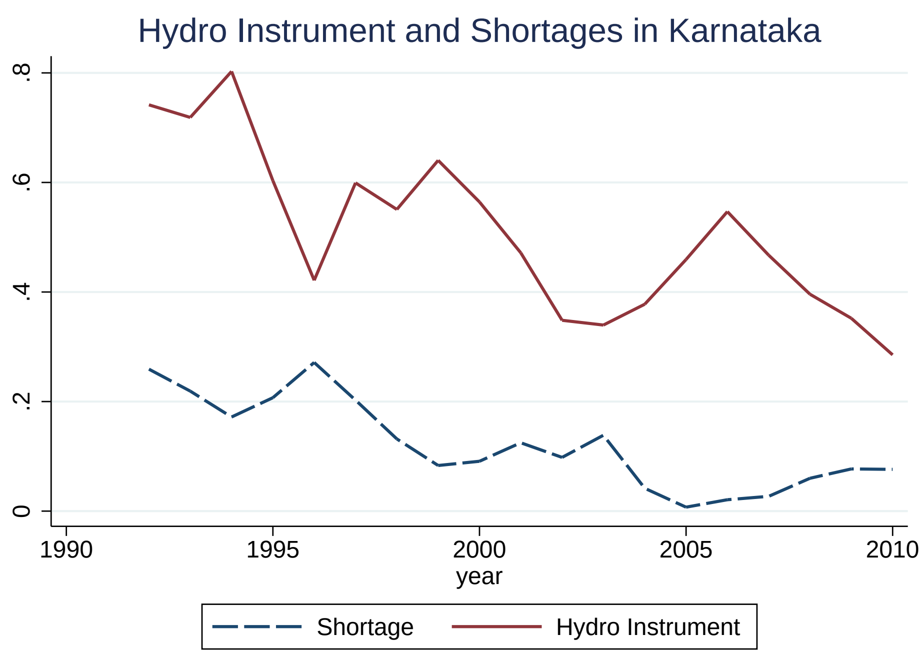
Notes: This figure presents the each state's mean ratio of hydroelectricity production to total consumption over 1992-2010. The graph includes only larger states with GDP larger than one billion Rupees in the year 2004 and with non-zero manufacturing production.

Figure 8: **Hydroelectricity Generation Over Time**



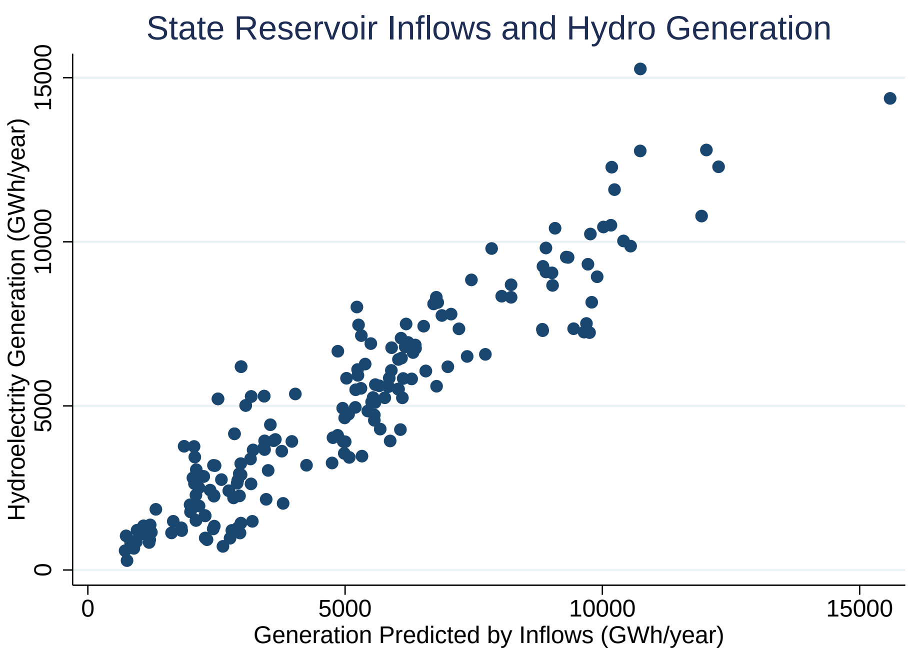
Notes: This figure presents hydroelectric generation over the study period for five large states that are relatively reliant on hydro.

Figure 9: **First Stage in Karnataka**



Notes: This figure presents shortages and the hydro instrument over the study period in the state of Karnataka.

Figure 10: **Correlation Between Reservoir Inflows and Hydro Production**



Notes: This is a scatterplot of hydroelectricity generation against the generation predicted using state-specific regressions of hydro generation on reservoir inflows.

Online Appendix: Not for Publication

How Do Electricity Shortages Affect Productivity? Evidence from India

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

A Appendix: Annual Survey of Industries Data Preparation

We extract a subset of variables from the raw data separately for each year and then stack all years of data together.[23] We correct accounts in 1993-94 to 1997-98 whose values have been supplied in "pre-multiplied" format from the Central Statistical Organisation's Ministry of Statistics and Planning Implementation (CSO/MOSPI). We then merge in state names based on the coding schemes provided with the Annual Survey of Industries (ASI) documentation, and we create a separate consistently-defined state variable which takes into account the creation of Jharkhand, Chhattisgarh and Uttaranchal (now Uttarakhand) in 2001 from Bihar, Madhya Pradesh and Uttar Pradesh, respectively.

India classifies manufacturing establishments with its National Industrial Classification (NIC), which resemble industrial classifications commonly used in other countries. The classifications were revised in 1987, 1998, 2004, and 2008. We convert all industry classifications to the NIC-1987 scheme using concordances provided by MOSPI with our data purchases. All financial amounts are deflated to constant 2004-05 Rupees. Revenue (gross sales) is deflated by a three-digit commodity price deflators as available in the commodity-based table "Index Numbers Of Wholesale Prices In India – By Groups And Sub-Groups (Yearly Averages)" produced by the GOI Office of the Economic Adviser-Ministry of Commerce & Industry (OEA 2013). Each three-digit NIC-1987 code is assigned to a commodity listed in this table. The corresponding commodity deflator is used to deflate revenues. To deflate material inputs, we construct the average output deflator of a given industry's supplier industries based on India's 1993-94 input-output table (CSO 2012). Fuels and total energy costs (fuels plus electricity) are deflated by the price index for "Fuel, Power, Light, and Lubricants." Capital is deflated by an implied national deflator calculated from "Table 13: Sector-wise Gross Capital Formation" from the Reserve Bank of India's Handbook of Statistics on the Indian Economy. Electricity costs are deflated using a national GDP deflator.

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

A.1 Determination of Base Sample

Appendix Table A21 details how the sample in Table 6 is determined from the original set of observations in the ASI. The 1992-2010 ASI dataset begins with 949,992 plant-year observations. Plants may still appear in the data even if they are closed or did not provide a survey response. We drop 172,697 of these plants reported as closed or non-responsive. We drop a trivial number of observations missing state identifiers and observations in Sikkim, which has only been included in the ASI sampling frame in the most recent years. We drop 45,664 observations reporting non-manufacturing NIC codes. We remove a small number of observations (primarily in the early years of our sample) which are exact duplicates in all fields, assuming these are erroneous multiple entries made from the same questionnaire form. Since we are concerned largely with revenue and productivity, we remove the 102,036 observations with missing revenues. We also drop the 9,095 observations with two or more input revenue share flags.

With this intermediate sample, we use median regression to estimate total factor productivity (TFP) under a full Cobb-Douglas model in capital, labor, materials, and energy. This full Cobb-Douglas productivity term is used only for the final sample restriction, which is to drop 464 plant-years which have log-TFP greater than 3.5 in absolute value from the sample median. Such outlying TFP values strongly suggest

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

misreported inputs or revenues. The final sample is comprised of 616,129 plant-years, of which 362,439 are from the sample scheme and 253,690 are from the census scheme.

A.2 Variable-Specific Sample Restrictions

After the final sample is determined, there may still be observations which have correct data for most variables but misreported data for some individual variable. When analyzing specific variables (such as self-generation share, energy revenue share, or output in Table 8), we therefore additionally restrict the sample using the following criteria:

- We generate "input revenue share flags" for labor and materials if input cost is more than two times revenues, and we generate input revenue share flags for electricity and fuels if input cost is greater than revenues.[24] Because we also observe physical quantities for labor and electricity, we generate analogous input revenue share flags by multiplying physical quantities by prices, resulting in an implied revenue share based on these physical quantities. For electricity, we use the median real price (in Rs/kWh) of purchased electricity in any given state and year. For labor, we assume a very conservative 1,000 Rs per person per annum wage rate. When using either of these inputs as an outcome, we omit observations with an input revenue share flag for that input.
- There are a trivial number of observations which report unrealistic count of persons engaged (greater than 200,000), which we make missing in those cases.
- We generate "within-plant outlier" flags for observations with unrealistically large year-to-year fluctuations in revenue, TFP, or any input. We flag observations if the change in logged value is more than 3.5 (or 1.5 in a robustness check) from both adjacent observations. For a plants' first or last year, it is flagged if the change is more than 3.5 (or 1.5) from the plant's one adjacent observation.

A.2.1 Cleaning Electricity Variables

We clean electricity variables in the following ways:

- We make electricity consumption missing for all observations that report zero electricity consumption (other than brick kilns).
- We make all electricity variables missing if the plant reports consuming more than 110 percent or less than 90 percent of the total amount of electricity they report purchasing and generating.
- We make missing the values of electricity purchased and sold if the implied price per kilowatt-hour is less than 2 percent or more than 5000 percent of the median grid electricity price calculated across plants in the same state and year. We also make missing the reported quantities of electricity purchased and sold if the respective price flag is triggered.

A.2.2 Production Function and Productivity Estimation

We recover production function coefficients given by Equations (10), (11), and (12) for each of the 143 three-digit industries in the dataset. (To ensure sufficient sample size in each three-digit industry, we adjust industry definitions slightly to ensure each three-digit industry has at least 100 plant-year observations.) We use separate median regression for each two-digit industry, allowing for a linear time trend and separate intercepts for each underlying three-digit industry. Consistent with the description in Section 3, the estimation sample includes only census scheme plant-year observations that report zero electricity generation. After calculating production function coefficients, we compute TFP from Equation (13).

We use several alternative methods for calculating production function coefficients and TFP for robustness checks:

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

- To check if our results are sensitive to assumptions about elasticity of demand, we calculate productivity terms for $\epsilon = -4$ and $\epsilon = -\infty$.
- We calculate an alternative materials term that adds the estimated cost of fuels not used for electricity generation. (To avoid relying on this estimated cost, our usual materials term does not include any fuels, so these costs do not enter the production function.)
- Omitting the linear time trend when estimating production coefficients, which amounts to taking the unconditional median by industry of the revenue shares for materials, labor, and electricity.
- Because in some industries plants with no self-generation may be unusual, we estimate production functions and productivity using all plants, i.e. including those that self-generate.

B Appendix: Additional Tables

B.1 Supporting Tables for Textile Case Study

Table A1: **Power Holidays**

Number of Plants	**State**	**Scheduled Power Holidays**
1	Gujarat	Saturday before Sept 26, 2008; Sunday between Sept 26, 2008 and July 10, 2009; Monday after July 10, 2009
1	Dadra and Nagar Haveli	Sunday
1	Gujarat	Saturday before July 10, 2009; Sunday after July 10, 2009
1	Maharashtra	Tuesday
1	Maharashtra	None
3	Gujarat	Saturday before July 10, 2009; Monday after July 10, 2009
14	Maharashtra	Friday

Notes: This table lists the scheduled power holidays for plants in the textile case study in Section 4

Table A2: **Textile Summary Statistics**

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Daily Data					
Production (1000s of Picks)	442	1455	0	9098	26,114
Percent Grade A	55.3	28.5	0	100	12,489
Quality Defect Index	4.02	5.0	0.13	56.6	13,223
1(Power Holiday)	0.14	0.34	0	1	46,288
Monthly Data					
Energy Costs (Rs 1000s)	300	282	8.88	1466	307
Labor (1000s of Hours)	32.3	20.7	4.39	148	575
Power Holiday Output Share	0.11	0.05	0	0.33	902
Diesel Price (Rs/liter)	35	1.67	31.4	38.3	902
Shortage	14.46	7.5	0	25.7	902

Notes: This table presents summary statistics for the textile case study in section 4. The top panel includes the variables observed for each day. There are two measures of quality: the percent of fabric graded quality level A, and the Quality Defect Index, a severity-weighted measure of the number of defects per meter of fabric. The bottom panel includes variables observed for each month. “Shortage” is the Central Electricity Authority’s monthly estimated electricity shortage percentage for the state where the plant is located. Diesel prices are the Mumbai prices recorded by the website mypetrolprice.com. All rupees are deflated to constant 2004-2005 values using the textile wholesale price index (Office of the Economic Advisor 2013).

B.2 First Stages

Table A3: **Base Estimates First Stages**

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Fuel Rev Share)	ln(Output)	ln(TFP)
Δ Hydro	-0.134	-0.139	-0.100	-0.101
	(0.020)***	(0.019)***	(0.016)***	(0.016)***
Number of Obs.	172,317	220,613	374,157	366,302
Number of Clusters	2,781	2,936	3,262	3,261
A-P F-Stat	43.98	52.51	39.36	39.6

Notes: This table presents the first stage estimates for the IV regressions in Panel B of Table 8. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Angrist-Pischke weak instruments test. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A4: **Additional First Stages for Self-Generation Share**

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable:	Δ Shortage	Δ Shortage	Δ Shortage	Δ Shortage	Δ Shortage x Elec Int	Δ Shortage x Self-Gen
Δ Hydro	-0.134	-0.110	-0.130	-0.062	0.018	-0.003
	(0.038)***	(0.027)***	(0.021)***	(0.016)***	(0.004)***	(0.002)*
Δ Rainfall			-0.007			
			(0.005)			
Δ Cooling Degrees			0.000			
			(0.002)			
Δ Hydro x Elec Intensive				0.009	-0.088	0.006
				(0.005)*	(0.015)***	(0.003)**
Δ Hydro x Self-Generator				-0.076	-0.036	-0.133
				(0.016)***	(0.008)***	(0.020)***
Number of Obs.	172,317	124,771	170,356	301,386	320,545	320,545
Number of Clusters	30	491	2,719	3,187	3,229	3,229
A-P F-Stat	12.51	16.1	40.28	43.95	32.44	43.95

Notes: This table presents the first stage estimates for alternative specifications with potentially weakest first stage identification, using the sample with self-generation share as the outcome variable. Column 1 clusters by state instead of state-by-year difference. All other columns cluster by state-by-year difference. Column 2 includes one-year differences only. Column 3 controls for rainfall and cooling degrees. Columns 4-6 are the three first stages for Table 9. Samples for columns 1, 2, and 3 are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Angrist-Pischke weak instruments test. Robust standard errors. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A5: **Additional First Stages for ln(Revenue)**

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable:	Δ Shortage	Δ Shortage	Δ Shortage	Δ Shortage	Δ Shortage x Elec Int	Δ Shortage x Self-Gen
Δ Hydro	-0.100	-0.091	-0.098	-0.064	0.021	-0.003
	(0.036)**	(0.023)***	(0.016)***	(0.015)***	(0.005)***	(0.002)**
Δ Rainfall			-0.002			
			(0.005)			
Δ Cooling Degrees			0.001			
			(0.002)			
Δ Hydro x Elec Intensive				0.010	-0.095	0.008
				(0.005)**	(0.016)***	(0.003)**
Δ Hydro x Self-Generator				-0.077	-0.037	-0.136
				(0.015)***	(0.008)***	(0.019)***
Number of Obs.	374,157	229,177	370,167	374,157	374,158	374,158
Number of Clusters	30	494	3,179	3,262	3,262	3,262
A-P F-Stat	7.52	15.52	36.5	51.21	37.63	51.21

Notes: This table presents the first stage estimates for alternative specifications with potentially weakest first stage identification, using the sample with ln(Revenue) as the outcome variable. Column 1 clusters by state instead of state-by-year difference. All other columns cluster by state-by-year difference. Column 2 includes one-year differences only. Column 3 controls for rainfall and cooling degrees. Columns 4-6 are the three first stages for Table 9. Samples for columns 1, 2, and 3 are limited to plants that ever self-generate electricity. F-statistic is for the heteroskedasticity and cluster-robust Angrist-Pischke weak instruments test. Robust standard errors. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

B.3 Robustness Checks for Table 8

Table A6: **Robustness Check: Omitting Industry-by-Year Controls**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.229	-0.014	0.036	0.136
	(0.024)***	(0.059)	(0.041)	(0.032)***
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	2,781	2,936	3,263	3,261

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.579	0.780	-0.815	-0.149
	(0.109)***	(0.249)***	(0.346)**	(0.166)
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	2,781	2,936	3,263	3,261

Notes: This table presents estimates of Equation (24), omitting the industry-by-year controls μ_{jt}. Panel B instruments for Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A7: **Robustness Check: Stricter Tolerance for Eliminating Within-Plant Outliers**
Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.201	-0.037	0.006	0.098
	(0.022)***	(0.054)	(0.036)	(0.027)***
Number of Obs.	153,764	197,125	362,273	358,923
Number of Clusters	2,727	2,860	3,237	3,251

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.530	0.608	-0.498	-0.003
	(0.097)***	(0.217)***	(0.280)*	(0.140)
Number of Obs.	153,764	197,125	362,273	358,923
Number of Clusters	2,727	2,860	3,237	3,251

Notes: This table presents estimates of Equation (24), using a within-plant outlier tolerance of 1.5 natural logs instead of 3.5. Panel B instruments for Shortage using the hydroelectric generation instrument. Robust standard errors, clustered by state-by-year difference. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A8: **Robustness Check: Including All Within-Plant Outliers**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.244	0.012	0.060	0.074
	(0.023)***	(0.056)	(0.043)	(0.034)**
Number of Obs.	226,244	228,572	376,019	366,943
Number of Clusters	2,962	2,964	3,273	3,262

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.602	0.732	-0.626	0.173
	(0.102)***	(0.244)***	(0.379)*	(0.198)
Number of Obs.	226,244	228,572	376,019	366,943
Number of Clusters	2,962	2,964	3,273	3,262

Notes: This table presents estimates of Equation (24), without dropping any within-plant outliers. Panel B instruments for Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A9: **Robustness Check: One-Year Lags Only**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.203	0.084	-0.102	0.038
	(0.037)***	(0.084)	(0.059)*	(0.050)
Number of Obs.	124,762	152,687	229,172	225,012
Number of Clusters	491	491	494	494

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.543	0.271	-0.574	-0.025
	(0.187)***	(0.325)	(0.387)	(0.253)
Number of Obs.	124,762	152,687	229,172	225,012
Number of Clusters	491	491	494	494

Notes: This table presents estimates of Equation (24), using the sample of one-year differences only. Panel B instruments for Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A10: **Robustness Check: Clustering by State**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.227	-0.035	0.020	0.095
	(0.029)***	(0.079)	(0.079)	(0.053)*
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	30	30	30	30

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.568	0.639	-0.682	-0.034
	(0.111)***	(0.278)**	(0.300)**	(0.283)
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	30	30	30	30

Notes: This table presents estimates of Equation (24), clustering by state instead of state-by-year difference. Panel B instruments for Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A11: **Robustness Check: Controlling for Rainfall and Cooling Degrees**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.229	-0.031	0.021	0.097
	(0.023)***	(0.055)	(0.040)	(0.030)***
Δ Rainfall	0.003	-0.016	0.008	0.001
	(0.003)	(0.009)*	(0.008)	(0.006)
Δ Cooling Degrees	0.001	-0.006	-0.001	-0.001
	(0.001)	(0.004)	(0.003)	(0.002)
Number of Obs.	170,358	218,029	370,179	362,411
Number of Clusters	2,719	2,865	3,180	3,179

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Shortage	0.582	0.683	-0.652	-0.003
	(0.110)***	(0.250)***	(0.342)*	(0.163)
Δ Rainfall	0.008	-0.007	0.004	0.001
	(0.004)*	(0.010)	(0.009)	(0.006)
Δ Cooling Degrees	0.001	-0.007	0.000	-0.000
	(0.001)	(0.004)*	(0.003)	(0.002)
Number of Obs.	170,358	218,029	370,179	362,411
Number of Clusters	2,719	2,865	3,180	3,179

Notes: This table presents estimates of Equation (24), also including controls for rainfall and cooling degrees. Panel B instruments for Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A12: **Robustness Check: Using Peak Shortage Instead of Average Shortage**

Panel A: Difference Estimator

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Peak Shortage	0.063	-0.013	-0.017	-0.001
	(0.013)***	(0.030)	(0.025)	(0.018)
Number of Obs.	172,319	220,622	374,168	366,319
Number of Clusters	2,781	2,936	3,263	3,261

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Δ Peak Shortage	0.630	0.735	-0.733	-0.004
	(0.179)***	(0.286)**	(0.403)*	(0.183)
Number of Obs.	170,358	218,029	370,179	362,411
Number of Clusters	2,719	2,865	3,180	3,179

Notes: This table presents estimates of Equation (24), using the CEA Peak Shortage estimate instead of (average) Shortage. Panel B instruments for Peak Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

B.3.1 Estimates with Alternative TFP Measures

Table A13: **Robustness Check: Estimates with Alternative TFP Measures**
Panel A: Difference Estimator

	(1)	(2)	(3)	(4)	(5)
	Perfect Comp. ($\epsilon = -\infty$)	ϵ=-4	Materials+ Non-Elec Fuels	No Year Controls	Include Self-Generators
Δ Shortage	0.150	0.251	0.095	0.090	0.098
	(0.031)***	(0.046)***	(0.029)***	(0.027)***	(0.025)***
Number of Obs.	365,945	364,852	338,742	366,097	366,176
Number of Clusters	3,259	3,260	2,950	3,261	3,261

Panel B: Instrumental Variables

	(1)	(2)	(3)	(4)	(5)
	Perfect Comp. ($\epsilon = -\infty$)	ϵ=-4	Materials+ Non-Elec Fuels	No Year Controls	Include Self-Generators
Δ Shortage	0.095	0.334	-0.029	-0.320	-0.232
	(0.176)	(0.251)	(0.146)	(0.177)*	(0.165)
Number of Obs.	365,945	364,852	338,742	366,097	366,176
Number of Clusters	3,259	3,260	2,950	3,261	3,261

Notes: This table presents estimates of Equation (24), using alternative measures of TFP described in Appendix A.2.2. Panel B instruments for Peak Shortage using the hydroelectric generation instrument. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

B.4 Fixed Effects Estimates

Table A14: **Fixed Effects Estimates with ASI Data**

Panel A: Fixed Effects

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Shortage	0.366	-0.174	-0.059	0.221
	(0.046)***	(0.073)**	(0.094)	(0.054)***
Number of Obs.	182,252	229,164	385,913	381,524
Number of Clusters	30	30	30	30

Panel B: Fixed Effects with Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Shortage	1.039	1.039	-0.986	-0.519
	(0.375)***	(0.575)*	(0.671)	(0.430)
Number of Obs.	169,453	215,072	345,230	340,569
Number of Clusters	30	30	30	30

This table presents estimates of Equation (24) using fixed effects instead of differences, also including state-specific linear trends. Panel B instruments for Peak Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A15: **Fixed Effects Robustness Check: Omitting Industry-by-Year Controls**

Panel A: Fixed Effects

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Shortage	0.287	-0.070	0.026	0.175
	(0.039)***	(0.100)	(0.074)	(0.050)***
Number of Obs.	276,510	312,684	609,140	600,655
Number of Clusters	30	30	30	30

Panel B: Fixed Effects with Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Shortage	0.807	0.883	-0.779	-0.466
	(0.262)***	(0.428)**	(0.439)*	(0.345)
Number of Obs.	240,794	284,431	501,541	494,230
Number of Clusters	30	30	30	30

This table presents estimates of Equation (24) using fixed effects instead of differences, also including state-specific linear trends. It is identical to Table A14, except omitting the industry-by-year controls μ_{jt}. Panel B instruments for Peak Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A16: **Fixed Effects Robustness Check: Stricter Tolerance for Eliminating Within-Plant Outliers**

Panel A: Fixed Effects

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Shortage	0.257	-0.109	0.063	0.108
	(0.041)***	(0.094)	(0.066)	(0.030)***
Number of Obs.	261,055	291,987	597,687	594,387
Number of Clusters	30	30	30	30

Panel B: Fixed Effects with Instrumental Variables

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Energy Rev Share)	ln(Output)	ln(TFP)
Shortage	0.737	0.577	-0.552	-0.251
	(0.239)***	(0.337)*	(0.293)*	(0.255)
Number of Obs.	223,176	259,927	485,209	486,097
Number of Clusters	30	30	30	30

This table presents estimates of Equation (24) using fixed effects instead of differences, also including state-specific linear trends. It is identical to Table A14, except using a within-plant outlier tolerance of 1.5 natural logs instead of 3.5. Panel B instruments for Peak Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A17: **First Stages for ASI Fixed Effects Regressions**

	(1)	(2)	(3)	(4)
Outcome Variable:	Self-Gen Share	ln(Fuel Rev Share)	ln(Output)	ln(TFP)
Hydro Generation	-0.120	-0.125	-0.105	-0.105
	(0.049)**	(0.049)**	(0.047)**	(0.047)**
Number of Obs.	240,781	284,422	501,527	494,207
Number of Clusters	30	30	30	30
A-P F-Stat	6	6.43	4.95	4.93

Notes: This table presents the first stage estimates for the IV regressions in Panel B of Table A14. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

B.5 Other Tables

Table A18: **Biggest Obstacle for Growth**

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

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

Table A19: **Separate Results for Self-Generators and Non-Self-Generators**

	(1)	(2)	(3)	(4)
Outcome Variable:	ln(Output)	ln(Output)	ln(TFP)	ln(TFP)
Δ Shortage	-0.246	-2.813	-0.087	-0.231
	(0.293)	(1.034)***	(0.142)	(0.418)
Number of Obs.	234,300	139,868	230,346	135,973
Number of Clusters	2,977	3,088	2,971	3,086
Self-Generators	Yes	No	Yes	No

Notes: This table presents estimates of 24, splitting the sample by self-generators versus plants that never self-generate. Panel B instruments for Shortage using the hydroelectric generation instrument. Samples for columns 1 and 2 are limited to plants that ever self-generate electricity. Robust standard errors, clustered by state-by-year difference. *,**, ***: Statistically different from zero with 90, 95, and 99 percent confidence, respectively.

Table A20: **Calibration Parameters**

Parameter		Note
Production Function Coefficients	β_m, β_l	From Input Cost Share of Non-Self Generators (see equations (11) and (12))
	β_k	Constant Return to Scale
	γ	Electricity Input Cost Share
Elasticity of Demand	ϵ	
Price of Grid Power	$p^{E,G}$=4.5	World Bank Survey
Price of Self-Generated Power	$p^{E,S}$=7	World Bank Survey
Shortages	δ	Data collected from Central Electric Authority
Prices of Output, Labor, Materials	$p = 1$, $p^L = 1$, $p_M = 1$	Normalization

Table A21: **Determination of Base Sample**

Step	**Dropped obs.**	**Sample Size**
Original ASI dataset		949,992
Closed plants	-172,697	777,295
Missing state codes or in Sikkim	-99	777,196
Non-manufacturing NIC codes	-45,664	731,532
Exact duplicates	-312	731,220
Missing revenues	-102,036	629,184
Multiple input revenue share outliers	-9,095	620,089
Productivity outliers	-3,960	616,129
Total observations		616,129

Notes: This table details how the sample in Table 6 is determined from the original set of observations in the Annual Survey of Industries.