Partial identification of power plant productivity

Zach Flynn*

University of Wisconsin — Madison

January 2, 2017

Most recent version of paper

Abstract

Traditionally, productivity, the part of output that is not a result of observed input use, is modeled as a controlled or uncontrolled Markov process. I allow it to be a function of unobserved choices; plants differ in their productivity because they do different things, not only because they receive different shocks. If we think giving incentives to plants to be more productive will cause them to be so—if we think policy can affect productivity—then we must have a model in mind where productivity is a choice or else plants can not respond to the incentive. When we allow it to be a choice, the problem of identifying productivity becomes more difficult because productivity and inputs are now chosen on the basis of the same underlying state variables. No instrument can shift input choice without shifting productivity choice. *But* a broad class of economic models predict productivity and inputs are positively related in a certain statistical sense even when productivity is a function of unobserved inputs or other choices. I show this economic theory can meaningfully partially identify the effect of policy on productivity. I use the theory to bound the effect of restructuring in the electricity industry on power plant productivity, finding the policy lowered power plant productivity and arguing that an approach that allows productivity to be chosen by the power plant is necessary to estimate the effect of restructuring on power plant productivity.

Keywords: Partial identification, evaluating how policy affects productivity, electricity industry restructuring, regulation, plant productivity, allocation.

*Email: zflynn@wisc.edu. I thank Amit Gandhi, Alan Sorensen, Ken Hendricks, Enghin Atalay, Jack Porter, Xiaoxia Shi, Daniel Quint, Michael Dickstein, and other participants at Wisconsin department seminars for great feedback.

1. Introduction.

Why do some plants produce more output than other plants that use the same inputs? Are there policies that would make the less productive plants more productive? The answer matters. Producing more output with fewer resources is the basis of economic growth, and, if there are policies that can cause growth, we should consider pursuing them. But if policy can affect productivity, then it must be the case that plants can choose their productivity—or else, how could they respond to the policy?—and this introduces new identification problems.

If productivity and inputs are both choices, then without further assumption, they are both made on the basis of the same underlying state variables; there is no instrument that can shift one choice without shifting the other[1].

My goal is to learn how policy affects productivity so I need to allow productivity to be a choice and deal with the identification problem. I could add restrictions to the productivity model until I create a state variable which changes productivity without changing inputs, usually requiring timing assumptions about when shocks to productivity are known by the plant, how and when investment in productivity is made, and how and when input decisions are made (the current approach in the literature and the most closely related work to mine, see Doraszelski and Jaumandreu 2013 and De Loecker 2013). Instead, I leave the problem as it is, resign myself to the lack of instruments, and develop a partial identification approach to bound the effect of policy on productivity. What I gain by doing so is the ability to avoid arbitrary timing assumptions about when plants know and do things, "invertibility" assumptions common in the proxy literature which require only a limited number of unobserved state variables, or assumptions on the nature of competition (I do not assume perfect or monopolistic competition, a common assumption in the current methods).

I assume only a weak form of positive dependence between productivity and inputs[2] and that the production function is increasing. The positive dependence between productivity and inputs comes from the fact that productivity increases the marginal product of inputs in the input neutral production function I use (and that is commonly used in the literature). By using the partial identification approach, I am able to put little restriction on how productivity is chosen, how plants compete, or, most importantly, on how a policy can affect productivity.

The current identification arguments in the literature are based on the popular proxy method of estimating productivity. They argue the identification problem can be solved by using lagged inputs to "instrument" for current inputs[3]. The method has been developed in a large literature,

[1]The more general point: when residuals (productivity) are endogenous, instrumental variable assumptions presume that there is some underlying state variable that changes covariate choices (input choices) but not productivity choice which requires a "separability" between covariates and residuals in the choice problem which is difficult to justify here.

[2]A statistical assumption called, "linear positive association", with the advantage over alternative notions of positive dependence that I can show it is true in a broad class of economic models.

[3]More precisely, the proxy methods argue lagged inputs proxy for lagged productivity and, if productivity is an uncontrolled Markov process, controlling for lagged productivity leaves only the unobserved innovation shock to productivity which is uncorrelated with any inputs chosen before the plant knows the shock. Assuming the plant decides on its input choices before the innovation shock then give identification of the production function, but see Gandhi, Navarro, and Rivers (2015) for identification problem when the timing assumption is not satisfied.

see Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg, Caves, and Frazer (2006), and Gandhi, Navarro, and Rivers (2015), and applied widely in the industrial organization and international trade literature (see Syverson 2011 for a survey of applications).

The traditional proxy method treats productivity as an uncontrolled Markov process and does not allow productivity to be a choice so policy can not affect plant productivity under its assumptions[4]. But Doraszelski and Jaumandreu (2013) and De Loecker (2013) modified the proxy method to allow productivity to be a controlled Markov process so productivity can depend on a plant's choices and be affected by policy. They require investment in productivity to be a function of observed variables chosen before any current unobserved shocks to productivity—we have to observe how plants invest in productivity, an unobserveable. Productivity must also be a *fixed* input and not respond immediately to unobserved shocks to productivity. Policy can then affect productivity if it affects observed investment—or if it changes the Markov process—but then a timing assumption is necessary for identification: the plants must adjust to the policy with a lag as long as the lag of the data because policy must be uncorrelated with current shocks to productivity.

These assumptions restrict the nature of productivity and the extent to which it can respond to policy. I show we can meaningfully partially identify the policy effect of interest without making them.

Relative to the controlled Markov proxy approach, I allow productivity to be a function of unobserved inputs or other choices (effort, the organization of production, unobserved capitals, etc.). Some of these may be fixed and respond to policy with a lag as long as the data, as allowed by the controlled Markov proxy approach, but other inputs or choices may be flexible. If we do not know what productivity is physically, it may be difficult to argue what observable variables determine investment in productivity or that we know whether productivity is a function of inputs fixed over whatever time interval we happen to observe our data or whether the inputs are flexible. I also do not need to assume either perfect or monopolistic competition as is common in many proxy approaches—especially in the presence of flexible inputs, see Gandhi, Navarro, and Rivers (2015).

I use the new identification approach and the empirical method I develop, the first contribution of the paper, to evaluate how restructuring in the electricity generation industry affected power plant productivity, allowing productivity to be a choice, the second contribution of the paper.

Historically, electric utilities were vertically integrated across the three stages of production: generation, transmission, and retail. The price they received for their electricity was tied to the costs they incurred—roughly, price was set to the average cost of the utility plus a regulated rate of return—so, if a utility lowered its costs, it would receive a lower price, giving it less of an incentive to reduce costs. In the mid-to-late 1990's, several US states restructured

[4]The Olley and Pakes (1996) proxy model *is* perfectly consistent with policy changing the distribution of surviving productivity or the extent to which more productive plants have greater market share, but it does not allow policy to change the productivity of a plant. Changes in the productivity of individual plants are changes in the production possibilities frontier; changes in the distribution of surviving productivity or in the allocation of output to more productive plants change the point inside the frontier that is selected. I focus on how policy changes the frontier while much of the previous literature has focused on allocation.

their electricity markets, splitting up the vertically-integrated utilities and letting competition among the firms generating electricity determine electricity price, while still regulating electricity transmission as a natural monopoly.

Did restructuring cause power plants to become more productive? Economic theory is ambiguous. Restructuring increased the incentive for a plant to be more productive because price was no longer tied to the plant's costs, but it also increased the competitiveness of the industry, encouraging entry, potentially reducing markups, and lowering the return to making investments in productivity. Which of these two opposing effects[5] dominates is an empirical question.

The two effects motivate the two contributions my method makes to the identification of productivity. The incentive effect requires that power plants can respond to the new incentive to be more productive, so I allow plants to make choices to become more productive. The nature of the competitive effect depends on how utilities compete, how positive of a markup they can expect to earn after becoming more productive. If I assumed perfect competition, there would be no markups to chase, limiting the incentive to become more productive. So I do not make strong assumptions on how plants compete, allowing the data to speak to how large the competitive effect is.

I find restructuring lowered power plant productivity. The competitive effects of restructuring dominated the incentive effects. I also find market size matters for productivity: when market size increases, power plants become more productive, a natural result in models where productivity depends on unobserved inputs but at odds with models where productivity is an exogenous process[6].

I also find market level productivity (output-share weighted productivity) fell as a consequence of restructuring. I do an Olley and Pakes (1996) decomposition of aggregate productivity into an average productivity term and a reallocation term measuring the covariance between market share and productivity. I find the negative average productivity effect of restructuring could not be overcome by the positive reallocation effect of the policy.

Fabrizio, Rose, and Wolfram (2007) also study the effect of restructuring on power plant productivity. They estimate the effect of restructuring on conditional input demand equations. In a conditional input demand equation, the endogeneity problem is that output is correlated with unobserved productivity. Fabrizio, Rose, and Wolfram (2007) use a proxy for demand (total electricity sales in a state, a measure of market size) to instrument for output in these equations, but demand is only uncorrelated with productivity if productivity is an exogenous process. If productivity is a function of unobserved inputs, a choice, then demand affects it, like demand affects observed input use and output; market size is not excluded.

When productivity is an input, that restructuring affects productivity and that the relationship

[5]Another potentially negative effect of restructuring on productivity is that vertical integration might have had some efficiencies that were lost with the forced split between electricity generation and transmission.

[6]When productivity is exogenous, market size either does not affect the distribution of surviving productivities (as in Melitz 2003) or it is negatively related to productivity because greater market size predicts a lower threshold productivity for entry (one of many examples: standard static model with perfect competition, unit mass of firms with a Pareto distribution of productivity, decreasing returns to scale, and demand has unit elasticity. The result is much more general than that model though.).

between market size and productivity is positive are natural results. But if productivity is an exogenous process, then policy can not affect it and there should not be a positive relationship between market size and productivity because greater market size should allow *less* productive plants to enter the market. I allow productivity to be affected by policy, to be based on unobserved choices, and I find results that are difficult to explain with a model where productivity is not based on unobserved choices.

To compare results: Fabrizio, Rose, and Wolfram (2007) find the effect of restructuring on productivity is positive[7], but I find productivity falls by at least 1% to 5% as a result of restructuring.

I start by describing the new empirical method I use to partially identify the effect of policy on productivity in Section 2. I then give background on the electricity industry, the restructuring policy, and the data I use in Section 3. In Section 4, I present reduced-form estimates that show how restructuring lowered plant size and output-to-input ratios. These alone do not answer the question of how restructuring affected productivity, so I build a structural measure of productivity by building a production function for electricity in Section 5. I take my method to the data and estimate the effect of restructuring on power plant productivity in Section 6 and its effect on market-level productivity in Section 7. I then conclude and give some additional identification assumptions to consider which can further narrow the bounds on parameters of interest.

2. Method.

I propose a new method to bound the effect of a policy on productivity which allows productivity to depend on current and unobserved choices, does not rely on difficult-to-justify timing assumptions, and which does not take a strong stance on plant conduct.

Let q_{it} be log megawatt-hours produced by power plant i in year t net of the electricity the power plant used itself, let h_{it} be log fuel use (in units of heat), let c_{it} be the log nameplate capacity of the power plant (in megawatts), let ℓ_{it} be log labor (in number of employees), and let e_{it} be log total non-fuel expenditures[8] (in dollars). Power plant productivity is a_{it},

$$q_{it} = f\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right) + a_{it}. \tag{1}$$

The difference between my approach to productivity and the approach commonly taken in the literature is that I view productivity as a function of unobserved inputs. The model I have in mind is that, for a vector of unobserved input choices (or choices more generally), y,

$$a_{it} = \psi\left(y_{it}\right) + b_{it}, \tag{2}$$

Where b_{it} is a stochastic process as in the literature but where y_{it} is a vector of unobserved

[7] They find a statistically insignificant but positive effect on fuel efficiency (the main result to compare to mine) and a positive effect on labor efficiency and the efficiency of non-fuel expenditures.

[8] Non-fuel expenditures are mostly management costs and maintenance expenditures; the definition is the same as in Fabrizio, Rose, and Wolfram (2007).

"input" choices the plant makes—potentially in response to a policy change—affecting productivity[9]. I assume nothing about the dynamics of y_{it} or b_{it}.

My empirical approach focuses directly on the question of interest, how the policy affects productivity, rather than on identifying the full distribution of productivity or the production function itself. The method I develop will be less informative about the output elasticities of the production function, but, as I will show in the empirical results, it is informative about the effect of policy on productivity.

To fix ideas, consider the problem of bounding the difference in productivity before and after restructuring occurred in the electricity generation industry. Of course, the difference in mean productivity between restructured and regulated states is not causal, and I will address that problem later in the application, but it will be useful to describe the general method with a specific problem in mind.

The difference in average log productivity between plants in restructured states and plants in non-restructured states is,

$$\begin{aligned}\tau &= \mathbb{E}[a_{it}|\text{Restruct}_{it}=1]-\mathbb{E}[a_{it}|\text{Restruct}_{it}=0]\\ &= \mathbb{E}[q_{it}|\text{Restruct}_{it}=1]-\mathbb{E}[q_{it}|\text{Restruct}_{it}=0]\\ &-\{\mathbb{E}[f(h_{it},c_{it},\ell_{it},e_{it})|\text{Restruct}_{it}=1]-\mathbb{E}[f(h_{it},c_{it},\ell_{it},e_{it})|\text{Restruct}_{it}=0]\}.\end{aligned} \tag{3}$$

The first difference we know, the difference in average log output between restructured and non-restructured states; it is observed in data. The second difference depends on the unknown production function $f(\cdot)$. I bound the second difference by putting restrictions on the production function.

2.1. Identification assumptions.

The first assumption I make is uncontroversial: I assume the production function is increasing,

$$\nabla f(h_{it},c_{it},\ell_{it},e_{it})\geq 0. \tag{4}$$

The assumption gives a lower bound on how $f(\cdot)$ changes with input use, but I still need an upper bound. I find one in economic theory.

Intuition: greater productivity increases the marginal product of input use so productivity and input choice have a supermodular relationship in the plant's profit-maximization problem. So productivity and input use should be *positively related* by Topkis (1978).

To turn that intuition into a statistical prediction, I assume productivity and inputs are what I will call linearly positively associated. Normalize $a \geq 0$ by choosing a constant for the production function[10]. Then, a and (h,c,ℓ,e) are linearly positively associated if and only if for any

[9] The separable form of the relationship between unobserved inputs and stochastic productivity is not important. Nothing would change if I wrote, $a_{it}=\psi(y_{it},b_{it})$.

[10] Not technically a normalization because it assumes that log productivity has some lower bound, but it is reasonable to assume no plant exists which can produce no output however many inputs it uses.

two positive increasing functions, g_1 and g_2,

$$\operatorname{cov}[ag_1(h,c,\ell,e), g_2(h,c,\ell,e)] \geq 0. \tag{5}$$

When I substitute, $a = q - f(h,c,\ell,e)$ into the above equation, I obtain a restriction on how much $f(\cdot)$ can increase with input use,

$$\operatorname{cov}[qg_1(h,c,\ell,e), g_2(h,c,\ell,e)] \geq \operatorname{cov}[f(h,c,\ell,e)\,g_1(h,c,\ell,e), g_2(h,c,\ell,e)]. \tag{6}$$

But why linear positive association instead of another statistical assumption (like Manski and Pepper 2000's monotone instrumental variable assumption)? Because I can show the statistical prediction holds in a broad class of economic models.

I start by showing the assumptions holds in the class of models where the only state variable is b_{it} and $a_{it} = \varphi(y_{it}) + b_{it}$ where y_{it} are unobserved input choices. Say that plants take output price as given (Appendix A gives the exact condition on demand I need which is much less restrictive than this; essentially, the elasticity of demand can not fall too quickly). Because greater b_{it} increases the marginal product of each input, input use is increasing in b_{it}—whether the input is observed or not—and because $\varphi(\cdot)$ is increasing,

$$a_{it}(b_{it}) = \varphi(y(b_{it})) + b_{it}, \tag{7}$$

is an increasing function of b_{it}. Because $a \geq 0$ (a normalization), $g_1 \geq 0$, and because the covariance of increasing functions of a random variable is positive,

$$\operatorname{cov}\{a_{it}(b_{it})\,g_1[h(b_{it}), c(b_{it}), \ell(b_{it}), e(b_{it})], g_2[h(b_{it}), c(b_{it}), \ell(b_{it}), e(b_{it})]\} \geq 0. \tag{8}$$

But we do not need to assume there is only one state variable. Any model with the following two features is in the positive association class of models:

(1) Let s be the vector of plant-specific state variables. s is positively associated in the sense of Esary, Proschan, and Walkup (1967) (the covariance of any two increasing functions of s is positive; plants that are large in one way are more likely to be large in other ways). Affiliation of s is sufficient for positive association.

(2) Let $a_{it} = a(s_{it})$ and $(h_{it}, c_{it}, \ell_{it}, e_{it}) = z(s_{it})$. $a(\cdot)$ is an increasing function and each element of $z(\cdot)$ is an increasing function.

Any model with these two features satisfies my (weaker) assumption that productivity and inputs are linearly positively associated[11]. The first condition is not a restriction at all when there is only one state variable.

See Appendix A for an example of a class of economic models where productivity is an unobserved input choice, productivity and inputs are positively associated, and there are multiple unobserved state variables.

[11]The monotone instrumental variable assumption, $d\mathbb{E}[a(s_{it})|h(s_{it}) = h]/dh \geq 0$, does not necessarily hold even if $a(\cdot)$ and $h(\cdot)$ are increasing and s is positively associated.

For a class of production models that does not satisfy the positive association assumptions, see models where the cross-partial of the production function with respect to an input and productivity is negative[12]. For example, say production uses two observed inputs, labor and capital, productivity is labor-augmenting, and greater capital implies a lower marginal product of labor, all else equal. In this model, the cross-partial for capital and productivity is negative so greater productivity might cause the plant to use less capital.

2.2. The identified set.

For estimation, assume $f(h,c,\ell,e)$ is a linear-in-parameters production function. That is, there exists a J-vector of known functions $r_j(h,c,\ell,e)$ such that,

$$f(h,c,\ell,e)=\sum_{j=1}^{J}\theta_j r_j(h,c,\ell,e), \tag{9}$$

for an unknown parameter vector $\theta\in\mathbb{R}^J$ where $r_1(\cdot)=1$. Returning to the example problem, we want to know τ, the average difference in productivity between power plants in restructured and non-restructured states,

$$\begin{aligned}\tau &= \mathbb{E}\left[q_{it}|\text{Restruct}_{it}=1\right]-\mathbb{E}\left[q_{it}|\text{Restruct}_{it}=0\right]\\ &-\sum_{j=1}^{J}\theta_j\times\left\{\mathbb{E}\left[r_j(h_{it},c_{it},\ell_{it},e_{it})|\text{Restruct}_{it}=1\right]-\mathbb{E}\left[r_j(h_{it},c_{it},\ell_{it},e_{it})|\text{Restruct}_{it}=0\right]\right\}.\end{aligned} \tag{10}$$

To find an upper bound on τ, we want to know which θ vector produces both a production function that is increasing, consistent with the linear positive association assumption, and gives the maximum value of τ subject to those constraints. We want to solve a constrained optimization problem.

What is nice about the optimization problem is that it turns out to be a linear program, which can be solved easily by well-established algorithms, even if the program has many constraints or controls. The linear positive association assumptions become, for any increasing functions g_1 and g_2,

$$\begin{aligned}&\text{cov}\left(q_{it}g_1(h_{it},c_{it},\ell_{it},e_{it}),g_2(h_{it},c_{it},\ell_{it},e_{it})\right)\\ \geq&\sum_{j=1}^{J}\theta_j\text{cov}\left(r_j(h_{it},c_{it},\ell_{it},e_{it})g_1(h_{it},c_{it},\ell_{it},e_{it}),g_2(h_{it},c_{it},\ell_{it},e_{it})\right).\end{aligned} \tag{11}$$

The covariance terms can be estimated by their sample analogues.

The program also contains non-estimated constraints which ensure the production function is increasing at any observed input choice. These too are linear in θ,

$$\sum_{j=1}^{J}\theta_j\frac{\partial}{\partial h}r_j(h_{it},c_{it},\ell_{it},e_{it})\geq 0\quad\forall\, it, \tag{12}$$

[12]Productivity must not be input-neutral for this to be true.

And likewise for the other inputs.

Lastly, I need to impose linear constraints to ensure $a_{it} \geq 0$ (my normalization). These are,

$$q_{it} \geq \sum_{j=1}^{J} \theta_j r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) \quad \forall\, it \tag{13}$$

If I had nothing to estimate, the lower and upper bound would simply be the values of two (semi-infinite) linear programming problems with a finite number of control variables (θ) and an infinite number of constraints because there are an infinite number of increasing functions we could use for the linear positive association constraints.

I could then compute the upper bound on τ by solving the following problem,

$$\begin{aligned}
&\overline{\tau} = \mathbb{E}[q_{it}|\text{Restruct}_{it} = 1] - \mathbb{E}[q_{it}|\text{Restruct}_{it} = 0] \\
&- \min_{\theta} \sum_{j=1}^{J} \theta_j \times \left(\mathbb{E}\left[r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) | \text{Restruct}_{it} = 1 \right] - \mathbb{E}\left[r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) | \text{Restruct}_{it} = 0 \right] \right) \\
&\text{ST:} \quad \operatorname{cov}(q_{it} g_1 (h_{it}, c_{it}, \ell_{it}, e_{it}), g_2 (h_{it}, c_{it}, \ell_{it}, e_{it})) \\
&\qquad \geq \sum_{j=1}^{J} \theta_j \operatorname{cov}\left(r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) g_1 (h_{it}, c_{it}, \ell_{it}, e_{it}), g_2 (h_{it}, c_{it}, \ell_{it}, e_{it}) \right) \quad \forall g_1, g_2 \text{ increasing} \\
&\qquad \sum_{j=1}^{J} \theta_j \frac{\partial}{\partial h} r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) \geq 0 \quad \forall\, it, \\
&\qquad \sum_{j=1}^{J} \theta_j \frac{\partial}{\partial c} r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) \geq 0 \quad \forall\, it, \\
&\qquad \sum_{j=1}^{J} \theta_j \frac{\partial}{\partial \ell} r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) \geq 0 \quad \forall\, it, \\
&\qquad \sum_{j=1}^{J} \theta_j \frac{\partial}{\partial e} r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) \geq 0 \quad \forall\, it, \\
&\qquad q_{it} \geq \sum_{j=1}^{J} \theta_j r_j (h_{it}, c_{it}, \ell_{it}, e_{it}) \quad \forall\, it
\end{aligned} \tag{14}$$

The lower bound on τ replaces the minimization problem above with a maximization problem.

But we do have to estimate the linear positive association constraints, and that becomes a non-trivial problem because, intuitively, if we estimate each constraint with error, each constraint must be satisfied, and there are an infinite number of constraints, the estimated bounds can be too narrow—even asymptotically—because of rare, large errors. The problem has been discussed in the econometric literature both in the context of the monotone instrumental variable assumption from Manski and Pepper (2000) and for the more general problem in Chernozhukov, Lee, and Rosen (2013). I showed, in Flynn (2016), that we can use an approximation to the above linear programming problem which is also a linear program to estimate the upper bound on τ; the estimator is computationally convenient and has a standard, asymptotically normal inference theory.

The estimator is simply the sample analogue of the above linear program after modifying the linear positive association constraints.

2.3. The empirical linear positive association constraints.

The linear positive association constraints are,

$$\begin{aligned} &\operatorname{cov}\left(q_{it} g_1\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right), g_2\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right)\right) \\ \geq \sum_{j=1}^{J} \theta_j \operatorname{cov}&\left(r_j\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right) g_1\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right), g_2\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right)\right) \quad \forall g_1, g_2 \text{ increasing} \end{aligned} \tag{15}$$

It turns out that many of the constraints are redundant. There is only one type of g we need to be concerned about, the "step functions" (proof in Appendix C),

$$g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u\right) = \mathbf{1}\left(h_{it} \geq u_1\right) \times \mathbf{1}\left(c_{it} \geq u_2\right) \times \mathbf{1}\left(\ell_{it} \geq u_3\right) \times \mathbf{1}\left(e_{it} \geq u_4\right). \tag{16}$$

We do not need to use every $u \in \mathbb{R}^4$, either. We only need to use the u in a countably dense subset of $\mathbb{R}^4$. Empirically, I apply a cumulative distribution function to h, c, ℓ, and e and then choose u in the rational numbers in $[0,1]^4$. So for some cumulative distribution functions, $\Phi_h, \Phi_c, \Phi_\ell, \Phi_e$, redefine g as,

$$g\left(h_{it}, c_{it}, \ell_{it}, e; u\right) = \mathbf{1}\left(\Phi_h\left(h_{it}\right) \geq u_1\right) \times \mathbf{1}\left(\Phi_c\left(c_{it}\right) \geq u_2\right) \times \mathbf{1}\left(\Phi_\ell\left(\ell_{it}\right) \geq u_3\right) \times \mathbf{1}\left(\Phi_e\left(e_{it}\right) \geq u_4\right). \tag{17}$$

So, now the linear positive association constraints are,

$$\begin{aligned} &\operatorname{cov}\left(q_{it} g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right) \\ \geq \sum_{j=1}^{J} \theta_j &\operatorname{cov}\left(r_j\left(h_{it}, c_{it}, \ell_{it}, e_{it}\right) g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right) \\ \forall u^1, u^2 &\in \left\{u \in \mathbb{R}^4 : u_p = \frac{k_p}{m} \text{ for } k_p < m \text{ and } k_p, m \in \mathbb{N} \text{ for } p = 1, \ldots, 4\right\}. \end{aligned} \tag{18}$$

But there are still an infinite number of constraints. I show in Flynn (2016) that if we use a subset of these constraints and if we add constraints slowly enough, the estimated bound ($\overline{\tau}$) is consistent and asymptotically normal (under some assumptions, of course). I limit the number of constraints by fixing the "m" in the above set of constraints. From Monte Carlo results in that paper, I settled on the following choice of constraints,

$$
\begin{aligned}
&\operatorname{cov}\left(q_{it} g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right) \\
&\geq \sum_{j=1}^{J} \theta_j \operatorname{cov}\left(r_j(h_{it}, c_{it}, \ell_{it}, e_{it}) g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right) \\
&\forall u^1, u^2 \in \left\{u \in \mathbb{R}^4 : u_p = \frac{k_p}{m} \text{ for } k_p < m = \lfloor 5 \times (n/1000)^{1/6} \rfloor \text{ and } k_p \in \mathbb{N} \text{ for } p = 1, \ldots, 4\right\}.
\end{aligned} \tag{19}
$$

I also add linear positive association constraints that directly restrict how correlated log productivity is with the basis functions of the production function. These constraints add nothing asymptotically but are useful in finite samples,

$$
\begin{aligned}
&\operatorname{cov}\Big(a_{it}(\theta) g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), \\
&\left[r_j(h_{it}, c_{it}, \ell_{it}, e_{it}) - \min_{it} r_j(h_{it}, c_{it}, \ell_{it}, e_{it})\right] \times g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\Big) \geq 0 \text{ for } j = 2, \ldots, J.
\end{aligned} \tag{20}
$$

There is still one more problem to solve. In every finite sample, some of the constraints will use few observations, and if those constraints bind, they will be badly estimated, which will make the estimator "jump around" and not be asymptotically normally distributed. To solve the problem, I need to weaken the constraints that use few observations. Define,

$$
n\left(u^1, u^2\right) = \min\left\{\sum_{it} g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), \sum_{it} g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right\} \tag{21}
$$

Given two sequences of positive tuning parameters, β_n and ϵ_n (the selection of which I discuss in Flynn 2016 and Appendix C.2), the empirical linear positive constraints are,

$$
\begin{aligned}
&\widehat{\operatorname{cov}}\left(q_{it} g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right) + \exp\left(\frac{\beta_n - n\left(u^1, u^2\right)/n}{\epsilon_n}\right) \\
&\geq \sum_{j=1}^{J} \theta_j \widehat{\operatorname{cov}}\left(r_j(h_{it}, c_{it}, \ell_{it}, e_{it}) g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^1\right), g\left(h_{it}, c_{it}, \ell_{it}, e_{it}; u^2\right)\right) \\
&\forall u^1, u^2 \in \left\{u \in \mathbb{R}^4 : u_p = \frac{k_p}{m} \text{ for } k_p < m = \lfloor 5 \times (n/1000)^{1/6} \rfloor \text{ and } k_p \in \mathbb{N} \text{ for } p = 1, \ldots, 4\right\}.
\end{aligned} \tag{22}
$$

The final estimator of the upper bound on τ, $\overline{\tau}$, then solves the linear programming problem in (14), replacing the linear positive association constraints with the empirical linear positive association constraints and replacing the rest of the parameters with their sample analogues.

When constructing the linear positive association constraints in practice, I also control for state-year fixed effects because the theory behind positive association is at its most robust when predicting that there is a positive relationship between productivity and inputs within a market, when all plants face the same aggregate shocks (in practice, adding state-year fixed effects had little effect on my empirical results). The covariance restrictions are the same as above except each variable has the state-year mean taken out of it (let st index state-year),

$$
\begin{aligned}
&\widehat{\text{cov}}\Big(q_{it}g\left(h_{it},c_{it},\ell_{it},e_{it};u^1\right)-\overline{q_{it}g\left(h_{it},c_{it},\ell_{it},e_{it};u^1\right)}_{st},\, g\left(h_{it},c_{it},\ell_{it},e_{it};u^2\right)-\overline{g\left(h_{it},c_{it},\ell_{it},e_{it};u^2\right)}_{st}\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \exp\left(\frac{\beta_n - n\left(u^1,u^2\right)/n}{\epsilon_n}\right) \\
&\geq \sum_{j=1}^{J}\theta_j\widehat{\text{cov}}\Big(r_j\left(h_{it},c_{it},\ell_{it},e_{it}\right)g\left(h_{it},c_{it},\ell_{it},e_{it};u^1\right)-\overline{r_j\left(h_{it},c_{it},\ell_{it},e_{it}\right)g\left(h_{it},c_{it},\ell_{it},e_{it};u^1\right)}_{st}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g\left(h_{it},c_{it},\ell_{it},e_{it};u^2\right)-\overline{g\left(h_{it},c_{it},\ell_{it},e_{it};u^2\right)}_{st}\Big) \\
&\qquad\qquad\qquad \forall u^1,u^2 \in \left\{u\in\mathbb{R}^4 : u_p = \frac{k}{m} \text{ for } k < m = \lfloor 5\times(n/1000)^{1/6}\rfloor \text{ and } k\in\mathbb{N}\right\}.
\end{aligned}
\tag{23}
$$

I describe how to compute standard errors, confidence intervals, and choose Φ in Appendix C.

The objective function is linear in θ in the constrained optimization problem, not only for difference in means (like for τ), but also for linear functions of linear regression coefficients. Let x_{it} be a vector of observables—say, $x_{it} = (\text{Restruct}_{it},\xi_i,\mu_t)$ where ξ are plant-level fixed effects and μ are year-level fixed effects to control for selection's influence on our estimate of the effect of restructuring via time-invariant plant characteristics and aggregate shocks,

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad a_{it} = x_{it}^{\top}\gamma + v_{it}, \quad \mathbb{E}[x_{it}v_{it}] = 0 \\
&\implies \gamma = \mathbb{E}\left[x_{it}x_{it}^{\top}\right]^{-1}\mathbb{E}[x_{it}a_{it}] = \mathbb{E}\left[x_{it}x_{it}^{\top}\right]^{-1}\mathbb{E}[x_{it}q_{it}] - \mathbb{E}\left[x_{it}x_{it}^{\top}\right]^{-1}\mathbb{E}\left[x_{it}r_j\left(h_{it},c_{it},\ell_{it},e_{it}\right)^{\top}\right]\theta.
\end{aligned}
$$

So elements of γ are linear functions of θ. For example, the first element of γ is the difference in mean productivity between restructured and regulated plants once we control for individual and year fixed effects. The statistics I estimate will all be linear regression coefficients or linear functions of those coefficients.

3. The Electricity Industry.

I now turn to how I apply the method to study the effect of restructuring on power plant productivity and why it makes sense to use the new method to do so, giving background on the industry, the restructuring policy itself, and on the data I use.

3.1. Industry and policy background.

Historically, electric utilities in America were vertically integrated across the three stages of production: generation, transmission, and retail to end consumers. The government regulated the utilities as natural monopolies under "rate-of-return" regulation: public service commissions set electricity price to the power plant's average cost plus a regulated rate of return[13].

For an example of the policy in practice, in Madison, Wisconsin, September 2016, Madison Gas and Electric had lower than expected average fuel costs. They provided all the electricity people wanted at current electricity prices, and they did it cheaper than expected. So they had to issue refunds[14] to all their consumers (see MGE 2016). They gained little by generating cheap electricity.

If the government lowers the price a utility can charge when the utility lowers its costs, the utility has less incentive to take any costly action not prescribed by the regulator to lower its cost of output. The regulator (the principal) can not perfectly regulate unobserved choices the power plant (the agent) makes that affect its productivity because those choices are hidden actions (see Laffont 1994 for a review of the literature on regulation as a principal-agent problem).

Recognizing that rate-of-return regulation weakens the incentive to minimize costs, several US states restructured their electricity markets, ending rate-of-return regulation in those stages of production and allowing electricity price to be set by competition[15]. The states still regulated electricity transmission as a natural monopoly to ensure electricity producers could reach end consumers. See Borenstein and Bushnell (2015) for a fuller history of the restructuring policy.

Restructuring provided an incentive to become more productive but also increased the level of competition so the policy has a theoretically ambiguous effect on productivity. I study how restructuring changed power plant productivity in practice. I am not the first to do so, but I am the first to allow power plant productivity to be a choice—which is the model of productivity we have in mind if we think restructuring might change power plant productivity (see Fabrizio, Rose, and Wolfram 2007 and my discussion of it in Section 1).

I use the difference in how restructuring affected investor-owned utilities and municipally-owned utilities to identify the causal effect of restructuring on productivity. Municipally-owned power plants were never subject to rate of return regulation. The idea was that as government-owned entities they would not use their monopoly power in the electricity market to charge higher prices to maximize profit. The municipalities served by these utilities are not served by investor-owned utilities. So only investor-owned utilities were affected by restructuring. I use municipal power plants as a control group to identify the effect of restructuring on productivity of the investor-owned utilities.

Before restructuring, investor-owned utilities and municipally-owned utilities faced similar incentives, both had mandates to price at roughly what was required to recover costs and earned regulated rates of return. The main difference between the two was by whom they were

[13] In some states, rewards were not based on average cost but on some approximation of marginal cost, but the main point is that the price was not set by market forces. See Knittel (2002) for more on these other policies.

[14] The refund was about 6 cents-per-kilowatt-hour. Average price of electricity in Wisconsin is about 15 cents-per-kilowatt-hour.

[15] In some states, the liberalization was only in the generation stage. This was a contributing factor to the California energy crisis (see Joskow 2001) where wholesale prices for electricity were set by market forces after restructuring, but utilities were required to provide retail services at a regulated rate (65$/megawatt-hour) for four years. But market prices eclipsed regulated retail price before that time expired.

regulated: investor-owned utilities were regulated by a state public service commission and municipally-owned utilities by local city councils. But after restructuring, the incentives of the two types of utilities diverged which helps me identify the effect of restructuring on power plant productivity.

3.2. The Data.

I use the same data as Fabrizio, Rose, and Wolfram (2007) for the years 1981 to 1999, which I downloaded from Wolfram's website[16]. The original data is from the Energy Information Administration (EIA), the Federal Energy Regulatory Commission (FERC), and the Rural Utilities Service (RUS).

The data is at the power plant-level and includes large (greater than 100 megawatt capacity), fossil fuel power plants from the years 1981 to 1999. The first restructuring policy changes were in 1996 so the data has information on power plants before and after a state restructured its electricity market.

I then extend the dataset from 2000 to 2003 using data from the EIA and FERC, hoping to capture additional information on the effect of the restructuring policy. It might have taken some time for market structure to adjust to the restructuring policy and for capacity to adjust so including additional years should be helpful in identifying the full effect of restructuring. It takes about 3 years to build a natural gas power plant and longer to build a coal plant so entry encouraged by restructuring might not have been picked up in the data ending in 1999 (see NEA 2016).

I measure output by millions of megawatt-hours of electricity generated net of electricity the power plant uses itself, fuel use by units of heat energy (millions of million British thermal units), capacity by the total nameplate capacity of the plant (in megawatts), labor by total employment (the quality of the labor input is in productivity, another reason to consider productivity an input because labor quality is a choice), and non-fuel expenditures in millions of dollars. The data is annual.

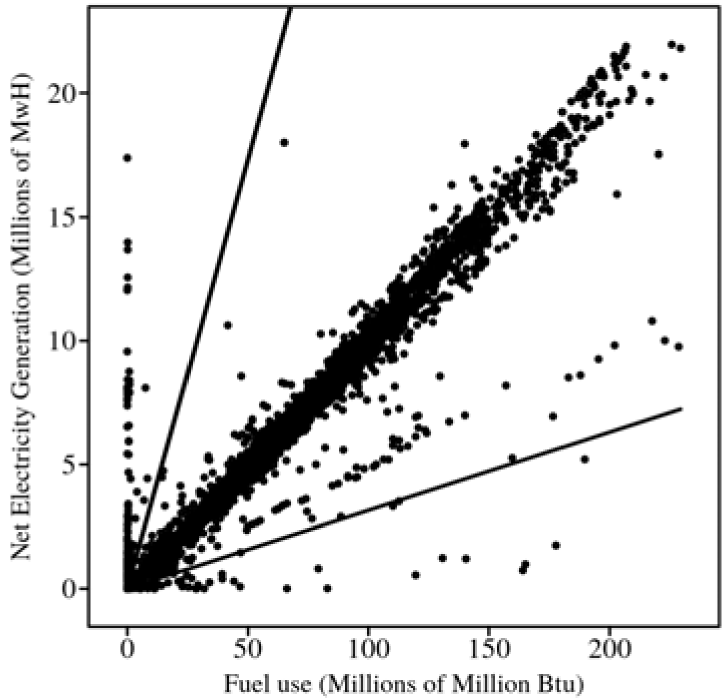
I include only power plants with positive employment and non-fuel expenses, and I remove some outliers. Some fuel data is clearly incorrect; some plants have impossibly large or small fuel-to-output ratios (likely because the utility filled out the form using different units). To deal with the measurement error, I use the average heat rates for different types of fossil fuel power plants published by the Energy Information Administration for the year 2004 and remove observations that have a heat rate more than three times the average heat rate or less than one-third the average rate (I check for robustness to the choice of extreme heat rates in Appendix B.2).

Figure 1 shows the joint distribution of fuel and power and which observations I consider outliers in the main results.

Table 1 gives descriptive statistics on the data with the outliers removed, the dataset I use to estimate the productivity effect of restructuring. The table shows that investor-owned power

[16] I thank Fabrizio, Rose, and Wolfram very much for making their data available and easy to use.

Figure 1: Relationship between fuel and electricity generated (with outliers), observations between the two lines make up the dataset I use



plants are larger than municipally-owned power plants so I will need to control for plant size as well before I can use municipally-owned power plants as a control group.

Table 1: Descriptive statistics of data (with outliers removed)

Statistic	Plants	Output	Fuel	Capacity	Labor	Non-fuel
Mean	All	3.36	34.83	796	152	14.28
	IOU	3.46	35.83	825	152	14.67
Median	All	2.09	21.94	574	112	9.79
	IOU	2.20	23.13	613	113	10.14
Standard Deviation	All	3.69	37.28	666	133	14.43
	IOU	3.76	37.89	679	130	14.83
Number of Plant-Years	11,485					
Range of years	1981 to 2003					

Notes. IOU stands for investor-owned utility. Output is in units of millions of megawatt hours, fuel is in units of millions of million British thermal units, capacity is in megawatts, labor is in number of employees, and non-fuel expenditures are in millions of dollars.

4. Effect of restructuring on input efficiency and plant size.

Before estimating the effect of restructuring on a structural measure of power plant productivity, I look to the data to estimate the effect of restructuring on input efficiency and plant size (output and input use). The first goal of this section is to document that restructuring made large power plants smaller, supporting the theory that there was a competitive effect of restructuring. The increased entry from independent power producers lowered the output of the large, utility-owned power plants that make up my dataset. The second goal is to build both a static and dynamic panel model to estimate the effect of restructuring on different measures of plant performance. In this section, I will use the two models to estimate the effect of restructuring on input efficiencies (input-to-output ratios). In the following sections, I will replace input efficiencies with structural productivity and see how my answer changes. The effect of restructuring on the different input efficiencies combined with the linear positive association assumption determine how restructuring effects productivity.

I start by defining when restructuring starts in a state. I consider two definitions: when the law actually passed ("Law Passed") and when formal meetings about implementing restructuring started ("Announcement").

The second measure acknowledges power plants may have started to improve their productivity as soon as it became clear restructuring would pass, in advance of the law officially passing. The first measure acknowledges that there was considerable uncertainty about whether restructuring would pass in many states and plants may not have made any changes until they knew for sure it would pass.

Because many states that announced they would restructure never actually did, I use Law Passed as the main definition for restructuring in the text and put the full results for the Announcement definition in Appendix B.1.

I run the following regression to examine the effect of restructuring on plant size, where Y is either output or one of the inputs,

$$\log Y_{it} = \tau_Y \text{Restruct}_{it}\text{IOU}_i + \beta_Y \text{Restruct}_{it} + \kappa_Y m_{rt} + \alpha_i^Y + \delta_t^Y + U_{it}^Y, \tag{24}$$

Where m_{rt} is log census region population and is included to control for demand effects that might have been contemporaneous with restructuring and α_i^Y and δ_t^Y are plant and year fixed effects.

The effect of restructuring on plant size is τ_Y. I look at the effect on power plants owned by investor-owned utilities because these utilities were the only utilities subject to rate-of-return regulation. I use municipally-owned power plants, who were not subject to rate of return regulation, as a control group to identify the effect of restructuring on plant size. I find the effects are negative when Y is power (-15%), fuel use (-13%), capacity (-12.2%), labor use (insignificant), and non-fuel expenditures (-24.3%). Power plants became smaller as a result of restructuring. See Table 2 for full details.

One explanation for power plants contracting in response to restructuring is that their residual demand curves became steeper; they had market power and began to use it. Additionally, increases in competition shifted their residual demand curves in, lowering their choice of output and input use.

Next, I estimate the effect of restructuring on measures of power plant efficiency, B. I evaluate the effect of restructuring in two models, a static panel model and a dynamic panel model. First, the static model,

$$\log B_{it} = \tau_B \text{Restruct}_{it}\text{IOU}_i + \beta_B \text{Restruct}_{it} + \kappa_B m_{rt} + \alpha_i^B + \delta_t^B + U_{it}^B, \tag{25}$$

I estimate τ_B for the input efficiencies, the ratio of output to each of the inputs: fuel, capacity, labor, and nonfuel expenditures. I find that the effect of restructuring on nonfuel expenditures was positive (8.8%), but the effect on fuel (-2.4%) efficiency was negative. If the production function was Cobb Douglas, had constant returns to scale, and power plants minimized static cost, then productivity can only increase if the ratio of output to inputs increases. Because I find fuel efficiency fell, I have some reduced form evidence for what I will show more formally: restructuring lowered power plant productivity at turning electricity into fuel.

But we might think increasing B_t is more difficult if B_{t-1} is small. It is more expensive to make large efficiency gains. So I also estimate a dynamic panel model,

$$\log B_{it} = \rho \log B_{it-1} + \tau_B \text{IOU}_i \text{Restruct}_{it} + \beta_B \text{Restruct}_{it} + \kappa_B m_{rt} + \alpha_i^B + \delta_t^B + U_{it}^B. \tag{26}$$

$\log B_{it-1}$ is correlated with e_{it} by standard arguments from the dynamic panel literature. I use the Anderson-Hsiao estimator to solve the problem. The Anderson-Hsiao estimator amounts to two stage least squares with B_{it-3} as an instrument for $\Delta \log B_{it-1}$ in the following regression[17],

$$\Delta \log B_{it} = \rho \Delta \log B_{it-1} + \tau_B \text{IOU}_i \Delta \text{Restruct}_{it} + \beta_B \Delta \text{Restruct}_{it} + \kappa_B \Delta m_{rt} + \Delta \delta_t^B + \Delta U_{it}^B \tag{27}$$

[17] The choice of instrument allows the residual to be an AR1 process.

Table 2: Effect of restructuring on plant-level efficiencies and input use (static panel model)

Dependent variable	Effect (Law Passed)	Effect (Announcement)
Log fuel efficiency	$-2.4\%^{**}$ (1.0%)	$-1.8\%^{***}$ (0.6%)
Log capacity efficiency	-3.3% (4.6%)	$-10.5\%^{***}$ (2.7%)
Log labor efficiency	$-12.8\%^{***}$ (4.8%)	$-5.5\%^{**}$ (2.8%)
Log non-fuel efficiency	$8.8\%^{*}$ (4.8%)	$7.2\%^{***}$ (2.8%)
Log power	$-15.5\%^{***}$ (4.9%)	$-17.6\%^{***}$ (2.8%)
Log fuel	$-13.1\%^{***}$ (4.6%)	$-15.8\%^{***}$ (2.7%)
Log capacity	$-12.2\%^{***}$ (1.4%)	$-7.1\%^{***}$ (0.8%)
Log labor	-2.7% (2.4%)	$-12.1\%^{***}$ (1.4%)
Log non-fuel	$-24.3\%^{***}$ (3.4%)	$-24.8\%^{***}$ (2.0%)

$*$ denotes statistical significance at the 10% level, $**$ at the 5% level, and $***$ at the 1% level.

I find restructuring has a negative effect on fuel efficiency in the dynamic panel model, but a positive effect on some non-fuel efficiencies.

The dynamic model results say that the static model understates the extent of the loss to fuel efficiency by a good two to three percent (although the estimate is more noisy).

I evaluate the policy using both the static and dynamic model because while the dynamic model allows productivity to have a more realistic adjustment process, that comes at the cost of adding additional parameters and requiring multiple lags of the data, lowering statistical power. The static model gives more precise answers, but they might be biased in some way by ignoring the dynamic nature of productivity adjustment. It will turn out that both models give the same signs of the effects I study, but the dynamic model suggests the magnitude of the effects are larger.

The problem with the estimates using input efficiencies for B is that I do not know if B is a measure of physical productivity or some combination of demand, the nature of competition, and productivity. I use the same regression models but replace B with A, a structural measure of productivity.

Table 3: Dynamic panel estimates of the effect of restructuring on input efficiencies

Performance measure	Effect (Law Passed)	Effect (Announcement)
Fuel efficiency	−6.8%*	−0.8%
	(3.7%)	(1.6%)
Log capacity efficiency	7.2%	−4.0%
	(7.6%)	(3.8%)
Log labor efficiency	13.2%	−0.1%
	(8.9%)	(4.5%)
Log non-fuel efficiency	28.3%**	5.4%
	(12.0%)	(6.0%)

$*$ denotes statistical significance at the 10% level, $**$ at the 5% level, and $***$ at the 1% level.

5. The electricity production function.

My structural measure of productivity comes from a model of electricity generation.

Power plants produce electricity (Q) using fuel (H), capacity (C), labor (L), and non-fuel expenditures (E).

Within an hour, electricity production is linear in fuel use: no other input can substitute for fuel and electricity is produced roughly in proportion to the fuel burned, until the power plant reaches its capacity constraint.

Say, within a short time period δ, production is,

$$Q^{\delta}_{it} = \min\left(\lambda_{it}H^{\delta}_{it}, \mathscr{C}^{\delta}_{it}\right), \tag{28}$$

Where $\mathscr{C}^{\delta}_{it}$ gives the plant's capacity available for production within-δ. I assume there is never a reason to burn more fuel than the power plant's capacity constraint, so:

$$Q^{\delta}_{it} = \lambda_{it}H^{\delta}_{it}. \tag{29}$$

Summing over the time periods δ for the entire year gives,

$$Q_{it} = \lambda_{it}H_{it} \implies q_{it} - h_{it} = \log\lambda_{it} \tag{30}$$

I am not assuming that the output elasticity for fuel use is 1, only that we can recover the plant's λ_{it} by dividing Q_{it} by H_{it}. The fuel output elasticity would only be 1 if increasing fuel use did not cause capacity constraints to bind.

I assume $\log\lambda_{it}$ is a function of labor, capacity, non-fuel expenditures, and productivity (a_{it}).

I assume a translog functional form[18],

$$\log \lambda_{it} = \theta_c c_{it} + \theta_{c\ell} c_{it} \ell_{it} + \theta_{ce} c_{it} e_{it} + \theta_{cc} c_{it}^2 + \theta_\ell \ell_{it} + \theta_{\ell e} \ell_{it} e_{it} + \theta_{\ell\ell} \ell_{it}^2 + \theta_e e_{it} + \theta_{ee} e_{it}^2 + a_{it}. \tag{31}$$

I call a_{it}, "fuel productivity", the power plant's productivity at turning fuel into electricity.

I use this model of production instead of a standard Cobb-Douglas production function for three reasons:

(1) I have not put a lower bound on the output elasticities except that they must be greater than zero, but I am confident the true output elasticity for fuel is much larger than zero. This model of electricity production ensures fuel has a non-negligible effect on output and that every production function in the identified set is reasonable (I could imagine, although I do not think it is true, that labor, capacity, and other non-fuel inputs do not causally increase the ratio of output-to-fuel use).

(2) It reduces the number of inputs in the production function. I can treat the only inputs as capacity, labor, and non-fuel expenses.

(3) I want to be able to compare my results to Fabrizio, Rose, and Wolfram (2007) who study the effect of restructuring on different conditional input demand equations. The above model of production leads to the same reduced-form as their model of conditional fuel demand (see Section 3.1 for their derivation of the same regression),

$$h_{it} - q_{it} = -\log \lambda_{it}, \tag{32}$$

Where they replace $\log \lambda_{it}$ with covariates and estimate how restructuring changed it. Taking the negative of my estimates then gives results that are directly comparable to theirs.

I assume log labor, log capacity, and log non-fuel expenditures are positively associated with fuel productivity and that the derivative of the production function is positive for capacity, labor, and non-fuel expenditures on the observed support of the inputs.

Fabrizio, Rose, and Wolfram (2007) are able to have a more flexible production model than I do by allowing for separate productivity shocks for each variable input (fuel, labor, and non-fuel expenses) in conditional input demand equations. I can think about what they do as allowing output elasticities for the different inputs to vary by plant. I focus on fuel productivity and leave productivity at other aspects of production (such as the maintenance of capacity) for future research.

6. The effect of restructuring on power plant productivity.

Restructuring might change a power plant's productivity by changing its incentive to be more productive. Economic theory is not clear on the direction of the change because there are two opposing effects. The *incentive* effect of restructuring increases the value of productivity to

[18] A Cobb-Douglas functional form would give me more statistical power, more of the estimated effects would be statistically significant, but it is in the spirit of the paper to be flexible.

the power plant by allowing them to keep cost reductions as profits. But restructuring also increased the competition power plants faced; it forced utilities to sell some of their power plants, encouraged entry by non-utility generators, and (possibly) raised the profitability of power plants in general, leading to increased competition, driving down markups, leading to less investment in productivity. I refer to these negative effects on productivity as the *competitive* effect of restructuring (see Vives 2008 for more theory on when greater competition implies less productivity).

I find the competitive effect dominates the incentive effect.

Earlier, I estimated the effect of restructuring on fuel efficiency but that effect includes both the effect of input and demand changes and the effect due to productivity changes. In this section, I use the same two panel models, one static and the other dynamic, to estimate the part of restructuring's effect on fuel efficiency that is due to a change in productivity.

The effect of restructuring on fuel efficiency in the static model was τ in the following regression,

$$q_{it} - h_{it} = \tau \mathrm{IOU}_i \mathrm{Restruct}_{it} + \beta \mathrm{Restruct}_{it} + \kappa m_{rt} + \alpha_i + \delta_t + \nu_{it}. \tag{33}$$

We can decompose the effect τ into an effect due to changes in input use versus the effect due to changes in productivity,

$$\tau = \tau_r^\top \theta + \tau_a, \tag{34}$$

Where τ_r is the vector of τ's estimated by replacing $q_{it} - h_{it}$ with the basis functions of the translog production function in the above regression. τ and τ_r can be estimated from data so τ_a is a linear function of θ which is partially identified by the requirements the production function is increasing and productivity is linearly positively associated with inputs.

As in Section 4, I use the fact that only investor-owned utilities were subject to rate-of-return regulation to identify the effect of restructuring. Municipally-owned power plants were never subject to rate-of-return regulation, and they do not compete with investor-owned utilities directly so restructuring should have no effect on them. β is essentially a restructured-state specific time trend because restructuring has no causal effect on municipally-owned power plants. I treat municipal power plants as a control group and use the contrast between the effect of restructuring on municipal power plants and investor-owned utilities to identify the true effect of restructuring, τ_a. The use of a control group is necessary because restructuring also acts as a dummy for later years in the data for a certain set of states so, without a control group, I might only be picking up a trend.

m_{rt} is log region population where region is one of nine Census regions (plus separate regions for Alaska and Hawaii), a measure of market size. Theories where productivity depends on unobserved input choices find market size plays a key role in determining productivity, in the same way that market size affects observed input choice.

τ_a is the main statistic of interest, the effect of restructuring. I find the effect is *negative* for fuel productivity (between −3% and −1%). The competitive effects of restructuring dominated the incentive effects, leading to lower fuel productivity. 1% to 3% changes to fuel productivity are economically significant effects; for each unit of fuel the average plant burns, it produces 1% to 3% less electricity after restructuring if it has the same capacity, labor, and non-fuel expenditures.

For comparison, Fabrizio, Rose, and Wolfram (2007) found a positive effect of restructuring on fuel productivity (although insignificant) and a positive effect of restructuring on labor and non-fuel productivity.

The effect of restructuring on fuel productivity is a policy concern beyond its effect on the power plant's cost curve because the amount of fuel required to produce electricity is determined by fuel productivity and fuel use determines how much pollution is caused by electricity production.

The effect of market size on productivity is positive and the estimated κ_a (defined in the same way as τ_a) is large (greater than 2%). Greater market size leads to higher productivity choice in line with many models of productivity choice but at odds with models where productivity is exogenous. When productivity is exogenous, greater market size only lowers the productivity threshold needed to enter so the sign of κ_a would be negative (or zero, as in the Melitz 2003 model).

The result gets right to the heart of the methodological point of the paper. Productivity acts like an input responding to a larger market as other inputs do, not as a parameter or a scientific process exogenous to economic choice[19].

I next use a dynamic panel model to study the policy which is likely more representative of the true productivity process but adds additional complications to identification and inference. I decompose the dynamic panel model of fuel efficiency into the effect of restructuring on fuel efficiency due to changes in input use and the effect due to changes in productivity and look to identify the second effect.

The dynamic model of fuel efficiency was,

$$q_{it} - h_{it} = \rho\,(q_{it-1} - h_{it-1}) + \tau \text{IOU}_i \text{Restruct}_{it} + \beta \text{Restruct}_{it} + \kappa m_{rt} + \alpha_i + \delta_t + v_{it}. \tag{35}$$

The parameter now of interest is τ_a, the contribution of productivity to the effect of restructuring on fuel productivity, which can be written as,

$$\tau = \tau_r^\top \theta + \tau_a \implies \tau_a = \tau - \tau_r^\top \theta, \tag{36}$$

a linear function of the production function parameters. I estimate τ and τ_r using the Anderson-Hsiao instrument, the third lag of log fuel efficiency and use the identification assumptions to bound θ.

The results from the dynamic model argue that the static model understates how much restructuring caused productivity loss which caused fuel efficiency to fall after restructuring. The

[19]The main difference between Fabrizio, Rose, and Wolfram (2007) and this paper: they use a measure of market size to *instrument* for productivity. I argue it is a determinant of productivity; it is not excluded (see the model in Appendix A).

Table 4: Effect of restructuring on plant fuel productivity

Parameter	Model	LB (90%)	LB	UB	UB (90%)
τ_a	Static	−3.86%	−2.62%	−1.12%	0.26%
τ_a	Dynamic	−11.82%	−7.02%	−5.83%	−1.29%
κ_a	Static	2.44%	3.64%	4.53%	5.71%
κ_a	Dynamic	1.64%	4.53%	5.30%	8.30%

Notes. Columns "LB (90%)" and "UB (90%)" give the upper and lower limits of the 90% confidence interval around the parameter. All results for the Law Passed definition of restructuring, see Appendix B.1 for the Announcement results.

dynamic model finds τ_a between −6% and −7% (a 90% confidence interval between −1% and −12%), a large negative effect on power plant productivity.

The main results of my estimates for the effect of restructuring on power plant productivity are:

(1) The effect of restructuring on power plant productivity is negative, lowering fuel efficiency.

(2) When market size increases, so does productivity.

For full results, see Table 4, where I give both the estimated lower and upper bounds on the parameters of interest and also a 90% confidence interval for the parameter of interest.

7. How restructuring changed market productivity.

But restructuring did not only change the productivity of the surviving power plants, it also changed market-wide productivity which is an important concern for a state deciding whether to implement the policy. I look at the effect of restructuring on a region's aggregate productivity and decompose it into a reallocation effect and an average effect using the Olley and Pakes (1996) decomposition.

Because bounds on the parameter of interest are easier to form when the statistic of interest is a linear function of the production function parameters (see Section 2 and Flynn 2016), I use an aggregate productivity measure which is linear in the production function parameters when I take the logarithm.

I call it, "Geometric Aggregate Productivity" (GAP) to differentiate it from the Olley and Pakes (1996) measure. I compute aggregate productivity at the Census region level where r indexes region,

$$\begin{aligned} GAP_{rt} &= \prod_{i \in rt} A_{it}^{o_{it}} \\ o_{it} &= \frac{Q_{it}}{\sum_{j \in rt} Q_{jt}}. \end{aligned} \tag{37}$$

Relative to the arithmetic aggregate productivity measure in Olley and Pakes (1996), the GAP measure will care more about power plants with low productivity but high output share because they are not as easily compensated for by another plant with high productivity and high output share[20]. So lowering the market share of less productive plants will have a more positive effect on GAP than on the Olley and Pakes (1996) measure.

Aggregate productivity increases both when average log productivity increases and when the correlation between a power plant's share of log output and its productivity increases. The Olley and Pakes (1996) decomposition is,

$$\log \text{GAP}_{rt} = \underbrace{\overline{a}_{rt}}_{\text{Average productivity}} + \underbrace{\sum_{i \in rt} (a_{it} - \overline{a}_{rt})(o_{it} - \overline{o}_{rt})}_{\text{Reallocation}}, \tag{38}$$

Where $\overline{a}_{rt}$ is average log productivity in region r at time t and $\overline{o}_{rt}$ is average market share. The first term is the log of the geometric mean of productivity and the second term measures the effect of allocation on aggregate productivity.

I estimate the effect of restructuring on aggregate productivity with the following regression,

$$\begin{aligned} \log \text{GAP}_{rt} &= \tau \times \text{Restruct}_{rt} + \delta_r + \mu_t + \nu_{rt} \\ \implies \Delta \log \text{GAP}_{rt} &= \tau \Delta \text{Restruct}_{rt} + \Delta\mu_t + \Delta\nu_{rt}. \end{aligned} \tag{39}$$

Where $\text{Restruct}_{rt} = 1$ when any state in the region has restructured.

I find regional aggregate productivity fell after restructuring (between −2% and −1%). Restructuring's effect on aggregate productivity can be decomposed into its effect on the average term and its effect on the reallocation term.

Restructuring caused the geometric mean of productivity in a region to fall between −3% and −2%. Average market productivity fell as plant productivity fell.

But the reallocation effect was positive, between 1 to 2%, restructuring caused the covariance between market share and productivity to increase.

The direction of the effect of restructuring on the two components of aggregate productivity gives further evidence that restructuring's main effect was to increase competition, not raise the incentive to become more productive. The negative average productivity effect comes directly from the negative effect on plant level productivity, but the positive reallocation effect reveals that the greater competition did allocate greater market share to the more productive plants.

[20] Arithmetic mean-preserving spreads decrease geometric means.

Table 5: Effect of restructuring on geometric aggregate fuel productivity (GAP)

Parameter	LB (90%)	LB	UB	UB (90%)
GAP	−2.43%	−1.61%	−1.41%	−0.63%
Average Productivity	−4.35%	−3.65%	−2.64%	−1.89%
Reallocation	0.82%	1.22%	2.09%	2.44%

Notes. LB (90%) and UB (90%) denote the lower and upper limits of the 90% confidence intervals. All results for the Law Passed definition of restructuring.

8. Conclusion and additional assumptions on the production function to consider.

We do not always need to make strong assumptions about the nature of productivity to identify how policy affects it. I learn how restructuring in the electricity industry affects power plant productivity—a critical policy question—without putting much structure on what kind of choices affect it, when those choices are made, or on how power plants compete before or after restructuring.

The main policy implication of the paper is that introducing competition into electricity generation did not increase productivity, but I find evidence that creating incentives to lower costs will increase productivity because greater market size (greater demand) increases productivity.

The empirical method I develop can be used in many other applications where productivity is a function of unobserved input choices. For example, in the executive compensation literature, the executive is modeled as providing an input to the firm's production problem; the input is unobserved by both the firm and the economist, creating both a principal-agent problem for the firm and an identification problem for the economist. We do not know whether the input the executive provides is flexible or fixed, and we can not observe investment in the input, in either case. But the empirical method I develop can be used to estimate how policy affects productivity when executives are providing an unobserved input (perhaps, a policy limiting executive compensation).

We can tighten the lower bound on the output elasticities if we are willing to assume at least one observed input is flexible and that plants have positive markups (I can not assume that in this paper because regulated power plants are often forced to price near average cost, and it's not clear what happens to markups after restructuring). Let θ_M be the output elasticity of M, a flexible input and let P_M be the price of M and P_Q be the price of output. If plants minimize variable cost and have positive markups, from the standard cost minimization problem, we learn:

$$\frac{P_M M}{P_Q Q} \leq \theta_M. \tag{40}$$

So the bounds are even tighter in unregulated industries where markups are always positive.

These extra bounds also let us use the approach to evaluate how policies change the distribution of markups or how markups differ across different types of plants (exporters versus non-exporters) while putting very little structure on how plants compete.

Linear-in-parameters shape restrictions can also be added to the empirical method. For example, we can require the production function has weakly decreasing output elasticities, a form of the intuition that using more of an input causes it to produce less, with the inequalities,

$$\sum_{j=1}^{J} \theta_j \frac{\partial^2 r_j}{\partial z_\ell^2}(z) \leq 0. \tag{41}$$

Or, we can assume the production function is log concave, another form of "decreasing returns", which amounts to the following assumption,

$$\sum_{j=1}^{J} \theta_j \left[\frac{\partial^2 r_j}{\partial z_\ell^2}(z) - \frac{\partial r_j}{\partial z_\ell}(z) \right] \leq 0. \tag{42}$$

When I tried imposing log concavity in my particular problem, the results were virtually identical (the extreme production functions under the linear positive association and increasing assumptions, happened to be log concave), but that might not be the case in other datasets.

Whether concavity of the production function in each input is a linear restriction depends on the form of the production function (it is a linear restriction for the Cobb Douglas production function).

We can also tighten the bounds by adding other assumptions common in the productivity literature (Markovian assumptions, assumptions about input demand) on top of the linear positive association and increasing production function assumption to learn what results depend on which assumptions.

While the method I develop in the paper tells us little about output elasticities or the level of marginal costs, it does tell us a lot about how productivity changes in response to policy or how it varies across different kinds of plants that face different incentives. These are often the ultimate questions of interest, and I show they can be answered with weak assumptions.

References.

Ackerberg, D. A., K. Caves, and G. Frazer (2006). Structural identification of production functions.

Borenstein, S. and J. Bushnell (2015). The u.s. electricity industry after 20 years of restructuring. *NBER Working Paper* (21113).

Chernozhukov, V., S. Lee, and A. Rosen (2013). Intersection bounds: estimation and inference. *Econometrica 81*(2), 667–737.

De Loecker, J. (2011). Recovering markups from production data. *International Journal of Industrial Organization 29*, 350–355.

De Loecker, J. (2013). Detecting learning by exporting. *American Economic Journal: Microeconomics 5*(3), 1–21.

Doraszelski, U. and J. Jaumandreu (2013). R&d and productivity: Estimating endogenous productivity. *The Review of Economic Studies 80*(4), 1338–1383.

Esary, J., F. Proschan, and D. Walkup (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics 38*(5), 1466–1474.

Fabrizio, K., N. Rose, and C. Wolfram (2007). Do markets reduce costs? assessing the impact of regulatory restructuring on us electric generation efficiency. *American Economic Review 97*(4), 1250–1277.

Flynn, Z. (2016). Inference on functions of parameters partially identified by the intersection of countably many linear inequalities. *WORKING PAPER*.

Gandhi, A., S. Navarro, and D. Rivers (2015). On the identification of production functions: how heterogeneous is productivity? *WORKING PAPER*.

Hicks, J. (1939). Value and capital. *Oxford: Clarendon Press*.

Joskow, P. (2001). California's electricity crisis. *Oxford Review of Economic Policy 17*(3), 365–388.

Knittel, C. (2002). Alternative regulatory methods and firm efficiency: Stochastic frontier evidence from the u.s. electricity industry. *The Review of Economics and Statistics 84*(3), 530–540.

Laffont, J.-J. (1994). The new economics of regulation ten years after. *Econometrica 62*(3), 507–537.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317–341.

Manski, C. and J. Pepper (2000). Monotone instrumental variables, with an application to the returns to schooling. *Econometrica 68*(4), 997–1012.

Marschak, J. and W. Andrews (1944). Random simulataneous equations and the theory of production. *Econometrica 12*, 143–205.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

MGE (2016). Residential electric rates - madison, wisconsin. https://drive.google.com/file/d/0B3lRF6Rj0MEuMzhZdWNDNG1LZ0E/view?usp=sharing. Accessed: 10-11-2016.

NEA (2016). Nuclear energy agency press kits - economics of nuclear power faqs. https://www.oecd-nea.org/news/press-kits/economics-FAQ.html. Accessed: 2016-08-19.

Olley, S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*, 1263–1297.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature 49*(2), 326–365.

Topkis, D. (1978). Minimizing a submodular function on a lattice. *Operations Research 26*(2), 305–321.

Vives, X. (2008). Innovation and competitive pressure. *The Journal of Industrial Economics 56*(3), 419–469.

A. A class of economic models that predict positive association between productivity and input use.

A power plant operator chooses output (Q), observed, and unobserved inputs (which result in productivity, A) to maximize profits net of the cost of those inputs and the future value of being more productive today.

The power plant is defined by the demand curve it faces (parametrized by ξ) and the cost of increasing productivity relative to the benefit it receives from doing so, λ. In terms of the general positive association class of models in the text, $s = (\lambda, \xi)$. Operators solve the following problem,

$$\begin{aligned} \text{Cost}\left(\frac{Q}{A}, W\right) = \quad & \min_{H,C,L,E} \quad W_H H + W_C C + W_L L + W_E E \quad \text{ST:} \quad F(H,C,L,E) \geq \frac{Q}{A} \\ & \max_{Q,A} \quad P(Q,\xi)Q - \text{Cost}\left(\frac{Q}{A}, W\right) - V(A,\lambda), \end{aligned} \tag{43}$$

Where V is the cost of productivity and the future value of becoming more productive now. λ depends on the non-unobserved input part of productivity (the "b_{it}" from the main text) and how costly the unobserved inputs are.

I think of capacity as flexible for now, but I will relax the assumption soon.

I make four assumptions.

Assumption 1. *$d\left[MR(Q,\xi)Q\right]/dQ \geq 0$ for all ξ.*
Assumption 2. *All inputs are normal; conditional input demand is increasing in output. An equivalent assumption: if the wage of any input increases, output decreases.*
Assumption 3. *Marginal revenue is increasing in ξ.*
Assumption 4. *The cross-partial of $V(A,\lambda)$ is negative. The marginal cost of productivity is decreasing in λ,*

$$\frac{\partial}{\partial A \partial \lambda} V(A,\lambda) \leq 0. \tag{44}$$

Assumption 1 is the worst of the four, but it is true for many common demand functions. It is true for any demand function with a weakly increasing (in absolute value) demand elasticity—examples: price-taking plants[21], linear demand, constant elasticity demand, and logit demand (discrete choice).

Theorem 1. *If the absolute value of the elasticity of the demand curve is increasing in price, the product of marginal revenue (as a function of output) and output is increasing in output.*

Proof. $\text{Rev}(Q) = P(Q)Q$, where $P(Q)$ is the inverse demand curve. Marginal revenue is:

$$\begin{aligned} \text{MR}(Q) &= P(Q) \times \left[P'(Q) \times \frac{Q}{P(Q)} + 1\right] \\ \implies \text{MR}(Q)Q &= P(Q)Q \times \left[P'(Q) \times \frac{Q}{P(Q)} + 1\right] = \text{Rev}(Q) \times \left[P'(Q) \times \frac{Q}{P(Q)} + 1\right] \end{aligned} \tag{45}$$

Taking the derivative gives:

$$\begin{aligned} \frac{d}{dQ}\left[\text{MR}(Q)Q\right] &= \text{MR}(Q) \times \left[P'(Q) \times \frac{Q}{P(Q)} + 1\right] + \text{Rev}(Q)\frac{d}{dQ}\left[P'(Q) \times \frac{Q}{P(Q)}\right] \\ &= \frac{\text{MR}(Q)^2}{P(Q)} + \text{Rev}(Q)\frac{d}{dQ}\left[P'(Q) \times \frac{Q}{P(Q)}\right] \end{aligned} \tag{46}$$

The second term is positive because a decreasing demand elasticity implies the inverse demand elasticity is increasing. The first term is also positive so:

$$\frac{d}{dQ}\left[\text{MR}(Q)Q\right] \geq 0. \tag{47}$$

□

Assumption 1 says costs limit the size of the plant, not demand. There is no interior solution to the revenue maximization problem.

Theorem 2 (Revenue maximization)**.** *If Assumption 1 is true and if there exists no $Q^\star$ such that $MR(Q^\star) = MR'(Q^\star) = 0$, there is no interior solution to the revenue maximization problem.*

Proof. Assume there exists a $0 < Q^\star < \infty$ such that revenue is maximized. Then:

$$\text{MR}(Q^\star) = 0. \tag{48}$$

Assume $\text{MR}'(Q^\star) < 0$, then:

$$\text{MR}(Q^\star) + Q^\star \text{MR}'(Q^\star) < 0, \tag{49}$$

Contradicting Assumption 1. So $\text{MR}'(Q^\star) \geq 0$ and, by assumption, $\text{MR}'(Q^\star) > 0$. So $Q^\star$ is a local minimum, not a maximum.

There is no interior solution to the revenue maximization problem. □

[21] Whether plants take price as given because of a lack of market power or if they take price as given because they are not allowed to charge any other price (rates set by regulation) is not important.

If Assumption 2 is not true, plants produce more output as inputs are more expensive. An old literature argues whether these "regressive" inputs exist, with Hicks (1939) being both the first to address them and the first to dismiss them. I follow Hicks (1939). Do greater fuel prices cause power plants to generate more electricity?

Assumptions 3 and 4 give the two plant parameters meaning.

Define $\widetilde{Q} = Q/A$. The four assumptions imply that the plant's objective function has increasing differences in the parameters (λ, ξ) and choices $(A, \widetilde{Q})$ and that the objective function is supermodular in $\left(A, \widetilde{Q}\right)$. By Topkis (1978), the choice of $\left(A, \widetilde{Q}\right)$ is increasing in (λ, ξ). Because input demand is normal, input demand is increasing in $\widetilde{Q}$. So the choice of (A, H, C, L, E) is increasing in (λ, ξ), which is Theorem 3, establishing the second property of the positive association class of models.

Theorem 3. *Given Assumptions 1, 2, 3, and 4, productivity choice and input use are increasing in the strong set order with respect to* (λ, ξ).

Proof. Let $\widetilde{Q} = Q/A$. Plants choose $\widetilde{Q}$ and A,

$$\max_{\widetilde{Q},A} P\left(\widetilde{Q} \times A, \xi\right)\left(\widetilde{Q} \times A\right) - \text{Cost}\left(\widetilde{Q}, W\right) - V(A, \lambda). \tag{50}$$

The objective function is supermodular in $\left(\widetilde{Q}, A\right)$ and has increasing differences in $\left(\widetilde{Q}, A; \lambda, \xi\right)$.

Supermodularity in $\left(\widetilde{Q}, A\right)$ requires the cross-partial of the objective function in $\widetilde{Q}$ and A is non-negative.

$$\begin{aligned}\frac{\partial}{\partial A}\left(P'\left(\widetilde{Q}A, \xi\right)\widetilde{Q}A^2 + AP\left(\widetilde{Q}A, \xi\right) - \text{Cost}'\left(\widetilde{Q}, W\right)\right) &= P''\left(\widetilde{Q}A, \xi\right)\widetilde{Q}^2A^2 + 2\widetilde{Q}AP'\left(\widetilde{Q}A, \xi\right) \\ &\quad + \widetilde{Q}AP'\left(\widetilde{Q}A, \xi\right) + P\left(\widetilde{Q}A, \xi\right) = \frac{d}{dQ}\left[\text{MR}(Q, \xi)Q\right] \geq 0,\end{aligned} \tag{51}$$

Which is Assumption 1.

Increasing differences in $\left(\widetilde{Q}, A; \lambda, \xi\right)$ says the cross-partial of any control $\left(\widetilde{Q}, A\right)$ with any parameter (λ, ξ) is non-negative.

The cross-partial of $\widetilde{Q}$ and λ is zero.

The cross-partial of $\widetilde{Q}$ and ξ is,

$$\begin{aligned}\frac{\partial}{\partial \xi}\left(P'\left(\widetilde{Q}A, \xi\right)\widetilde{Q}A^2 + AP\left(\widetilde{Q}A, \xi\right)\right) &= \frac{1}{A}\frac{\partial}{\partial \xi}\left(P'\left(\widetilde{Q}A, \xi\right)\widetilde{Q}A + P\left(\widetilde{Q}A, \xi\right)\right) \\ &= \frac{1}{A}\frac{\partial}{\partial \xi}\text{MR}\left(\widetilde{Q}A, \xi\right) \geq 0,\end{aligned} \tag{52}$$

Which is Assumption 3.

The cross-partial of A and λ is,

$$-\frac{\partial}{\partial A \partial \lambda} V(A, \lambda) \geq 0, \tag{53}$$

Which is Assumption 4.

The cross-partial of A and ξ is,

$$\begin{aligned} \frac{\partial}{\partial \xi}\left(P'\left(\widetilde{Q}A, \xi\right) \widetilde{Q}^2 A + \widetilde{Q} P\left(\widetilde{Q}A, \xi\right)\right) &= \widetilde{Q} \times \frac{\partial}{\partial \xi}\left(P'\left(\widetilde{Q}A, \xi\right) \widetilde{Q}A + P\left(\widetilde{Q}A, \xi\right)\right) \\ &= \widetilde{Q} \times \frac{\partial}{\partial \xi} \mathrm{MR}\left(\widetilde{Q}A, \xi\right) \geq 0, \end{aligned} \tag{54}$$

Which is Assumption 3.

By Topkis (1978), the choice of $\left(\widetilde{Q}, A\right)$ is increasing in ξ and λ.

The choice of (Z, A) is increasing in ξ and λ as well.

For any input ℓ:

$$\frac{\partial Z_\ell}{\partial \lambda} = \frac{\partial}{\partial \lambda} Z_\ell\left(\widetilde{Q}(\lambda, \xi), W\right) = \frac{\partial}{\partial Q} Z_\ell\left(\widetilde{Q}, W\right) \times \frac{\partial \widetilde{Q}}{\partial \lambda} \geq 0, \tag{55}$$

Because, by Assumption 2, I have $\partial Z_\ell(Q, W)/\partial Q \geq 0$, and, as I have just shown, $\partial \widetilde{Q}/\partial \lambda$ is positive.

The same is true for ξ. □

Theorem 3 makes it clear why demand shifters can not instrument for productivity; productivity is increasing in demand shifters, ξ.

So the model satisfies the second property of the positive association class of models: productivity and input use are increasing in the states (λ, ξ). For the model to be in the positive association class, we need ξ and λ to be positively associated.

There are two reasons why ξ and λ will be positively associated:

1. Greater ξ raises the future marginal value of productivity (higher λ).
2. Greater ξ increases productivity choice (Theorem 3) so plants with higher ξ are likely to have greater past productivity, lowering the cost to current productivity (increasing λ).

Both of those reasons require me to say something about dynamics. In the interest of simplicity, say $\xi_t = \xi$ for all t (allowing it to be a Markov process or to grow with A_{t-1} does not change the results). λ_t, on the other hand, will change over time,

$$\lambda_t = \Lambda_t(\lambda_{t-1}, A_{t-1}, \upsilon_t), \tag{56}$$

Where Λ_t is increasing in all arguments and υ_t is an exogenous shock to λ_t, independent of (λ_{t-1}, ξ). Λ_t is increasing in A_{t-1} because it is cheaper to move from a moderate A_{t-1} to a high A_t than from a low A_{t-1} to a high A_t. We can write λ_t as,

$$\lambda_t = \Lambda_t\left(\lambda_{t-1}, A_{t-1}\left(\lambda_{t-1}, \xi\right), \upsilon_t\right) = \widetilde{\Lambda}_t\left(\lambda_{t-1}, \xi, \upsilon_t\right). \tag{57}$$

Where $\widetilde{\Lambda}_t$ is increasing in all arguments because $A_{t-1}\left(\lambda_{t-1}, \xi\right)$ is increasing in both arguments. So greater ξ will, holding λ_{t-1} fixed, increase λ_t, so in the long run λ_t will tend to be greater if ξ is larger, suggesting the two are positively associated.

I formalize the result in Theorem 4.

Theorem 4. *If λ_0 is independent or, more generally, positively associated with ξ, then, for all t, λ_t and ξ will be positively associated.*

Proof. The argument is by induction on t.

If (λ_0, ξ) are positively associated, then because υ_1 is independent of (λ_0, ξ), $(\upsilon_1, \lambda_0, \xi)$ are positively associated (see Esary, Proschan, and Walkup 1967). Because increasing functions of positively associated random variables are also positively associated, and,

$$\lambda_1 = \widetilde{\Lambda}_1\left(\lambda_0, \xi, \upsilon_1\right) \tag{58}$$

is increasing in all arguments, (λ_1, ξ) is positively associated.

Now, assume (λ_{t-1}, ξ) are positively associated. λ_t is,

$$\lambda_t = \widetilde{\Lambda}_t\left(\lambda_{t-1}, \xi, \upsilon_t\right). \tag{59}$$

$(\lambda_{t-1}, \xi, \upsilon_t)$ are positively associated because (λ_{t-1}, ξ) are positively associated and υ_t is independent of (λ_{t-1}, ξ). So (λ_t, ξ) are positively associated because increasing functions of positively associated random variables are positively associated.

So, by induction, (λ_t, ξ) are positively associated.

□

Because increasing functions of positively associated random variables are positively associated, Theorem 4 implies productivity and input use are positively associated.

So I have established the second property.

Theorem 5. (a, h, c, ℓ, e) *are positively associated.*

Proof. $a(\lambda, \xi)$, $h(\lambda, \xi)$, $c(\lambda, \xi)$, $\ell(\lambda, \xi)$, and $e(\lambda, \xi)$ are increasing functions of λ and ξ (Theorem 3).

(λ, ξ) are positively associated (Theorem 4).

(a, h, c, ℓ, e) are positively associated because increasing functions of positively associated vectors are positively associated (see Esary, Proschan, and Walkup 1967). □

For a less abstract example of a model that satisfies these assumptions,

$$\max_{Q,A} \kappa \times \left(\frac{Q}{M_t}\right)^{\eta} Q - \text{Costs}\left(\frac{Q}{A}, W\right) - V\left(A - A_{t-1}\right), \tag{60}$$

Where M_t is market size ("ξ") and A_{t-1} is the previous period's productivity ("λ"). If the following assumptions are true, then inputs and productivity are positively associated:
Assumption 5. $0 \geq \eta \geq -1$*; demand is elastic.*
Assumption 6. $V' \geq 0, V'' \geq 0$*; convex adjustment cost of productivity.*
Assumption 7. *All inputs are normal.*
Assumption 8. (A_0, M_1) *are positively associated.*
Assumption 9. $M_{t+1} = \mathscr{M}_t\left(M_t, \upsilon_{t+1}\right)$ *where* $\mathscr{M}_t$ *is increasing in both arguments, can depend on time, and* υ_{t+1} *is independent of both* M_t *and* A_{t-1}.

The positive association prediction survives realistic deviations from the model; it is a robust prediction.

Two deviations are particularly likely for the electricity generation industry:

1. Power plant capacity is a fixed input, another state variable.

2. Fuel prices vary geographically and by fuel type.

If capacity (C) reduces the marginal cost of $\widetilde{Q}$,

$$\frac{\partial}{\partial \widetilde{Q} \partial C} \text{Cost}\left(\widetilde{Q}, C\right) \leq 0, \tag{61}$$

Then $A(\lambda, \xi, C)$, $H(\lambda, \xi, C)$, $E(\lambda, \xi, C)$, and $L(\lambda, \xi, C)$ are all increasing functions. If C is positively associated with (λ, ξ), then the model is still in the class of positively associated models. C is positively associated with (λ, ξ) if higher capacity plants tend to face better demand curves (that is why the plant is large in the first place) and have greater future values to productivity (because capacity increases the marginal product of productivity).

If the negative of fuel prices $(-W_H)$ is positively associated with (λ, ξ, C) (say, because fuel prices are independent of λ and ξ and negatively related to plant size), then the model allowing fuel price variation is in the positive association class because,

$$\frac{\partial}{\partial W_H \partial \widetilde{Q}} \text{Cost}\left(\widetilde{Q}, C, W\right) \geq 0, \tag{62}$$

By Assumption 2.

So, from the Topkis (1978) theorem, we have $A(\lambda, \xi, C, -W_H)$, $H(\lambda, \xi, C, -W_H)$, $E(\lambda, \xi, C, -W_H)$, and $L(\lambda, \xi, C, -W_H)$ are increasing in all their arguments, and we have assumed those states are positively associated so the model is in the positive association class.

B. Robustness checks.

Table 6: Effect of restructuring on plant fuel productivity (Announcement definition)

Parameter	Restructuring	Model	LB (90%)	LB	UB	UB (90%)
τ_a	Announcement	Static	−2.48%	−1.77%	−0.51%	0.26%
τ_a	Announcement	Dynamic	−2.93%	−0.84%	−0.49%	1.59%
τ_a	Law Passed	Static	−3.86%	−2.62%	−1.12%	0.26%
τ_a	Law Passed	Dynamic	−11.82%	−7.02%	−5.83%	−1.29%
κ	Announcement	Static	2.24%	3.43%	4.44%	5.60%
κ	Announcement	Dynamic	1.30%	4.17%	5.08%	8.06%
κ	Law Passed	Static	2.44%	3.64%	4.53%	5.71%
κ	Law Passed	Dynamic	1.64%	4.53%	5.30%	8.30%

Notes. LB (90%) and UB (90%) give the upper and lower limits of the 90% confidence interval.

Table 7: Effect of restructuring on geometric aggregate fuel productivity (GAP) (Announcement definition)

Parameter	Restructuring	LB (90%)	LB	UB	UB (90%)
GAP	Announcement	−0.35%	0.43%	0.66%	1.42%
GAP	Law Passed	−2.43%	−1.61%	−1.41%	−0.63%
Average Productivity	Announcement	−1.00%	−0.32%	0.69%	1.46%
Average Productivity	Law Passed	−4.35%	−3.65%	−2.64%	−1.89%
Reallocation	Announcement	−0.57%	−0.10%	0.80%	1.12%
Reallocation	Law Passed	0.82%	1.22%	2.09%	2.44%

Notes. LB (90%) and UB (90%) give the lower and upper limits of the 90% confidence intervals.

B.1. Robustness to restructuring definition.

In the main text results, I used the Law Passed definition of restructuring which, I argue, is likely a better definition than the Announcement definition because many states that started planning to restructure never actually did.

In this appendix, I give the results for the Announcement definition. The results have mostly the same sign as for the Law Passed definition but are less precisely estimated. The only difference is that the point estimates of the bounds on the effect of restructuring on geometric aggregate productivity are positive under the Announcement definition but the magnitude of the effect is less than 1% and the lower bound is not statistically different from zero.

A simple theory that explains the noisier estimates under the Announcement definition: some plants believed restructuring would happen and other plants did not.

Table 8: Effect of restructuring on plant fuel productivity (remove more extreme heat rates)

Parameter	Restructuring	Effect	LB (90%)	LB	UB	UB (90%)
τ_a	Announcement	Static	−1.74%	−1.29%	−0.16%	0.36%
τ_a	Announcement	Dynamic	−1.23%	−0.30%	0.04%	0.95%
τ_a	Law Passed	Static	−3.05%	−2.26%	−0.85%	0.05%
τ_a	Law Passed	Dynamic	−1.11%	0.95%	1.47%	3.54%
κ_a	Announcement	Static	−1.26%	−0.51%	0.32%	1.03%
κ_a	Announcement	Dynamic	−2.59%	−1.21%	−0.57%	0.85%
κ_a	Law Passed	Static	−1.19%	−0.44%	0.39%	1.10%
κ_a	Law Passed	Dynamic	−2.30%	−0.92%	−0.30%	1.11%

Notes. LB (90%) and UB (90%) give the upper and lower limits of the 90% confidence interval.

B.2. Robustness to outlier selection method.

In the main results, I removed all observations with a fuel to power ratio more than 3 times the average rate published by the Energy Information Administration or less than (1/3) of the average rate. To check for robustness to outliers, in this section, I remove all observations with a fuel to power ratio more than 1.75 times the average rate or less than (1/1.75) times the average rate. 1.75 is chosen to omit the lower line of points that is included in Figure 1.

The estimates of the effects are noisier (because we remove more observations), but the sign of the effect of restructuring does not change. I lose 195 observations when moving to this more restricted sample. There are 898 investor-owned power plant-years in restructured states (by the Law Passed definition) in the original data and 867 in the restricted sample data.

The estimates of the aggregate effect of restructuring are qualitatively the same as the results in the main text.

The only true deviation from the main results is that the κ estimated for plant fuel productivity is small and insignificant.

The issue with limiting the number of power plant heat rates that we include in the dataset is that it reduces the variation in the dependent variable, power to heat ratios so this will move the coefficients toward zero mechanically. The dataset I use in the main text of the paper is a better balance of being concerned with measurement error and with removing useful variation.

Table 9: Effect of restructuring on geometric aggregate fuel productivity (GAP) (remove more extreme heat rates)

Parameter	Restructuring	LB (90%)	LB	UB	UB (90%)
GAP	Announcement	−0.78%	0.04%	0.23%	1.02%
GAP	Law Passed	−2.44%	−1.58%	−1.34%	−0.44%
Average Productivity	Announcement	−0.92%	−0.19%	0.76%	1.51%
Average Productivity	Law Passed	−3.64%	−2.88%	−1.91%	−1.14%
Reallocation	Announcement	−0.89%	−0.57%	0.29%	0.52%
Reallocation	Law Passed	0.22%	0.52%	1.39%	1.65%

Notes. LB (90%) and UB (90%) give the lower and upper limits of the 90% confidence intervals.

C. Econometric details.

The econometric details are mostly excerpted from Flynn (2016) with the exception of Theorem 6.

C.1. Proof that we only need step functions in the linear positive association constraints.

Theorem 6. *Let z be the L-vector of inputs. The parameter set identified by the linear positive association can be written by only using, g_1 and g_2 such that,*

$$g_j(z;u) = \mathbf{1} \;\; \Phi(z) \geq u_j \qquad u_j \in \mathbb{Q}^L \cap [0,1]^L \qquad \textit{for } j = 1, 2, \tag{63}$$

Where $\Phi(z) \in [0,1]^L$ and $\Phi_\ell(z) = \Phi_\ell(z_\ell)$ where Φ_ℓ is a strictly increasing function for $\ell = 1, \ldots, L$.

Proof. The parameter set defined by the intersection only over the step functions in Theorem 6 clearly contains the parameter set identified by the linear positive association because the step functions are increasing functions. I show it is also contained by the set identified by linear positive association, establishing equality of the sets.

The proof draws somewhat to Theorem 3.4 of Esary, Proschan, and Walkup (1967).

Let $t_1, \ldots, t_k$ be a finite set of scalars which approach the rational numbers in $[0,1]$ as $k \to \infty$.

Let g_1 and g_2 be two non-negative, bounded, continuous, increasing functions.

Define $h_k(z) = g_1(\eta)$ where,

$$\eta_i = \Phi_i^{-1} \max_{j \in \{1,\ldots,k\}} t_j : \mathbf{1} \;\; \Phi_i(z_i) \geq t_j \qquad \text{for } i = 1, \ldots, L. \tag{64}$$

Define $v_k(z) = g_2(\eta)$.

Assume,

$$\operatorname{cov}(a\mathbf{1}(\Phi(z_i) \geq u), \mathbf{1}(\Phi(z_i) \geq v)) \geq 0. \tag{65}$$

For all u and v in the rational numbers between $[0, 1]$.

There is a positive function κ_h such that,

$$h_k(z) = \sum_{t_1=1}^{k} \cdots \sum_{t_L=1}^{k} \kappa_h(t_1, \ldots, t_L) \times \mathbf{1}(\Phi_1(z_1) \geq t_1, \ldots, \Phi_L(z_L) \geq t_L). \tag{66}$$

There is such a κ_v for v_k as well.

Therefore, by the linearity of the covariance,

$$\begin{aligned} \sum_{t^1} \sum_{t^2} \kappa_h \;\; t^1 \;\; \kappa_v \;\; t^2 \;\; \operatorname{cov} \;\; a\mathbf{1} \;\; \Phi(z) \geq t^1 \;\; , \mathbf{1} \;\; \Phi(z) \geq t^2 \;\; &= \\ \operatorname{cov}(ah_k(z), v_k(z)) &\geq 0. \end{aligned} \tag{67}$$

Because g_1 and g_2 are continuous, $0 \leq h^k \to g_1$, $0 \leq v^k \to g_2$, and $0 \leq v^k h^k \to g_1 g_2$ point-wise (as $k \to \infty$, as $\{t_1, \ldots, t_k\} \to \;\; \cap [0, 1]$). Because g_1 and g_2 are bounded and because the sequences in k are increasing, by the monotone convergence theorem,

$$\begin{aligned} \mathbb{E}(ah_k(z)) &\to \mathbb{E}(ag_1(z)) \\ \mathbb{E}(v_k(z)) &\to \mathbb{E}(g_2(z)) \\ \mathbb{E}(ah_k(z) v_k(z)) &\to \mathbb{E}(ag_1(z) g_2(z)). \end{aligned} \tag{68}$$

So I have what I wanted to show,

$$0 \leq \lim_{k \to \infty} \operatorname{cov}(ah_k(z), v_k(z)) = \operatorname{cov}(ag_1(z), g_2(z)). \tag{69}$$

□

C.2. Tuning parameter selection.

I first choose what increasing function to use for Φ_h, Φ_c, Φ_ℓ, and Φ_e. I fit a mixture of two normal distributions to each input's marginal distribution and use that distribution function for Φ. Visually, the fit is a close approximation to the empirical distribution of each of the inputs. Intuitively, using the true distribution for Φ would produce good results because the observations would be evenly distributed (no clumping just above or below certain thresholds).

I choose ϵ_n using the data-driven method in Flynn (2016), but β_n still needs to be chosen.

$\exp(\beta_n/\epsilon_n)$ sets the scale of the penalty function so the first thing to do is to normalize the scale of the constraints. I multiply the penalty function by the maximum absolute value of the coefficients, effectively normalizing all the coefficients to be between 0 and 1,

$$w_k = ||(d_k, p_k)||_\infty \tag{70}$$

$$d_k^\top \theta \leq p_k \iff (d_k/w_k)^\top \theta \leq (p_k/w_k) \tag{71}$$

$$\rightarrow \frac{\widehat{d}_k}{\widehat{w}_k}^\top \theta \leq \frac{\widehat{p}_k}{\widehat{w}_k} + \exp \quad \frac{\beta_n - (n_k/n)}{\epsilon_n} \tag{72}$$

$$\iff \widehat{d}_k^\top \theta \leq \widehat{p}_k + \widehat{w}_k \exp \quad \frac{\beta_n - (n_k/n)}{\epsilon_n} \tag{73}$$

I follow Flynn (2016) and choose,

$$\exp(\beta_n/\epsilon_n) = (\log\log n)^{1/r}, \tag{74}$$

For a parameter r. Monte Carlo simulations in Flynn (2016) showed this parameter did not matter much, but I select $r = 4$ because the results were slightly better for that parameter

ϵ_n chooses which constraints bind (the other factors are constant across constraints) so it is the most important parameter. I choose ϵ_n by minimizing an approximation to the mean squared error of the estimator,

$$\epsilon_n = \arg\min_{\epsilon} n \left[\sum_{k=1}^{K_n} \widehat{\lambda}_n(\epsilon) \widehat{w}_k (\log\log n)^{1/r} \exp\left(-\frac{n_k}{n\epsilon}\right. \right]^2 + \widehat{\sigma}_n(\epsilon)^2,$$

Where $\widehat{\sigma}_n$ is the standard error of the lower bound and $\widehat{\lambda}_n$ are the Lagrange multipliers. While I could select different tuning parameters for each statistic I bound, I only select the tuning parameters for the upper bound on the average sum of the output elasticities and use those tuning parameters for all the statistics I bound. That choice of statistic makes sure as many linear positive association constraints bind as possible so it should be informative about which tuning parameters to choose.

The second term is the variance of the estimator and the first term is a first order approximation to the bias from using the penalty function. Let $\widehat{t}_n$ be the estimated, penalized function and $\widehat{t}_{n,0}$ be the estimated, non-penalized program (but still truncated with $k \leq K_n$). Then, the first order Taylor expansion gives,

$$\widehat{t}_{n,0} \approx \widehat{t}_n + \sum_{k=1}^{K_n} \widehat{\lambda}_n(\epsilon) \widehat{w}_k (\log\log n)^{1/r} \exp\left(-\frac{n_k}{n\epsilon}\right. ,$$

Which is the source of the first term.

The two terms capture the basic trade-off with choosing ϵ_n: low ϵ_n makes bias fall but allows lower probability constraints to bind, increasing variance.

Choose a grid for ϵ_n and compute the objective function over each point on the grid. Plot ϵ versus the value of the objective function. You will see the objective is discontinuous in ϵ around the points where the basis of the program switches. The idea is to choose the region of ϵ with the smallest value of the objective function but to choose ϵ a safe distance from the switching point (where the program would be non-differentiable). The easiest way to do this is to choose the ϵ that minimizes the objective function subject to the constraint that it is a reasonable distance from a point where the basis switches. Because the objective function (and the estimate t_n) will be continuous in ϵ in this neighborhood, it is not so important what distance you pick.

Choosing ϵ_n in this way ensures that the t_n program is differentiable which, in large samples, should ensure the t_n program is, as well.

C.3. Step-by-step computation instructions.

Theorem 7. *Let t_n^u be the estimate of the upper bound, t^u, of some statistic, and let t_n^ℓ be the estimate of the lower bound, t^ℓ. Given the assumptions in Flynn (2016),*

$$pr\left(t_n^\ell - c_n\frac{\sigma_\ell}{\sqrt{n}} \leq t \leq t_n^u + c_n\frac{\sigma_n^u}{\sqrt{n}}\right) \geq \alpha. \tag{75}$$

For $\alpha \in \left(\frac{1}{2}, 1\right)$. Where c_n solves:

$$\Phi\left(c_n + \sqrt{n}\frac{t_n^u - t_n^\ell}{max\left(\sigma_n^\ell, \sigma_n^u\right)}\right) - \Phi(-c_n) = \alpha, \tag{76}$$

Where Φ is the standard normal CDF.

Proof. Proof in Flynn (2016). □

Figure 2: Step-by-step procedure to estimate bounds on linear statistics

1. Fit a mixture of two normal distributions to the marginal distribution of each element of z, the vector of inputs. Let φ be the vector of the estimated cumulative distribution functions.
2. Set $m_n = 5 \times (n/1000)^{1/6}$ (see Flynn (2016) for more details).
3. Estimate all the coefficients of the linear programming problem.
4. Choose β and ϵ using the method in Flynn (2016) (details in Appendix C.2).
5. Say $T[a]$ is a linear statistic (like a linear regression coefficient) so that,

$$T[a] = T[y] - \sum_{j=1}^{b} \theta_j T \quad r_{b,j}(c,\ell) \quad . \tag{77}$$

6. Solve the linear programming problem to find the maximum and minimum value of the statistic $T[a]$, a linear function of θ, that satisfy the linear inequalities constructed. Let $\overline{\theta}$ be the primal solution to the upper bound and $\underline{\theta}$ be the primal solution to the lower bound. Let $\underline{\lambda}$ and $\overline{\lambda}$ be the dual values of the upper and lower bound linear programs that correspond to the positive association constraints that bind.
7. Compute $\underline{\Sigma}$, the variance-covariance matrix of the covariance estimates, $n(u_1, u_2)/n$ in the binding positive association constraints, and the estimator of $-T \quad r_{b,j}(c,\ell)$ for the lower bound. Similarly, compute $\overline{\Sigma}$ for the upper bound.
8. Let B be the indexes of the binding constraints for the lower bound,

$$\upsilon_n = \quad \underline{\lambda}_k \underline{\theta} \quad {}_{k \in B}, \quad -\underline{\lambda}_k \quad {}_{k \in B}, \quad \frac{\underline{\lambda}_k}{\epsilon} \exp \quad \frac{\beta - (n_k/n)}{\epsilon} \qquad {}_{k \in B}, \underline{\theta} \tag{78}$$

$$\underline{\sigma}_n = \sqrt{\upsilon_n^{\top} \underline{\Sigma} \upsilon_n} \tag{79}$$

   Do a similar computation for the upper bound.
9. Then, compute the confidence interval using the formula in Theorem 7.