Blackouts:

The role of India's wholesale electricity market

Supplementary appendix: For online publication

Akshaya Jha ⓡ Louis Preonas ⓡ Fiona Burlig*

Appendix A provides additional results and sensitivity analysis related to our demand estimation.

Appendix B provides details on our simple structural model, as well as robustness and supplementary results.

Appendix C provides further details on data sources and data construction.

∗. Jha: H. John Heinz III College, Carnegie Mellon University, and NBER. Email: akshayaj@andrew.cmu.edu. Preonas: Department of Agricultural and Resource Economics, University of Maryland. Email: lpreonas@umd.edu. Burlig: Harris School of Public Policy and Energy Policy Institute (EPIC), University of Chicago, and NBER. Email: burlig@uchicago.edu.

A Demand estimation appendix

A.1 States with high versus low generation costs

As discussed in the main text, we re-estimate our demand regressions splitting the sample by high versus low average-variable-cost states. To do this, we take the sample-wide average of average variable costs across power plants in each state s, weighting plants by their average capacity. Then, we split the sample to states above versus below the median of this capacity-weighted average variable cost. We present these split sample results in Appendix Table A.1: states facing higher average variable costs of production exhibit a higher cost elasticity of demand on average. This suggests that utilities that may be losing more money on the margin are more sensitive to wholesale cost increases—supporting our central hypothesis that utilities *choose* to purchase less electricity when procurement costs are higher.

Table A.1: Demand regressions splitting states by high versus low average variable costs

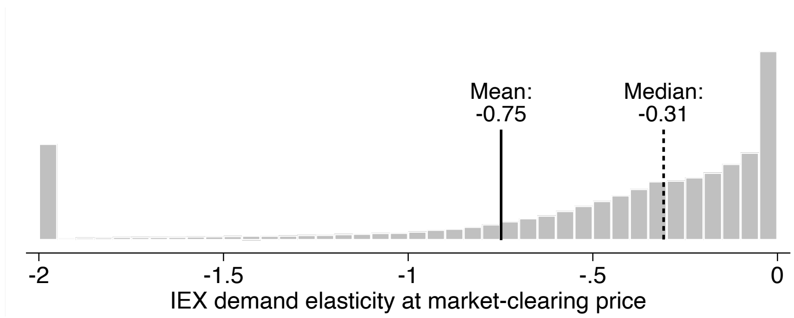
	Outcome: log (Quantity)			
	(1) OLS	(2) IV	(3) OLS	(4) IV
Equipment outage rate	-0.14^{***}		-0.02	
	(0.04)		(0.05)	
log (Average variable cost)		-0.54^{***}		-0.09
		(0.18)		(0.27)
Split on states' average variable cost	High AVC	High AVC	Low AVC	Low AVC
State + date FEs	Yes	Yes	Yes	Yes
Region × year, region × month FEs	Yes	Yes	Yes	Yes
Mean demand met (in GWh)	68.76	68.76	118.12	118.12
State-day observations	20,993	20,993	21,219	21,219
First-stage estimate		0.26^{***}		0.19^{***}
		(0.06)		(0.02)
Kleibergen-Paap F-statistic		22.44		94.87

Notes: This table presents additional results from estimating Equation (2). The dependent variable is the natural logarithm of total GWh of energy purchased in the wholesale sector in state s on date t. We split the sample on states' average variable cost of generation (i.e., above/below-median of the capacity-weighted state-level average variable costs over our sample period). Columns (1) and (3) are otherwise identical to Column (1) of Table 2. Columns (2) and (4) are otherwise identical to Column (4) of Table 2. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.2 Price elasticity of demand in the IEX

As discussed in the main text, Appendix Figure A.1 plots the elasticity of demand for electricity in the IEX at the market-clearing price over all 15-minute intervals in our sample. The mean elasticity in this market is -0.75, with a median of -0.31, reinforcing that demand for power in India's wholesale market is downward sloping.

Figure A.1: Histogram of observed demand elasticities in the IEX day-ahead market



Notes: We extract the price elasticity of IEX demand from observed aggregate bid curves for 201,012 separate 15-minute intervals. We bottom-code this distribution at -2 for ease of presentation. The solid (dashed) line reports the mean (median) elasticity.

A.3 Robustness: Demand estimation

Appendix Table A.2 provides sensitivity analysis for our demand elasticity estimates in Table 2. Column (1) accounts for local (within-state) transmission constraints. We assign plants to "load pockets" (i.e, high-demand areas)—proxied by being located within one of India's 20 most populous districts. Then, we construct the instrument using only equipment outages outside of these load pockets. Column (1) also restricts the sample to state-days with idle capacity outside of load pockets, which are unlikely to have faced a local transmission constraint.

Next, Appendix Table A.2 provides two alternate versions of Column (6) in Table 2, which jointly removes potential short-outage effects and anticipation effects by instrumenting using only equipment outages already on at least their fifth day. Column (2) instruments using all days of 5-day-or-longer equipment outages (i.e., not omitting days 1–4), which isolates potential differences in outage type (i.e., long versus short equipment outages). Column (3) lags our preferred instrument based on all equipment outages by 4 days, which removes any potential anticipation effects. Both sensitivities yield similar elasticity estimates.

Table A.2: Sensitivities for demand regressions – sample restrictions

	Outcome: log (Quantity)			
	(1) IV	(2) IV	(3) IV	(4) IV
log (Average variable cost)	-0.55^{***}	-0.54^{**}	-0.39^{*}	-0.50^{**}
	(0.20)	(0.22)	(0.21)	(0.21)
Equip. outages outside load pockets	Yes			
Idle capacity available outside load pockets	Yes			
Equip. outages that span 5+ days		Yes		
Equip. outages lagged 4 days			Yes	
Include state-days with < 50% coverage				Yes
State + date FEs	Yes	Yes	Yes	Yes
Region × year, region × month FEs	Yes	Yes	Yes	Yes
Mean demand met (in GWh)	150.12	90.25	90.26	88.70
State-day observations	12,093	42,212	38,842	43,044
First-stage estimate	0.31^{***}	0.19^{***}	0.17^{***}	0.18^{***}
	(0.05)	(0.04)	(0.03)	(0.03)
Kleibergen-Paap F-statistic	37.49	27.79	24.13	30.24

Notes: Column (1) shows that our demand elasticity estimates are robust to transmission constraints around load pockets: the instrument only includes equipment outages at plants not located in one of India's 20 most populous districts, and restricts the sample to state-days where a "non-load-pocket" plant reported idle capacity. Column (2) is similar to Column (6) of Table 2, except that the instrument also includes days 1–4 of equipment outages that (eventually) persist for at least 5 consecutive days. Column (3) is identical to Column (4) of Table 2, except that the instrument is lagged by 4 days. Column (4) is identical to Column (4) of Table 2, except that it includes the 2% of observations where our cost and outage data cover less than 50% of total generating capacity (thermal + hydro) in that state-day cell. Regressions are otherwise identical to Table 2. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Finally, Column (4) of Appendix Table A.2 relaxes our preferred sample restriction, where we omit observations for which our state-day equipment outage rate and average variable cost are calculated using less than 50% of total (thermal + hydro) capacity. In the main text, we impose this restriction to account for the incompleteness of our plant-day level outage and cost data: if less than 50% of capacity is represented in our right-hand-side variables, they might have a weaker relationship with state-day level demand. However, we estimate an even larger demand elasticity when we include these state-days in the regression.

Appendix Table A.3 conducts sensitivity analysis on our definition of "idle capacity available." Columns (2) and (5) of Table 2 report results based on our preferred definition of idle capacity: if any capacity (> 0 MW) in state s on day t was available to generate (i.e., was not on outage) but did not. For this specification, we estimate a reduced-form effect of -0.12 and an IV estimate of -0.42. Columns (1)–(2) of Appendix Table A.3 strengthen

Table A.3: Sensitivities for demand regressions – idle capacity available

	Outcome: log (Quantity)			
	(1) OLS	(2) IV	(3) OLS	(4) IV
Equipment outage rate	-0.16^{***}		-0.16^{*}	
	(0.06)		(0.09)	
log (Average variable cost)		-0.62^{**}		-0.39
		(0.28)		(0.29)
Idle capacity available	Yes	Yes	Yes	Yes
Definition of idle capacity	$>$ 250 MW idle in state		$\geq$ 2 plants idle in state	
State + date FEs	Yes	Yes	Yes	Yes
Region $\times$ year, region $\times$ month FEs	Yes	Yes	Yes	Yes
Mean demand met (in GWh)	184.59	184.59	218.43	218.43
State-day observations	10,625	10,625	5,564	5,564
First-stage estimate		0.26^{***}		0.41^{***}
		(0.06)		(0.11)
Kleibergen-Paap F-statistic		21.63		14.10
Mean of equipment outage rate	0.08	0.25	0.08	0.25
SD of equipment outage rate	0.06	0.06	0.26	0.26
Mean potential GWh (idle cap.)	35.00	35.00	55.77	55.77

Notes: This table presents sensitivity analysis on our preferred definition of "idle capacity avaiable". (i.e. any plant in the state with any positive capacity that was avaiable to generate but did not). Columns (1)–(2) strengthen this definition from ≥ 0 MW to ≥ 250 MW of idle capacity in the state.Columns (3)–(4) apply an alternate definition: at least 2 plants in the state must have capacity that was avaiable to generate but did not.Columns (1) and (3) are otherwise identical to Column (2) of Table 2. Columns (2) and (4) are otherwise identical to Column (5) of Table 2. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

this definition to at least ($>$ 250 MW) of idle capacity in state s on day t. This yields similar cost-elasticity of demand estimates. In Columns (3)–(4), we apply the stronger sample restriction that at least 2 plants in state s had idle capacity on day t. This shrinks the sample considerably, causing us to lose statistical power. Nevertheless, we still recover similar point estimates.

Appendix Table A.4 conducts sensitivity analysis to assuage concerns that our demand elasticities are driven by transmission constraints, as opposed to utilities choosing to purchase less power. Using data on IEX market-clearing prices at the subregion/15-minute-interval level (Indian Energy Exchange (2014–2019)), we identify region-days where there were zero price wedges between constituent subregions.[1] Columns (1)–(2) reveal that our

1. Each of the India's 5 market regions comprises 2–3 subregions. IEX price wedges between subregions

Table A.4: Sensitivities for demand regressions – no intra-regional IEX price wedges

	Outcome: log (Quantity)			
	(1) OLS	(2) IV	(3) OLS	(4) IV
Equipment outage rate	-0.08^{**}		-0.07^{*}	
	(0.03)		(0.04)	
log (Average variable cost)		-0.46^{**}		-0.33^{*}
		(0.23)		(0.17)
No IEX price wedges within region	Yes	Yes	Yes	Yes
Idle capacity available in region			Yes	Yes
State + date FEs	Yes	Yes	Yes	Yes
Region × year, region × month FEs	Yes	Yes	Yes	Yes
Mean demand met (in GWh)	87.09	87.09	104.24	104.24
State-day observations	38,792	38,792	32,424	32,424
First-stage estimate		0.16^{***}		0.22^{***}
		(0.04)		(0.03)
Kleibergen-Paap F-statistic		21.36		49.96
Mean of equipment outage rate	0.10	0.30	0.10	0.30
SD of equipment outage rate	0.09	0.09	0.34	0.34
Mean potential GWh (idle cap.)	8.64	8.64	10.34	10.34

Notes: This table tests whether our demand estimates are driven by trading frictions by including only the subset of state-days with zero intra-regional wedges in the market-clearing IEX price. We construct this subsample using IEX prices IEX at the subregion by 15-minute interval level, classifying a region-day as having "no wedges" if within-interval prices were identical for all subregions, for all 96-minute intervals of the the day. Columns (1)–(2) apply this sample restriction, and are otherwise identical to Columns (1) and (4) of Table 2. Columns (3)–(4) are analogous, but impose an additional restriction of there being idle capacity available in the region. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

demand elasticity estimates are robust to dropping state-days with intra-regional IEX price wedges. Columns (3)–(4) find similar results when we impose the additional restriction of idle generating capacity being available within the region. These results support the interpretation that our estimates in Table 2 indeed reflect downward-sloping utility demand.

In Appendix Table A.5, we estimate the wholesale demand elasticity with respect to marginal cost, rather than average variable cost. Columns (1)–(2) use the 95th and 98th percentiles (respectively) of the realized marginal cost distribution for operating plants in each state-day. This yields similar demand elasticity estimates, despite the fact that using marginal cost does not align with the incentives implied by cost-of-service regulation in this setting (i.e., prices are set by regulators based on average costs rather than market-based

indicate that transmission constraints prevented unconstrained power trade between subregions.

Table A.5: Demand regressions using marginal cost instead of average variable cost

	Outcome: log (Quantity)		
	(1) IV	(2) IV	(3) IV
log (95th percentile of marginal cost)	-0.25^{**}		
	(0.10)		
log (98th percentile of marginal cost)		-0.39^{**}	
		(0.19)	
log (Maximum marginal cost)			-0.61
			(0.40)
State + date FEs	Yes	Yes	Yes
Region × year, region × month FEs	Yes	Yes	Yes
Mean demand met (in GWh)	90.25	90.25	90.25
State-day observations	42,212	42,212	42,212
First-stage estimate	0.35^{***}	0.22^{***}	0.14^{*}
	(0.06)	(0.06)	(0.07)
Kleibergen-Paap F-statistic	40.86	14.37	3.61

Notes: These regressions estimate the wholesale demand elasticity with respect to alternate constructions of marginal cost (as opposed to our preferred average variable costs). Columns (1) and (2) use the 95th and 98th percentiles of marginal costs among operating plants within each state-day; and Column (3) uses the maximum marginal cost for each state-day. Regressions are otherwise identical to Column (4) of Table 2. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

dispatch). Column (3) uses the maximum marginal cost among operating plants in each state-day, which is most analogous to a conventional competitive market equilibrium where the aggregate marginal cost curve intersects demand. This yields an underpowered first stage, which makes sense given that the maximum is particularly prone to outliers induced by measurement error in our constructed marginal costs.

We next conduct sensitivity analysis on our preferred log-log functional form of demand. Appendix Table A.6 presents a version of Table 2 under the assumption of linear demand. This yields reduced-form estimates in Columns (1)–(3) that are larger in magnitude than in our main specification: the 25.32 GWh/day reduction in quantity supplied implied by the point estimate in Column (1) corresponds to a 17% change in quantity supplied. Columns (4)–(6) likewise imply larger demand elasticities than our preferred log-log model. For Column (4), a 10% increase in average variable cost (i.e., a 0.14 Rs/kWh increase) causes wholesale quantity to fall by 21 GWh/day (i.e., a 14% decrease)—implying a cost elasticity of demand of −1.4. However, we interpret this estimate with caution due to the low first-stage F-statistic. For Column (5) where the F-statistic is larger, a 10% increase in AVC (i.e.,

Table A.6: Demand regressions assuming linear demand instead of log-log demand

	Outcome: Quantity (GWh)					
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Equipment outage rate	-25.32^{***}	-37.61^{***}	-28.42^{***}			
	(4.60)	(9.84)	(4.83)			
Average variable cost (Rs/kWh)				-149.48^{**}	-104.92^{***}	-141.49^{***}
				(59.45)	(36.67)	(53.52)
Idle capacity available		Yes			Yes	
Equip. outages already 5+ days			Yes			Yes
State + date FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region-year, region-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean demand met (in GWh)	145.71	202.72	145.71	145.71	202.72	145.71
State-day observations	42,215	13,722	42,215	42,215	13,722	42,215
First-stage estimate				0.17^{***}	0.36^{***}	0.20^{***}
				(0.06)	(0.07)	(0.07)
Kleibergen-Paap F-statistic				7.37	30.16	7.66
Mean of avg variable cost (Rs/kWh)				1.41	1.33	1.41
SD of avg variable cost (Rs/kWh)				0.59	0.49	0.59
Mean of equipment outage rate	0.10	0.09	0.08	0.10	0.09	0.08
SD of equipment outage rate	0.09	0.07	0.08	0.09	0.07	0.08
Mean potential GWh (idle cap.)	8.92	27.43	8.92	8.92	27.43	8.92

Notes: This table is identical to Table 2, except that rather than estimating a log-log model, we estimate a linear model. The outcome variable is GWh of energy purchased by in the wholesale sector in state s on date t (in levels). The endogenous right-hand-side variable is the average variable cost of generation in Rs/kWh in state s on date t (in levels). Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. See notes under Table 2 for further details.

0.13 Rs/kWh) causes wholesale quantity to fall by 14 GWh/day (i.e., a 6.7% decrease)—implying an elasticity of -0.67. This is within the 95% confidence interval of the elasticity estimate implied by our log-log specification (see Column 5 of Table 2).

A.4 Robustness: Equipment outages and lagged demand

Dynamics in power plant maintenance present a potential concern for our estimates in Table 1. It is possible that technical failures are more likely after periods of high demand (e.g., if plants have run "too hot" for many consecutive days, or if high demand causes them to forgo their planned maintenance). We address these concerns by adding lags to our regressions of equipment outage rates on demand shocks.

Appendix Table A.7 conducts robustness on Column (1) of Table 1, adding daily lags of temperature and the realized quantity of wholesale electricity demanded. We cannot control for contemporaneous quantity demanded because of simultaneity—the equilibrium quantity demanded is directly impacted by equipment-outage-induced cost shocks. Appendix Table A.7 reveals no meaningful dynamics related to temperature. While the coefficient on contemporaneous daily temperature is statistically significant at the 10% level, its magnitude is extremely small: we can reject that a 1°C change is associated with a 0.1 pp change in equipment outage rates.

We find a statistically significant correlation between equipment outage rates and the 3-day lag of state-level demand, but statistically insignificant estimates on the corresponding 1-, 2-, 4-, and 5-day lags. The magnitude of the coefficient on the 3-day lag is extremely small: we can reject that a 10% increase in 3-day-lagged quantity demanded is associated with a 0.1 pp increase in equipment outage rates. Given the lack of impacts for the other lags, this correlation likely stems from sampling variability.

Appendix Table A.8 conducts robustness on Column (3) of Table 1, adding 2 monthly lags of temperature, forecasted demand (only observed at the monthly level), and realized quantity demanded. We find no evidence that equipment outage rates are correlated with lagged temperature or lagged forecasted demand. While we recover a statistically significant correlation between equipment outage rates and the 1-month lag of realized quantity demanded, this magnitude is not economically meaningful: we can reject that a 10% increase in 1-month-lagged equilibrium quantity demanded is associated with a 1 pp increase in equipment outage rates. Again, this correlation with the 1-month lag likely stems from sampling variability given the small magnitude and lack of impacts from other lags.

Table A.7: Equipment outage rates do not respond to elec. demand shocks (daily lags)

	Outcome: Share of plant's capacity on equip. outage	
	(1)	(2)
Mean temperature in state (°C)	-0.0005^{*}	-0.0004^{*}
	(0.0003)	(0.0003)
Mean temperature in state (°C), 1-day lag	0.0001	0.0003
	(0.0002)	(0.0002)
Mean temperature in state (°C), 2-day lag	-0.0002	-0.0002
	(0.0002)	(0.0002)
Mean temperature in state (°C), 3-day lag	-0.0000	-0.0002
	(0.0002)	(0.0002)
Mean temperature in state (°C), 4-day lag	0.0000	-0.0000
	(0.0002)	(0.0002)
Mean temperature in state (°C), 5-day lag	-0.0003	-0.0003
	(0.0002)	(0.0003)
log (Quantity demanded in state), 1-day lag		-0.0064
		(0.0043)
log (Quantity demanded in state), 2-day lag		0.0042^{*}
		(0.0025)
log (Quantity demanded in state), 3-day lag		0.0064^{***}
		(0.0022)
log (Quantity demanded in state), 4-day lag		0.0008
		(0.0035)
log (Quantity demanded in state), 5-day lag		0.0006
		(0.0050)
Plant + month-of-sample FEs	Yes	Yes
Region-year, region-month FEs	Yes	Yes
Mean of dep. var.	0.1110	0.1110
Plant-day observations	400,711	400,676

Notes: Regressions are identical to Column (1) of Table 1, except that we add 5 daily lags of temperature and 5 daily lags of the realized quantity of wholesale electricity demanded in state s. We omit the contemporaneous realized quantity demanded because equipment-outage-related cost shocks directly affect the market equilibrium. The dependent variable is plant i's equipment outage rate (i.e. the daily share of its total capacity on equipment outage). Both regressions control for the total number of dispatchable plants in each state. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.8: Equipment outage rates do not respond to elec. demand shocks (monthly lags)

	Outcome: Share of plant's capacity on equip. outage	
	(1)	(2)
Mean temperature in state (°C)	−0.0007	0.0001
	(0.0014)	(0.0014)
Mean temperature in state (°C), 1-month lag	−0.0007	−0.0019
	(0.0014)	(0.0015)
Mean temperature in state (°C), 2-month lag	−0.0015	−0.0019
	(0.0012)	(0.0012)
log (State's forecasted energy req.)	−0.0028	−0.0229
	(0.0145)	(0.0139)
log (State's forecasted energy req.), 1-month lag	0.0063	
	(0.0158)	
log (State's forecasted energy req.), 2-month lag	−0.0080	
	(0.0119)	
log (Quantity demanded in state), 1-month lag		0.0401**
		(0.0191)
log (Quantity demanded in state), 2-month lag		−0.0180
		(0.0163)
Plant + month-of-sample FEs	Yes	Yes
Region-year, region-month FEs	Yes	Yes
Mean of dep. var.	0.1137	0.1140
Plant-month observations	19,023	18,440

Notes: Regressions are identical to Column (3) of Table 1, except that we add 2 monthly lags of temperature, forecasted energy requirement, and realized quantity of wholesale electricity demanded in state s. We omit contemporaneous realized quantity demanded because equipment-outage-related cost shocks directly affect the market equilibrium. The dependent variable is plant i's equipment outage rate (i.e. the daily share of its total capacity on equipment outage). All regressions control for the total number of dispatchable plants in each state. Standard errors are clustered by sample month. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B Structural model appendix

B.1 Further details on our simple structural model

This section provides details on the simple structural model sketched out in Section 5. We discuss the demand side and the supply side of the model in turn.

Demand side For our preferred specification, we assume constant-elasticity wholesale demand, setting a cost elasticity of demand of -0.49 (based on Column (4) of Table 2). For state s on day t, quantity demanded is:

$$Q_{st} \;=\; \alpha_{st}\, AVC_{st}^{-0.49}$$

where average variable cost is defined as:

$$AVC_{st} \;\equiv\; \frac{\sum_{i\in s} MC_{st}\, Q_{st}}{\sum_{i\in s} Q_{it}}$$

We calibrate α_{st} to the observed quantity supplied Q^{*}_{st} and average variable cost AVC^{*}_{st}:

$$\alpha_{st} \;\equiv\; \frac{Q^{*}_{st}}{(AVC^{*}_{st})^{-0.49}}$$

We conduct sensitivity analyses assuming linear demand:

$$Q_{st} \;=\; a_{st} \;+\; b_{st}\, AVC_{st}$$

To set the slopes b_{st}, we assume that the cost elasticity of demand is -0.49 at the observed quantity supplied Q^{*}_{st} and average variable cost AVC^{*}_{st} (i.e., $b_{st} = -0.49\dfrac{Q^{*}_{st}}{AVC^{*}_{st}}$). Then, we calibrate the intercepts $a_{st} = Q^{*}_{st} - b_{st}AVC^{*}_{st}$.

These intercepts play an important role across specifications: for both constant-elasticity and linear demand, we set the choke price (where $Q_{st} = 0$) equal to a_{st}. This means that, even under constant-elasticity demand, we assume that $Q_{st} = 0$ for all $AVC_{st} \geq a_{st}$, in order to avoid consumer surplus being overly-sensitive to the curvature of demand close to $Q_{st} = 0$.

We also conduct sensitivity analyses assuming utilities respond to marginal cost (MC) as opposed to AVC. In this case, the demand curve passes through observed quantity Q^*_{st} and the 98th percentile of marginal cost MC^{98*}_{st}, and we assume an elasticity of -0.39 (based on Column (2) of Appendix Table A.5):

$$\begin{aligned} Q_{st} &= \alpha_{st} \, (MC^{98}_{st})^{-0.39} \\ \alpha_{st} &\equiv \frac{Q^*_{st}}{(MC^{98*}_{st})^{-0.39}} \end{aligned}$$

As above, we set the choke point based on the intercept implied by linear demand with the same elasticity: $b_{st} = -0.39 \times \dfrac{Q^*_{st}}{MC^{98,*}_{st}}$, $a_{st} = Q^*_{st} - b_{st} MC^{98,*}_{st}$. Then, we set $Q_{st} = 0$ for all $MC^{98}_{st} \geq a_{st}$.

For the supply-side interventions considered in Section 5.3, we assume that retail electricity prices do not respond to counterfactual changes in wholesale procurement costs. In other words, we hold the wholesale demand curve fixed when applying these supply-side interventions.

Supply side For each date t, we clear the market for each state s separately, conservatively assuming no interstate trade.[2] We stack all generating capacity that is available (i.e., not on outage) within each state from lowest to highest marginal cost, and then dispatch power plants in this order. Inframarginal plants produce at capacity, where we calculate capacity based on the 98th percentile of generation over the past year. The marginal plant produces output between zero and its capacity, such that quantity demanded implied by the relevant cost (either AVC or MC, depending on specification) is equal to the implied total quantity supplied.

B.2 Sensitivity analysis: Supply-side interventions

Appendix Tables B.1 and B.2 present robustness for our counterfactual simulations. In Appendix Table B.1, we model utilities as responding to average variable cost, but assuming linear demand rather than constant elasticity demand. In Appendix Table B.2, we model

2. By ignoring the potential benefits from reallocating output across states, we likely understate the increases in power supply that would result from decreases in wholesale procurement costs.

utilities as responding to marginal cost rather than average variable cost (using an elasticity of -0.39 from Column 2 of Table A.5), and assuming constant elasticity demand.

Table B.1: Quantity impacts of different supply-side interventions: Linear demand

	(1)	(2)	(3)	(4)
Supply curve scenario:	Least-cost dispatch (LC)	LC + US efficiency	LC + US outage rate	LC + 4 new UMPPs
Quantity increase relative to observed dispatch (GWh/day)	43.3	43.3	43.3	43.3
Incremental quantity increase relative to LC (GWh/day)		57.5	36.5	26.1
Incremental HHs shifted to 24×7 power		123.7M	78.6M	56.2M

Notes: This table reports the daily average national-level quantity impacts of different hypothetical interventions to the supply side of the Indian wholesale electricity market. In contrast with Table 4 in the main text, we assume linear demand rather than constant-elasticity demand. Column (1) reports the change in quantity supplied implied by switching from observed dispatch to dispatching plants from lowest-to-highest marginal cost within state/day. The first row is the same across all three incremental hypothetical supply-side interventions, each of which builds upon this "least-cost" scenario. The second row reports the incremental increase in quantity supplied under least-cost dispatch from: improving the thermal efficiency of Indian coal-fired plants to the average levels observed in the United States (following Chan, Cropper, and Malik (2014); Column (2)); reducing the forced outage rate of Indian coal-fired power plants to U.S. levels (Column (3)); and adding four 4,000 MW Ultra Mega Power Plants (UMPPs) with relatively low marginal costs to the supply curve (Column (4)). The bottom row reports the number of households that could be shifted to 24×7 power from each incremental increase in quantity supplied, assuming that Indian households currently face 3.4 hours per day of blackouts on average (Agrawal et al. (2020)). See text for further details.

As in Table 4, Column (1) presents the difference in quantity supplied between observed dispatch and least-cost dispatch. In our main specification, this difference is 44.9 GWh/day. This difference is very close to our main estimate under linear demand (43.3 GWh/day), and larger when utilities respond to marginal cost (94.1 GWh/day). Columns (2)–(4) present incremental impacts of improving the thermal efficiency of the Indian coal fleet to U.S. levels, applying the U.S. outage rate, and building four new UMPPs, respectively. These yield increases in quantity supplied of 57.5 GWh/day, 36.5 GWh/day, and 26.1 GWh/day under linear demand, and increases of 27.3 GWh/day, 163.1 GWh/day, and 103.5 GWh/day when utilities respond to marginal cost. These increases are broadly similar to the main results in Table 4.

Table B.2: Quantity impacts of different supply-side interventions: Utilities respond to MC

Supply curve scenario:	(1) Least-cost dispatch (LC)	(2) LC + US efficiency	(3) LC + US outage rate	(4) LC + 4 new UMPPs
Quantity increase relative to observed dispatch (GWh/day)	94.1	94.1	94.1	94.1
Incremental quantity increase relative to LC (GWh/day)		27.3	163.1	103.5
Incremental HHs shifted to 24×7 power		58.8M	350.9M	222.7M

Notes: This table reports the daily average national-level quantity impacts of different hypothetical interventions to the supply side of the Indian wholesale electricity market. In contrast with Table 4 in the main text, we assume that utilities respond to the marginal cost of the marginal unit rather than average variable cost. Column (1) reports the change in quantity supplied implied by switching from observed dispatch to dispatching plants from lowest-to-highest marginal cost within state/day. The first row is the same across all three incremental hypothetical supply-side interventions, each of which builds upon this "least-cost" scenario. The second row reports the incremental increase in quantity supplied under least-cost dispatch from: improving the thermal efficiency of Indian coal-fired plants to the average levels observed in the United States (following Chan, Cropper, and Malik (2014); Column (2)); reducing the forced outage rate of Indian coal-fired power plants to U.S. levels (Column (3)); and adding four 4,000 MW Ultra Mega Power Plants (UMPPs) with relatively low marginal costs to the supply curve (Column (4)). The bottom row reports the number of households that could be shifted to 24×7 power from each incremental increase in quantity supplied, assuming that Indian households currently face 3.4 hours per day of blackouts on average (Agrawal et al. (2020)). See text for further details.

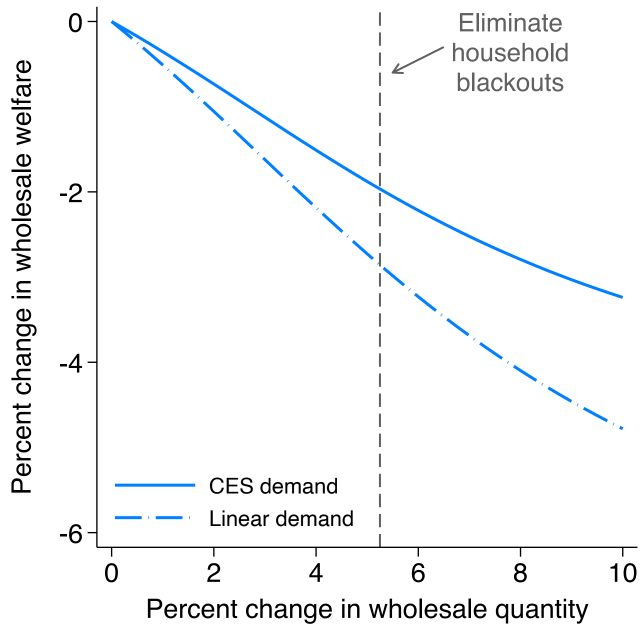
B.3 Additional results: Quantity mandate

Here we present additional results pertaining to the simulated impacts of a quantity mandate (as presented in Section 5.4). Appendix Figure B.1 plots the percentage decrease in wholesale welfare associated with moving from our baseline least-cost scenario to a mandate that $X\%$ more quantity must be supplied (varying X between 0–10%). The vertical line denotes the percentage increase in quantity supplied that would bring all households to 24×7 power (i.e., 105 GWh/day $=$ 266 million households $\times \frac{2.8\text{ kWh}}{24\text{ hours}} \times 3.4\frac{\text{hours}}{\text{day}}$).[3]

This figure shows that even a roughly 5% increase in quantity supplied would be sufficient to bring households to 24×7 power. In addition, we see that wholesale welfare falls at a slower rate at greater increases in quantity supplied. This stems primarily from capacity constraints limiting the amount of quantity that can be produced in a given state-day. For this reason, we focus our discussion on the relatively smaller quantity increases implied by bringing households to 24×7 power.

3. Our model distributes this daily national increase in quantity supplied proportionately across states.

Figure B.1: Changes in wholesale welfare from increases in wholesale quantity



Notes: This figure plots the percentage decrease in wholesale welfare associated with shifting from equilibrium to a quantity mandate that $X\%$ more quantity must be supplied, varying X between 0–10%. We do so under the assumption that wholesale demand is constant-elasticity (solid line) versus under the assumption that wholesale demand is linear (dashed line). The vertical line denotes the percentage increase in quantity supplied that would bring all households to 24×7 power under the constant-elasticity assumption (i.e., 105 GWh/day = 266 million households $\times \frac{2.8 \text{ kWh}}{24 \text{ hours}} \times 3.4 \frac{\text{hours}}{\text{day}}$); this daily national increase in quantity is distributed proportionately across the states. See Sections 5.2 and 5.4 for more details on the model and discussion of the findings, respectively.

Appendix Table B.3 presents the percentage changes in wholesale welfare, consumer surplus, and producer surplus from moving from our least-cost baseline scenario to a quantity mandate that requires 24×7 power for households. Column (1) presents our preferred specification, assuming that demand is constant elasticity (CE), markets clear and power plants are dispatched at the state/day level, and wholesale prices are set based on average variable cost (AVC) of supply. Producer surplus (i.e., generator profits) is zero by construction under AVC pricing (hence the label "NA"); consequently, wholesale welfare is equal to consumer (i.e., utility) surplus.

Column (2) conducts sensitivity analysis assuming linear demand, which has minimal effects on the wholesale welfare losses from a quantity mandate (1.7% versus 2.1% in our preferred specification with CE demand). If utilities were to respond to marginal cost (MC) rather than AVC, as in Column (3)'s sensitivity analysis, wholesale welfare would fall by

Table B.3: Wholesale market impacts of a quantity mandate to eliminate HH blackouts

	(1)	(2)	(3)
Consumer surplus (% change)	−2.1	−1.7	−11.7
Producer surplus (% change)	NA	NA	108.9
Wholesale welfare (% change)	−2.1	−1.7	−0.2
Functional form of demand	CE	Linear	CE
Prices = AVC or MC?	AVC	AVC	MC

Notes: This table reports the changes in consumer surplus, producer surplus, and wholesale welfare from moving from our baseline least-cost scenario to a quantity mandate sufficient to eliminate blackouts among households. We calculate the increase in national daily total quantity required to bring all households to 24×7 power as follows: 105 GWh/day = 266 million households $\times \frac{2.8 \text{ kWh}}{24 \text{ hours}} \times 3.4 \frac{\text{hours}}{\text{day}}$); the daily national increase in quantity supplied is distributed proportionately across the states. Our preferred specification in Column (1) assumes that demand is constant elasticity (CE), markets clear and power plants are dispatched at the state/day level, and wholesale prices are set based on average variable cost (AVC) of supply. Column (2) assumes linear demand (as opposed to CE demand), while Column (3) models wholesale prices set equal to marginal cost (MC, as opposed to AVC). When setting prices based on AVC, producer surplus (i.e, generator profits) is zero by construction (hence "NA" in Columns (1)–(2)). See Sections 5.2 and 5.4 for more details on the model and discussion of the findings, respectively.

much less from a quantity mandate (0.2% versus 2.1% in our preferred specification with AVC pricing). This is because a much larger 11.7% reduction in utility surplus would be almost fully offset by a 109% increase in profits earned by inframarginal power plants.

C Further details on the data

C.1 Constructing marginal costs

For fossil-fuel power plants, we follow the electricity economics literature (Fabrizio, Rose, and Wolfram (2007); Cicala (2022)) in approximating marginal costs as:

$$MC_{it} = \text{Fuel price}_{it} \cdot \text{Heat rate}_{it}$$

We first discuss where we obtain data on heat rates, and then proceed to describe how we construct fuel prices (inclusive of transportation costs and relevant taxes) separately for each type of plant.

Heat rates: A plant's heat rate is defined as the amount of heat input (in kcal) required to produce one MWh of electricity. For coal and lignite plants, we obtained heat rate data from the CEA's annual *Review of Performance of Thermal Power Stations*. We digitized the 2012–2014 *Reviews* (the most recent years available), and we obtained the 1997–2009 data from Chan, Cropper, and Malik (2014).[4] Since our analysis spans 2013–2019, we assign each plant its most recent heat rate observed in our data. For the 16 plants whose most recent heat rate was reported prior to 2012, we obtained more recent heat rate data from tariff petitions to the Central Electricity Regulatory Commission.

For natural gas-fired power plants, we assign heat rates based on the CEA's *Monthly Gas Reports*. These reports are only available for 2012, 2016, and 2017; we assign each plant its average observed heat rate. We follow the Ministry of Natural Gas and Petroleum in assuming that 10,000 kCal of heat energy is contained in one standard cubic meter of natural gas. These data enable us to assign heat rates for 58 of the 62 gas plants in our daily CEA sample.

Coal plants: We construct marginal costs for each coal-fired power plant as follows. We collect grade-specific coal prices reported aperiodically by Coal India Limited and Western

4. We thank the authors for sharing these data.

Coalfields Limited (prices reported in rupees per kg).[5] "Grades" refer to the kilocalories (kcal) of heat energy per ton of coal. We assign "minemouth" coal prices to each power plant based on the grades of coal mined from the coalfield and the geographic proximity of the plant to the coalfield. Nearly all of India's coal-fired power plants buy their coal at grade-specific prices set by the Ministry of Coal through long-term Fuel Supply Agreements.[6]

For geographic proximity, we calculate the distance by rail between coal plants and coalfields. To do so, we combine hand-coded plant latitude/longitude with geospatial data on India's coalfields from the U.S. Geological Survey (Trippi and Tewalt (2011)). Data on the rail network in India is from ML ML Infomap (2008)).

We approximate the grade of coal burned by the plant as follows, using data from the CEA's *Monthly Coal Reports*. First, we divide annual total quantity of electricity produced by each plant (in kWh) by the annual total quantity of coal consumed by each plant (in kg). This annual ratio is multiplied by each plant's heat rate in each year (in kcal/kWh). The resulting quantity is the annual aggregate amount of kcal of input heat energy obtained by the plant from one kg of coal. Taking the mean of this quantity gives us the approximate grade of coal burned by the plant, which ranges from 1,118 to 8,254 kcal/kg for non-lignite coal plants.[7]

Having assigned minemouth coal prices to plants, we next multiply these prices by one plus the royalty rate, the value-added tax, the excise tax, and a cess specific to West Bengal. The royalty rate is 14% for coal mined from all states other than West Bengal; in West Bengal, the royalty adder is applied in rupees per kg rather than percentage.[8] The value-added tax is 2% if the coal comes from out of state but 5% if the coal comes from the same state as the plant. The excise tax is 6% across the nation. West Bengal also charges a 25% tax on coal mined in its state.

5. Coal prices for Coal India Limited are available at: https://www.coalindia.in/Manage/ViewDocumentModule.aspx.

6. These are regulated "pithead" prices, which do not include the cost of transporting coal from mines to plants. The government implemented the "Scheme to Harness and Allocate Kolya (Coal) Transparently in India" policy (a.k.a. Shakti) in September 2017, which allocates *new* coal contracts to privately owned generating units based on an auction mechanism. There were two auctions during our sample period; the winning coal plants made up a very small share of the overall coal-fired capacity in our sample (Chirayil and Sreenivas (2010)).

7. We have heat rate and coal grade data for 84 coal-fired plants and 7 lignite-fired plants, representing approximately 50% and 80% of each fuel's respective generating capacity in CEA's daily generation data.

8. The royalty adder in West Bengal differs based on the grade of coal, ranging from Rs 4.5/1,000 kg to Rs 8.5/1,000 kg; further details are available upon request.

We next add transportation charges, additional taxes, stowing duty, and the West Bengal specific royalty adder to the minemouth price. Transportation charges, assessed in rupees per kg, vary both over time and by distance between mine and plant. We collect rail rates from the Indian Railway website, calculating the relevant distance between plant and coalfield as discussed above.[9] The majority of power plants receive coal from trains. The remaining two major categories are "pithead" plants colocated next to a mine (for whom transportation charges are zero) and plants who burn imported coal. In the absence of high quality data on the coal prices paid by plants burning imported coal, we assign these plants a domestic coal price based on the grade of coal closest to the one they actually burn.

India also charged a "clean energy" cess per kg of coal purchased, which we add to the minemouth price.[10] Finally, the Ministry of Coal charges a Rs 10/1,000 kg "stowing excise duty" related to the "assessment and collection of excise duty levied on all raw coal..."[11]

To convert coal prices from rupees per kg to rupees per kWh, we multiply the relevant price by the plant's aggregate quantity of electricity produced (in kWh) and divide by the plant's aggregate quantity of coal consumed (in kg).

Lignite plants: We obtain the lignite coal price per kg from the Central Electricity Regulatory Commission.[12] All lignite plants in India are colocated next to their source mine, so transportation costs are zero. After multiplying or adding the relevant royalties, taxes, and clean energy cess discussed above for coal plants, we multiply by an estimate of the heat content of lignite coal (in kcal per kg) from the same source as the price. Finally, we multiply the lignite coal price (now in rupees per kcal) by the plant's heat rate to obtain the marginal cost (in rupees per kWh) for each lignite plant.

Gas plants: For natural gas plants, we use gas prices originally reported in rupees per 1,000 cubic meters. We assume that 1 cubic meter of natural gas contains 10,000 kcal of

9. For example, the freight rate relevant for dates after November 1, 2018 is available here: http://www.indianrailways.gov.in/railwayboard/uploads/directorate/traffic_comm/downloads/Freight_Rate_2018/RC_19_2018.PDF

10. The Clean Energy Cess was replaced by the GST Compensation Cess in July 2017. Information on the history of the Clean Energy Cess is available at: http://iisd.org/sites/default/files/publications/stories-g20-india-en.pdf

11. Many of the taxes and subsidies relevant to the coal sector in India are discussed here: https://www.eria.org/uploads/media/07_RPR_FY2018_15_Chapter_6.pdf

12. The data are here: http://cercind.gov.in/2017/orders/255.pdf

heat energy, using this conversion factor to obtain gas prices in rupees per kcal. Finally, we multiply this price by each plant's heat rate (in kcal per kWh) to get each gas plant's marginal cost. Though this marginal cost does not include the costs associated with transporting gas, they are in line with the estimates reported by the Ministry of Power, which do include these costs.[13]

Nuclear plants: We assign each of the 7 nuclear plants in our sample a marginal cost based on tariff documents.[14]

Hydro, wind, and solar plants: Non-dispatchable run-of-river hydroelectric, wind, and solar resources have near-zero marginal cost. Dispatchable hydro generators face a complex dynamic optimization problem, as generation today may come at the expense of generation tomorrow due to a finite supply of water (Archsmith (2024)). Thus, we exclude hydro, wind, and solar resources from the analysis, implicitly assuming that they are inframarginal.

C.2 Power plant outages

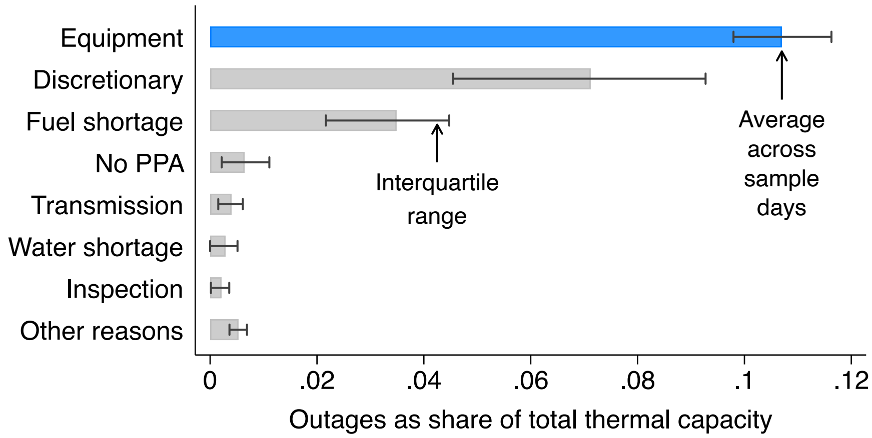
We string parse the CEA's *Daily Outage Reports* to classify eight mutually exclusive outage categories: equipment, discretionary, fuel shortage, no power purchase agreement (PPA), transmission problem, water shortage, inspection, and other. Figure C.1 shows the frequencies of these categories, plotting the share of total thermal plant capacity on each type of outage on the average day. Our empirical analysis only uses the equipment category.

Appendix Figure C.2 characterizes the duration of equipment outages during our sample period. The left panel shows that the median equipment outage lasts just 2 days, while 95% of equipment outages are shorter than 33 days long. This supports our assumption that equipment outages are short-lived exogenous shocks to utilities' wholesale procurement costs.

13. The average marginal cost per kWh we construct using data on gas prices is 2.09 while the corresponding average for the marginal costs reported by the Ministry of Power is 2.42.

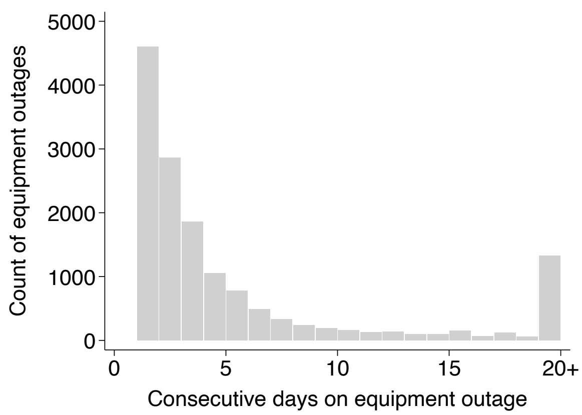
14. These data are reported in the following article by the chairman of an expert committee for the Department of Atomic Energy: http://www.thehindu.com/todays-paper/tp-opinion/Why-India-should-opt-for-nuclear-power/article14850892.ece

Figure C.1: Descriptive statistics on daily plant outages



Notes: Each bar shows the average share of total thermal capacity that reported an outage of a specific category. Bars report averages across 2,453 sample days, while whiskers report the interquartile range of daily outage shares. We manually classify outages into these categories using the reasons listed in the CEA's *Daily Outage Reports*.

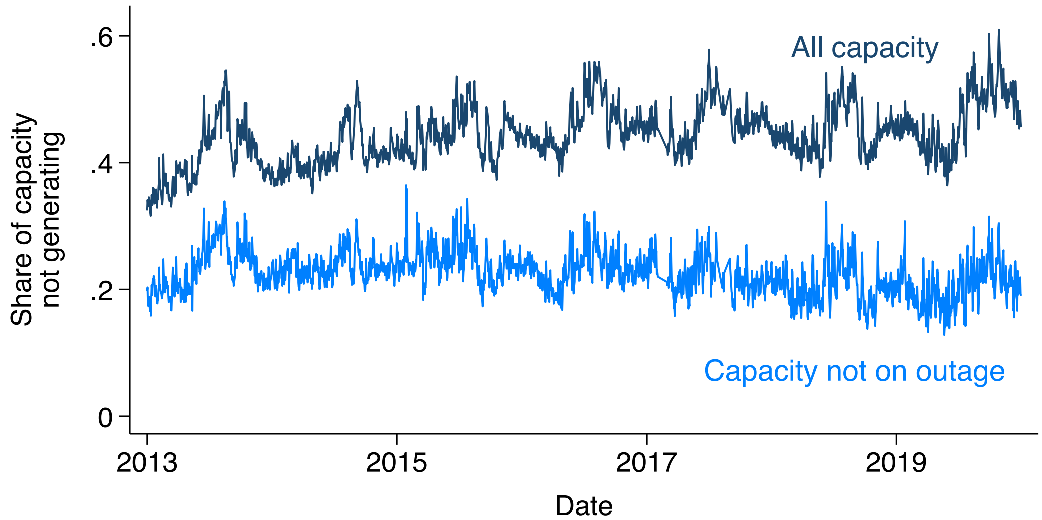
Figure C.2: Distribution of outage durations



Notes: This histogram summarizes the length of equipment outages; each observation is a set of consecutive days where a plant reports some capacity on equipment outage. During our sample period, the median equipment outage lasted 2 days.

Appendix Figure C.3 plots unused power plant generating capacity over our sample period. The navy line reports 1 − (generation / nameplate capacity), while the blue line reports 1 − (generation / nameplate capacity less outages). On average, 44% of all generating capacity is not being used, with 22% of capacity not on outage sitting idle. This illustrates that India has substantial excess generating capacity.

Figure C.3: Time series of unused generating capacity



Notes: This figure plots unused generating capacity in the Indian electricity sector. The navy line plots 1 − (generation / nameplate capacity), and the light blue line plots 1 − (generation / nameplate capacity less declared outages). On average, 44% of all capacity, or 22% of capacity not on outage, is unused on the average day.

C.3 Indian Energy Exchange (IEX) data

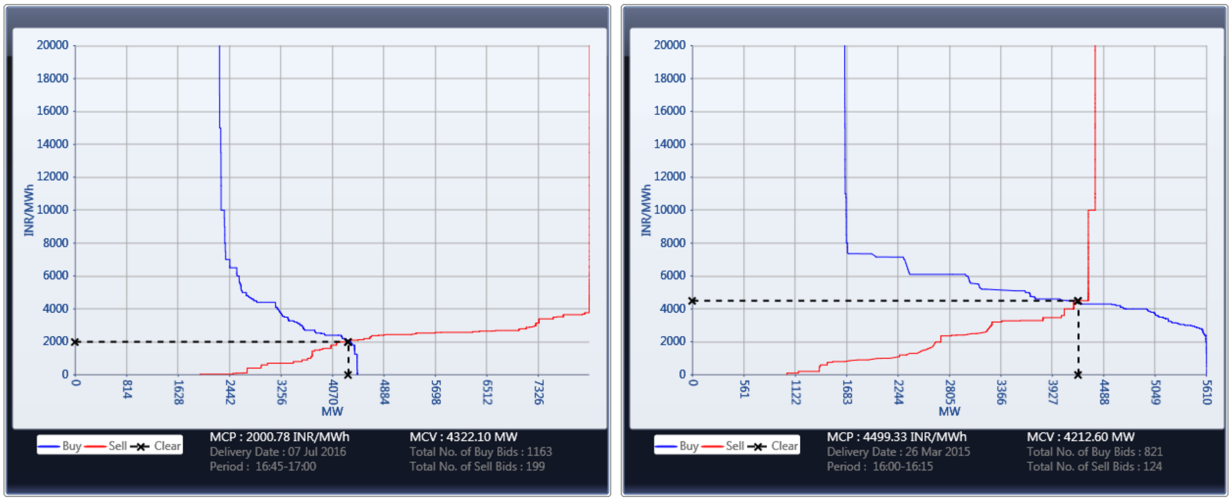
The Indian Energy Exchange (IEX) runs uniform-price auctions, where electricity suppliers submit offer curves, buyers (e.g., utilities) submit demand bid curves, and the market clears by aggregating supply and demand. Prices and quantities from the unconstrained market clearing process are adjusted to reflect transmission constraints. This results in separate prices and quantities for each 15-minute interval for each of India's five transmission regions.

The IEX publishes `.jpeg` images of the aggregate supply and demand curves for each 15-minute interval-of-sample. We downloaded these data from April 1st, 2014 through December 31st, 2019. We converted these images into data using the online WebPlotDigitizer tool (https://automeris.io/WebPlotDigitizer/). To do this, we upload the image and then label four points, which allows the software to convert the image into data on the price-quantity steps displayed for the aggregate supply and demand curves.[15] Appendix Figure C.4 presents two of the 201,012 15-minute intervals in our dataset.

The IEX also provides market clearing price and quantity data for each 15-minute in-

15. These images are available from the following link: https://www.iexindia.com/marketdata/demandsupply.aspx.

Figure C.4: Example IEX demand and supply curves



Notes: This figure displays two examples of the raw data we obtained from the Indian Energy Exchange. The left image shows the aggregate demand and supply curves for the 16:00–16:15 interval on March 26, 2015. The right image shows the same curves for the 16:45–17:00 interval on July 7, 2016. We digitized these images, originally in JPEG format, using OCR software.

terval for each of India's five transmission regions.[16] Across our sample, the average IEX market clearing price was Rs 3,121/MWh, while the average volume cleared was 1,128 MWh per 15-minute interval. We compare the equilibrium outcomes implied by our converted images to those provided by the IEX. The correlation between the two is extremely high—99.8%—which gives us confidence that the image conversion is working properly. We use these digitized interval-specific demand curves to calculate the price elasticity of IEX demand.[17]

C.4 Inflation adjustment

When relevant, all magnitudes are reported in 2016 constant rupees. We adjust for inflation using the monthly consumer price index for all items for India reported by the Organization for Economic Co-operation and Development.[18]

16. The price data are available from https://www.iexindia.com/marketdata/areaprice.aspx. The quantity data are available from https://www.iexindia.com/marketdata/areavolume.aspx.

17. To construct the elasticity at a given price-quantity point for each interval-specific demand curve, we smooth the demand curve and compute the "finite central difference" elasticity implied by moving Rs 5/MWh up versus moving Rs 5/MWh down the demand curve.

18. Data can be accessed here: https://fred.stlouisfed.org/series/INDCPIALLMINMEI

Appendix references

Agrawal, Shalu, Sunil Mani, Abhishek Jain, and Karthik Ganesan. 2020. *State of electricity access in India: Insights from the India Residential Energy Survey (IRES) 2020.* Council on Energy, Environment, and Water.

Archsmith, James. 2024. "Dam spillovers: Direct costs and spillovers from environmental constraints on hydroelectric generation." Working paper.

Chan, Hei Sing (Ron), Maureen L. Cropper, and Kabir Malik. 2014. "Why are power plants in India less efficient than power plants in the United States?" *American Economic Review* 104 (5): 586–590.

Chirayil, Maria, and Ashok Sreenivas. 2010. *Coal auctions still far from transparent.* https://www.thehindubusinessline.com/opinion/coal-auctions-still-far-from-transparent/article29281446.ece.

Cicala, Steve. 2022. "Imperfect markets versus imperfect regulation in U.S. electricity generation." *American Economic Review* 112 (2): 409–441.

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram. 2007. "Do markets reduce costs? Assessing the impact of regulatory restructuring on US electric generation efficiency." *American Economic Review* 97 (4): 1250–1277.

Indian Energy Exchange. 2014–2019. *Aggregate Demand Supply, April 2014–December 2019.* https://www.iexindia.com/market-data/real-time-market/aggregate-demand-supply.

ML Infomap. 2008. *Railroads: India, 2001.*

Trippi, M.H., and S.J. Tewalt. 2011. *Geographic Information System (GIS) Representation of Coal-Bearing Areas in India and Bangladesh,* U.S. Geological Survey Open-File Report 2011-1296.