Appendices E–G for: Market Power in Coal Shipping and Implications for U.S. Climate Policy

Louis Preonas

February 6, 2023

Appendices A–D available at: https://www.louispreonas.com/s/preonas_coal_appxABCD.pdf

E Empirical appendix for markup regressions

E.1 Nearest-neighbor matching

I match each captive plant to the k non-captive plants with the closest geographic proximity, enforcing a maximum distance of 200 miles between matched plants. I force exact matches on the preferred (modal) coal rank that each plant consumed between 2002–2006, in order to ensure that matched plants do not purchase coal from predominantly opposite sides of the country (i.e., bituminous coal from the east vs. sub-bituminous coal from the west). I also exclude the few plants with covariates that do not overlap with the opposite group (e.g., a captive plant that is older than *all* non-captive plants).

Formally, let $D_j = 1$ if plant j is captive and $D_j = 0$ if plant j is non-captive. Let $\mathcal{N}_j(k)$ be the set of non-captive matches for captive plant j, given a maximum number of matches $k \in \{1, 3, 5\}$. I construct nearest-neighbor weights $w_j(k)$ for non-captive plants by summing the inverse of the number of matches $|\mathcal{N}_l(k)|$ across all captive plants l:

$$w_j(k) = \begin{cases} 1 & \text{if } \; D_j = 1\,, \; |\mathcal{N}_j(k)| \in \{1, \ldots, k\} \\ 0 & \text{if } \; D_j = 1\,, \; |\mathcal{N}_j(k)| = 0 \\ \displaystyle\sum_{l \,|\, D_l = 1\,,\, j \in \mathcal{N}_l(k)} \frac{1}{|\mathcal{N}_l(k)|} & \text{if } \; D_j = 0\,, \; j \in \mathcal{N}_l(k) \text{ for some } l \\ 0 & \text{if } \; D_j = 0\,, \; j \notin \mathcal{N}_l(k) \text{ for all } l \end{cases}$$

All matched captive plants have weights $w_j(k) = 1$, and all unmatched plants have weights $w_j(k) = 0$. I adjust weights for matched non-captive plants to account for multiple matches.

Pruning and re-weighting the distribution of coal plants enables me to causally estimate the differential markups faced by captive plants. Using the potential outcomes notation of Rubin (1974), let $P_j(0)$ represent plant j's delivered coal price if it were not captive. Then,

$$P_j \;=\; P_j(0) \;+\; \tau_j D_j \tag{E1}$$

where τ_j is the causal effect of captiveness on markups. Since I do not observe $P_j(0)$ for captive plants, the empirical analog of Equation (E1) may suffer from selection bias:

$$P_{ojms} \;=\; \underbrace{E[P_{ojms}(0)|D_j = 0]}_{\beta_0} + \tau D_j \;+\; \underbrace{\Big(E[P_{ojms}(0)|D_j = 1] - E[P_{ojms}(0)|D_j = 0]\Big)}_{\text{unobserved selection bias}} + \varepsilon_{ojms}$$

I can use regression adjustment to reduce selection bias, by controlling for coal commodity characteristics, shipping costs, coal county fixed effects, and month-of-sample fixed effects:

$$P_{ojms} = \tau_j D_j + \beta_C \mathbf{C}_{ojms} + S(\mathbf{T}_{ojms}; \beta_T) + \eta_o + \delta_m + \ldots$$
$$\ldots \underbrace{\Big(E[\tilde{\mu}_{ojms}(0)|D_j = 1] - E[\tilde{\mu}_{ojms}(0)|D_j = 0] \Big)}_{\text{unobserved selection bias}} + \varepsilon_{ojms} \quad \text{(E2)}$$

If these cost controls are not misspecified, the remaining unobserved selection bias only pertains to the markup component of the price, $\tilde{\mu}_{ojms}$.[1]

To identify τ and recover causal estimates of differential markup levels, I applying nearest-neighbor regression weights $w_j(k)$. Under the assumptions of the matching estimator, these weights adjust the distribution of plants to eliminate the unobserved selection bias term in Equation (E2). Equation (4) also directly controls for plant covariates to eliminate selection bias in observables, in the style of a doubly robust estimator (Wooldridge (2007)).

Table E1 reports captive vs. non-captive summary statistics for the full sample of coal plants, separately for plants west vs. east of the Mississippi River. This split reduces the imbalances in the first three columns of Table 1, which are driven by broad geographic factors.

Table E1: Summary statistics – western vs. eastern coal plants

	All plants west of Mississippi River			All plants east of Mississippi River		
	Captive	Non-captive	Difference	Captive	Non-captive	Difference
Coal-fired capacity (MW)	765.01	722.87	42.14	842.23	772.65	69.57
	(663.85)	(672.90)	[0.71]	(800.34)	(714.69)	[0.46]
Number of coal units	1.98	2.05	−0.07	2.69	2.79	−0.10
	(1.08)	(1.08)	[0.70]	(1.43)	(1.74)	[0.62]
Vintage (year)	1973.10	1966.04	7.07	1965.07	1961.89	3.18
	(11.95)	(14.01)	[0.00]***	(14.46)	(13.01)	[0.06]***
Annual capacity factor	0.71	0.66	0.04	0.56	0.57	−0.01
	(0.13)	(0.16)	[0.08]*	(0.17)	(0.17)	[0.59]*
Heat rate (MMBTU/MWh)	11.09	11.46	−0.37	11.09	10.94	0.15
	(1.05)	(1.83)	[0.13]	(1.65)	(1.39)	[0.42]
Scrubber installed (1/0)	0.47	0.38	0.09	0.26	0.26	0.00
	(0.50)	(0.49)	[0.30]	(0.44)	(0.44)	[0.97]
Electricity market participant (1/0)	0.36	0.51	−0.14	0.60	0.77	−0.17
	(0.48)	(0.50)	[0.09]*	(0.49)	(0.42)	[0.00]*
Coal bought (million MMBTU/year)	52.12	49.99	2.13	45.81	42.08	3.73
	(44.79)	(49.75)	[0.79]	(50.63)	(41.56)	[0.52]
Sulfur content (lbs/MMBTU)	0.59	0.71	−0.12	1.13	1.13	0.00
	(0.39)	(0.74)	[0.21]	(0.66)	(0.78)	[0.97]
Ash content (lbs/MMBTU)	8.40	10.03	−1.63	8.51	8.61	−0.10
	(4.31)	(7.77)	[0.11]	(4.15)	(8.38)	[0.91]
Share spot market	0.13	0.07	0.06	0.23	0.23	0.01
	(0.27)	(0.15)	[0.13]	(0.31)	(0.26)	[0.88]
Share sub-bituminous	0.75	0.65	0.10	0.10	0.20	−0.11
	(0.40)	(0.44)	[0.15]	(0.27)	(0.35)	[0.01]
Average rail distance (miles)	661.86	712.73	−50.87	464.42	594.64	−130.22
	(400.95)	(405.32)	[0.50]	(350.00)	(418.89)	[0.01]
Non-rail plants	11	8	19	6	6	12
Utility plants	79	50	129	69	126	195
Total plants	90	58	148	100	182	282

Notes: This table presents summary statistics for all coal plants west vs. east of the Mississippi River, without imposing any sample restrictions or matching weights. It is otherwise identical to the left three columns of Table 1. Standard deviations are in parentheses, and p-values are in brackets. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

1. The tilde on $\tilde{\mu}_{ojms}$ signals that other unobservables may create selection bias in prices, besides markups.

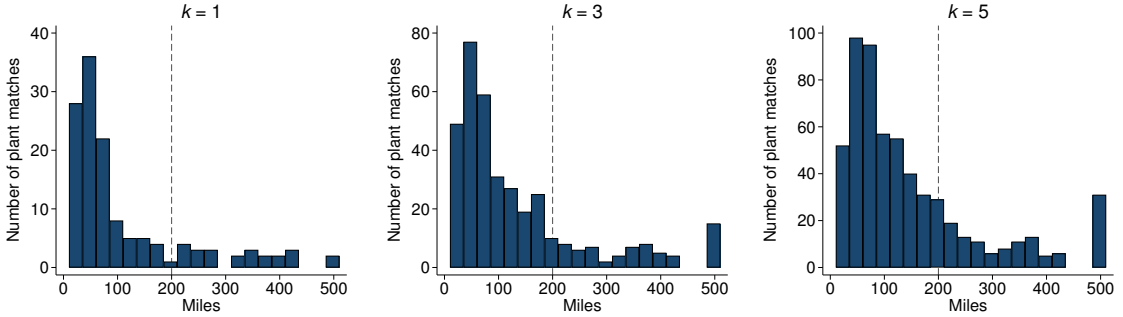
Table E2 supplements the right three columns of Table 1 using matched samples with $k = 1$ and $k = 5$. Figure E1 shows that most captive plants have multiple non-captive neighbors within 200 miles, and that most nearest-neighbor matches are less than 200 miles apart.

Table E2: Summary statistics – sensitivity to the number of matches

	Matched sample ($k = 1$)			Matched sample ($k = 5$)		
	Captive	Non-captive	Difference	Captive	Non-captive	Difference
West of Mississippi River (1/0)	0.44	0.44	−0.00	0.44	0.37	0.07
	(0.05)	(0.08)	[1.00]	(0.05)	(0.06)	[0.38]
Coal-fired capacity (MW)	802.37	831.74	−29.37	802.37	801.21	1.16
	(67.07)	(138.95)	[0.85]	(67.07)	(86.62)	[0.99]
Number of coal units	2.58	2.62	−0.04	2.58	2.53	0.04
	(0.15)	(0.22)	[0.88]	(0.15)	(0.15)	[0.83]
Vintage (year)	1966.24	1961.00	5.24	1966.24	1963.29	2.95
	(1.39)	(1.84)	[0.02]**	(1.39)	(1.38)	[0.13]
Annual capacity factor	0.63	0.63	0.00	0.63	0.63	−0.00
	(0.01)	(0.02)	[0.99]	(0.01)	(0.01)	[0.88]
Heat rate (MMBTU/MWh)	10.98	10.94	0.04	10.98	10.76	0.22
	(0.14)	(0.24)	[0.89]	(0.14)	(0.12)	[0.22]
Scrubber installed (1/0)	0.26	0.25	0.01	0.26	0.27	−0.01
	(0.05)	(0.06)	[0.88]	(0.05)	(0.05)	[0.88]
Electricity market participant (1/0)	0.45	0.49	−0.05	0.45	0.47	−0.02
	(0.05)	(0.08)	[0.64]	(0.05)	(0.06)	[0.77]
Coal bought (million MMBTU/year)	46.63	45.27	1.35	46.63	43.23	3.40
	(4.13)	(7.73)	[0.88]	(4.13)	(4.75)	[0.59]
Sulfur content (lbs/MMBTU)	0.79	0.86	−0.07	0.79	0.88	−0.09
	(0.06)	(0.10)	[0.56]	(0.06)	(0.07)	[0.35]
Ash content (lbs/MMBTU)	8.00	7.98	0.02	8.00	7.68	0.32
	(0.36)	(0.53)	[0.98]	(0.36)	(0.27)	[0.47]
Share spot market	0.19	0.15	0.03	0.19	0.17	0.02
	(0.03)	(0.03)	[0.44]	(0.03)	(0.02)	[0.63]
Share sub-bituminous	0.42	0.40	0.02	0.42	0.38	0.04
	(0.05)	(0.08)	[0.87]	(0.05)	(0.05)	[0.63]
Average rail distance (miles)	573.73	600.32	−26.59	573.73	586.29	−12.56
	(40.26)	(58.91)	[0.71]	(40.26)	(39.66)	[0.82]
Non-rail plants	0	0	0	0	0	0
Utility plants	87	55	142	87	108	195
Total plants	87	55	142	87	108	195

Notes: This table presents summary statistics for matched samples with $k = 1$ and $k = 5$ nearest-neighbor matches. It is otherwise identical to the right three columns of Table 1. Standard deviations are in parentheses, and p-values [in brackets] are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure E1: Distance to nearest k neighbors



Notes: Each panel displays a histogram of the distances to each captive plant's k nearest neighbors, with exact matches on coal rank, and removing non-utility and non-rail plants. My analysis excludes all matches greater than 200 miles.

E.2 Robustness: markup levels

Table E3 estimates markup levels by interacting rail captiveness with an indicator for a coal-by-barge option. I recover differential markups of up to \$5/ton for the least competitive group (captive, no barge option) compared to the most competitive group (non-captive, barge option). Columns (4)–(6) use a "balanced panel" of plants who report at least 1 coal delivery in each year from 2002 to 2015.[2] This yields larger point estimates, and shows that endogenous plant retirement is not biasing my results upwards. Olley and Pakes (1996) demonstrate that bias due to endogenous exit may remain even after balancing the panel, if exit is correlated with unobserved firm productivity. This is not a concern in my setting, as I directly control for each plant's productivity (i.e. its inverse heat rate) in $\mathbf{X}_{jm}$. The magnitudes in Table E3 are consistent with Table 2: we should expect larger markup estimates when the comparison group is more competitive. They also imply that markup levels could be as high as 13% of delivered prices and 34% of railroad's average costs per ton (i.e., the price gap from origin to destination).

Figure E2 reports sensitivity analysis for my estimates of markup levels. Each sensitivity analysis alters a single element of Equation (4), for my preferred specification reported in Column (2) of Table 2. Panel A uses a "balanced panel" of non-retiring coal plants, which increases the magnitude of my differential markup estimates for captive plants.

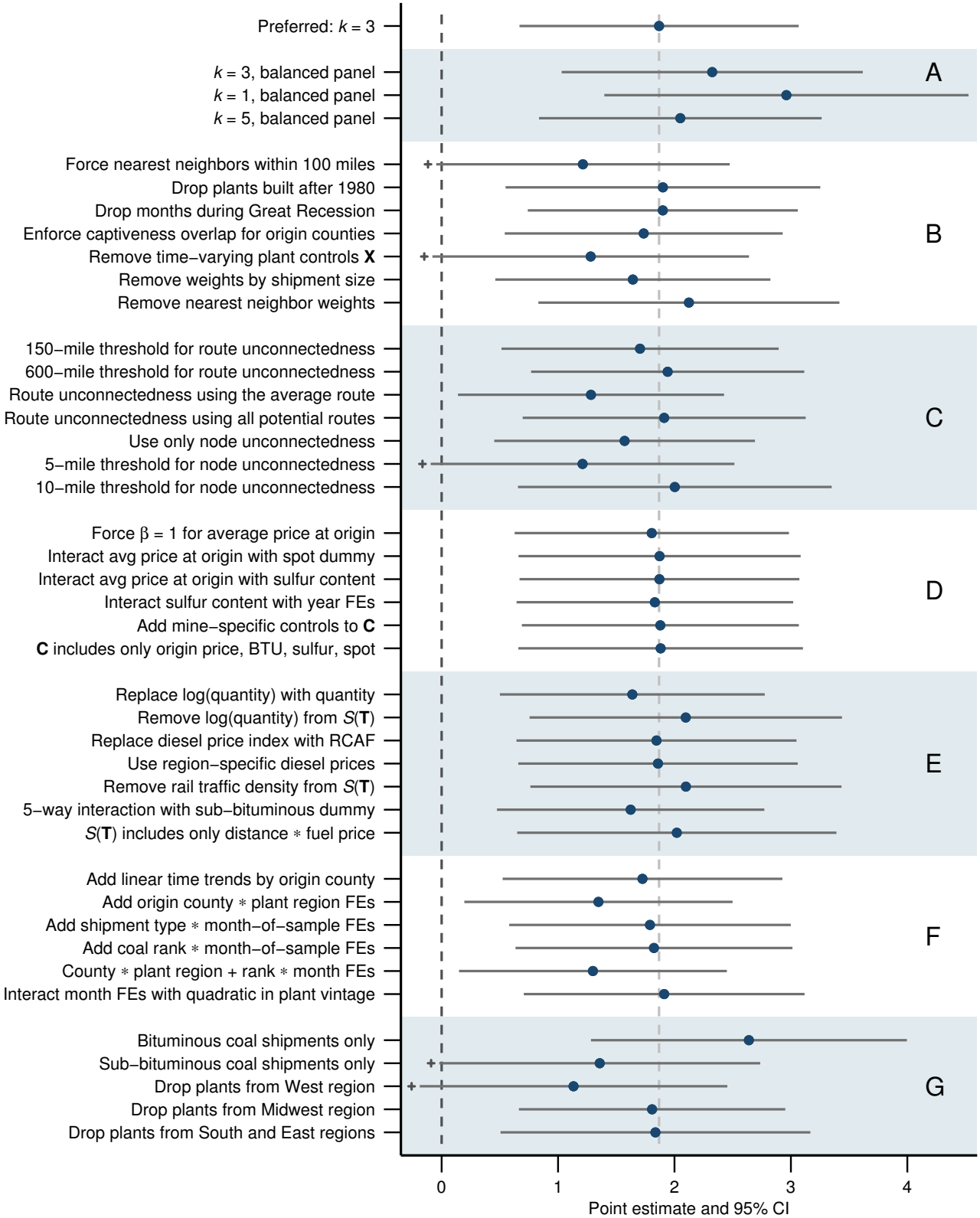
Table E3: Markup levels – interacting rail captiveness with coal-by-barge option

	Outcome: delivered coal price (\$/ton)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}$[Captive, no barge option]$_j$	3.89***	2.28***	2.25***	4.82***	2.89***	2.74***
	(1.04)	(0.83)	(0.76)	(1.08)	(0.89)	(0.82)
$\mathbf{1}$[Captive, barge option]$_j$	2.40**	0.86	0.82	3.22***	1.44	1.29
	(1.03)	(0.82)	(0.78)	(1.04)	(0.90)	(0.85)
$\mathbf{1}$[Non-captive, no barge option]$_j$	2.16*	0.01	0.28	2.71*	0.27	0.51
	(1.28)	(0.94)	(0.92)	(1.42)	(1.02)	(1.00)
Markup as % of delivered price	10.1%	6.1%	5.9%	12.7%	7.9%	7.4%
Markup as % of spatial price gap	27.2%	15.7%	15.5%	33.9%	20.0%	18.9%
k nearest neighbors	1	3	5	1	3	5
Balanced panel of plants				Yes	Yes	Yes
Avg price in omitted group (\$/ton)	38.59	37.31	37.94	38.13	36.54	37.11
Plants	142	184	195	96	126	133
Observations	66,336	88,529	93,968	57,829	76,708	81,547

Notes: Regressions are identical to Table 2, except that I interact rail captiveness with an indicator for whether plant j has the option to receive waterborne deliveries. The omitted group has the most competitive shipping regime: multiple rail carriers plus a barge option. Columns 4–6 use only plants that report at least 1 coal delivery in each sample year (2002–2015). Rows 4–5 divide point estimates for $\mathbf{1}$[Captive, no barge option]$_j$ by the omitted groups' average delivered price and spatial price gap. See notes under Table 2 for further details. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

2. Coal shipments are lumpy, and many active plants do not report deliveries in each month; I "balance" the panel to mitigate any confounding effects from plant exit, not to take a stand on the timing of coal deliveries. Throughout my analysis, I drop the few plants constructed after 1999.

Figure E2: Sensitivities – markup levels



Notes: This figure plots sensitivity analyses for my markup levels regressions. Each dot represents the estimated coefficient on the captiveness dummy in Equation (4), for a separate sensitivity. Unless otherwise noted, each sensitivity uses $k = 3$ nearest neighbor matches as in Column (2) of Table 2 (reported at the top of this chart). See surrounding text for a description of each sensitivity, following the order of the vertical axis from top to bottom.Whiskers denote 95% confidence intervals, with standard errors clustered by plant. Pluses denote statistical significance at the 10% level, but not at the 5% level.

Panel B conducts six sensitivities relating to my identifying assumptions for Equation (4). First, I tighten the nearest-neighbor matching threshold to 100 miles, removing the 14 matched captive plants with the greatest distance to their non-captive counterparts; the resulting point estimate is slightly attenuated but still weakly significant. Second, I restrict the sample to plants built before 1980, when the Staggers Act effectively legalized price discrimination; this yields nearly identical results. Third, I drop sample months during the Great Recession, a period of large commodity price fluctuations; this yields nearly identical results. Fourth, I enforce overlap in captiveness by coal county $\times$ month, dropping the 10% of *om* cells lacking both matched captive and non-captive plant; the resulting point estimate is quite similar, assuaging concerns that county-level commodity controls in $\mathbf{C}_{ojms}$ are mistakenly capturing localized supply effects (as opposed to commodity cost). Fifth, I remove plant-specific controls $\mathbf{X}_{jm}$ from the estimation sample, which are not strictly necessary for identification under the assumptions of nearest-neighbor matching; while this attenuates the point estimate and slightly reduces precision (as expected), it does not meaningfully alter my findings. Finally, I test the importance of my (interacted) regression weights: alternatively, (i) removing coal quantity weights, leaving only nearest-neighbor weights and treating large and small coal transactions equally; and (ii) removing nearest-neighbor weights, leaving only coal quantity weights and including *all* coal plants in Figure 4. In each case, my point estimate changes only slightly.

Panel C estimates Equation (4) using alternative definitions of D_j. Appendix F describes how I construct rail captiveness, which necessitates imposing an arbitrary cutoff of 300 miles for "route unconnectedness", or the distance a plant's shortest route must increase for the route to become "unconnected" after removing the modal carrier along the route. I classify a plant as captive if it becomes route-unconnected (after removing *each* route's modal carrier one-by-one) from *all* observed trading partners. My markup estimates are robust to halving this threshold to 150 miles, or doubling it to 600 miles. Next, I weaken my definition of captiveness, such that only the *average* shipment of coal need become route-unconnected; this prevents largely un-utilized routes (i.e. for singleton coal shipments) from influencing a plant's captiveness designation, yet has little effect my results. My results are also robust to strengthening the definition of captiveness to include routes from all *potential* originating coal counties with similar coal attributes to a plant's observed purchases; this prevents a plant from being designated as captive if it *could have* purchased coal from a county from which it does *not* become route-unconnected.[3] I find similar results if I redefine captiveness based only on "node unconnectedness", a more straightforward (though less nuanced) distinction that ignores coal routes; here, a plant is captive if removing any single Class I rail carrier renders that plant unconnected to any rail node. Finally, I conduct sensitivity analyses for my threshold for "node unconnectedness". My preferred cutoff is 6.6 miles, which is the 95th percentile of the distribution of plants' distance to their nearest terminal rail node; I define a plant as captive if all rail nodes within a 6.6-mile radius are controlled by a single Class I rail carrier. My results are similar if I strengthen this threshold to 5 miles, or weaken it to 10 miles.

Panel D conducts sensitivity analysis on coal commodity controls (i.e. $\mathbf{C}_{ojms}$ in Equation (4)). First, I force the coefficient on the county-year average mine-mouth price to be 1. Next,

3. One might worry that my preferred definition of captiveness would create many false positives. If a coal plant only purchases coal from county A (from which it becomes route-unconnected), but it *could have* purchased identical coal from county B (from which it does *not* become route-unconnected), then I would want to classify this plant as non-captive.

I interact this price with a spot market dummy, to let mine-mouth prices vary for contract vs. spot transactions. Since sulfur content drives coal price dispersion, I interact the county-year average mine-mouth price with each shipment's average sulfur content. I also estimate a different sulfur coefficient for each year, to allow for changes in the shadow price of SO_2 emissions. While Equation (4) includes coal county fixed effects, time-varying factors relating to county-specific coal production might create unobserved trends in commodity cost; I test for this possibility by adding several time-varying coal production controls.[4] Finally, I test a parsimonious version of $\mathbf{C}_{ojms}$ including only average mine-mouth price, BTU content, sulfur content, and a spot market dummy. Each of these sensitivities yields virtually identical results.

Panel E conducts sensitivity analysis on shipping cost controls (i.e. $S(\mathbf{T}_{ojms})$ in Equation (4)). My preferred specification flexibly controls for the 4-way interaction of: (i) shortest rail shipping distance (from my rail graph algorithm in Appendix F), (ii) monthly diesel fuel cost index (to capture time series variation in shipping costs), (iii) the log quantity of coal shipped (to allow for non-constant returns to scale in rail freight), and (iv) the share of route-miles reporting high rail traffic density (to account for congestion). Rail capacity constraints could create high shipping costs for very large shipments; to test for this, I replace log quantity with shipment quantity in levels (short tons). I also remove shipment size controls, in case markups are nonlinear in quantity. Next, I replace the AAR diesel cost index with the STB's Rail Cost Adjustment Factor (RCAF), which incorporates changes in non-fuel variable costs of rail shipping.[5] I also replace the AAR diesel cost index with region-specific diesel prices, in order to allow for geographic variation in transport fuel costs. Next, I remove rail traffic density from $S(\mathbf{T}_{ojms})$, in case it mismeasures congestion costs. Finally, I add a fifth interaction with coal rank (i.e. a sub-bituminous dummy), which effectively eliminates any potential unobservable differences between western vs. eastern rail shipping costs. My results are quite robust to each of these alternative versions of $S(\mathbf{T}_{ojms})$.

Taken together, Panels D–E demonstrate that the cost controls in Equation (4) are unlikely to be misspecified in a way that biases my estimates of average markups. Because my estimates are not sensitive to changes in either $\mathbf{C}_{ojms}$ or $S(\mathbf{T}_{ojms})$, this supports my interpretation of the estimated coefficient $\hat{\tau}$ as the average difference in *markups*, rather than simply the average difference in conditional coal price.

Panel F reports sensitivities using alternative fixed effects. My results are robust to adding (i) county-specific linear time tends, (ii) origin-county by plant-region fixed effects, (iii) month-of-sample fixed effects interacted with shipment type (i.e. contract vs. spot), (iv) month-of-sample fixed effects interacted with coal rank (i.e. bituminous vs. sub-bituminous), and (v) both (iii) and (iv). Since plant vintage is the only (weakly) imbalanced covariate after nearest-neighbor matching (see Tables 1 and E2), I also estimate Equation (4) interacting month-of-sample fixed effects with vintage and vintage squared—which yields nearly identical results.

Finally, Panel G estimates split samples based both on coal rank (bituminous vs. sub-bituminous),and by plant region (West, Midwest, and South/East). Average differential markup

4. These county-by-year controls include mine age, seam thickness and depth (which influence extraction costs), the share of coal produced from (more expensive) underground mines, the share of mine employees working underground (which increases labor costs), and hours worked per ton of coal produced. I weight-average each variable based on quarterly production across all mines in each county. As the composition of production shifts across mines, cause controls like seam depth will become time-varying.

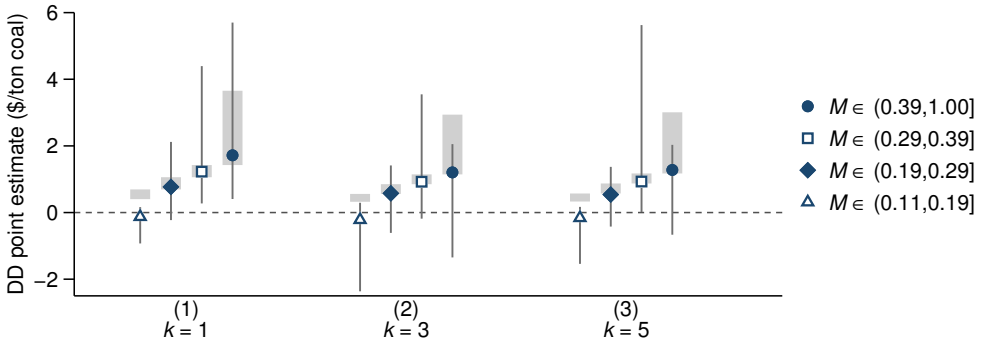
5. Busse and Keohane (2007) use the RCAF to control for time series variation in shipping costs. The RCAF uses the AAR fuel price index as an input. Unlike the monthly fuel price index, it is only published quarterly.

estimates are larger for bituminous shipments than for sub-bituminous shipments, scaling with their difference in average delivered coal price ($55.45 for bituminous vs. $23.58 for sub-bituminous). Still, my results are broadly consistent and retain at least weak statistical significance across all five split samples.

E.3 Robustness: markup DD

When estimating Equation (5) using $TREAT_j = \hat{M}_j$, I parameterize a linear relationship between $\hat{M}_j$ and P_{ojms}. This could be problematic, since $\hat{M}_j$'s predictions are not quantitatively accurate and may deviate from actual markup changes in a nonlinear manner. Figure E3 relaxes the linearity assumption by defining $TREAT_j$ using five discrete bins for the quintiles of $\hat{M}_j$'s positive support. I report DD estimates and bootstrap confidence intervals for quintiles 2–5, with the omitted category is the 64% of plants with $\hat{M}_j \leq 0.11$ (most of which are non-captive or have a coal-by barge option).[6] The resulting DD estimates increase almost linearly across bins of $\hat{M}_j$. They are consistent with estimates using linear $\hat{M}_j$: the light shaded rectangles multiply the analogous $\hat{\tau}$ from Table 3 by the end points of each bin, aligning closely with binned estimates for quintiles 3–5. The binned estimates for quintile 2 diverge slightly from the linear estimates, likely because they are differencing relative to the adjacent omitted bin. Figure E3 assuages concerns about imposing linear DD effects across the support of $\hat{M}_j$.

Figure E3: Markup DD using discrete bins of $\hat{M}_j$



Notes: This figure reports the results of 3 regressions. Each regression estimates Equation (5) defining $TREAT_j$ as four indicator variables for quintiles 2–5 of $\hat{M}_j$'s positive support. Coefficients represent the cumulative change in markups (through 48 months) caused by a $1/MMBTU increase in gas price, relative to the omitted group (quintile 1, plus all plants with $\hat{M}_j \leq 0$). Regression are identical to Table 3, except that $\hat{M}_j$ is binned instead of linear. Light gray rectangles indicate the corresponding range of $\hat{\tau}$ point estimates for $TREAT_j = \hat{M}_j$, multiplying the endpoints of each bin by the analogous linear estimate from Table 3. Whiskers denote bias-corrected accelerated bootstrapped 95% confidence intervals. See notes under Figure 6 and Table 3 for further details.

Table E4 reports numeric results corresponding to my DD regressions in Figure 6. Panel A includes results for all shipments, while Panels B–C use subsamples of long-term contract transactions and spot-market transactions. I omit Panel C estimates from Figure 6, since they are statistically imprecise (likely due to their smaller sample size). Table E5 reports the analogous contract vs. spot-market splits for Table 3, where DD effects are similarly imprecise for the smaller spot-market subsample. The regressions in Column (2) correspond to the navy dots in Figure E4. This figure reveals a similar pattern in the DD lag structure when I estimate Equation (5) using $L = \{24, 36, 48\}$ monthly lags.

6. Each "quintile" includes 8% of plants, because $\hat{M}_j \leq 0$ for 56% of plants. As in Table 3, Figure E3 omits 2 plants with $\hat{M}_j > 1$, which I address in Appendix E.6.

Figures E5–E6 convert my main specification into an "event study" around the drop in natural gas prices that began in July 2008. Rather than control for lags of the gas price, this model interacts $\hat{M}_j$ with quarter-of-sample dummies (indexing quarters by q):

$$P_{ojms} = \sum_q \tau_q \hat{M}_j \cdot \mathbf{1}[m \in q] + \beta_C \mathbf{C}_{ojms} + \beta_T S(\mathbf{T}_{ojms}) + \beta_X \mathbf{X}_{jm} + \eta_{oj} + \delta_m + \varepsilon_{ojms} \qquad \text{(E3)}$$

These plots reveal a persistent differential drop in markups after 2009 for plants with higher $\hat{M}_j$. These effects lag the decrease in gas prices (in light blue) by about 12 months, which aligns with Figure E4—where DD effects start to accumulate about 12 months after a gas price shock. I also find suggestive evidence of differential markup increases that lag the 2007 increase in markups, which is also qualitatively consistent with the predictions of my oligopoly model.

Table E4: Markup DD results – "treatment" interacted with gas price

	Outcome: delivered coal price ($/ton)				
	(1)	(2)	(3)	(4)	(5)
"Treated" group	Captive	Captive no water	Captive $\hat{\lambda}_{0j} \geq 0.16$	Captive no water $\hat{\lambda}_{0j} \geq 0.16$	$\hat{M}_j \geq 0.29$
A. All shipments					
DD estimate	0.12	0.41	0.73	0.92	1.41
($\hat{\tau}$, for 48 months)	$[-0.46, 0.71]$	$[-0.10, 0.92]$	$[-0.54, 1.38]$	$[-0.27, 1.56]$	$[0.48, 3.74]$
Avg price (untreated)	38.06	35.59	33.87	33.79	33.55
Plants	184	184	182	182	182
Observations	88,505	88,505	88,038	88,038	88,038
B. Contract shipments					
DD estimate	0.19	0.54	0.66	0.97	1.52
($\hat{\tau}$, for 48 months)	$[-0.39, 0.76]$	$[0.04, 1.04]$	$[-0.83, 1.19]$	$[-0.28, 1.47]$	$[0.44, 4.72]$
Avg price (untreated)	37.49	35.33	33.23	33.30	33.00
Plants	181	181	179	179	179
Observations	69,154	69,154	68,812	68,812	68,812
C. Spot-market shipments					
DD estimate	−0.58	−0.13	0.44	0.19	0.36
($\hat{\tau}$, for 48 months)	$[-2.11, 0.94]$	$[-1.47, 1.20]$	$[-2.71, 3.84]$	$[-2.33, 3.23]$	$[-0.80, 1.88]$
Avg price (untreated)	42.08	37.12	38.52	37.03	37.41
Plants	171	171	170	170	170
Observations	19,341	19,341	19,216	19,216	19,216

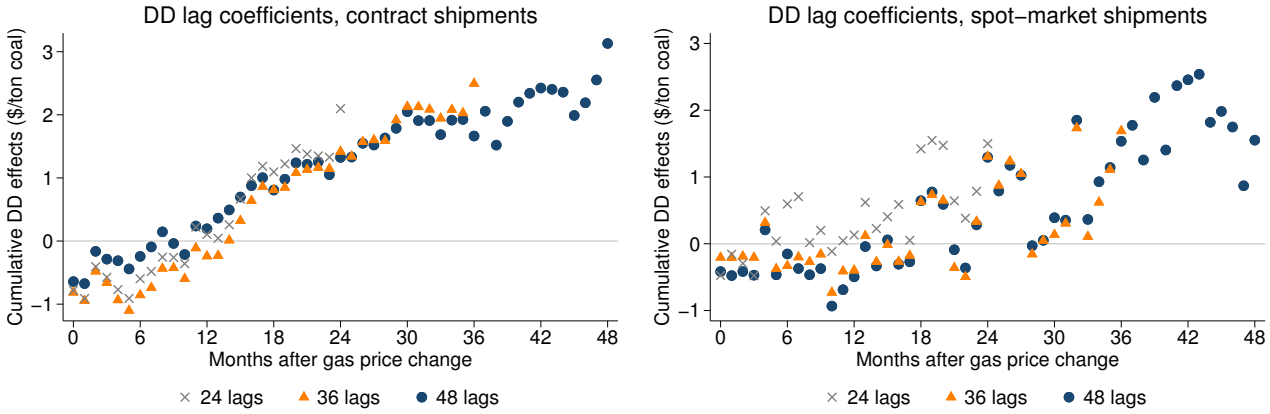
Notes: Panels A–B report numerical results corresponding to Figure 6. Panel C reports analogous results for spot market transactions. Brackets denote 95% confidence intervals, clustering by plant in Columns (1)–(2), and applying the bias-corrected accelerated bootstrap procedure in Columns (3)–(5). All regressions include month-of-sample and plant-by-origin-county fixed effects; commodity, shipping cost, and plant controls; and $k = 3$ nearest neighbor weights. See notes under Figure 6 for further details.

Table E5: Markup DD results – linear $\hat{M}_j$, contract vs. spot shipments

	Outcome: delivered coal price ($/ton)		
DD estimates with 48 lags	(1)	(2)	(3)
A. Contract shipments			
$\hat{M}_j \times (\text{Gas Price})_m$	3.91	3.13	3.20
	[1.72, 10.82]	[0.76, 9.06]	[0.88, 9.67]
k nearest neighbors	1	3	5
Mean of dep var ($\hat{M}_j = 0$)	30.75	31.59	31.88
Plants	138	179	190
Observations	51,506	68,812	73,113
B. Spot-market shipments			
$\hat{M}_j \times (\text{Gas Price})_m$	2.47	1.55	1.59
	[−1.44, 18.12]	[−3.47, 8.89]	[−2.89, 7.91]
k nearest neighbors	1	3	5
Mean of dep var ($\hat{M}_j = 0$)	34.20	35.23	35.55
Plants	130	170	177
Observations	14,346	19,216	20,354

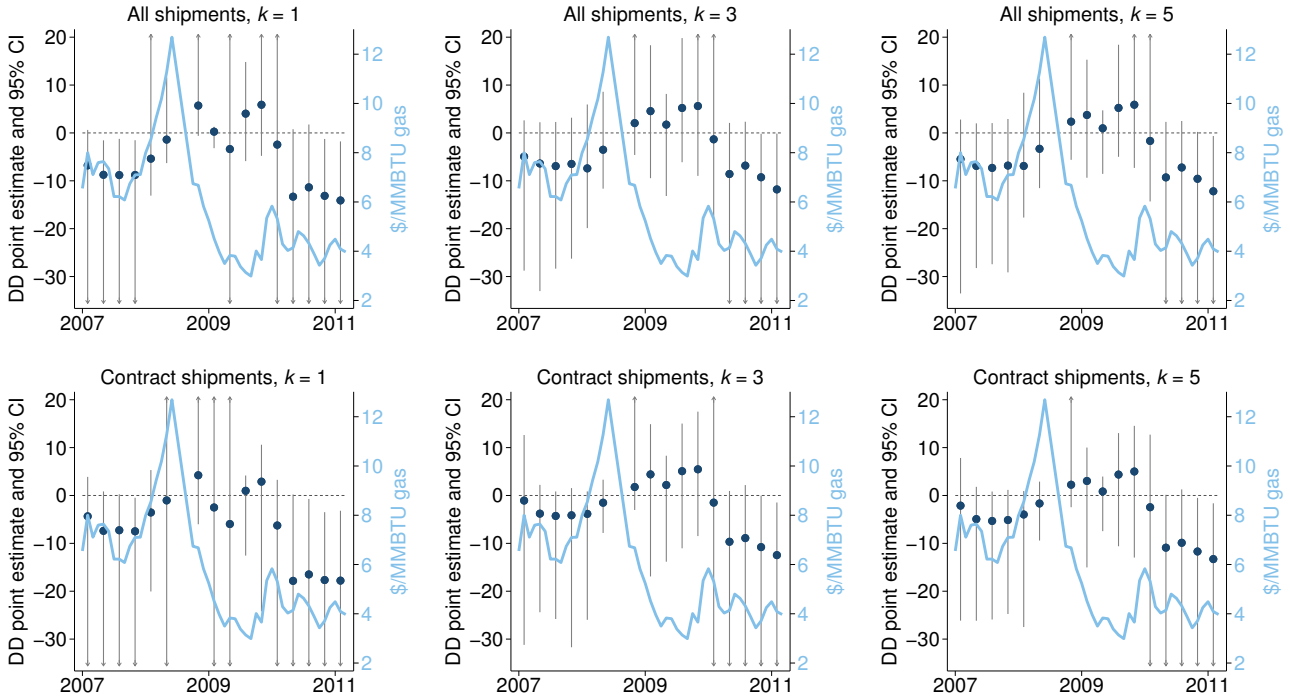
Notes: Regressions estimate Equation (5) with $TREAT_j = \hat{M}_j$, separately for contracts vs. spot-market shipments. Besides splitting the sample, they are identical to the regressions in Table 3. Figure E4 reports monthly lagged DD coefficients from the regressions in Column (2). Brackets report bias-corrected accelerated bootstrapped 95% confidence intervals. Regressions include month-of-sample and plant-by-origin-county fixed effects; commodity, shipping cost, and plant controls; and $k \in \{1, 3, 5\}$ nearest neighbor weights. Regressions drop two plants with extreme outlier values of $\hat{M}_j$; I address these outliers in Appendix E.6. See notes under Figure 6 for further details.

Figure E4: Cumulative DD effects using linear $\hat{M}_j$, varying the number of monthly lags



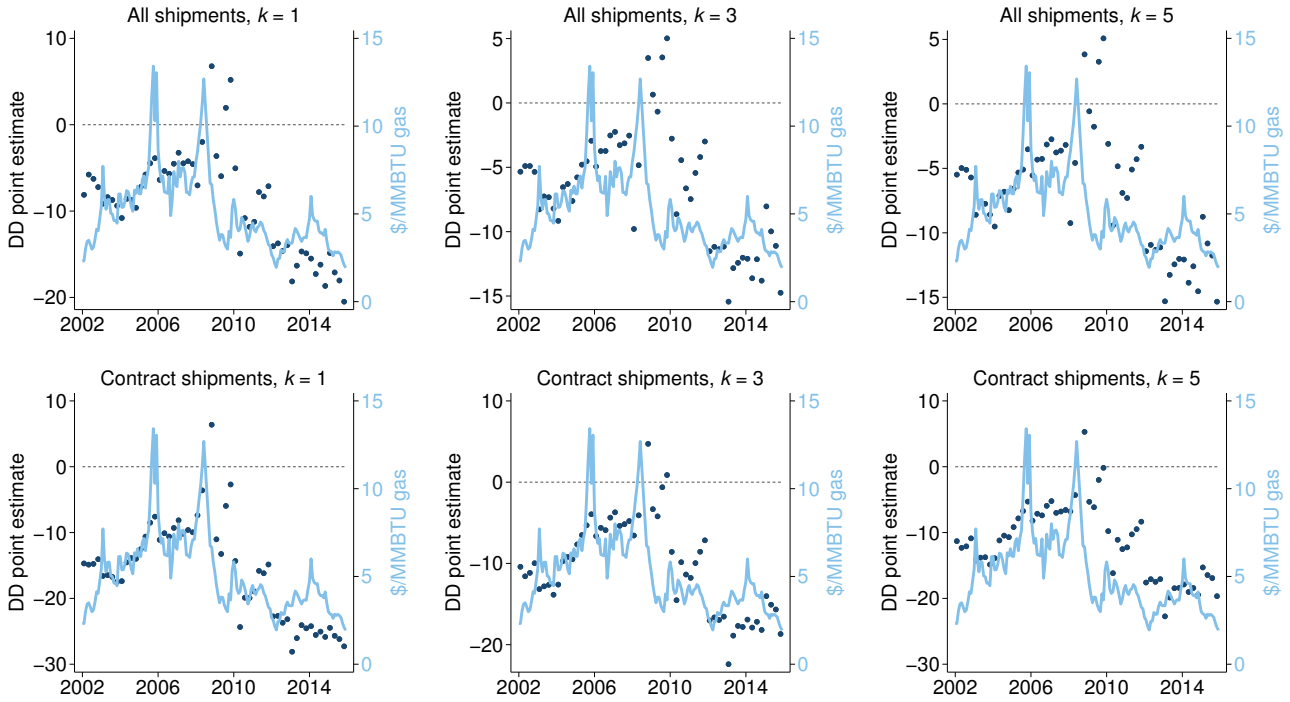
Notes: This figure plots lag-differenced DD coefficients from estimating Equation (5) with $TREAT_j = \hat{M}_j$, splitting the sample into contract vs. spot shipments. Each panel estimates three separate regressions for $L \in \{24, 36, 48\}$ monthly lags. I plot $(\hat{\tau}_0, \ldots, \hat{\tau}_{L-1})$ and $\hat{\tau}$; each coefficient represents cumulative DD effects through ℓ (or L) months. All regressions use $k = 3$ nearest neighbors. 48-month effects (the rightmost navy dot) correspond to Column (2) from Table E5. See notes below Figure 6 and Table 3 for details on the estimation.

Figure E5: DD event studies interacting $\hat{M}_j$ with quarter dummies (restricted event window)



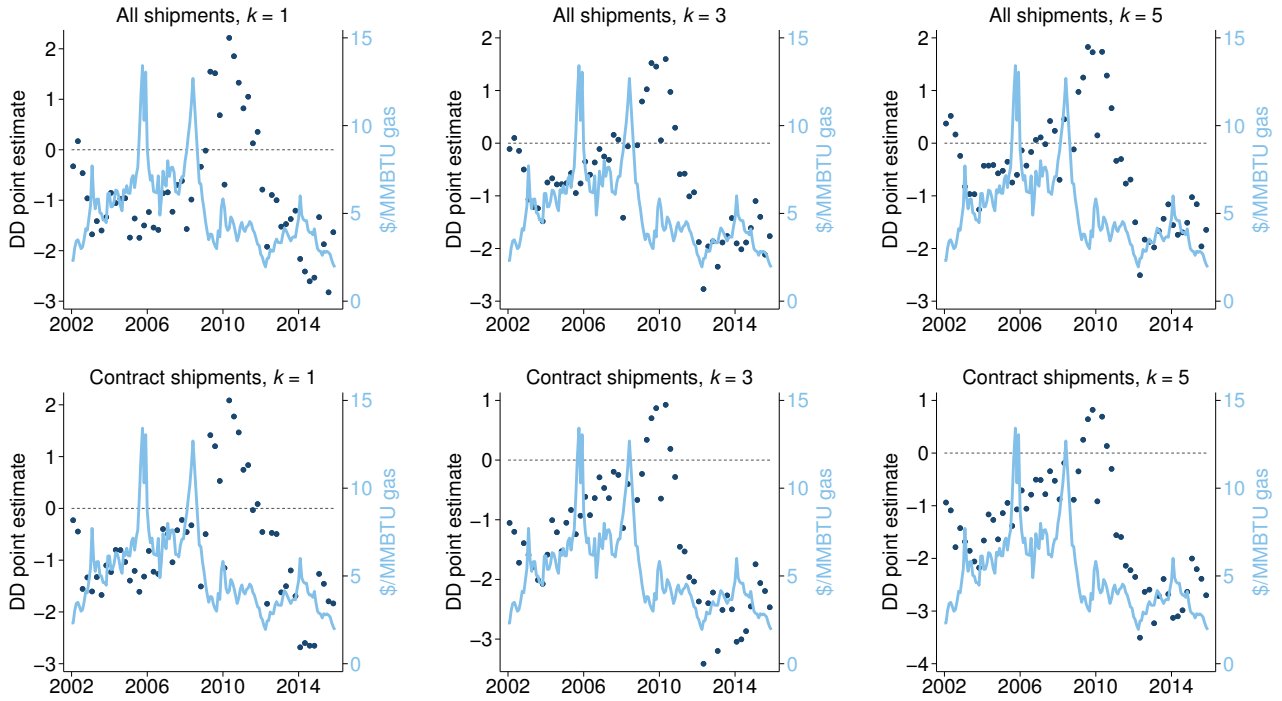
Notes: This figure converts Equation (5) from a lagged DD specification into an event-study specification (Equation (E3)). I interact $\hat{M}_j$ with quarter-of-sample dummies, and plot quarter-specific DD coefficients against the gas price time series (in light blue). I omit quarter 3 of 2008 (the last quarter of high gas prices), following the convention of omitting period $t-1$ in an event study. I also restrict the sample to 7 quarters before and after the gas price drop. Regressions include the same controls and fixed effects as in Table 3, plus basin-specific trends. Whiskers denote bias-corrected accelerated bootstrapped 95% confidence intervals.

Figure E6: DD event studies interacting $\hat{M}_j$ with quarter dummies (full sample)



Notes: This figure plots results from event-study regressions that are identical to Figure E5, but using the whole 2002-2015 sample period (and replacing basin-specific trends with basin-by-year fixed effects). I suppress confidence intervals for ease of presentation.

Figure E7: DD event studies interacting $D_j(1 - W_j)$ with quarter dummies (full sample)



Notes: This figure is identical to Figure E6, except that I define $TREAT = D_j(1 - W_j)$ instead of $TREAT = \hat{M}_j$.

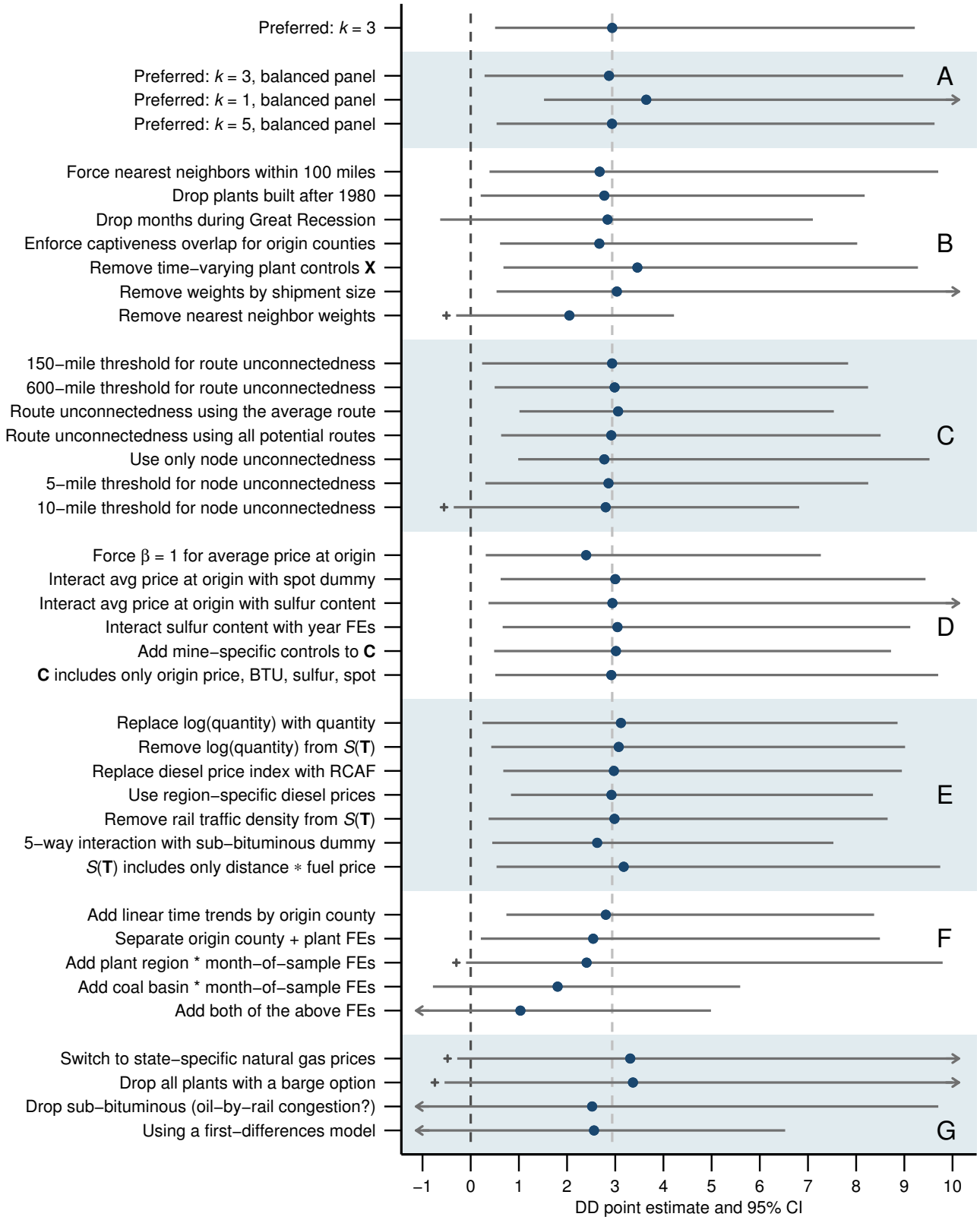
Figure E7 recovers a similar patter of quarterly DD estimates (albeit with smaller magnitudes) replacing $TREAT_j = \hat{M}_j$ in Equation (E3) with $TREAT_j = D_j(1 - W_j)$.

Figure E8 reports sensitivity analysis for my preferred DD estimate with $TREAT_j = \hat{M}_j$, from Column (2) of Table 3. This mirrors Figure E2, and I discuss nearly all robustness tests in detail in Section E.2 above.

Panel G includes four additional robustness tests specific to Equation (5), each yielding similar DD results. First, I replace the Henry Hub natural gas time series (Z_m) with state-specific average monthly gas prices for the electric power sector. These prices include some cross-sectional variation due to regional heterogeneity in gas production and pipeline capacity—which may more accurately competition from nearby gas plants, but could also complicate identification (e.g. if localized gas prices are endogenous to coal competition). Second, I drop plants with $W_j = 1$, in case barge shipping is not competitive and my assumption of $\frac{d\mu_j}{dZ} = 0$ is unrealistic. Third, I drop sub-bituminous coal shipments, which are more likely to trend in parallel with oil-by-rail congestion out of North Dakota (Covert and Kellogg (2018)). Fourth, I estimate Equation (5) in first-differences, redefining each variable $\Delta X_{ojm} = \bar{X}_{ojm} - \bar{X}_{oj(m-1)}$ (averaging across purchases s, and dropping observations where $\bar{P}_{oj(m-1)}$ is not populated).

Panel F includes four fixed effects sensitivities that differ from Figure E2, since Equation (5) controls for interacted county-by-plant fixed effects. First, I additively separate county fixed effects and plant fixed effects, to test for changes in the composition of shipping routes: perhaps average markups on route oj decreased, but plant j switched to a more expensive coal variety from a different county. This has little effect on my results, suggesting that my preferred estimates do no mask meaningful composition-induced markup changes. Second, I interact month-of-sample fixed effects with dummies for each plant's electricity market region,

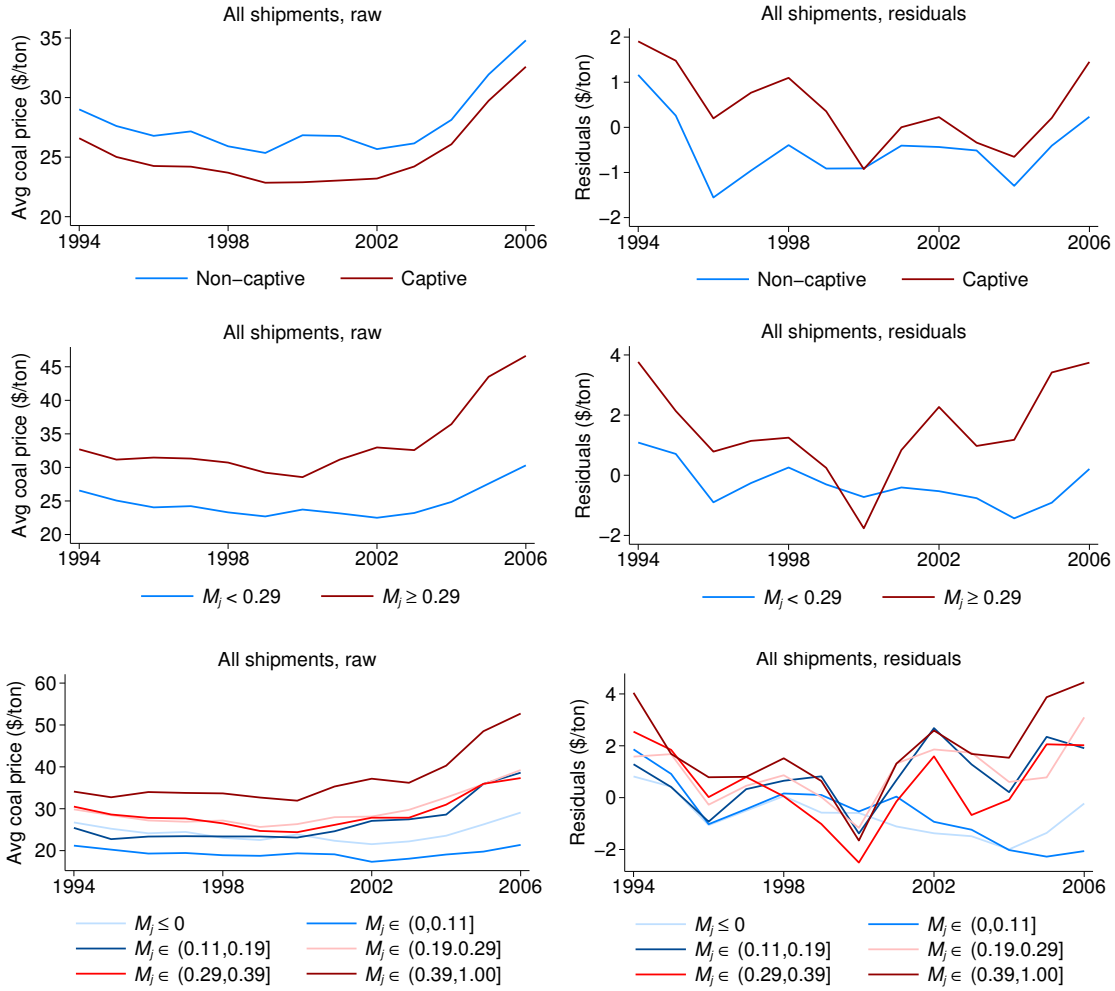
Figure E8: Sensitivities – markup DD using linear $\hat{M}_j$



Notes: This figure plots DD sensitivity analyses, defining $TREAT_j = \hat{M}_j$ and using both contract and spot shipments, with $k = 3$ nearest-neighbor matches (as in Column (2) of Table 3, reported at the top of this chart). Each dot reports $\hat{\tau}$ from estimating Equation (5) with 48 lags, for a separate sensitivity. See surrounding text for details, as well as text in Appendix E.2. Whiskers denote bias-corrected accelerated bootstrapped 95% confidence intervals; whiskers are truncated at −1 and 10 for ease of presentation. Pluses denote estimates where the 95% confidence interval includes zero, but the 90% confidence interval does not.

to allow for differential trends in regional electricity markets.[7] This yields similar results. Third, I interact month-of-sample fixed effects with coal basin fixed effects, to allow for differential trends across regional coal markets.[8] This attenuates my results, but also removes potentially key variation in $\hat{M}_j$. Fourth, when I include both sets of interacted month-of-sample fixed effects, my $\hat{\tau}$ estimate becomes further attenuated. Still, these robustness checks bolster my assumption of parallel counterfactual trends for plants with high vs. low $\hat{M}_j$. Figure E9 also reveals parallel trends in delivered coal prices prior to 2002 (the start of my sample period); a slight divergence of trends in $\hat{M}_j$ in 2004–06 makes sense, since gas prices increased immediately prior to the fracking boom.

Figure E9: Pre-2007 trends in delivered coal prices



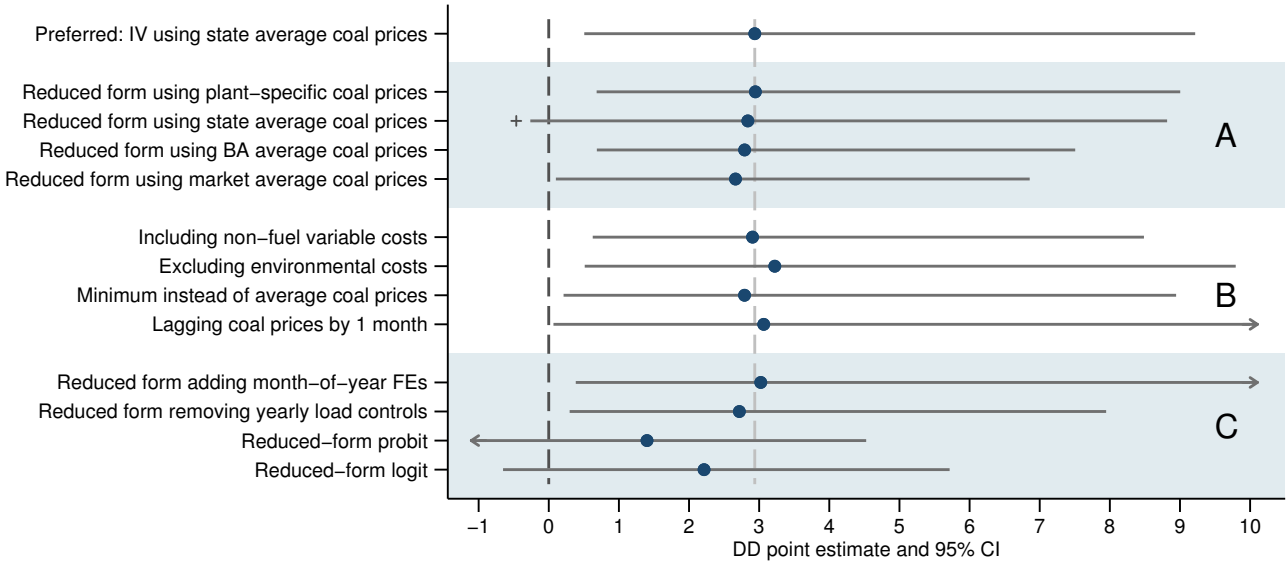
Notes: This figure reports annual average delivered coal prices for 1994–2006, prior to the fracking boom. The top panels split by captive vs. non-captive plants, as in regression (1) of Figure 6. The middle panels split on $\mathbf{1}[\hat{M}_j \geq 0.29]$, as in regression (9) of Figure 6. The bottom panels split plants into 6 groups: the 5 quintiles of $\hat{M}_j$'s positive support, and $\hat{M}_j \leq 0$ (as in Figure E3). The left panels report group-year averages of raw delivered coal prices. The right panels partial out all linear controls (but not fixed effects) used in my DD specification. All panels weight average using $k = 3$ nearest neighbor weights.

7. I define market regions as ISOs and NERC regions (see Appendix G.2.1).

8. I define 8 coal basins, following EIA's coal mine datasets: Northern Appalachia, Central Appalachia, Southern Appalachia, Illinois Basin, Interior, Powder River Basin, Uinta Region, and Western.

Figure E10 shows that my DD estimates are robust to alternative assumptions in my demand estimation procedure, which I used to construct $\hat{M}_j$. Panels A, B, and C correspond to the robustness checks in Figures A5, A6, and A7 (described in Appendix A.2). Finally, Table E6 shows that constructing $\hat{M}_j$ under the assumption of perfect collusion (rather than Cournot competition) yields nearly identical DD estimates. This is unsurprising, since switching from Cournot to collusion only slightly alters the expression for $\frac{d\mu_j}{dZ}$ (see derivations at the end of Appendix B.1), which only slightly changes $\hat{M}_j$ for the 21% of plants with $D_j = 0$ and $W_j = 0$.

Figure E10: Demand estimation sensitivities – markup DD using linear $\hat{M}_j$



Notes: This figure supplements Figure E8 with sensitivities on my demand estimation procedure. All regressions define $TREAT_j = \hat{M}_j$, use both contract and spot shipments, and use $k = 3$ nearest neighbor matches (as in Column (2) of Table 3, reported at the top of this chart). Each dot reports $\hat{\tau}$ from estimating Equation (5) with 48 lags, for a separate sensitivity. See surrounding text for details, as well as text in Appendix A.2. Whiskers denote bias-corrected accelerated bootstrapped 95% confidence intervals; whiskers are truncated at −1 and 10 for ease of presentation. Pluses denote estimates where the 95% confidence interval includes zero, but the 90% confidence interval does not.

Table E6: Markup DD results – constructing $\hat{M}_j$ assuming rail carriers perfectly collude

	Outcome: delivered coal price ($/ton)		
DD estimates with 48 lags	(1)	(2)	(3)
$\hat{M}_j \times (\text{Gas Price})_m$	3.66	2.94	3.03
	[1.67, 9.90]	[0.45, 9.28]	[0.38, 9.53]
DD effect for $\hat{M}_j = 0.29$	1.06	0.85	0.88
DD effect for $\hat{M}_j = 0.44$	1.61	1.29	1.33
k nearest neighbors	1	3	5
Mean of dep var ($\hat{M}_j = 0$)	31.34	32.21	32.53
Plants	140	182	192
Observations	65,856	88,038	93,476

Notes: Regressions are identical to Table 3, except that I construct $\hat{M}_j$ assuming perfect collusion ($\theta_j = N_j$) rather than Cournot competition ($\theta_j = 1$). At the end of Appendix B.1, I show how this alternate assumption on market conduct alters my comparative static $\frac{d\mu_j}{dZ}$. Brackets report bias-corrected accelerated bootstrapped 95% confidence intervals. See notes under Figure 6 for details.

E.4 Linking my DD specification to my oligopoly model

How does Equation (5) relate to my theoretical framework? If I directly observed coal markups, I would estimate a DD model resembling:

$$\mu_{ojms} = \tau M_j \cdot Z_m + \eta_j + \delta_m + \varepsilon_{ojms} \tag{E4}$$

Here, $\widehat{\tau}$ would capture the extent to which $\hat{M}_j$ predicts differential changes in markups, controlling for unit and time fixed effects (recall that $\hat{M}_j \sim \left[\frac{d\mu}{dZ}\right]_j$). Stated differently, $\widehat{\tau} > 0$ would mean markup changes are heterogeneous in $\hat{M}_j$, just as theory would predict.

While I don't observe μ_{ojms}, I can modify Equation (4) to residualize prices using cost controls. These residuals form the dependent variable in a lagged DD version of Equation (E4):

$$P_{ojms} = \beta_C \mathbf{C}_{ojms} + \beta_T S(\mathbf{T}_{ojms}) + \eta_o + \upsilon_{ojms} \tag{E5}$$

$$\hat{\upsilon}_{ojms} = \tau M_j \cdot Z^{\text{HH}}_{m-L} + \sum_{\ell=0}^{L-1} \tau_\ell M_j \cdot \Delta Z^{\text{HH}}_{m-\ell} + \eta_j + \delta_m + \varepsilon_{ojms} \tag{E6}$$

By the assumptions of Equation (4), the residuals $\hat{\upsilon}_{ojms}$ should capture variation in markups—i.e., coal price variation not explained by commodity cost controls, shipping cost controls, or origin fixed effects. This two-step procedure is nearly identical to Equation (5), except: (a) the covariates in Equation (E5) are not partialed out of the DD interaction terms, (b) plant fixed effects η_j replace route fixed effects η_{oj}, and (c) it omits time-varying plant controls $\mathbf{X}_{jm}$. Table E7 reports results for Equations (E5)–(E6), which are similar to my results from estimating Equation (5).

Table E7: Markup DD results – linear $\hat{M}_j$, using two-step estimator

	Outcome: delivered coal price ($/ton)		
DD estimates with 48 lags	(1)	(2)	(3)
$\hat{M}_j \times (\text{Gas Price})_m$	2.88	2.39	2.37
	[1.03, 8.59]	[0.55, 6.86]	[0.73, 7.07]
k nearest neighbors	1	3	5
Mean for $\hat{M}_j = 0$	31.37	32.24	32.57
Plants	140	182	193
Observations	65,915	88,143	93,582

Notes: Regressions are identical to Table 3, except using Equations (E5)–(E6) instead of my preferred DD specification (Equation (5)). Regressions use all shipments (contract and spot market), and vary the number of nearest-neighbor matches. Brackets denote bias-corrected accelerated bootstrapped 95% confidence intervals. See notes under Figure 6 and Table 3 for further details.

E.5 Bootstrap confidence intervals

Most of my DD regressions use generated regressors to define $TREAT_j$ in Equation (5). To conduct proper inference on $\hat{\tau}$ (my coefficient of interest), I must account for estimation error in $\langle \hat{\lambda}_{0j}, \hat{\lambda}_{1j}, \hat{\lambda}_{2j} \rangle$. I do this by using the "pairs cluster" bootstrap procedure outlined in Cameron,

Gelbach, and Miller (2008). I bootstrap pairs of regressands and regressors (i.e., drawing $\langle \mathbf{y}, \mathbf{X} \rangle$ jointly with replacement), and block bootstrap by plant (i.e., clusters, the unit of inference). Within each bootstrap iteration, I incorporate uncertainty in the generated regressor by simulating a new value of $TREAT_j$ and re-estimating Equation (5). Then, I use the bootstrapped distribution of point estimates to construct confidence intervals for $\hat{\tau}$ (and each $\hat{\tau}_\ell$).

I start by simulating random draws of $\langle \hat{\lambda}_{0j}, \hat{\lambda}_{1j}, \hat{\lambda}_{2j} \rangle$. For each plant j, for each $\hat{\lambda}$ parameter, I make 1 million draws (indexed by S) from the j-specific sampling distribution that aligns with Equations (7)–(9). This yields the j-specific simulated distributions $\{\hat{\lambda}_{0j}^{(S)}\}$, $\{\hat{\lambda}_{1j}^{(S)}\}$, and $\{\hat{\lambda}_{2j}^{(S)}\}$. I use these simulated distributions to construct the j-specific simulated distributions $\{\hat{M}_j^{(S)}\}$:

$$\hat{M}_j^{(S)} \equiv \begin{cases} \dfrac{\hat{\lambda}_{0j}^{(S)} \left[D_j + \hat{\lambda}_{2j}^{(S)} (2 - D_j)^{-1} \right] - \hat{\lambda}_{1j}^{(S)}}{2 + \hat{\lambda}_{2j}^{(S)} (2 - D_j)^{-1}} & \text{if} \quad W_j = 0 \\ 0 & \text{if} \quad W_j = 1 \end{cases} \tag{E7}$$

I winsorize each $\hat{M}_j^{(S)}$ draw at $[-2, 2]$ to reduce the influence of extreme draws with a near-zero denominator.[9]

The $\hat{M}_j$ that I use in my analysis is technically the mean of $\{\hat{M}_j^{(S)}\}$:

$$\hat{M}_j \equiv \frac{1}{1,000,000} \sum_{S=1}^{1,000,000} \hat{M}_j^{(S)} \tag{E8}$$

Since Equation (10) is nonlinear in estimated parameters, constructing $\hat{M}_j$ using $\langle \hat{\lambda}_{0j}, \hat{\lambda}_{1j}, \hat{\lambda}_{2j} \rangle$ point estimates would yield a biased estimate of $\hat{M}_j$ for plants with $W_j = 0$ and $\hat{\lambda}_{2j} \neq 0$ (due to Jensen's inequality).

Armed with these simulated sampling distributions, I bootstrap Equation (5) and construct confidence intervals using the following steps:

1. Estimate Equation (5) using $TREAT_j$, and store $\hat{\tau}$, $\hat{\tau}_0$, …, $\hat{\tau}_{L-1}$.

2. For each of 1000 bootstrap iterations indexed by B:

   (a) Construct a bootstrap sample B by drawing plants with replacement from within the regression sample from Step 1.

   (b) Define treatment for bootstrap sample B as $TREAT_j^{(S)}$, where $S = B$. For example, if $TREAT_j = \hat{M}_j$ in Step 1, then $TREAT_j^{(74)} = \hat{M}_j^{(74)}$ for iteration $B = 74$.[10]

   (c) Estimate Equation (5) using bootstrap sample B and $TREAT_j^{(S)}$, and store $\hat{\tau}^{(B)}$, $\hat{\tau}_0^{(B)}$, …, $\hat{\tau}_{L-1}^{(B)}$.

9. Unwinsorized draws with $\hat{\lambda}_{2j}^{(S)} \approx 2D_j - 4$ can be orders of magnitude too large! Only 0.7% of unwinsorized draws fall outside $[-2, 2]$. The resulting $\hat{M}_j$ estimates are nearly identical if I winsorize at $[-5, 5]$ or $[-100, 100]$.

10. I construct the analogous versions of $TREAT_j^{(S)}$ for different definitions of $TREAT_j$: $TREAT_j^{(S)} = D_j(1 - W_j)\mathbf{1}[\hat{\lambda}_{0j}^{(S)} \geq 0.16]$, $TREAT_j^{(S)} = \mathbf{1}[\hat{M}_j^{(S)} \geq 0.29]$, etc.

3. Using the distributions $\{\hat{\tau}^{(B)}\}$, $\{\hat{\tau}_0^{(B)}\}$, ..., $\{\hat{\tau}_{L-1}^{(B)}\}$ and correcting for measurement error, construct bias-corrected accelerated (BCA) bootstrap confidence intervals following Cameron, Gelbach, and Miller (2006).

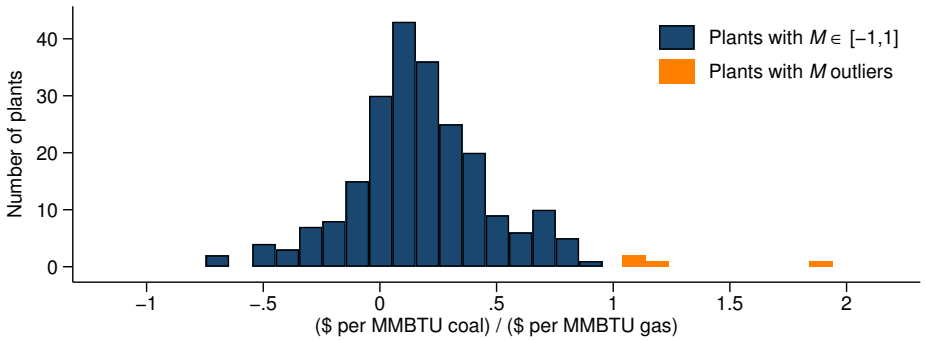
I construct percentile-based confidence intervals instead of bootstrap standard errors, because they capture asymmetry in the bootstrap distribution (created by nonlinearity in $\hat{M}_j^{(S)}$). Efron (1987) first developed the BCA refinement on bootstrap confidence intervals. The "bias correction" adjusts the percentile endpoints on the confidence intervals to improve coverage accuracy, while the "acceleration" adjusts for skewness (Hall (1992), pp. 128–141). While Cameron, Gelbach, and Miller (2008) show that the BCA adjustment can perform poorly for pairs cluster bootstraps with fewer than 20 clusters, my DD regressions all have at least 89 plants with non-zero nearest-neighbor weights.

E.6 Outliers in $\hat{M}_j$

Figure E11 presents the full distribution of $\hat{M}_j$, without winsorizing extreme values or restricting the sample with nearest-neighbor weights (with the same bin width as the bottom-right panel of Figure 5). I predict $|\hat{M}_j| > 1$ for 4 plants, which is implausibly large—implying that a \$1/MMBTU decrease in gas price would cause coal markups to fall by over \$1/MMBTU ($\approx$ \$20/ton). The presence of outliers is not surprising, since (i) $\hat{M}_j$ includes measurement error in $\langle \hat{\lambda}_{0j}, \hat{\lambda}_{1j}, \hat{\lambda}_{2j} \rangle$; (ii) measurement error in $\hat{\lambda}_{2j}$ enters the denominator of $\hat{M}_j$; and (iii) $\hat{M}_j$ adheres to the comparative static $\frac{d\mu_j}{dZ}$ (Equation (3)), which mischaracterizes markup responses by ignoring rail regulation (among other factors).

2 of these 4 outlier plants have non-zero nearest-neighbor weights, and I omit these 2 plants from all DD regressions where $TREAT_j$ is defined using $\hat{M}_j$ or $\hat{\lambda}_{0j}$. Columns (1)–(3) of Table E8 reproduce Table 3, while Columns (4)–(6) estimate these same regressions including the 2 outlier plants. As expected, including these extreme values attenuates my $\hat{\tau}$ estimates. Figure E12 shows that outliers have less influence on my DD estimates when $\hat{M}_j$ is binned (rather than linear). Regression (2) extends the highest bin to include outliers, while regression (3) adds an extra bin for plants with $\hat{M}_j > 1$; both sets of binned DD estimates are nearly identical to regression (1), which reproduces regression (2) from Figure E3 (without outliers).

Figure E11: Histogram of $\hat{M}_j$ estimates, including outliers



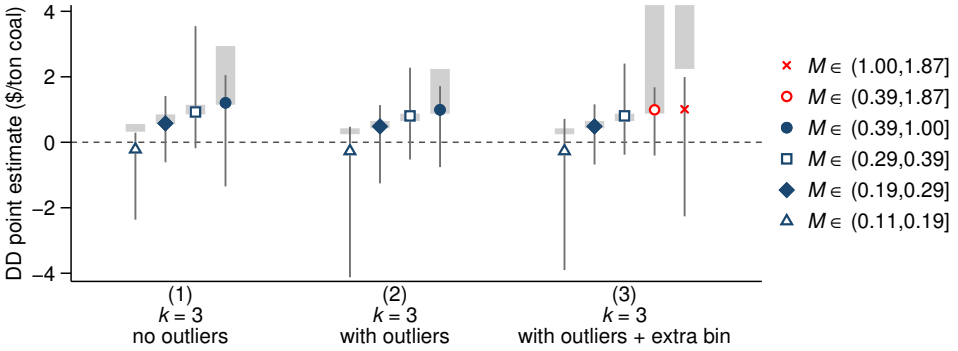
Notes: This histogram is analogous to the bottom-right panel in Figure 5, except that (i) it includes 408 coal plants (without restricting to nearest-neighbor matches), and (ii) it reports the full distribution of $\hat{M}_j$ without winsorizing. I color-code the 4 plants with outlier $\hat{M}_j$ predictions (i.e., $|\hat{M}_j| > 1$). Only 2 of these 4 plants have non-zero nearest-neighbor weights. I drop these 2 plants from all DD regressions where $TREAT_j$ is defined using $\hat{M}_j$ or $\hat{\lambda}_{0j}$, except for Table E8 and Figure E12 below.

Table E8: Markup DD results – linear $\hat{M}_j$, with vs. without outliers

	Outcome: delivered coal price ($/ton)					
DD estimates (48 lags)	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{M}_j \times (\text{Gas Price})_m$	3.66	2.94	3.01	2.24	2.24	2.45
	[1.69, 9.73]	[0.51, 9.22]	[0.43, 9.44]	[−3.60, 3.90]	[−0.94, 5.16]	[−0.23, 6.35]
k nearest neighbors	1	3	5	1	3	5
Including $\hat{M}_j$ outliers				Yes	Yes	Yes
Plants	140	182	192	142	184	194
Observations	65,856	88,038	93,476	66,323	88,505	93,943

Notes: Columns (1)–(3) reproduce Table 3. Columns (4)–(6) re-estimate identical regressions including two plants with $\hat{M}_j > 1$ (outliers that I omit from Table 3). Brackets denote bias-corrected accelerated bootstrapped 95% confidence intervals.

Figure E12: Markup DD – discrete bins of $\hat{M}_j$ with vs. without outliers



Notes: Regression (1) reproduces regression (2) from Figure E3, where the upper-most bin stops at $\hat{M}_j = 1$ (two outlier plants are excluded from the regression). Regression (2) includes these two outlier plants ($\hat{M}_j = 1.22$, $\hat{M}_j = 1.87$) in the upper-most bin, but is otherwise identical. Regression (3) likewise includes these two outlier plants, but creates a separate DD bin for $\hat{M}_j > 1$ (and is otherwise identical). Light gray rectangles indicate the corresponding range of $\hat{\tau}$ point estimates for $TREAT_j = \hat{M}_j$, multiplying the endpoints of each bin by the analogous linear estimate from Table E8. Whiskers denote bias-corrected accelerated bootstrapped 95% confidence intervals. See notes under Figure E3 and Table E8 for further details.

Table E9: Markup DD results – "treatment" interacted with gas price, including outliers

	Outcome: delivered coal price ($/ton)					
"Treated" group	Captive, $\hat{\lambda}_{0j} \geq 0.16$		Captive, no water, $\hat{\lambda}_{0j} \geq 0.16$		$\hat{M}_j \geq 0.29$	
	(1)	(2)	(3)	(4)	(5)	(6)
DD estimate	0.73	0.66	0.92	0.96	1.33	1.40
($\hat{\tau}$, for 48 months)	[−0.52, 1.38]	[−0.54, 1.14]	[−0.31, 1.56]	[−0.31, 1.48]	[0.44, 3.64]	[0.40, 4.25]
Shipments	All	Contract	All	Contract	All	Contract
Plants	184	181	184	181	184	181
Observations	88,505	69,154	88,505	69,154	88,505	69,154

Notes: Odd (even) columns are identical to Panel A (B) of Table E4, except that they include two plants with $\hat{M}_j > 1$. Brackets denote bias-corrected accelerated bootstrapped 95% confidence intervals.

F Rail graph algorithm and constructing captiveness

F.1 Shortest distance between counties and plants

I begin with a time-invariant GIS dataset of all active rail lines and terminal nodes in the contiguous U.S. (see Appendix G.3). I overlay all coal plant coordinates, and the production-weighted centroids of all coal-producing counties.[11] Then, I calculate the closest (as the crow flies) rail node to each coal plant, and to each coal county's production-weighted centroid.

Next, I calculate the shortest distance along the rail network between each pairwise combination of origin nodes (i.e. rail nodes matched to county centroids) and destination nodes (i.e. rail nodes matched to plant coordinates). I convert the rail network from a GIS dataset into a graph object defined by three elements: (i) a list of nodes (i.e. rail nodes); (ii) a list of edges (i.e. rail lines); and (iii) a distance weight corresponding to each edge (i.e. mileage of each rail line). Using Dijkstra's algorithm, I calculate the shortest path along all possible edges that connect each pair of origin and destination nodes, weighting edges by their distance.[12] This shortest distance for each route oj enters all of my markup regressions in $\mathbf{T}_{ojms}$.

This algorithm does not account for rail line ownership, and may calculate a shortest route that is owned/operated by multiple Class I carriers. An alternate approach would force each shortest path to have uniform ownership on all segments. However, this approach would impose excessive (unrealistic) structure on the rail network: in practice, rail carriers have track-sharing agreements. The Surface Transportation Board prevents carriers from extracting excessive rents from the use of short, pivotal track segments (U.S. Government Accountability Office (2006)).

I iterate the above algorithm 7 times, for 7 restricted rail networks. Each restricted network removes all rail nodes and lines that are owned or operated by 1 of the 7 Class I carriers. Comparing unrestricted vs. restricted networks, I define two concepts: "node unconnectedness" and "route unconnectedness". A plant becomes "node-unconnected" if removing any single Class I carrier leaves the plant unconnected from the network. A plant becomes "route-unconnected" if it becomes unconnected from all observed trading partners (i.e. origin counties) after removing the modal Class I carrier along each (unrestricted) shortest route. I define "captiveness" as the union of the sets of plants that become node-unconnected and route-unconnected.

F.2 Node unconnectedness

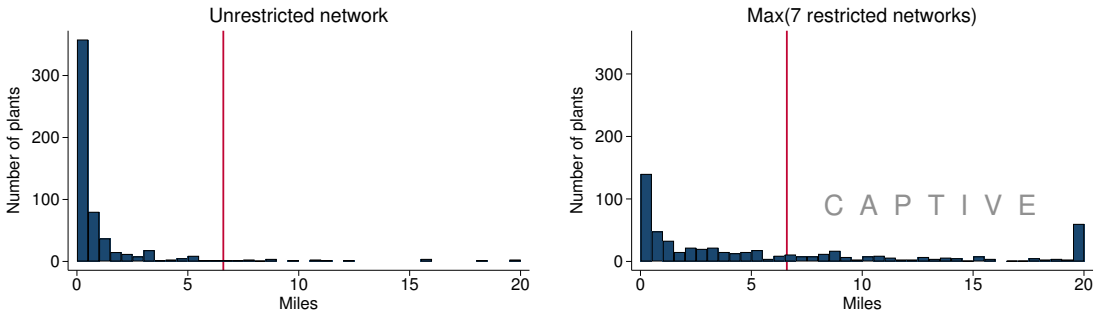
A plant becomes node-unconnected if a single Class I carrier controls all rail nodes from which it can potentially receive coal. I set a threshold of 6.6 miles for node unconnectedness, which is the 95th percentile of plants' distance to the nearest node on the *unrestricted* network.[13] I assume that nodes farther than 6.6 miles from a plant are not feasible delivery points.

11. I average the geographic coordinates of all coal mines in a given county, weighting by total mine production during my sample period. These time-invariant coordinates for more accurately approximates the location of coal production than the county's geographic centroid. Since EIA data do not report mine identifiers prior to 2008, matching rail nodes to individual coal mines would not provide a useful refinement.

12. Hughes (2011) uses a similar algorithm to calculate estimated shortest-distance paths along BNSF's network. Compared to BNSF's own reported distances, and the ratio of estimated-to-actual distances has a mean of 0.96 and a standard deviation of 0.03. This suggests that GIS-derived shortest distances closely approximate (yet slightly understate) actual rail shipping distances.

13. In other words, using the full network of active rail lines, 95% of coal plants are within 6.6 miles of a rail node. If its nearest node is controlled by a smaller non-Class I carrier, a plant cannot become node-unconnected.

Figure F1: Distance to plants' nearest rail node



Notes: The left histogram summarizes plants' distances to the nearest rail node on the unrestricted rail network. The red line denotes the 95th percentile of this distribution, which is 6.6 miles. The right histogram summarizes plants' maximum distances to the nearest rail node across all 7 restricted rail networks (since the nearest unrestricted node can only be controlled by 1 of the 7 Class I rail carriers, I plot maximum across all 7 restricted networks). For plants in the "captive" region of the right panel, *all* rail nodes within a 6.6-mile radius are controlled by the same Class I carrier. I top-code both histograms at 20 miles.

This threshold is exceedingly conservative. It is possible that a plant's flue gas stack (which corresponds to EPA's reported coordinates) is far from the plant's physical stock pile of coal—but not 6.6 miles apart. It is also possible, albeit highly unlikely, that a plant's rail node is actually over 7 miles away from its physical coal offloading point.[14] Measurement error aside, 6.6 miles is a conservative distance buffer: it would almost certainly be cost prohibitive to regularly transport multi-ton carloads of coal 6.6 miles over land by a mode *other* than rail.

Figure F1 plots histograms of plants' distance to the nearest rail node, with vertical lines at 6.6 miles. The left panel shows that virtually all plants are within 3 miles of the nearest node on the unrestricted network. The right panel shows the maximum distance to the nearest rail node across all 7 restricted rail networks—illustrating how removing a single rail carrier would render many plants unfeasibly far from the nearest rail node. I consider plants to the right of the 6.6-mile threshold in the right panel to be captive.[15] In Panel C of Figures E2 and E8, I show that my results are robust to node unconnectedness thresholds of 5 and 10 miles.

F.3 Route unconnectedness

A plant becomes route-unconnected if each of its coal-by-rail shipping routes is controlled by one Class I carrier. This need not the same carrier for each route, and a carrier need not control 100% of the rail lines along the (shortest) route. I construct a route unconnectedness indicator by comparing lengths of shortest paths. For each oj pair, I compare its unrestricted shortest path to its shortest path on the restricted network removing the modal Class I carrier along the unrestricted shortest path.[16] I base this comparison on each route's modal carrier, as the firm most likely to transact coal deliveries. An alternate strategy would consider how shortest distance changes after removing *any* Class I carrier along an o-to-j shortest route. However, this would increase the dimensionality of route unconnectedness, while potentially

14. Although a few plants have constructed small tracks to ferry coal from the rail carrier's node directly to their coal offloading points, their commercial rail node is virtually always within 6.6 miles.

15. In the left histogram, 11 plants are "node-unconnected" even in the unrestricted network. Since this likely results from geographic measurement error, I classify these plants as non-captive throughout my analysis.

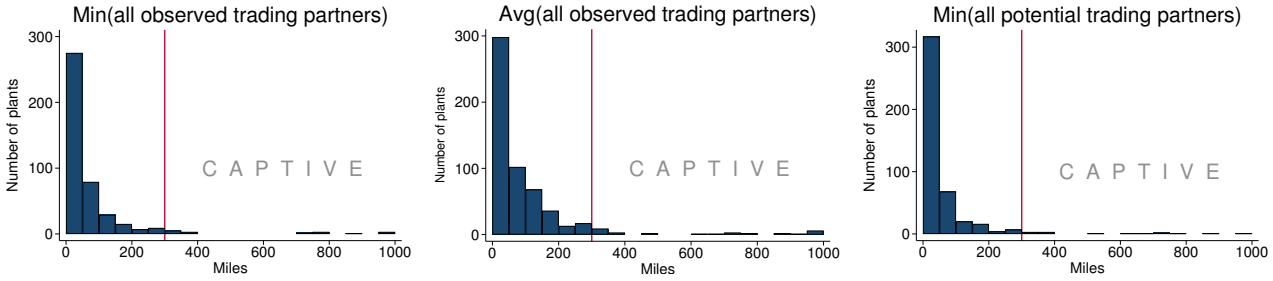
16. For example, suppose the shortest o-to-j path has 4 rail nodes, ordered A-B-C-D. Carrier 1 owns segments AB and BC, and Carrier 2 owns segment CD. If AB+BC is longer than CD, that makes Carrier 1 the modal carrier. Then, I compare AB+BC+CD to the shortest path on the restricted network removing Carrier 1.

over-weighting very short rail segments.[17] I prefer the more straightforward approach that uses a single unrestricted-vs.-restricted comparison per route.

I apply a 300-mile threshold to determine route unconnectedness: an origin-destination pair becomes route-unconnected if the length of its shortest connecting path increases by at least 300 miles after removing the modal carrier along its unrestricted shortest path. This threshold is arbitrary: a 300-mile increase in rail shipping distance would imply roughly a 20% increase in the median delivered coal price in my sample.[18] In many cases, route-unconnectedness is not sensitive to this mileage threshold, since removing the modal carrier eliminates all possible o-to-j paths—rendering the restricted shortest path infinitely long.[19] In Panel C of Figures E2 and E8, I show that my results are robust to doubling or halving this 300-mile threshold.

I consider plant j to be "captive" if it becomes route-unconnected from *every* county from which it purchased coal between 2002 and 2015. The left histogram in Figure F2 plots the minimum increase in mileage for each coal plant along its observed routes. The middle histogram weakens the definition of captiveness, based on route-unconnectedness of the *average* coal shipment to plant j; this prevents a seemingly captive plant from being classified as non-captive due to a few underutilized routes. Alternatively, the right histogram strengthens the definition of captiveness, based on route-unconnectedness across all of plant j's *potential* trading partners; this allows for a seemingly captive plant to be classified as non-captive if an un-utilized route does not become unconnected. All three definitions yield a similar set of "captive" plants, and similar results (see Panel C of Figures E2 and E8).

Figure F2: Plant's minimum increase in route distance from removing modal carrier



Notes: The left panel reports a histogram of plants' minimum increase in shortest shipping path, across all observed trading partners. The red line denotes the 300-mile threshold. I consider the few plants to the right of this 300-mile threshold to be rail captive. The middle and right panels show alternative captiveness definitions, either based on the average increase in the length of the shortest path (weight-averaged across all observed coal deliveries) or the minimum increase across all potential shortest paths (i.e. counties that produce coal with similar attributes to plant j's purchased coal, but which are not observed trading partners).

17. Using the same A-B-C-D example, suppose that removing Carrier 1 increases the shortest route by 50 miles, while removing Carrier 2 increases the shortest route by 500 miles. Under this alternate strategy, I would need to trade-off the 50 vs. 500 miles against the relative importance of each carrier on the unrestricted path.

18. During my sample period, the EIA reported a median rail shipping rate reported of \$0.025/ton-mile. At this rate, a 300-mile increase in rail distance implies an additional \$7.50/ton, or 20% of the median delivered coal price of \$38/ton.

19. The elimination of all connecting routes could result from a plant's nearest node becoming unconnected from a portion of the rail network, *or* from the removal of a plant's nearest node. I incorporate my threshold for node unconnectedness (discussed above) in determining whether plants j's closest node in the restricted network is indeed close enough to be a feasible coal delivery point. I also use a similar concept to define "origin-node-unconnectedness", whereby a coal-producing county becomes unconnected from the rail network. I apply an extremely conservative threshold of 42 miles, which is the 99th percentile of unrestricted nearest-node distance and exceeds the diameter of virtually all coal counties.

G Data

G.1 Coal transaction and production data

G.1.1 Coal shipment data

The core data for my analysis come from the Energy Information Administration (EIA) database of monthly fossil fuel deliveries to power plants. These survey data are at the monthly "order" level, and plants must report each purchase order or supplier contract separately. Per EIA's official documentation, "aggregation of coal receipt data into a single line item is allowed if the coal is received under the same purchase order or contract and the purchase type, fuel, mine type [i.e., surface vs. underground], state of origin, county of origin, and supplier are identical for each delivery." I refer to observations in this dataset as "purchases" (indexed by s).[20]

Since 2008, EIA has collected fossil fuel delivery data on Form 923, with monthly reporting required for all plants larger than 50 MW in generating capacity. Form 923 also reports monthly fuel receipts for a sample of small plants (1–50 MW), and annual fuel receipts for all plants larger than 1 MW. Before 2008, monthly data were collected on two separate forms, each with a 50 MW minimum reporting requirement: the Federal Energy Regulatory Commission's (FERC) Form 423 (for utility-owned plants between 1983–2007); and EIA Form 423 (for non-utility-owned plants between 2002–2007).[21] Since my sample period spans the 2008 changeover, I include only coal plants larger than 50 MW; this represents over 99% of U.S. coal-fired electricity generation. Prior to 2002, non-utility-owned plants were not required to report fuel deliveries; the vast majority of such plants were divested by utilities between 1997 and 2002.

My sample period is 2002–2015. Starting in 2002 affords me five years of data prior to the 2007 start of the fracking boom (Hausman and Kellogg (2015)), while minimizing any confounding effects from electricity market deregulation and coal plant divestments (which mostly occurred before 2002; Fabrizio, Rose, and Wolfram (2007)). Moreover, Linn and Muehlenbachs (2018) also document substantial data irregularities in coal deliveries prior to 2001.

EIA 423/923 data report average prices, total quantities, and average attributes for each fuel "shipment". Average prices are in \$/ton (for coal, i.e. P_{ojms} in Equations (4)–(5)), \$/barrel (for oil), and \$/thousand cubic feet (for gas); prices are inclusive of commodity costs, shipping costs, and markups. EIA withholds price data for deliveries to non-utility-owned plants; as I do not observe these prices, my analysis focus on utility-owned plants. Total delivered quantities are reported in tons (for coal), barrels (for oil) and thousand cubic feet (for gas); I convert physical quantities into energy content using the the average MMBTU content for each shipment (e.g. MMBTU per ton of coal). Besides BTU content, these data report average sulfur content and ash content for coal deliveries; these three attributes influence coal's commodity value, and I control for them in $\mathbf{C}_{ojms}$. The data also report each observation's fuel type—e.g., bituminous vs. sub-bituminous coal; fuel oil vs. kerosene; natural gas vs. liquefied petroleum gas.

Each fossil fuel purchase is classified as either spot market or long-term contract. Most contract shipments report expiration dates, which are inconsistently coded across years; hence, I control for a coarser measure of contract length in $\mathbf{C}_{ojms}$: a dummy for contracts expiring within 2 years. Longer coal contracts tend to have higher coal prices, since plants trade off higher

20. See https://www.eia.gov/electricity/monthly/pdf/technotes.pdf for descriptions of these survey data.

21. EIA Form 923 data are available at https://www.eia.gov/electricity/data/eia923/. FERC Form 423 and EIA Form 423 are each available at https://www.eia.gov/electricity/data/eia423/.

expected costs for lower cost variance and more reliable deliveries (Jha (2022)). I include a spot market dummy in $\mathbf{C}_{ojms}$, as spot shipments tend to have lower prices (for the same reason).

The 423/923 data report the originating county of each coal shipment, coal supplier names (reported with complete coverage only after 2006), and originating mine names and identifiers (reported since 2008). Since supplier and mine names are not reported through my 2002–2015 sample period, I treat the originating county as the unit of origin for each coal shipment. Starting in 2008, EIA began reporting each delivery's primary and secondary mode of transportation. I extrapolate backwards to assign transportation modes for 2002–2007 shipments using observed modes within each origin-destination pair.[22] I exclude all non-rail shipments from my main regression analysis on coal markups.

I also use monthly average delivered fuels prices from EIA 423/923 data to construct the cost ratio CR_{ud} in Equation (6). For coal prices, I use the BTU-weighted average monthly price received by each utility-owned coal plant, linearly interpolating prices for any missing months (P_{jm} in Equation (A1)). However, averaging delivered prices by month would obscure day-to-day variation in gas prices. I use prices from natural gas trading hubs to construct daily prices for each gas plant, comparing monthly average hub prices with monthly average 423/923 prices to add retail distribution costs into Z_{gd} in Equation (A2). (See Appendix G.5 below.)

G.1.2 Aggregated coal prices

I use aggregate coal price data from two sources. First, I use average mine-mouth sales prices at the county-year level, published in EIA's *Annual Coal Report* (ACR).[23] EIA discloses the average annual price for open market coal sales for counties with a sufficient number of mining firms to disclose aggregate prices, which corresponds to 62% of coal production. The ACR also reports average prices at the state-year level, separately for surface and underground mines, which I combine with coal production data (described below) to algebraically infer average prices for withheld counties. I control for these county-by-year average prices in $\mathbf{C}_{ojms}$, allowing me to better isolate delivered coal markups by capturing within-county changes in coal price.[24]

Second, I use monthly average prices for coal delivered to electric power plants, published in EIA's *Electric Power Monthly* (EPM). These prices are state-by-month aggregates, and EIA reports separate average prices for utility vs. non-utility plants (a.k.a. independent power producers).[25] This lets me populate average delivered coal prices for non-utility plants, for which 423/923 prices are masked. EPM withholds prices for state-months with too few firms, and I assign average prices for withheld cells by algebraically inferring missing prices from aggregate quantities (where possible), or using region-month average prices. I only use these prices in the fuel cost ratio for my counterfactuals (CR_{ud} in Equation (D5)). I also use average EPM prices for natural gas deliveries (constructed analogously) for sensitivity analysis in Figure C6 and in Figure E8 (first row of Panel G).

22. I assign the missing 2002–2007 transportation modes for the same oj pair as "rail" with a high degree of confidence. This backwards extrapolation is unambiguous for the vast majority of oj pairs.

23. I extract average coal sales prices from the data tables: https://www.eia.gov/coal/annual/.

24. Month fixed effects control for changes in the global commodity price, while coal county fixed effects control for each county's average mine-mouth price. County-year average prices control for cross-county differences in coal price that aren't captured by controlling for BTU, sulfur, and ash content.

25. EPM data tables are available at https://www.eia.gov/electricity/monthly/. Tables 4.10.A and 4.10.B report average delivered coal prices, while Tables 4.6.A and 4.6.B report average delivered coal quantities.

G.1.3 Coal production and mine characteristics

I use several publicly available datasets on coal mining and production, published by the Mine Safety and Health Administration (MSHA).[26] The "Mines" dataset reports mine-specific characteristics, all linked to a unique longitudinal mine identifier. For each mine, I observe its name; status (e.g., active, abandoned); county; latitude and longitude; and average seam thickness (or height).[27] The "Quarterly Mine Employment and Coal Production Report" dataset (MSHA Form 7000-2) reports coal production, average number of employees, and total employee-hours worked, disaggregated by mine subunit (i.e. underground vs. surface operations).

Using MSHA mine identifiers, I merge these two datasets with EIA's Form 7A, or "Annual Survey of Coal Production and Preparation".[28] These data allow me to cross-validate mine-specific production and employment for each year. EIA also classifies each mine as either "surface" or "underground" (consistent with definitions in the Annual Coal Report). EIA splits mines into 8 distinct basins/regions: Appalachia Northern, Central, and Southern; Illinois Basin; Powder River Basin; Uinta (Utah, Colorado, and southern Wyoming); Interior (between Illinois and Wyoming/Colorado); and Western (non-Powder River Basin, non-Uinta).

Following Cicala (2015), I use stratigraphic data from the U.S. Geological Survey (USGS) to calculate the depth of mine seams, as coal closer to the surface tends to be less expensive to extract.[29] I rasterize both USTRAT and COALQUAL data in order to assign each coal mine a time-invariant depth, basic on its geographic coordinates.

I use mine coordinates and production data to determine the time-invariant production-weighted geocoordinates for each coal producing county. This serves as an input into my rail graph algorithm (see Appendix F.1). I use the above datasets to construct several time-varying controls, in order to add mine-specific controls to $\mathbf{C}_{ojms}$ (see Panel D of Figures E2 and E8).

G.2 Power plant data

G.2.1 Plant characteristics and operations

EIA's Form 860 (a.k.a. "Annual Electric Generator Report") surveys all U.S. electric power plants with at least 1 MW in total generating capacity.[30] For each plant-year, these data report the plant name and identifier; county; parent utility name and identifier; regulatory status; a cogeneration indicator; and the plant's primary purpose (i.e., to sell electricity, for all plants in my sample). I also observe characteristic of each plant's constituent generating units, including each generator's type (e.g., steam turbine, combined-cycle, combustion turbine); status (e.g., operating, retired); year constructed; nameplate capacity (i.e., the maximum megawatts a generator is built to produce); and primary fuel consumed (e.g., natural gas, bituminous coal).

I use data on power plant operations, collected by the following EIA forms, in reverse chronological order: 923, 906/920, 906, and 759.[31] My primary variables of interest are monthly heat input by fuel (i.e. fuel consumption in MMBTUs), and net electricity generated from each

26. These data are available at http://www.msha.gov/OpenGovernmentData/OGIMSHA.asp.

27. Thicker mine seams, or coal strata, are associated with less expensive coal extraction costs.

28. EIA Form 7A datasets are available at http://www.eia.gov/coal/data.cfm#production.

29. Data are here: https://ncrdspublic.er.usgs.gov/ncrds_data/, https://ncrdspublic.er.usgs.gov/coalqual/.

30. EIA Form 860 data are available at http://www.eia.gov/electricity/data/eia860/.

31. Since 2008, Form 923 has collected both fuel receipts and plant operations. For 2001–2007, Forms 906 and 920 collected generation data. Prior to 2001, these data were collected separately for utility (Form 759) vs.

fuel (i.e. MWh sold to the electricity grid), which are both reported at the plant-month level. The data also report monthly fuel consumption disaggregated to the boiler level, and monthly net generation disaggregated to the generator level. These sub-plant units often map 1-to-1, where a single unit boils water and generates electricity; however, in many cases, multiple boilers map to a single generator or vice versa.[32] These operations data allow me to calculate plants' utilization rates and heat rates, or inverse thermal efficiency in units of MMBTU/MWh.

EIA Forms 767, 860, and 923 collect detailed data on plants' pollution abatement costs and pollution control devices.[33] At the plant level, these forms report annual capital expenditures on pollution abatement, and annual operation and maintenance costs of pollution control devices such as scrubbers (flue gas desulfurization units) and flue gas particulate collectors. They also report revenues from selling plant byproducts (e.g. gypsum from scrubbers). At the boiler level, they report detailed data on scrubber characteristics and operations, while providing a crosswalk to match boiler identifiers with generator identifiers. I use these data to control for the presence of a scrubber in $\mathbf{X}_{jm}$, the main pollution control technology that influences plants' coal purchases. I also use variable environmental costs (O&M net of byproduct revenues) in constructing unit-specific marginal costs (i.e. MC^{env} in Equations (A1)–(A2)).

Finally, I use the Environmental Protection Agency's (EPA) Emissions & Generation Resource Integrated Database (eGRID) to assign plant geographic coordinates and electricity markets regions.[34] EPA assigns plants to regions of the electricity grid based at three distinct hierarchies. First, North American Electric Reliability Corporation (NERC) regions define 8 contiguous reliability regions of the transmission grid; there is substantial trade of electricity within but not across NERC regions. Second, many plants participate in wholesale electricity markets, and eGRID data assign these plants to a particular market, or Independent System Operator (ISO). I combine these two definitions for my demand estimation, giving preference to the ISO market regions while using NERC regions for plants that do not sell to an ISO.[35]

Third, eGRID assigns most plants to a power control area (PCA), or the plant's region on the transmission grid, over which a single Balancing Authority dispatches plants to instantaneously meet demand. Applying a consistent PCA definition across plants is not trivial, as PCA boundaries evolve over time. Cicala (2022) identifies 98 major PCAs in the U.S., excluding Alaska and Hawaii. I use a multi-step PCA matching procedure to ensure that PCA definitions are consistent across plants and sample years, similar to the procedure used by Linn and Muehlenbachs (2018). I cross-verify these definitions with two additional data sources: (i)

non-utility plants (Form 906). All years of data are available at http://www.eia.gov/electricity/data/eia423/, https://www.eia.gov/electricity/data/eia923/, or https://www.eia.gov/electricity/data/eia923/eia906u.html.

32. Form 906/920 boiler and generator data are both missing for 2006–2007.

33. Prior to 2005, EIA collected these data on Form 767 (available at http://www.eia.gov/electricity/data/eia767/). Since 2007, most of these information has been collected on Forms 860 and 923.

34. EPA's eGRID data are available at http://www.epa.gov/energy/egrid/, however these data only exist for the following years: 1996–2000, 2004, 2005, 2007, 2009, 2010, 2012, and 2014.

35. The 7 ISOs are California ISO (CAISO); Electric Reliability Council of Texas (ERCOT); ISO New England (ISONE); Midcontintent ISO (MISO, formerly Midwest ISO); New York ISO (NYISO); PJM (formerly Pennsylvania-New Jersey-Maryland); and Southwest Power Pool (SPP). I define two market regions as NERC regions: Florida Reliability Coordinating Council (FRCC); and SERC Reliability Corporation (formerly Southeast Electric Reliability Council). My remaining two market regions subsets of NERC's Western region that exclude California: Northwest Power Pool (NWPP); and the Southwest Reserve Sharing Group (SRSG). I define market regions to be time-invariant. Following Linn and Muehlenbachs (2018), I make minor adjustments to ISO/NERC boundaries such that all PCAs lie within a single market region.

Federal Energy Regulatory Commission (FERC) Form 714, "Annual Electric Balancing Authority and Planning Area Report"; and (ii) electricity supply curves from SNL Financial.[36] I construct marginal cost ratios by averaging marginal costs across plants within each PCA.

G.2.2 EPA Air Markets Program data

My demand estimation and counterfactuals use high-frequency Continuous Emissions Monitoring Systems (CEMS) data on power plant operations.[37] CEMS data include all fossil generating units that are either larger than 25 MW in capacity, or whose emissions are regulated under an EPA program. For each unit, for every hour since 2000, the data report fuel input (in MMBTUs), gross generation (in MWh), and tons of SO_2, NO_x, and CO_2 emitted.

CEMS data report primary fuel, which I use to classify coal vs. gas units.[38] Units are classified as boilers, combined cycle, or combustion turbines. I take these definitions as given, treating the CEMS unit identifier as unit u throughout my analysis. For most coal plants, CEMS units correspond to boilers (not generators); I match 96% of CEMS units to an EIA boiler identifier. My main estimates include the subset of plants meeting all of the following criteria: (1) appearing in CEMS data; (2) at least one CEMS unit has coal as primary fuel; (3) categorized by CEMS as "electric utility"; (4) appearing as coal-consuming in EIA's 860 data; (5) receiving coal deliveries in EIA's 423/923 data; (6) at least 50 MW in total capacity.

CEMS data report *gross* electricity generation, or total power generated for each unit-hour. However, a power plant's relevant unit of economic output is *net* electricity generation, which subtracts power that is not sold to the grid.[39] EIA data report net generation at the unit-month level, and EIA's boiler-to-generator crosswalk lets me compare net vs. gross generation for most unit-months. I calculate average net-to-gross ratios to rescale hourly CEMS generation and more precisely measure electricity sold onto the grid. For plants that I cannot merge across EIA-CEMS datasets, I use linear projection to populate missing unit-months. I drop extreme outliers with net-to-gross ratios below 0.2 or above 2, following Cicala (2022).[40] Nearly all of my assigned net-to-gross ratios are between 0.91 and 0.94; the median plant sells (on average) 93% of its gross MWh onto the grid (see the left panel of Figure G1).

I also assign heat rates at the unit-month level, dividing each unit's MMBTU of fuel consumed per month (from EIA boiler-level data) by the unit's monthly net MWh (from EIA

36. FERC Form 714 data are available at https://www.ferc.gov/industries-data/electric/general-information/electric-industry-forms/form-no-714-annual-electric/data; they report names of all plants within each Balancing Authority (which are often close to isomorphic with PCAs). SNL (now S&P Global) data are proprietary, and available at https://www.spglobal.com/commodityinsights/en/commodities/.

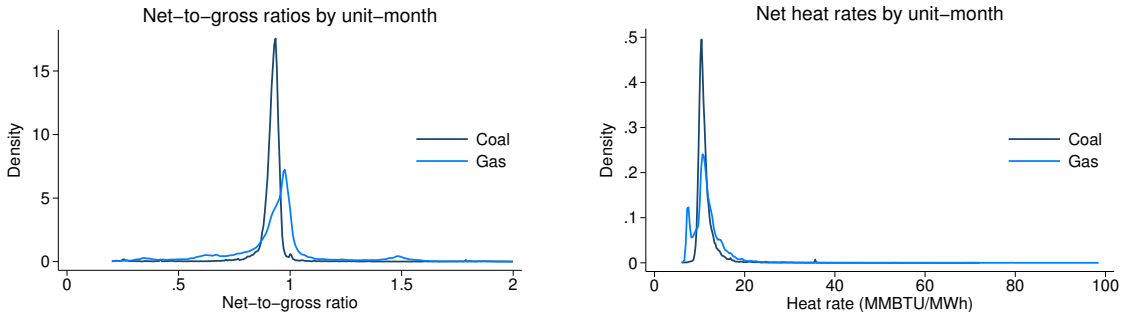
37. See http://ampd.epa.gov/ampd/. These data files are available for bulk download at both daily and hourly temporal resolutions, at ftp://ftp.epa.gov/dmdnload/emissions/daily/ and ftp://ftp.epa.gov/dmdnload/emissions/hourly/. For a detailed (archived) factsheet on CEMS data protocols, see https://web.archive.org/web/20090211082920/http://epa.gov/airmarkets/emissions/continuous-factsheet.html.

38. I assume that units are 100% coal when their primary fuel is listed as coal, and 0% coal units otherwise. My demand estimation omits the portions of these units' time series during which their primary fuel is gas.

39. Plants use some generation on-site (e.g., to run scrubbers). Net generation is the unit of the electricity supply curve. CEMS data report gross generation, since all MWh contribute pollution.

40. Some combined-cycle gas plants have net-to-gross ratios as large as 1.4, implying that they sell 140% of gross generation onto the grid. This reflects the fact the CEMS data include only the steam portion of the combined-cycle unit, while the turbine cycle does not report to CEMS. I abstract from plants' startup and shutdown periods (as do Davis and Hausman (2016); and Cicala (2022)); my net-to-gross calculations smooths generation expended (but not sold) during startup and shutdown periods across each unit-month.

Figure G1: Kernel densities: Net-to-gross ratios and heat rates



Notes: I plot kernel densities for net-to-gross ratios and heat rates, where each CEMS unit has a separate monthly observation. The left panel reveals greater dispersion in net-to-gross ratios for gas units, and that coal units tend to expend more electricity on site (i.e. lower net-to-gross ratios). The right panel shows how virtually all heat rates are between 6–20 MMBTU/MWh, with combined-cycle gas units surpassing coal units in efficiency (i.e. lower heat rates, which is the inverse of thermal efficiency).

generator-level data).[41] Following Cicala (2022), I remove outliers with heat rates below 6 MMBTU/MWh and above 100 MMBTU/MWh. I also use linear projection to populate missing monthly heat rates within each plant. For both net-to-gross ratios and heat rates, a substantial share of CEMS units do not match to EIA generator-level data (due to incompleteness in the boiler-to-generator crosswalk). For these units, I assign net-to-gross ratios and heat rates by plant-unit-type-month and then by plant-month. Virtually all heat rates are between 6–20 MMBTU/MWh (see the right panel of Figure G1).

The dependent variable in Equation (6) is unit u's CEMS hourly generation divided by its capacity, or its capacity factor CF_{uh}. Since nameplate capacity is inconsistently reported in EIA unit-level data, I assign capacity as maximum observed hourly generation for unit u in month m. This accommodates seasonal differences in plant capacity due to temperature variation, and other operational constraints. It also ensures that $CF_{uh} \in [0, 1]$ by construction.

To construct marginal costs, I multiply unit-specific heat rates by plant-specific fuel prices—from monthly EIA coal delivery data for coal units (see Appendix G.1.1), and from daily hub-specific prices for gas units (see Appendix G.5). I add marginal environmental costs using EIA 767 data (see Appendix G.2.1), merged at the unit-level where possible. I also add the implied marginal costs of SO_2, NO_x, and CO_2 emissions, for units covered by allowance trading programs. I use EPA Air Markets Program Data to assign monthly participation dummies for the 7 major allowance trading programs listed in Table G1. To monetize the implied costs of emissions under each of these programs (as part of MC^{env} in Equations (A1)–(A2)), I multiply these unit-month-specific participation dummies by unit-month-specific emissions rates for each relevant pollutant (in tons per net MWh, as calculated from CEMS SO_2, NO_x, and CO_2 data), and by average monthly allowance prices for each program (see Appendix G.6).

G.2.3 Non-fuel variable costs

Both coal- and gas-fired electricity production are Leontief in fuel inputs (Fabrizio, Rose, and Wolfram (2007)). I do not account for additional variable input costs (e.g. labor, maintenance) in my coal demand estimation, mainly because reliable data are not widely available. I am less concerned with characterizing plants' production functions than with predicting generation conditional on the cost of one input to production (i.e. fuel). If I could credibly control for

41. Heat rates are inverse thermal efficiency: MMBTUs in per MWh out. I use fuel consumption from EIA, since CEMS fuel use data appear less precisely measured than CEMS generation and emissions data.

Table G1: Allowance trading programs for SO_2, NO_x, and CO_2

Program	Years in place	Geographic coverage	Pollutants traded
Acid Rain Program (ARP)	1995–present	48 states + DC	SO_2
Ozone Transport Commission (OTC) NO_x Budget Program	1999–2002	10 eastern states	NO_x (May–Sept)
State Implementation Plan (SIP) NO_x Budget Trading Program (NBP)	2003–2008	23 eastern states	NO_x (May–Sept)
Clean Air Interstate Rule (CAIR)	2009–2015	26 eastern states 26 eastern states 26 eastern states	SO_2 NO_x (May–Sept) NO_x (annual)
Cross-State Air Pollution Rule (CSAPR)	2015–present	23 eastern states 25 eastern states 28 eastern states	SO_2 NO_x (May–Sept) NO_x (annual)
Regional Greenhouse Gas Initiative (RGGI)	2009–present	10 eastern states	CO_2
California Cap and Trade Program	2013–present	California	CO_2

Notes: EPA implemented 5 major allowance trading programs during my sample period: ARP, OTC, NBP, CAIR, CSAPR. I also include 1 regional program (RGGI) and 1 state-level program (CA cap and trade). Many plants must purchase two separate allowances for NO_x emissions, to comply with the annual NO_x requirements and also with the more stringent ozone-season NO_x requirement (from May to September). For information on each program see (from top to bottom of the table): https://www.epa.gov/airmarkets/acid-rain-program , https://www.epa.gov/airmarkets/ozone-transport-commission-nox-budget-program , https://www.epa.gov/airmarkets/nox-budget-trading-program , https://archive.epa.gov/airmarkets/programs/cair/web/html/index.html , https://www.epa.gov/csapr , https://www.rggi.org/ , https://www.arb.ca.gov/cc/capandtrade/capandtrade.htm

variation in these non-fuel, non-environmental costs (as a part of CR_{ud} in Equation (6)), this might increase the precision of my $\langle \hat{\lambda}_{0j}, \hat{\lambda}_{1j}, \hat{\lambda}_{2j} \rangle$ estimates. Ignoring this variation is unlikely to induce bias, as it is unlikely to be correlated with a price-taking plant's fuel costs.

Cicala (2022) omits non-fuel, non-environmental costs, arguing that these costs are second-order. By contrast, Davis and Hausman (2016) include time-invariant, technology-specific estimates of plants' variable operations and maintenance costs. I conduct sensitivity analysis in Figures A6 and E10 that similarly incorporates technology-specific non-fuel cost estimates, using SNL's default cost assumptions for each CEMS unit type: coal boilers ($2.67/MWh), gas boilers ($3.04/MWh), combined-cycle gas ($1.26/MWh), and gas turbines ($6.81/MWh).[42]

G.3 Rail data

G.3.1 GIS rail data

I use GIS data on the U.S. rail network published by the Bureau of Transportation Statistics (BTS). I use four BTS shapefiles of the rail network, published in 2014, 2012, 2011, and 2006.[43] Each shapefile includes a rail line-specific attribute dataset including: identifiers for the line's two terminal nodes; line length in miles; line type (e.g. mainline, non-mainline, abandoned); an

42. These defaults average across 2009–2015 SNL supply curves, available at https://www.spglobal.com/commodityinsights/en/commodities/ (subscription required). SNL reports yearly non-fuel, non-environmental cost data for a subset of unit-years, reportedly from FERC Form 1. However, SNL data are only moderately correlated with FERC Form 1 data (https://www.ferc.gov/docs-filing/forms/form-1/data.asp), and each dataset covers only a fraction of CEMS units. Hence, I choose to assign default values consistently across all units.

43. BTS GIS datasets are available at https://data-usdot.opendata.arcgis.com/search?tags=Rail. The U.S. Geological Survey (USGS) has also published less complete shapefiles of the U.S. rail network (available at https://data.usgs.gov/datacatalog/data/USGS:ad3d631d-f51f-4b6a-91a3-e617d6a58b4e) Pre-2005 shapefiles are not publicly available online.

indicator for freight (as opposed to passenger) lines; each line's primary, secondary, and tertiary owners (where applicable); and a list of rail carriers with trackage rights on each line.

I focus exclusively on Class I rail carriers, with annual operating revenues greater than $453 million. This revenue threshold associated with the Class I designation has increased gradually over time, as defined by the Surface Transportation Board (STB). 7 Class I rail carriers currently operate in the U.S.: BNSF and Union Pacific (UP) in the West; CSX (CSXT) and Norfolk Southern (NS) in the East; and Canadian National (CN), Canadian Pacific (CP), Kansas City Southern (KCS), three carriers with smaller geographic footprints.

I merge node and line identifiers across annual shapefiles. 98% of mainline rail lines merge across all four shapefiles, with identical latitudes and longitudes. 99% of Class I track mileage maintained constant ownership between 2006 and 2014. This let me treat the rail network as static. A similar analysis in the 1980–90s would need to account for consolidation of Class I (and non-Class I) carriers; fortunately for my application, the last Class I merger was in 1999.[44]

I calculate the shortest rail distance between each originating coal county and each coal plant (see description in Appendix F). As an additional shipping cost control, I proxy for rail network congestion using the reported rail traffic density of each line. BTS measures traffic density in million gross tons (MGT), and the 2014 (2011) shapefile reports densities for 2011 (2009) in discretized categories. I construct a high-density indicator equal to 1 if a rail line has density greater than 50 MGT in 2011 *or* greater than 40 MGT in 2009, and equal to 0 otherwise. This classifies 11% of total track-miles as "high-density", and 18% of mainline track-miles as "high-density". I integrate this indicator across all track-miles on each oj shortest route to control for the fraction of each shipping route on high-density lines (a component of $\mathbf{T}_{ojms}$).

G.3.2 Rail shipping costs

I use the Association of American Railroads (AAR) fuel price index to control for changes in the cost of rail freight. This index summarizes changes in the average price per gallon of No. 2 diesel fuel paid by rail carriers, based on monthly surveys of rail operators. It is a single industry-wide number published monthly, inclusive of federal excise taxes and transportation/handling costs. The AAR fuel price index is a component of the Surface Transportation Board's (STB) Rail Cost Adjustment Factor (RCAF). The RCAF is published quarterly (not monthly), and combines 7 different price indices into a single number summarizing changes to rail freight costs: diesel fuel; labor; materials and supplies; equipment rents; depreciation; interest; and other expenses. The STB uses the RCAF as a basis for adjudicating rail rate cases.[45]

I constructed a monthly fuel price index time series using publicly available data from the AAR website.[46] I similarly construct a quarterly RCAF time series from several historic documents published on both the STB and AAR websites.[47] I use the AAR fuel price index in

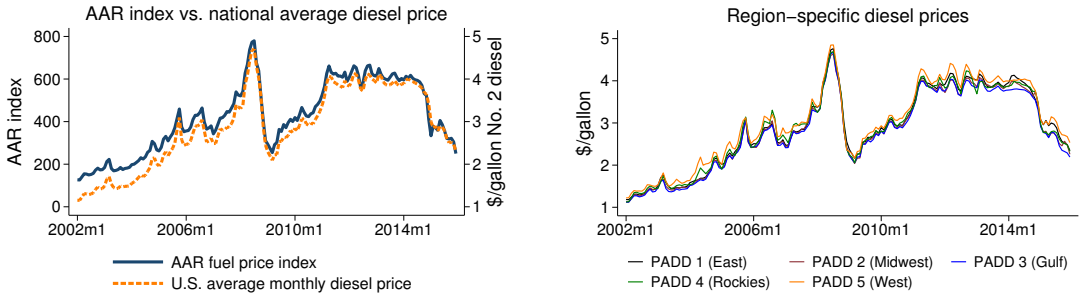
44. See: https://en.wikipedia.org/wiki/Timeline_of_Class_I_railroads_(1977-present).

45. For more details, see https://www.aar.org/wp-content/uploads/2018/03/Index_RCAFDescription.pdf.

46. AAR's website has changed since I built this time series. I used a spreadsheet of historic fuel price indices (2003–2012), formerly available at https://www.rita.dot.gov/bts/publications/multimodal_transportation_indicators/2013_02/fuel_prices/railroad_fuel. For 2013–2015, I digitized PDFs published by AAR. For 2002, I used average monthly U.S. diesel prices to extrapolate backwards (see next paragraph).

47. 2005–2015 RCAF indices were available at https://www.aar.org/data-center/rail-cost-indexes. For remaining years, I digitize PDFs on AAR's website (no longer available) and queried STB decisions (https://www.stb.gov/Decisions/), query "quarterly rail cost adjustment factor").

Figure G2: AAR fuel price index and diesel prices



Notes: The left panel compares the time series of the monthly AAR fuel price index to the U.S. national average No. 2 diesel price. The correlation between the two series is 0.99. The right panel plots region-specific diesel prices for the 5 PADD regions, revealing virtually no cross-sectional variation in diesel prices.

$\mathbf{T}_{ojms}$ (rather than the RCAF), since the RCAF provides only quarterly variation. Figures E2 and E8 (Panel E) show that my DD results are not sensitive to this choice.

The left panel of Figure G2 shows that the AAR survey-based price index closely tracks U.S. monthly average diesel prices.[48] PADD-specific price series are highly correlated, with all pairwise correlations greater than 0.99 (see the right panel of Figure G2).[49] Hence, the AAR time series does not obscure important cross-sectional variation in diesel prices. Figures E2 and E8 (Panel E) show that my results robust to using PADD-specific diesel prices in $\mathbf{T}_{ojms}$.

G.4 Distance to navigable waterway

I use GIS data to identify the subset of coal plants with access to barge shipments. These data come from three sources. First, I use the U.S. Army Corps of Engineers Waterway Mile Marker Database, which reports the locations of all navigable inland U.S. waterways (i.e., rivers and the Gulf Intracoastal Waterway).[50] Second, I use a shapefile of the Great Lakes from *Natural Earth*.[51] Third, I use a shapefile of U.S. coastlines, also from *Natural Earth*.[52] I calculate the minimum distance of each coal plants to the nearest navigable river, Great Lake, or coastline. I set the "water option" indicator $W_j = 1$ for plants with a minimum distance less than 1.5 miles.

The vast majority of coal-by-barge deliveries are to plants within 1.5 miles of a navigable river, Great Lake, or coastline. While a 0-mile threshold might seem appropriate, the Army Corp of Engineers data record navigable rivers at 1-mile intervals. This discreteness necessitates a non-zero threshold. A 1.5-mile threshold correctly assigns several plants that receive exclusively waterborne deliveries despite being non-adjacent to navigable water—these plants have constructed long conveyor belts to carry coal overland from barge offloading points.

I cross-check this GIS-derived indicator with transportation modes listed in EIA's 423/923 data. 15 coal plants receiving a substantial share of coal deliveries by water are not located within 1.5 miles of a river, Great Lake, or coastline. After manually checking each of these plants in Google Earth, I find that they are all located a few miles from a Great Lake or

48. Available at https://www.eia.gov/dnav/pet/PET_PRI_GND_A_EPD2D_PTE_DPGAL_M.htm.

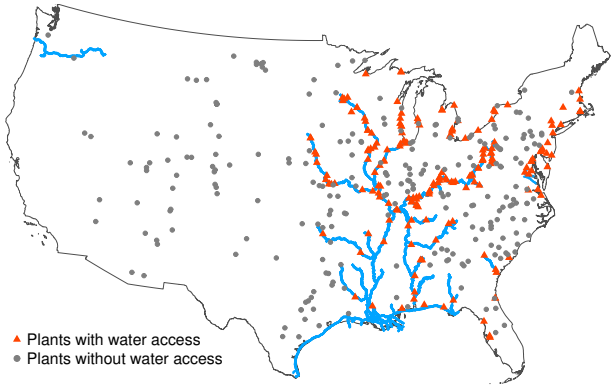
49. The five Petroleum Administration for Defense Districts (PADDs) are defined as East Coast (PADD 1), Midwest (PADD 2), Gulf Coast (PADD 3), Rocky Mountains (PADD 4), and West Coast (PADD 5).

50. https://geospatial-usace.opendata.arcgis.com/datasets/604cdc08fe7d43cb90a0584a0b198875_0/explore

51. These data are a subset of a global lakes shapefile, available at http://www.naturalearthdata.com/http//www.naturalearthdata.com/download/10m/physical/ne_10m_lakes.zip.

52. These data are a subset of a global coastlines shapefile, available at http://www.naturalearthdata.com/http//www.naturalearthdata.com/download/10m/physical/ne_10m_coastline.zip.

Figure G3: Coal plants with access to barge shipments



Notes: I plot all 430 coal plants in my unrestricted sample. Red triangles denote plants with a barge option (i.e., $W_j = 1$); gray circles denote plants without a barge option (i.e., $W_j = 0$). Blue lines map all navigable rivers and the Gulf Intracoastal Waterway.

coastline on small inlets (which do not appear in my GIS data). I switch $W_j = 1$ for each of these plants. Figure G3 maps the full sample of 430 coal plants, splitting on $W_j = 1$ vs. $W_j = 0$. I also map navigable inland waterways, which represent only the largest U.S. rivers.[53]

G.5 Natural gas prices

I use natural gas price data from SNL, which reports prices for 104 trading hubs at locations throughout the U.S. pipeline network.[54] The primary hub is Henry Hub in Erath, Louisiana, which establishes a standard commodity price for U.S. gas markets. The Henry Hub average monthly gas price (i.e. Z_m) provides identifying time series variation in Equation (5).

Compared to coal, natural gas behaves more like a uniform-price commodity market—at least within pipeline network spanning the continental U.S. However, variation in pipeline costs creates some regional variation in U.S. gas prices: plants that are farther from gas-producing regions tend to pay higher prices. Pipeline capacity constraints also lead to gas price dispersion, especially in winter months in New England (due to spikes in demand for gas heating).

I leverage both cross-sectional and time-series variation in gas prices. For each of SNL's 104 trading hubs, I manually assign geographic coordinates using SNL's mapping interface. Then, I match each gas-fired CEMS unit to its nearest trading hub (by straight-line distance). Figure G4 maps gas trading hubs and CEMS gas-fired generating units; I can assign very localized prices in some regions (i.e. Oklahoma) but not others (i.e. Florida). Figure G5 illustrates how daily gas prices can wildly diverge across trading hubs. SNL does not report complete daily time series for all hubs, and I populate missing values in each series via linear projection.[55]

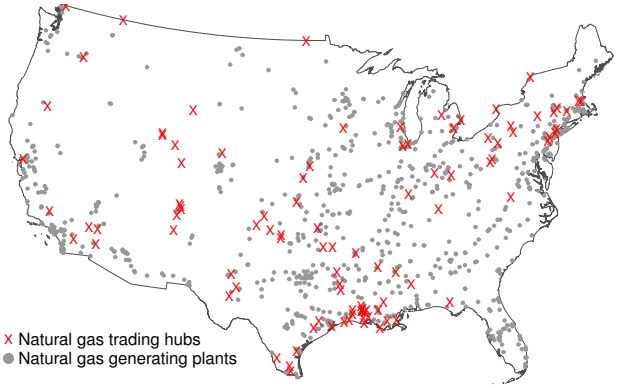
While SNL reports hub-specific *wholesale* gas prices, power plants pay *retail* gas prices that reflect additional pipeline fees for the final portion of the gas distribution (i.e., the smaller pipelines that connect hubs to plants). Electricity regulators and utilities also fund gas pipeline

53. Water is a key input to thermal electricity generation—fossil fuel combustion creates heat, which boils water to create steam Hence, most coal plants are located adjacent to some water source, such as a river or lake. Figure G3 shows that many such water sources are not navigable, and could not feasibly convey coal barges.

54. These data are in SNL's (now S&P Global's) "Market Prices" menu (Advanced Search → Commodity), and require a subscription. Henry Hub prices are available at http://www.eia.gov/dnav/ng/hist/rngwhhdd.htm.

55. Geographically proximate hubs tend to have highly correlated prices. In fact, using roughly 20 gas hubs (from representative regions) would be sufficient to characterize nearly all cross-sectional dispersion in daily gas prices. My results would be quite similar if I matched gas plants only to hubs with complete time series.

Figure G4: Natural gas trading hubs and gas plants



Notes: I plot all natural gas trading hubs with available daily price data from SNL (red Xs). I also map all power plants with a CEMS gas-fired generating unit that was active during my sample period (gray dots).

construction and investment cost recovery by raising plants' marginal fuel prices. I compare SNL monthly average hub prices to EIA 423/923 monthly average delivered gas prices; I estimate wholesale-to-retail price adjustment factors at the plant-month level (where possible) and then at the state-year level, using linear interpolation to populate missing plant-months. I then add each plant-month-specific adjustment factor to each plant's matched daily hub price to arrive at Z_{gd} in Equation (A2), or the gas price paid by gas unit g on day d.

G.6 Allowance prices for SO_2, NO_x, and CO_2

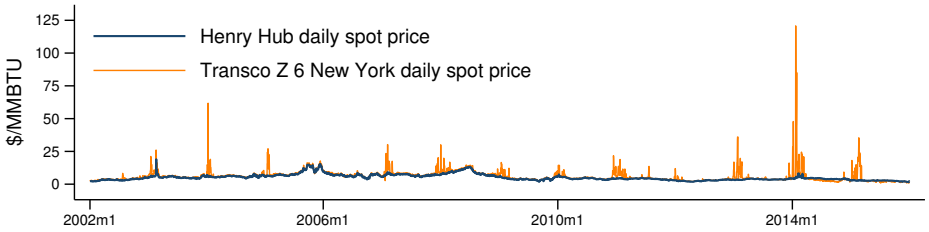
I construct time series of monthly allowance prices in order to monetize each fossil generating unit's marginal environmental costs. My primary data source for allowance prices is SNL, which reports average monthly prices for each of the emissions trading programs listed in Table G1.[56] However, these data do not cover all months in my 2002–2015 sample. I supplement SNL allowance prices with several additional data sources, to populate full monthly time series:

- ARP SO_2 allowance prices for 2002–2005 from BGC Environmental Consulting.[57]
- OTC seasonal NO_x allowance prices for 2002 from BGC Environmental Consulting.[58]
- NBP seasonal NO_x allowance prices for 2003–2005 from EPA annual progress reports.[59]
- CAIR annual NO_x allowance prices for 2009–2011 from EPA annual progress reports.[60]
- CSAPR prices for all three allowance types for 2015 from Evolution Markets.[61]
- RGGI quarterly CO_2 allowance auction results for 2008–2012 from RGGI.[62]

I use linear interpolation to populate months that are still missing.

56. Available for download in SNL's (now S&P Global's) "Market Prices" menu (subscription required).
57. http://www.bgcebs.com/registered/aphistory.htm
58. http://www.bgcebs.com/registered/apnx0623.htm
59. https://www.epa.gov/sites/production/files/2015-08/documents/noxreport03.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/ozonenbp-2004.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/2006-nbp-report.pdf
60. https://www.epa.gov/sites/production/files/2015-08/documents/cair09_ecm_analyses.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/arpcair10_analyses.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/arpcair11_analyses_0.pdf
61. http://www.evomarkets.com/content/news/reports_10_report_file.pdf; http://www.evomarkets.com/content/news/reports_12_report_file.pdf
62. https://www.rggi.org/Auctions/Auction-Results/Prices-Volumes

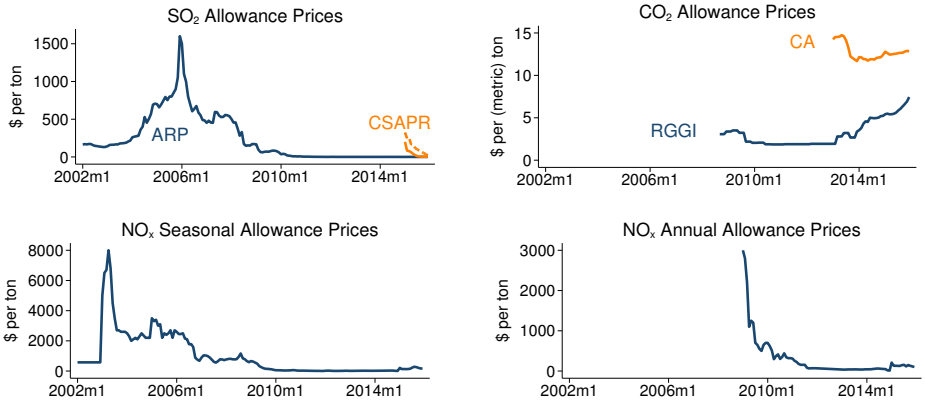
Figure G5: Natural gas daily hub-specific price variation (example)



Notes: This figure plots daily SNL prices for two gas trading hubs: Henry Hub (in navy), and Transco Z 6 hub in New York (in orange). This illustrates how on a given day, gas prices faced by plants in the Mid-Atlantic region may diverge by over $100/MMBTU from those faced by gas plants in Louisiana. Transco Z 6 price spikes occur in winter months due to regional supply constraints.

The resulting allowance price time series are imperfect. Allowance markets were thin during parts of my sample period, and low trading volumes likely explain many data gaps. In using prevailing allowance prices to monetize plants' marginal pollution costs, my goal is to approximate the contemporaneous price signal to which plants optimized generation decisions. Even if SNL time series were complete across all months, these prices would likely still mismeasure plants' (interpretation of their) shadow costs of SO_2, NO_x, and CO_2 emissions. Figure G6 plots allowance price time series for SO_2, CO_2, seasonal NO_x, and annual NO_x. There has been tremendous variation in SO_2 and NO_x prices: spikes/drops reflect adjustments in expectations immediately following (announced) policy changes (Schmalensee and Stavins (2013)).

Figure G6: Allowance price time series



Notes: This figure plots allowance price time series for the four tradable allowance types. Orange lines in the top-left panel plot SO_2 allowance price for two groups of plants under CSAPR. The orange line in the top-right panel plots California cap-and-trade allowance prices (in $ per metric ton), while the blue line plots RGGI allowance prices (in $ per short ton). For both type of NO_x allowance price series, I plot a single prevailing allowance price series that spans 2–4 distinct policy periods. All prices are nominal.

G.7 Temperature data

The PRISM Climate Group at Oregon State University maintains daily (and monthly) spatial datasets of temperature, precipitation, dew point, and vapor pressure, for the conterminous United States. Each daily dataset incorporates readings across 20 separate networks of weather stations, and applies a spatial interpolation algorithm to produce gridded rasters at 4-kilometer resolution.[63] I project plant coordinates onto each day's maximum temperature raster to construct a panel of daily maximum temperature at each plant location.

63. I use PRISM's "AN81d" product, which exists at two resolutions: 4 km (for free) and 800m (for a fee). PRISM data are available at http://www.prism.oregonstate.edu, along with extensive documentation.

Supplemental appendix references

Busse, Meghan R., and Nathaniel O. Keohane. 2007. "Market Effects of Environmental Regulation: Coal, Railroads, and the 1990 Clean Air Act." *RAND Journal of Economics* 38 (4): 1159–1179.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2006. "Bootstrap-Based Improvements for Inference with Clustered Errors." Working Paper.

———. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–427.

Cicala, Steve. 2015. "When Does Regulation Distort Costs? Lessons from Fuel Procurement in U.S. Electricity Generation." *American Economic Review* 105 (1): 411–444.

———. 2022. "Imperfect Markets versus Imperfect Regulation in U.S. Electricity Generation." *American Economic Review* 112 (2): 409–41.

Covert, Thomas R., and Ryan Kellogg. 2018. "Crude by Rail, Option Value, and Pipeline Investment." NBER Working Paper 23855.

Davis, Lucas, and Catherine Hausman. 2016. "Market Impacts of a Nuclear Power Plant Closure." *American Economic Journal: Applied Economics* 8 (2): 92–122.

Efron, Bradley. 1987. "Better Bootstrap Confidence Intervals." *Journal of the American Statistical Association* 82 (397): 171–185.

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram. 2007. "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on U.S. Electric Generation Efficiency." *American Economic Review* 97 (4): 1250–1277.

Hall, Peter. 1992. *The Bootstrap and Edgeworth Expansion.* New York: Springer-Verlag.

Hausman, Catherine, and Ryan Kellogg. 2015. "Welfare and Distributional Implications of Shale Gas." *Brooking Papers on Economic Activity* Spring:71–125.

Hughes, Jonathan E. 2011. "The Higher Price of Cleaner Fuels: Market Power in the Rail Transport of Fuel Ethanol." *Journal of Environmental Economics and Management* 62 (2): 123–139.

Jha, Akshaya. 2022. "Regulatory Induced Risk Aversion in Coal Contracting at U.S. Power Plants: Implications for Environmental Policy." *Journal of the Association of Environmental and Resource Economists* 9 (1): 51–78.

Linn, Joshua, and Lucija Muehlenbachs. 2018. "The Heterogeneous Impacts of Low Natural Gas Prices on Consumers and the Environment." *Journal of Environmental Economics and Management* 89:1–28.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263–1297.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688–701.

Schmalensee, Richard, and Robert N. Stavins. 2013. "The SO_2 Allowance Trading System: The Ironic History of a Grand Policy Experiment." *Journal of Economic Perspectives* 27 (1): 103–122.

U.S. Government Accountability Office. 2006. *Freight Railroads: Industry Health Has Improved, but Concerns about Competition and Capacity Should Be Addressed.* Report to Congressional Requesters GAO-07-94.

Wooldridge, Jeffrey M. 2007. "Inverse Probability Weighted Estimation for General Missing Data Problems." *Journal of Econometrics* 141 (2): 1281–1301.