Energy Transitions in Regulated Markets*

Gautam Gowrisankaran† Ashley Langer‡ Mar Reguant§

November 15, 2023

Abstract

Increasing concerns with damages from electricity emissions have coincided with two major energy transitions: from coal to natural gas and then to renewable sources. In many U.S. states, these energy transitions are occurring in regulated electricity markets. This paper develops and estimates a model of rate-of-return regulation for electric utilities which choose investment and retirement capacities in the long-run and generation quantities in each hour. Standard rate-of-return regulation limits utilities to earning a "fair" return on their capital, or "rate base." We adapt this framework to our empirical context by assuming that the maximum rate of return decreases with variable costs and that the rate base may omit capital that is not adequately used. Utilities therefore face a tradeoff between using legacy generation technologies and newer, cheaper ones. We find that cost-minimizing utilities generate 47% less electricity from coal, because they do not face usege incentives; a social planner who considers carbon emissions generates 95% less. Over 10 years, cost-minimizing utilities facing low natural gas prices retire approximately half of their coal capacity, with the social planner retiring 15 percentage points more. Feasible regulatory structures that altered usage incentives or penalties for high costs would not have approached cost minimizing outcomes.

JEL Codes:
Keywords: electricity regulation, used and useful, stranded assets.

*We thank Steve Cicala, Lucas Davis, Meredith Fowlie, Catie Hausman, Erin Mansur, and seminar participants for their helpful comments. We are grateful to Miguel Borrero, Kelli Marquardt, Nicholas Moshkriz, Filippo Paternollo, and Claire Petersen for excellent research assistance.

†Columbia University, CEPR, and NBER

‡University of Arizona and NBER

§Northwestern University, BSE, CEPR, and NBER

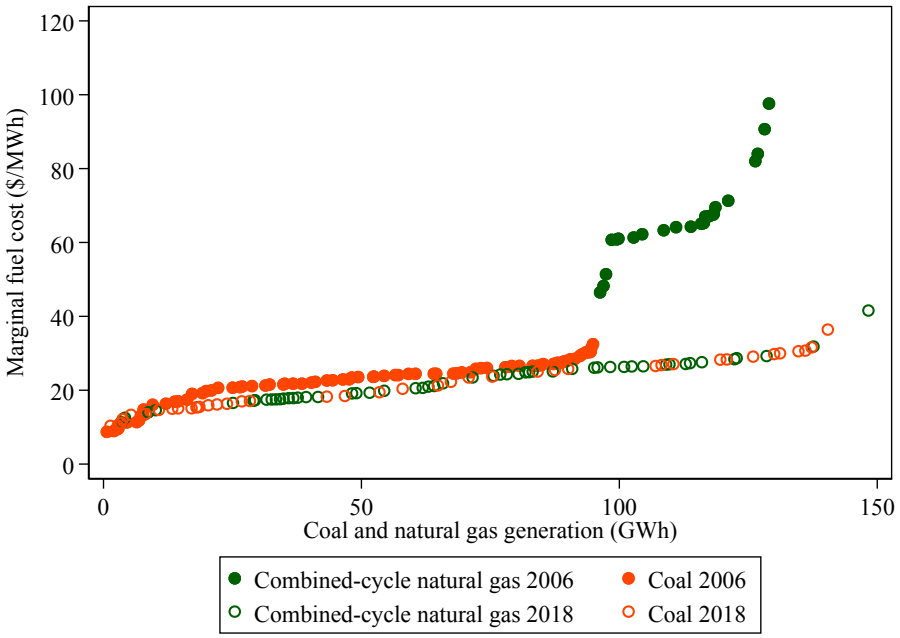
1 Introduction

Electricity generation is a critically important component of the economy and modern life. However, it also generates substantial negative externalities. In particular, electricity generation in the U.S. contributed 31% of U.S. CO_2 emissions in 2019 (Energy Information Agency, 2020). Beyond its climate change impacts, electricity generation also emits substantial local pollutants that harm human health, with damages for the U.S. estimated at $57.3 billion in 2017 (Holland et al., 2020).

At the same time that society is increasingly concerned about the damages from electricity emissions, two major transitions in electricity generation are underway. First, over the past two decades, combined-cycle technology and hydraulic fracturing (fracking) substantially lowered the cost of generating electricity with natural gas. Figure 1 shows marginal electricity fuel costs in 2006 and 2018 for combined-cycle natural gas and coal generators in our sample. The figure sorts generators by dispatch order—i.e., in order of increasing marginal cost—with capacities displayed cumulatively. In 2006, combined-cycle natural gas generators (in green) had uniformly higher fuel costs than coal. Yet, by 2018, there were substantially more combined-cycle natural gas generators, and their costs were similar to or below coal. This shift to natural gas has led to a 28% reduction in CO_2 emissions between 2005 and 2018 and lower local air pollution (U.S. Energy Information Administration, 2018). Second, the costs of solar panels and batteries have both dropped over 80% since 2010 (International Renewable Energy Agency, 2020; Cole and Frazier, 2019; Goldie-Scot, 2019). Given these cost declines and the substantially lower pollution externalities from renewables, a second transition, to renewable energy, has started.

Since electricity generation has high fixed costs and low marginal costs, electric utilities were historically regulated as natural monopolies. The 1990s saw extensive restructuring in the U.S. and Europe, where electricity generation—and sometimes distribution—was opened to competition. In the U.S., this deregulatory push ended with the California electricity crisis of the early 2000s, leaving a patchwork system where states have substantially different levels of regulatory control. The impact of falling natural gas generation costs varied across

Figure 1: Marginal Fuel Costs Over Time



Note: Generator level data on marginal fuel costs and capacities from analysis sample.

regulated and restructured markets. For instance, between 2006 and 2018, 30.3% of coal capacity exited in restructured states, whereas only 19.7% exited in regulated states.[1] This suggests that it is important to study whether electricity regulation increases social costs by slowing the transitions to new energy sources.

This paper develops and estimates a model of rate-of-return regulation for electric utilities. In our model, a utility chooses investment and retirement capacity levels in the long run and generation quantities by fuel-technology and imports each hour. Rate-of-return regulation of electricity creates a structure under which the utility is incentivized to meet demand (electricity load) while encouraging low cost generation and limiting underused capital (Joskow, 1974).[2] We estimate our model using publicly available data on hourly electricity generation, load, and revenues by utility, and capacity by fuel-technology and utility. We use our model to evaluate how both the current and alternative regulatory structures affect generation costs, transitions to new energy sources, and pollution.

[1] Authors' calculations based on analysis data, discussed in Section 2.2.

[2] The regulator may also seek to minimize the environmental impact of electricity generation.

Rate-of-return regulation limits a utility to earning a "fair" rate of return on its capital, referred to as its "rate base" (Viscusi et al., 2018). While rate-of-return regulation aims to create incentives for the utility to make efficient decisions, it is well understood that it is difficult to structure natural monopoly regulations appropriately. In particular, a literature has shown that rate-of-return regulation can lead to over-investment in capital because regulation limits the utility's return per unit of capital and the utility can endogenously respond by increasing capital (Averch and Johnson, 1962). Regulators therefore have combined rate-of-return regulation with the requirement that only "prudent" capital investments be included in the rate base. Energy transitions complicate the regulator's task of determining prudence, because the regulator may not know the current and future costs of alternative utility investment and operation decisions. One common approach to determining whether an investment is prudent is whether it is "used and useful" for generation (Gilbert and Newbery, 1994; Fisher et al., 2019), but this may cause the utility to operate capital inefficiently.

Our model captures these key features of rate-of-return regulation of electricity. The regulator in our model accomplishes its objectives via two instruments. First, it offers a maximum rate of return that is declining in the utility's variable costs. This instrument encourages the utility to invest in low-cost generators and use low variable cost sources. Second, the regulator determines whether capital should be fully included in the rate base based on its utilization.

The utility optimizes against this regulatory structure in its operations and investment decisions. When choosing imports and which fuel-technologies to operate to meet load, a utility has two potentially conflicting incentives. First, to increase its allowable rate of return, the utility seeks to have low total variable costs, which include fuel, ramping, import, and operations and maintenance costs.[3] Second, particularly after the decline on natural gas generation costs, it may want to use relatively expensive coal generators to ensure that they are considered used and useful.

Each long-run period, which we model to be three years long, the utility chooses how much coal capacity to retire and how much combined-cycle natural gas capacity to invest

[3]The existence of ramping costs links these operations decisions across hours.

in. The utility faces a quadratic cost adjusting each type of capacity, which includes an idiosyncratic shock to marginal adjustment costs in each period.

Our model depends on structural parameters that underlie the regulatory framework: the maximum rate of return the regulator offers to the utility, the determinants of coal generation being considered used and useful, and the contribution of each fuel/technology to the rate base. Further, the structural parameters include operating and retirement/investment costs.

We estimate the operations parameters with a nested fixed-point indirect inference approach. Specifically, we run regressions on our actual data that capture key features such as utilities' revenues as a function of costs, ramping behavior, and utilization in periods when extra utilization is likely to increase the rate base. We then search across structural parameters. For each structural parameter vector, we simulate utilities' optimal decisions. We search for the structural parameter vector that yields the most similar regression coefficients between simulated data and real data. The indirect inference approach estimates parameters that match important data correlations.

We also estimate the investment and retirement costs with a nested fixed-point approach, in this case using maximum likelihood. In each three-year period, we compute the likelihood of observed retirement and investment quantities for each potential parameter value and search over structural parameters to maximize the likelihood. For a given parameter value, we solve each utility's dynamic programming problem and calculate the likelihood of observing the investment and retirement decisions in the data. We model continuous choices of retirement and investment and an unobservable shock to marginal cost, rather than, as is often done, a model with discrete choices and unobservable terms on each choice. We follow Gowrisankaran and Schmidt-Dengler (2023), which develops an algorithm that facilitates the computation of this type of model.

We use our structural parameter estimates to conduct several counterfactual policy experiments, which compare variable costs, capacity utilization, investment, and pollution across environments. To begin, we investigate a benchmark where a social planner minimizes the dynamic expected costs of electricity provision. We compare the planner solution to the baseline predictions from the current regulatory environment to evaluate the potential limitations

of rate-of-return regulation. To understand the impact of adjusting different components of the current regulatory structure, we then compare the baseline to counterfactual regulatory environments that vary the penalty for high total variable costs and the used-and-useful criterion. These counterfactuals reveal the extent to which the regulator faces tradeoffs between incentivizing the utility to make efficient operations decisions and capacity investment decisions. Finally, we estimate the impact of carbon taxes on outcomes, allowing carbon taxes to change utilities' incentives by affecting total variable costs. Comparing their impact under both the current regulatory environment and the planner solution.

Literature: This paper relates to three broad literatures. First, we build on a long-standing literature on the theory of regulation. A number of papers have examined the optimal design of rate-of-return regulation (e.g., Averch and Johnson, 1962; Baumol and Klevorick, 1970; Klevorick, 1971, 1973; Joskow, 1974; Gilbert and Newbery, 1994); and Joskow (2007). A different strand of the literation has focused on setting incentives correctly in the presence of asymmetric information (Laffont and Tirole, 1986).

Second, we extend the empirical literature on the impact of electricity regulation, which includes Fowlie (2010); Davis and Wolfram (2012); Cicala (2015); Abito (2020); Lim and Yurukoglu (2018); MacKay and Mercadal (2019); Cicala (2022); and Dunkle Werner and Jarvis (2022). The closest paper in this literature to ours is Abito (2020), which structurally estimates a model of an electric utility in order to understand the operating inefficiencies created by rate-of-return regulation due to asymmetric information. We add to these two literatures by specifying and structurally estimating a model of regulation that incorporates utilities' investment, retirement, and operations decisions when faced with rate-of-return regulation.

Finally, we contribute to the growing empirical literature on the dynamics of investment and exit in electricity markets, which includes Myatt (2017); Eisenberg (2019); Linn and McCormack (2019); Abito (2020); Elliott (2022); Butters et al. (2023); and Gowrisankaran et al. (2023). Gowrisankaran et al. (2023), written by an overlapping set of co-authors, is similar to the other papers cited here in considering coal plant retirements for independent power producers. This paper differs by modeling decisions of regulated utilities.

The next section presents our model of electricity regulation. Section 2 introduces the data and background. Section 3 presents the model. Section 4 presents reduced-form evidence on operational and investment decisions by utilities, which also helps clarify how we interpret the data through the lens of our model. Section 5 presents our estimation approach, and Section 6.1 discusses our estimation and counterfactual results. Section 7 concludes.

2 Industry Background and Data

2.1 Industry Background

Rate-of-return regulation often exists in industries that are considered "natural monopolies" (Viscusi et al., 2018). The goal of this regulatory approach is to ensure the provision of the good—in electricity, to literally keep the lights on—while avoiding monopoly prices. In the U.S., state regulatory agencies, generally called Public Utility Commissions (PUCs), implement electricity rate-of-return regulation with these goals in mind.[4]

Specifically, in our context, PUCs make three different types of decisions.[5] First, they approve utility capital investments and retirements which, together with depreciation rates, ultimately affect what is included in the "rate base," which is defined as the capital stock on which the regulator gives the utility an allowable rate of return. Second, they determine which of utilities' reported non-capital costs are reimbursable. Third, they decide on the allowable rate of return which, together with the first two decisions, determine the profits that utilities can earn.

PUCs collect information to make these decisions largely via Integrated Resource Plans (IRPs) and rate hearings. Utilities propose IRPs that lay out the capital investments and retirements which they believe will best allow them to meet regulator objectives. PUCs then

[4]For instance, the Minnesota statute governing public utilities states "[t]he commission, [..] shall give due consideration to the public need for adequate, efficient, and reasonable service and to the need of the public utility for revenue sufficient to enable it to meet the cost of furnishing the service, including adequate provision for depreciation of its utility property used and useful in rendering service to the public, and to earn a fair and reasonable return upon the investment in such property." (Minn. Stat. § 216B.16, subds. 6).

[5]Our discussion of the regulatory process draws heavily from a guide to electricity regulation written by an independent think tank (Lazar, 2016) and a classic textbook on regulation (Viscusi et al., 2018).

decide which of these investments are likely to be prudent and therefore reimbursable in the rate base.

PUCs also hold periodic rate hearings, which are opportunities for them to adjust consumer rates (Joskow, 2014; Abito, 2020). Before these hearings, utilities submit documentation of their recent performance—including usage of existing generators, costs, and revenues—as well as expected future performance, including from investments. PUCs use rate hearings to decide on the allowable rate of return on the rate base and then set consumer electricity price schedules so that the utility can expect to cover its reimbursable variable costs and earn the allowable rate of return.

Beyond changing electricity rates, PUCs can respond to varying circumstances by making discretionary adjustments to the rate base, e.g., for construction in progress, investments in terminated projects, and fuel stocks. They may explicitly look at metrics such as usage when making these discretionary adjustments. As (Lazar, 2016, p. 52) explains: "Generally, to be allowed in rate base, an investment must be both used and useful in providing service and prudently incurred. The utility has the burden of proving that investments meet these well-established tests, but often enjoys presumption of use and usefulness, and prudence in the absence of evidence to refute it."

This regulatory structure encourages utilities to invest in capital to increase their rate base. In fact, regulated utilities own most of the generation capacity within their states. For instance, using Energy Information Agency (EIA) data, the Rocky Mountain Institute found that on average utilities owned 76% of generation capacity within regulated states, while the comparable figure for restructured states was 15% (Shwisberg et al., 2020). At the same time, most utilities also trade electricity with outside generation sources either bilaterally or through regional electricity markets. For instance, Missouri is a regulated state, but it lies within the Midcontinent Independent System Operator (MISO) electricity market, and utilities in Missouri trade electricity through this platform.

Advocacy groups and research organizations have argued that this regulatory structure leads to inefficient deviations from utility cost minimization, despite the potential for trades. Numerous groups have found that utilities that trade in wholesale electricity markets often

choose to "self-commit" (or mandate that their own generators must run) even when these generators' costs exceed the market price (Fisher et al., 2019; Daniel et al., 2020; Potomac Economics, 2020). Further, regulated utilities may have a preference to build their own capacity rather than signing power purchase agreements with third parties who can produce electricity at lower cost (Cross-Call et al., 2018; Wilson et al., 2020). This inclination has extended to recent decisions concerning renewable energy, where there is also concern that regulated utilities are underinvesting (Bottorff et al., 2022; Biewald et al., 2020; Daniel, 2021).

The potential for these types of inefficiencies is supported by economic theory. In general, rate-of-return regulation will not yield long-run cost minimization because utilities have an incentive to over-invest in capital (Averch and Johnson, 1962). Further, rate-of-return regulation establishes a system of incentives against which utilities optimize, which is designed to limit the potential for regulator time inconsistency or renegotiation.[6] Rate-of-return regulation was developed in an environment with relatively stable technologies, but electricity generation over the past 25 years has been characterized by changing technologies, fuel prices, and perceptions of environmental harm. These changing market conditions can lead to further inefficiencies if, for instance, utilities run inefficient generation technologies out of dispatch order to justify that they are prudent and merit continued capital investment.

2.2 Data

We use data on the electricity industry in the U.S. from a variety of publicly available sources. Our data include both annual measures—such as generator capacity, fuel prices, and utility revenues—and hourly measures, such as load, generation, and wholesale electricity prices. Our main estimation sample extends from 2006 to 2018. We now describe each of our data sources in turn.

Our primary annual data on the electricity derive from the EIA. We merge together information from three EIA forms. First, EIA Form 861 provides annual total revenue for electric utilities that are obligated to report this information. We deflate revenues and prices

[6]Nonetheless, a regulator's preferences may change when political control of the PUC shifts, which may cause inefficiencies such as underinvestment (Lim and Yurukoglu, 2018).

(discussed below) to January, 2006 dollars using the CPI net of food and energy. Form 860 records information about each power plant's capacity, fuel-technology type, and U.S. state. We retain information on plants with three fuel-technology types: coal (COAL), combined-cycle natural gas (CCNG), and other (non combined-cycle) natural gas turbines (NGT). This form also provides a utility identifier, which allows us to merge these data with the revenue data. Finally, Form 923 has annual plant-level data on fuel energy input in MMBtus and electricity generation in MWhs. We combine these data to recover each plant's heat rate, which is its fuel consumption per unit of generation.

We merge these data with hourly plant-level data on the quantity of electricity generated from the Environmental Protection Agency's (EPA's) Continuous Emissions Monitoring System (CEMS) data. The CEMS data also include the state in which the plant is located, which we use to convert all reported hours into Eastern Standard Time. We then collapse the combined EIA/EPA data across generators of the same fuel-technology type within a utility-hour. We keep data from states defined as regulated in Cicala (2022) in the Eastern Interconnection.[7] We limit our data geographically because regulated utilities in the Eastern Interconnection have nearby Independent System Operators (ISOs), and we can use data from these ISOs to define import prices, as we note below.

Our last major data source is the Federal Energy Regulatory Commission (FERC) Form 714 data. These data include hourly load served by each utility as well as a crosswalk that allows us to match these data to the EIA utility identifiers. We then use these data to define hourly imports into the utility as load net of our three primary generation fuel-technology types. Thus, imports will include both renewable (which is relatively moderate during this time-frame) and nuclear generation. We keep only those utilities with at least five years of revenue and load data and drop three utilities with excessive exports or very low capacity.

We supplement these primary data sources with data on wholesale electricity prices and coal and natural gas fuel prices. We use wholesale electricity prices to understand the tradeoffs that regulated utilities face in importing electricity instead of generating it. Even regulated

[7]While we restrict the data used in our structural estimation to regulated states, we provide reduced-form evidence comparing regulated and restructured states.

utilities can purchase electricity from restructured markets—and some markets such as the Midcontinent Independent System Operator (MISO) include regulated states within their boundaries. We assume that each regulated utility faces an import price that is equal to this wholesale electricity price in the nearest restructured state. We collect nodal prices from ISOs and then construct an average hourly wholesale electricity price for each U.S. state and hour. As we describe in Section 5.1, to estimate import supply curves, we pair these wholesale electricity price data with average daily temperature at the state level from PRISM.[8]

Finally, we use EIA Form 423 to construct natural gas and coal fuel prices by utility-year. These data record annual contracted fuel prices at the plant level. We aggregate the information to the state-year level by taking the mean, weighting by annual generation at each generator, to capture the opportunity cost of fuel faced by utilities. Also, we assume that coal fuel prices do not change over time, because there is little variation in these prices in our data and what variation there is may not reflect changes in opportunity cost given that coal is often sold with long-term contracts.

2.3 Summary Statistics

Table 1 presents summary statistics of our analysis data at the utility-year level. The first column presents overall averages and standard deviations while the second and third columns present the values for the first year of our data (2006) and the last year (2017), respectively. We see that coal capacity declines substantially over our analysis sample—from 3.77 TW to 2.86 TW per utility—while CCNG capacity increases from 1.95 TW to 2.97 TW per utility. Average coal fuel prices are \$2.45/MMBtu over our sample, and marginal costs of coal generation are about \$20/MW over our sample (with variation over time coming from changes in the states sampled over time). In contrast, natural gas fuel prices fall from a high of nearly \$8/MMBtu in 2006 to only \$3.12/MMBtu in 2017. This drop in fuel prices caused CCNG marginal costs to fall by 46% over our time period. Finally, our data record information on 26 unique utilities. The average annual revenues of these utilities is approximately \$2 billion

[8]We downloaded these data from Prof. Wolfram Schlenker's website, `http://www.columbia.edu/~ws2162/links.html`.

per year (in January, 2006 dollars), a figure that is consistent across years.

Table 2 presents similar summary statistics for our hourly-level analysis data. Utilities in our data serve an average of a little more than 4 terawatt hours of load per hour. In 2006, the majority (66%) of this load was met by coal and only a small amount (13%) was met by CCNG. By 2017, this situation had changed substantially, with 34% of load met by coal on average and 33% met by CCNG. The remainder of load is generally met with imports,[9] with NGT consistently producing only a small percentage of load. This is consistent with many NGT generators being used as "peakers" that only generate in times with high load. Finally, import prices reflect the overall decrease in natural gas prices, displaying a 43% drop between 2006 and 2017.

3 Model of Electricity Regulation

We lay out the details of our model in four parts. We begin with a broad overview of the regulatory environment and then provide details of the regulator's instruments. We then turn to the utility's long-run investment and retirement decisions and finally its hourly operations decisions.

3.1 Overview

We model a regulator that uses rate-of-return regulation to provide incentives to a natural monopoly electric utility. Following Viscusi et al. (2018), the regulator has two primary objectives: it wants enough generation and imports to meet consumer electricity demand (referred to as load) in every hour, and it wants to keep consumer rates low.[10] The regulator could also potentially be concerned about mitigating environmental harm.

The regulator observes the utility's costs and usage decisions. However, asymmetric information about the costs of alternative decisions may keep it from prescribing the utility's

[9] We allow exports to be represented as negative imports, so some utilities will have more generation from Coal, CCNG, and NGT than total load in particular hours.

[10] Regulated utilities also provide transmission and distribution services, but this is not the focus of our study. See for example Lim and Yurukoglu (2018).

Table 1: Summary Statistics from Data at Utility/Year Level

	Overall	2006	2017
Capacity (TW):			
Coal	3.51	3.77	2.86
	(4.57)	(5.03)	(3.15)
CCNG	1.95	1.07	2.97
	(3.84)	(2.94)	(5.08)
NGT	0.78	0.69	1.12
	(1.14)	(1.07)	(1.43)
Fuel Price ($/MMBtu):			
Coal	2.45	2.02	2.37
	(0.79)	(0.65)	(0.58)
Natural Gas	5.35	7.97	3.12
	(2.27)	(1.02)	(0.42)
Marginal Cost ($/MWh):			
Coal	21.39	18.34	19.19
	(11.75)	(9.02)	(11.80)
CCNG	30.25	38.68	20.83
	(23.66)	(33.65)	(7.96)
NGT	49.82	74.65	57.54
	(57.78)	(52.08)	(154.63)
Utility Revenues (Billions of Dollars):			
	1.98	1.92	2.05
	(2.39)	(2.61)	(2.42)
Number of Unique Utilities:			
	26	25	20

Notes: The first column reports summary statistics over the whole 2006-2017 period. The second and third columns report summary statistics for the years 2006 and 2017, respectively. For all variables we report the sample mean and standard deviation over the relevant period, with the exception of the number of unique utilities. We report standard deviations in parentheses.

Table 2: Summary Statistics from Data at Utility/Hour Level

	Overall	2006	2017
Load Served (TWh):			
	4.16	4.10	4.58
	(5.03)	(5.05)	(5.14)
Production (TWh):			
Coal	2.16	2.72	1.55
	(2.61)	(3.35)	(1.46)
CCNG	1.01	0.53	1.52
	(1.76)	(1.23)	(2.34)
NGT	0.10	0.07	0.21
	(0.24)	(0.20)	(0.36)
Import Quantity (TWh):			
	1.49	1.34	1.84
	(2.63)	(2.60)	(2.76)
Import Price ($/MWh):			
	33.05	40.99	23.25
	(19.33)	(21.12)	(8.24)
Number of Observations:			
	2,476,657	214,955	175,194

Notes: The first column reports summary statistics over the whole 2006-2017 period. The second and third columns report summary statistics for the years 2006 and 2017, respectively. For all variables we report the sample mean and standard deviation over the relevant period. We report standard deviations in parentheses.

optimizing actions (Joskow, 2007). For this reason, we assume that, instead of dictating choices, the regulator imposes a fixed incentive structure that encourages the utility to take actions that meet the regulator's goals. In response, the utility makes two types of decisions, both with the goal of maximizing expected discounted profits. In the long run, it chooses the investment levels of different generation technologies. Given these technologies, it makes hourly operations decisions to meet load.

These interactions face changing technologies and market conditions. Specifically, fuel prices change which generation technologies are most cost effective, technological change creates new generation options, and perceptions of environmental harm change policy priorities. These conditions contribute to the regulator's asymmetric information problem as the regulator must now predict current and future costs of multiple generation technologies. In combination with the long-lived nature of generation capital, these conditions also require the utility to continuously reevaluate its operations, investment, and retirement decisions.

In the model, the regulator uses two instruments to generate appropriate incentives. First, it offers a maximum rate of return, which specifies the profits the utility can earn as a function of its capital stock, also known as the rate base. In order to ensure that the utility invests in and uses capital efficiently, this maximum rate of return is declining in operating costs. Second, the regulator only includes generators that it deems to be "used and useful" in the rate base. This combines the idea that only "prudent" capital investments should be included in the rate base (Viscusi et al., 2018) with the idea that a generator's utilization is one objective way to measure the prudence of an investment (Gilbert and Newbery, 1994).

The two instruments serve distinct roles. Having the maximum rate of return decline with operating costs incentivizes the utility to invest in and use low variable cost technologies. Having only used-and-useful capital enter in the rate base limits the utility's capital investments. Together, these instruments lead the utility to make decisions that further the regulator's objectives.

Nonetheless, these instruments are not perfect, allowing multiple inefficiencies to persist. First, utilities may operate generators out of dispatch order to increase their effective capital. We expect this to be of particular concern for utilities that operate coal generators rendered

economically inefficient by the decrease in the marginal cost of generating electricity with natural gas. Second, utilities in our model will still have an incentive to over-invest in capital (Averch and Johnson, 1962), although this is limited by the used-and-useful standard. Estimating our model will allow us to quantify these distortions.

3.2 Regulatory Instruments

As discussed above, the regulator needs to ensure that the utility meets load, ℓ_{yh}, in each hour, $h = 1, \dots H$, of each year, y. We assume that load varies across hours, but that hourly load is perfectly inelastic. The utility meets load using generation capacity, K_y^f, from different fuel-technologies, f, and imports from electricity interchanges outside of its service area. We let $f \in \{COAL, CCNG, NGT\}$, represent coal, combined-cycle natural gas, and (non combined-cycle) natural gas turbines respectively. We denote the quantity of generation from each fuel-technology q_{yh}^f and import choices q_{yh}^m.

The regulator's first instrument is the maximum rate of return that the utility can earn in each year, $\bar{s}_y$. In order to encourage the utility to use its capital efficiently—operating generators from lowest to highest marginal cost—the regulator sets $\bar{s}_y$ to be a decreasing function of total variable (generation and import) costs, $TVC(\vec{q}_y)$:

$$\bar{s}_y \equiv \left(\frac{TVC(\vec{q}_y)}{CostBasis} \right)^{-\gamma} \quad \text{with} \quad \gamma > 0, \tag{1}$$

where $\vec{q}_y$ is the vector of all generation and import quantities over the year, and $CostBasis$ is a baseline, observable cost metric that varies across utilities.[11]

The regulator's second instrument is the rate base that it determines to be "used and useful," B_y. We define B_y to be a weighted sum of MWs of useful capital across fuel-technologies, expressed in dollar terms:

$$B_y \equiv \alpha \left[\alpha^{CCNG} K_y^{CCNG} + \alpha^{NGT} K_y^{NGT} + \alpha^{COAL} UU \left(\overline{Q}_y^{COAL} / K_y^{COAL} \right) K_y^{COAL} \right], \tag{2}$$

[11] This declining rate of return is observed in other regulated sectors such as natural gas (Hausman, 2019).

where the α^f terms weight capital of different fuel-technologies, α converts MWs of useful capital into dollars of rate base, $\overline{Q}_y^{COAL}/K_y^{COAL}$ is the annual utilization rate of coal, $\overline{Q}_y^{COAL} \equiv \frac{1}{H}\sum_{h=1}^{H} q_{th}^{COAL}$, and $UU\left(\overline{Q}_y^{COAL}/K_y^{COAL}\right)$ is the function determining the share of coal capital that the regulator deems used and useful. Coal's contribution to the rate base, its *effective capital*, is then the product $UU(\cdot) \times K_y^{COAL}$, and is analogous to physical capital for the other fuel-technologies. The α and α^f terms are parameters. To avoid collinearity, we normalize $\alpha^{CCNG} = 1$, which allows us to interpret α as the rate base contribution (in dollars) of one MW of combined-cycle natural gas capacity.

We assume that the used-and-useful standard only binds for coal generators. As shown in Figure 1, falling natural gas prices during our sample period meant that combined-cycle natural gas generators both decreased costs and demonstrated their usefulness for baseload generation. Non-combined-cycle natural gas generators have higher average heat rates than combined-cycle ones, but often serve as "peakers" used primarily in periods of exceptionally high load. For this reason, they would generally not need to prove usefulness via high utilization rates. In contrast, coal generation frequently became uneconomical during our sample period, implying that utilities need to balance lower generation costs against proving that coal was used and useful.

We specify the coal used-and-useful standard with a simple logit functional form of utilization:

$$UU(\overline{Q}_y^{COAL}/K_y^{COAL}) = \frac{\exp\left(\mu_1 + \mu_2 \frac{\overline{Q}^{COAL}}{K^{COAL}}\right)}{1 + \exp\left(\mu_1 + \mu_2 \frac{\overline{Q}^{COAL}}{K^{COAL}}\right)}, \tag{3}$$

where μ_1 and μ_2 are additional parameters. In this formulation, μ_1 determines coal capacity's contribution to the rate base when it is unused, and μ_2 determines how this contribution rises with utilization.

Given the utility's choices, its rate of return is:

$$s_y \equiv (Revenues_y - TVC(\vec{q}_y))/B_y. \tag{4}$$

The regulator sets $Revenues_y$ so that the utility earns its maximum rate of return, i.e. $s_y = \overline{s}_y$. The regulator ensures this rate of return by setting the average consumer electricity price to $Revenues_y$ divided by the total load.

3.3 Investment and Retirement Decisions

The utility facing this regulatory structure must make decisions over capacity investment and retirement by fuel-technology and how to meet load in each hour. Focusing in this subsection on capacity investment and retirement, we assume that these decisions take three years to implement, consistent with the long time horizons necessary to build or decommission fossil fuel generators. We therefore define a period, t, to be a three-year window.

Every period, the utility makes the decision of how much to invest, separately by fuel-technology, x_t^f, and incurs costs based on these decisions. The utility can choose $x_t^f > 0$, which corresponds to investment, or $x_t^f < 0$, which corresponds to retirement. After this decision, every year, the utility decides which fuel-technologies or imports to use to supply electricity for every hour of the year. At the end of the year, the regulator observes utility costs, determines the rate base, and provides the utility with its variable profits given the rate of return in (4).

We treat each of the three fuel-technologies differently, reflecting their characteristics during our sample period. We assume that the utility chooses only (non-negative) investment for combined-cycle natural gas capacity and only whether to retire existing coal capacity. We make these choices since the vast majority of entry decisions are for combined-cycle natural gas capacity, and the vast majority of exit decisions are for coal capacity.[12] Finally, to limit the complexity of our model, we do not endogenize the utility's choice of non-combined-cycle natural gas capacity, which we view as less central to utility regulation.

In each year, the utility makes optimizing operations decisions given its long-run state, Ω, and earns variable profits given these decisions, $\pi^*(\Omega)$. Since Ω needs to include any factor that affects expected current or future profits, it could, in principle, include many variables.

[12]While we observe a few instances of entry of coal generators in the data, the decision to undertake these investments largely occurred before our sample period.

For tractability we restrict the time-varying component of Ω to four variables. In addition to t itself, we include natural gas fuel price, p_t^{NG}, which we assume follows an exogenous AR(1) process. The state Ω also includes the coal and combined-cycle natural gas capacities, both of which vary with the utility's decisions, $K_{t+1}^f = K_t^f + x_t^f$. Beyond the time-varying state, Ω includes a number of fixed states: the utility's heat rates for all fuel-technologies, non-combined-cycle gas capacity, coal fuel price, $CostBasis$, and hourly import supply curves and load.

The utility makes investment decisions to maximize the expected discounted long-run sum of operating profits net of investment costs in an infinite horizon model with an annual discount factor of $\beta = 0.95$. We only allow the utility to make these decisions over the 30 years after 2005, and assume that the long-run state remains fixed after that point since we view predictions after this time horizon as overly uncertain.

Investment costs build on Ryan (2012) and Fowlie et al. (2016) (R/FRR). Each fuel-technology's investment costs include time-invariant fixed, linear, and quadratic terms and a stochastic cost shock:

$$InvCosts^f(x_t^f|\varepsilon_t^f) = \delta_0^f \mathbb{1}\{x_t^f \neq 0\} + x_t^f(\delta_1^f + x_t^f\delta_2^f + \sigma^f\varepsilon_t^f), \tag{5}$$

where $(\delta_0^f, \delta_1^f, \delta_2^f, \sigma^f)$ for $f \in \{COAL, CCNG\}$ are parameters to estimate. Unlike in R/FRR, where the stochastic shock is on the fixed cost of investment, we assume that each period's shocks ε_t^f increase *marginal* investment costs and that they are distributed *i.i.d.* with a standard normal density. Many capacity investment models specify *i.i.d.* shocks for each candidate investment level (e.g., Collard-Wexler, 2013). We instead specify unobservable cost shocks as increasing the marginal cost of investment because a shock to each investment level may yield unrealistic substitution patterns as the number of levels grows. Our approach generates a distribution of capacity changes in any state, which allows us to match the randomness in the data.

Focusing on the timing of the investment decisions, each period the utility first observes the natural gas fuel price shock. It then observes its shock to the coal marginal cost of

retirement, ε_t^{COAL}, and makes its coal retirement decision. Next, it observes the shock to the combined-cycle natural gas marginal cost of investment, ε_t^{CCNG}, and makes its combined-cycle natural gas investment decision. It then makes operations decisions and earns variable profits over the three year period (which are a function of its state at the time of the coal investment decision). At the end of the period, capacity adjusts to reflect investment and retirement decisions.

We can write utility i's coal retirement decision Bellman equation for $t \leq 10$ (during the 30 year decision period) as:

$$V_i^{COAL}(K^{COAL}, K^{CCNG}, p^{NG}, t, \varepsilon^{COAL}) = \pi_i^*(K^{COAL}, K^{CCNG}, p^{NG}) + \max_{x^{COAL} \leq 0} \left\{ -InvCosts^{COAL}(x^{COAL}|\varepsilon^{COAL}) + EV_i^{CCNG}(K^{COAL} + x^{COAL}, K^{CCNG}, p^{NG}, t) \right\}, \quad (6)$$

where we include an index i in π^*—and therefore in the value functions—to account for the effect of the utility's fixed states on profits. Further, EV_i^{CCNG} is the expectation of the value function at the start of the combined-cycle natural gas investment decision, before the ε^{CCNG} investment cost shock is realized. We include variable profits here since they are a function of the state at this stage.

For its gas investment decision, the Bellman equation is:

$$V_i^{CCNG}(K^{COAL'}, K^{CCNG}, p^{NG}, t, \varepsilon^{CCNG}) = \max_{x^{CCNG} \leq 0} \Big\{ -InvCosts^{CCNG}(x^{CCNG}|\varepsilon^{CCNG}) + \beta \int EV_i^{COAL}(K^{COAL'}, K^{CCNG} + x^{CCNG}, p', t+1) dg(p'|p^{NG}) \Big\}, \quad (7)$$

where $K^{COAL'}$ is the coal capacity after the coal retirement decision, $g(p'|p^{NG})$ is the conditional density of the next period's fuel prices, and EV_i^{COAL} is the expectation of the value function at the start of the next period, integrating over the ε^{COAL} investment cost shock.

Equations (6) and (7) show that the utility can adjust its next period's capital deterministically but is faced with a stochastic evolution of fuel prices and cost shocks. For $t > 10$, both Bellman equations look similar to these equations except that the utility does not make investment or retirement decisions and natural gas fuel prices do not evolve.

3.4 Operations Decisions

The investment and retirement decisions depend critically on the annual profits π^* that the utility would earn at any long-run capacity and fuel price state. Within a year, the utility maximizes variable profits with its hourly choices of generation by fuel-technology and imports, $\vec{q_y}$.

The regulator requires the utility to meet load in every hour. In return, it compensates the utility by giving it a rate of return based on its rate base and its total variable costs over the year. These costs are the sum of variable costs across the three fuel-technologies, f, and imports:

$$TVC(\vec{q_y}) = \sum_h \left[\sum_f \left[q^f_{yh} \times (heat^f \times p^f_y + om^f) + \rho^f \times Ramp(q^f_{yh-1}, q^f_{yh}) \right] + \int_0^{q^m_{yh}} S^m_{yh}(q) \right].$$

For each fuel-technology, variable costs sum fuel, operation and maintenance (O&M), and ramping costs. Each fuel-technology has a "heat rate," $heat_f$, and fuel price per MMBtu, p^f_y, which together determine the marginal fuel cost. We model O&M costs, om^f, as constant per MWh of generation, and ramping costs, ρ^f, as constant per MW of generation increase, i.e., $Ramp(q^f_{yh-1}, q^f_{yh}) = q^f_{yh} - q^f_{yh-1}$ in the case where $q^f_{yh} > q^f_{yh-1}$ and zero otherwise.[13] We assume that natural gas turbine plants do not have ramping costs, so $\rho^{NGT} = 0$.

The utility faces an inverse import supply curve $S^m_{yh}(q^m_{yh})$. We assume that it imports energy from various sources with separate contracts and hence it pays different sources different amounts. Hence, and following the literature (Bushnell et al., 2008; Gowrisankaran et al., 2016; Reguant, 2019), its import costs are the integral under the inverse supply curve rather than the maximum import price times quantity imported, $q^m_{yh} \times S^m_{yh}(q^m_{yh})$.

The utility chooses $\vec{q_y}$, to maximize its variable profits given the regulatory structure presented in Section 3.2. The existence of ramping costs creates a dynamic linkage between hours, which implies that we need to consider profit maximization jointly across hours of the year. To simplify the operations decision problem, we assume that the utility observes all hourly loads and import supply curves at the beginning of the year. Given these assumptions,

[13] We do not model ramping costs when $h = 1$.

optimized profits are:

$$\pi_i^*(K^{COAL}, K^{CCNG}, p^{NG}) = \max_{\vec{q}_y} \overbrace{\left(\frac{TVC_i(\vec{q}_y, p^{NG})}{CostBasis_i}\right)^{-\gamma}}^{\text{Max. rate of return}} \overbrace{B(\vec{q}_y, K^{COAL}, K^{CCNG})}^{\text{Rate base}} \tag{8}$$

$$\text{subject to:} \quad \underbrace{\sum_{f=1}^{F} q_h^f + q_h^m = \ell_h \ \forall \ h}_{\text{Generation and imports meet load}} \quad \text{and} \ \underbrace{0 \leq q_h^f \leq K^f \ \forall \ f, h}_{\text{Capacity constraints}},$$

where the i subscript makes explicit which terms depend on the fixed state. As defined in Equation (2), the rate base, $B(\vec{q}_y, K^{COAL}, K^{CCNG})$, is a function of the utility's quantity choices, $\vec{q}_y$, because coal's utilization determines its contribution to the rate base. Equation (8) incorporates the tension the regulatory structure creates for the utility: for a given rate base, the utility wants to minimize TVC, but it also needs to generate with coal in order for coal to be considered used and useful.

Equation (8) presents profit maximization over the year as a constrained optimization problem. However, we conceptualize this problem as a discretized finite-horizon Bellman equation for use in our estimation. In this formulation, the utility receives the regulatory profit in the terminal hour. This profit is a function of the realized costs and coal utilization rate. Because of ramping, realized costs depend upon the hourly sequence of coal and combined-cycle natural gas generation. Thus, in any hour, h, the utility keeps track of four states (in addition to h): the cumulative TVC and coal utilization prior to this hour, and last hour's coal and combined-cycle natural gas generation. These five variables are sufficient for the utility to evaluate the impact of its actions on its state-contingent value starting at hour $h+1$. Having solved the Bellman equations backwards to hour $h = 1$, we then forward simulate—using the calculated state-contingent optimal policies—to recover the optimal actions.

4 Reduced Form Evidence

We now present reduced-form evidence that illustrates the identification of the structural model and motivates its specification, using data from the Eastern Interconnection of the U.S. We first provide support for our assumption that the regulator permits lower rates of return when variable costs relative to the base year are higher and then for our assumption that the effective capital depends on whether the capital is used and useful. Finally, we provide evidence on entry and exit.

4.1 Determinants of the allowable rate of return

In order to understand whether the regulator adjusts the maximum allowable rate of return, $\bar{s}$, as TVC_y/TVC_0 changes, we regress the observed rate of return on measures of TVC_y/TVC_0. We construct total variable generation and import costs using heat rate and fuel price data from EIA and calculating import costs following Bushnell et al. (2008). Ideally, we would directly measure the observed rate of return as equal to the allowable rate of return, so $\bar{s} = (Revenues(\vec{q}_y) - TVC(\vec{q}_y))/K_y^e$, at the utility's profit maximizing choice of $\vec{q}_y$.

The central issue with this approach is how to measure K_y^e and TVC_y. There are at least two important measurement concerns. First, we do not observe effective capital, K_y^e, directly in the data. One substitute for the effective capital is to measure the utility's actual generation capital by summing the MWs of capacity across different fuel-technologies. However, the dollar cost of capital will vary across and within types—because of different capital costs and depreciation—implying that this does not appropriately capture the value of a utility's capital. In addition, this measure does not incorporate the fact that the regulator may decrease effective capital if it does not deem some generators to be used and useful.

Second, we also only observe the variable generation and import costs, and do not observe other variable costs such as those for transmission and distribution. This is problematic because utilities may have underlying differences in these variable costs. For instance, utilities that serve a large geographic area may have higher marginal transmission costs. These extra

variable costs are an omitted variable that varies across utilities and may lead to biased estimates of the effect of TVC_y/TVC_0 on the maximum allowable rate of return.

Thus, to motivate our assumption that the regulator permits lower rates of return when TVC_y/TVC_0 is high, we construct measures of s_y using revenues minus TVC_y divided by two different proxies for effective capital K^e: (1) total period load and (2) total MWs of capacity. The structural model directly specifies the determinants of the rate base, so these measures serve to provide evidence on its underlying identification. Further, we take two approaches to controlling for the omitted variable bias. We run the regressions with utility fixed effects—which are difficult to incorporate into a structural model—and with both the dependent and independent variables differenced from their average in an initial set of years. The fracking boom provides substantial within-utility variation in costs that will help identify these specifications. We also report results without fixed effects or differencing.

Table 3 presents the results of these 6 regressions. For ROR defined with load, the coefficient on $TVC_y/load$ in the regression without fixed effects or differencing is not statistically significant. Once we add utility fixed effects, TVC_y/K^e_y has a negative association with ROR that is highly statistically significant, using both definitions. The results that difference total variable costs from the first year of the data show similar patterns to the results with fixed effects, and are more directly tied to the identification of our structural estimates.

4.2 Determinants of the effective capital

Beyond the determinants of $\bar{s}$, we also motivate our assumption that generators must be used and useful to be included in the measure of effective capital, K^e_y. In principle, we could observe these incentives by evaluating whether utilities operate generators in hours when operation would be uneconomical in the absence of regulatory constraints, and in particular for generators where operation would be pivotal in convincing the regulator that the generator is used and useful, such as coal generators in the post-fracking era.

Because of complex issues in generator scheduling, such as ramping costs and operations and maintenance costs, we identify the used-and-useful effect by comparing the scheduling of

Table 3: Regressions of Rate of Return on Total Variable Cost per Effective Capital

	Dependent Variable: Variable Profits per MW of Capacity					
Variable Costs per Capacity (1e3$/MW)	−89.7	−360.1***				
	(94.5)	(59.3)				
Variable Costs per High Load (1e6$/MWh)			0.057	−0.462***		
			(0.127)	(0.059)		
Variable Costs (1e6$)					−0.017***	−0.026***
					(0.005)	(0.007)
Utility FE	N	Y	N	Y	N	Y

Note: Variable costs include fuel and import costs but exclude O&M and ramping costs. Variable profits are revenues net of these variable costs. High load is the 95th percentile of hourly load for the utility-year. *** indicates significance at the 1% level.

generators in regulated markets to restructured markets. This allows us to tie this potential out-of-dispatch-order generation to the presence of regulation. Thus, we examine the likelihood of coal generators operating when wholesale electricity prices are low relative to the generator's marginal fuel costs, and compare this between regulated and restructured markets. We also examine whether this likelihood varies with generator usage in the previous year.

We therefore run regressions at the generator-hour level of whether a generator is operating on an indicator variable for the generator's marginal cost being above wholesale price and this indicator interacted with whether the generator is in a restructured state, controlling for state and year fixed effects.[14] Because there are no wholesale markets in regulated states, we recover the effective wholesale price using the closest market price from a restructured state, which proxies for the price the utility pays for imports. Since our structural model aggregates generators by fuel-technologies, we also run our regressions at this level.

Table 4 shows the results of these regressions for both coal and combined-cycle natural gas generators. The first column shows that coal in regulated states is 3.2 percentage points

[14]Regulatory status is completely determined by state during our sample years, so we do not include a separate control for regulatory status.

less likely to operate when its marginal fuel costs exceed the wholesale (import) price, although this difference is not statistically significant. This effect is 12.2 percentage points larger in restructured markets. The second column shows that the differential usage between regulated and restructured states is not present for combined-cycle natural gas generators, with both types of states operating substantially less when fuel costs are above wholesale prices. This suggests that coal in regulated states is less responsive to price signals than coal in restructured states, but that this same pattern does not exist for CCNG. This could be explained by regulated utilities wanting to operate coal generators in order to prove that they are used and useful.

Figure 2 repeats the regressions from Table 4 individually at the state level. Panel (a) shows the usage when fuel costs are above wholesale prices for coal (vertical axis) and CCNG (horizontal axis). Green states are regulated and red states are restructured. While there is little pattern based on regulatory status for CCNG usage when fuel costs are above wholesale prices, the change in usage when coal fuel costs exceed wholesale prices is clearly related to regulatory status. The seven states with the lowest differential coal utilization—which are at the bottom of the graph—are all regulated states. This reinforces the idea that regulatory status dramatically impacts utilization for expensive, older technologies.

Panel (b) of Figure 2 again plots the response of coal to low wholesale prices, but plots this against the share of generation owned by electric utilities in the state. Regulated states generally have utility ownership shares over 60%, whereas all restructured states other than Michigan have utility ownership shares under 30%. The best fit line through this relationship shows that coal's responsiveness to low wholesale prices does correlate strongly with utility ownership share.

5 Estimation Approach

We estimate the model developed in Section 3 in three parts. First, we estimate each utility's hourly import supply curve using generation, load, weather, and price data. We then combine these same data with the estimated import supply curves and data on utilities' revenues to

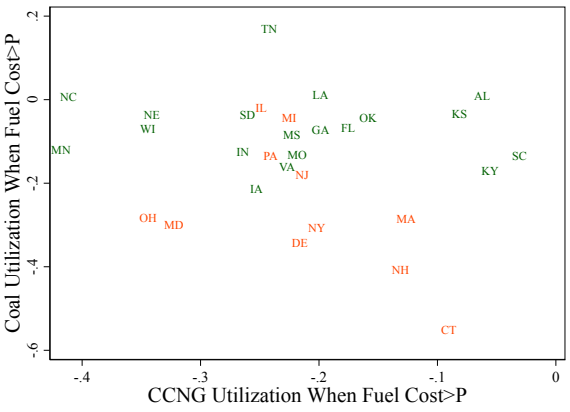
Table 4: Generator Use and Used-and-Useful Standards

	$\mathbb{1}\{$Fuel-Technology Operating$\}$	
	Coal	Combined Cycle Natural Gas
$\mathbb{1}\{$MC > price$\}$	-0.032	-0.203^{***}
	(0.031)	(0.030)
$\mathbb{1}\{$MC > price$\}\times$ Restructured	-0.122^{**}	0.004
	(0.050)	(0.029)
R^2	0.089	0.132
N	19,782,462	20,723,458

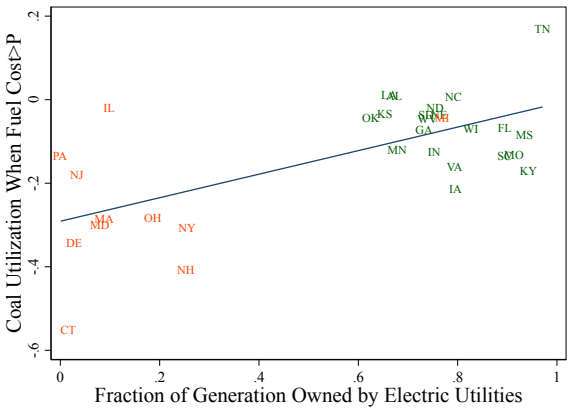
Note: Regressions are linear probability models that include state and year fixed effects. Standard errors are clustered at the state and year level. *** indicates significance at the 1% level and ** indicates significance at the 5% level.

Figure 2: Utilization When $FuelCosts > P$ in Regulated Versus Restructured Markets

(a) Usage Out of Dispatch Order for Coal and CCNG



(b) Coal Out of Dispatch Order Usage vs. Utility Ownership Share



Note: Panel (a) shows results from a regression of whether a fuel is generating on whether the fuel costs are above the import price, by state. Green states are regulated and red states are restructured. Panel (b) shows the same regression coefficients for coal generation by state, plotted against the fraction of generation in the state owned by electric utilities.

estimate the parameters governing the regulatory incentive structure and utilities' costs of operation within a year. Finally, we combine investment and retirement data with profits across states as predicted by our estimated operations parameters to recover the distribution of investment and retirement costs for combined-cycle natural gas and coal respectively. We use full solution approaches to estimate both the operations and investment/retirement parameters, which allows us to evaluate the impact of counterfactual regulatory structures with the same techniques that we use to estimate the structural parameters.

Broadly, identification of the model follows from the intuition that the observed sharp decline in natural gas fuel prices had different implications across utilities, depending upon the utilities' capital mixes. For instance, consider a utility with substantial coal and combined-cycle natural gas capacity. Early in our sample, when natural gas prices were high, this utility would have generated with coal first, and only used natural gas when electricity demand was high. After natural gas prices fell, the utility faced conflicting incentives: it wanted to run natural gas generators to keep fuel costs low and be allowed higher $\overline{s}$, but it also wanted to use—and not retire—coal generators to increase its effective capital. This contrasts with a utility with predominantly coal capacity that needed to meet load with coal even after natural gas prices fell. These differences help identify parameters that determine both operations and investment and retirement decisions.

We discuss our estimation approach and the specific identification arguments for the import, operations, and investment/retirement parameters in turn.

5.1 Import Supply Curve

Each hour, a utility u chooses the share of load to meet with its own generation and the share to import from facilities it does not own. To understand these decisions, we follow Bushnell et al. (2008), Gowrisankaran et al. (2016), and Reguant (2019) and estimate a linear import supply curve that models the quantity of electricity imported to the utility as a function of import price and controls:

$$q^m_{uyh} = (\psi_{u0} + \psi_{u1} p^{NG}) p^m_{yh} + \psi_{u2} p^{NG} + \psi_{u3} X_{uyh} + \varepsilon^m_{uyh}. \tag{9}$$

We allow all parameters to vary across utilities and allow the intercept and slope of the import supply curve to vary with the natural gas fuel price, because the supply curves in exporting regions will change as fuel prices change.[15] The controls, X_{uyh}, capture demand shocks in the exporting region, and include cooling degree days, heating degree days, and their squares for every state in the nearest ISO, interacted with hour of the day. We also include fixed effects for the day of week, month of year, and hour of day.

Recovering the import supply curves requires understanding the causal impact of import price on import quantity, but an OLS regression of (9) would not consistently estimate the supply curve because the data reflect variation in both demand and supply. Therefore, we identify the import supply curve using instruments that plausibly shift the demand for imports without affecting the import supply curve. Specifically, as in the literature discussed above, we instrument for the import price with the utility's local load, after controlling for X_{uyh}.[16] In many contexts, demand shifters are used as instruments for price in supply estimation. In electricity markets, since local load is nearly perfectly inelastic, load itself instruments for price in supply curve estimation. This instrument is valid if, in addition to local load being perfectly inelastic, it is unaffected by local supply shocks and uncorrelated with shocks to demand (conditional on X_{uyh}) in exporting regions.

We use the estimates from (9) to recover a supply curve for each utility-year-hour. We follow the above papers and specify intercepts of these curves as including the residual from (9), i.e. $\hat{\psi}_{u2}p^{NG} + \hat{\psi}_{u3}X_{uyh} + \hat{\varepsilon}^{m}_{uyh}$ where the hats indicate estimated values.

5.2 Operations Parameters

We use utilities' hourly generation choices within a year to estimate the structural parameters that govern the regulator's maximum rate of return, how capital is deemed used and useful, and utilities' generation costs. The regulator's maximum rate of return is governed by γ—the penalty for high variable costs, α—the conversion from MW of CCNG capital to dollars

[15] As discussed in Section 2.2, we approximate the import price with the wholesale market price at the closest wholesale electricity node and define import quantity as the difference between load and generation.

[16] Given that we interact import price with natural gas fuel price, we also use the interaction the natural gas price with local load as an additional instrument.

of rate base, and α^{NGT} and α^{COAL}—these fuel-technologies' capacity contributions to the rate base relative to CCNG. The terms governing used and useful are the logit intercept and return to coal utilization, μ_1 and μ_2. Finally, generation costs depend on ramping cost parameters—ρ^{COAL} and ρ^{CCNG} and O&M parameters—om^{COAL}, om^{CCNG}, and om^{NGT}. We do not need to estimate parameters for fuel costs, because we recover them by multiplying fuel price by heat rate, both of which are observed in the data.

Each utility in each year makes hourly operations decisions in a finite-horizon dynamic model as discussed in Section 3.4. This optimization results in hourly quantities of imports and generation by fuel-technology, annual variable costs, and annual coal utilization. Conditional on the vector of parameters, these outcomes then determine the utility's annual profits.

We estimate the structural parameters via a nested fixed-point indirect inference approach (Gourieroux et al., 1993). This involves a non-linear search over the structural parameters discussed above in order to find the parameters that most closely match coefficients from regressions run on data simulated from the model to those run on the actual data. For each structural parameter vector and utility-year, we solve for the state-contingent optimal policies and use these to recover optimizing quantities, costs, and coal utilization. We discuss our approach to the utility's optimal operations policies before turning to our indirect inference approach.

We solve for these policies using backward induction. At the terminal hour of the year, the utility earns its profits, which depend upon the major time-invariant observable characteristics of variable costs, $CostBasis$.[17] We define $CostBasis$ as fuel costs and a measure of import costs in the first year the utility appears in the data, because the other components of TVC—including O&M and ramping costs—are not directly observable in the data.

In earlier hours, the utility knows its current state and future hourly load. For each hour, we solve for optimal policies across states, which include the previous hour's coal and CCNG generation, cumulative average variable costs, and coal utilization up to this point.

[17]Empirically, we include $CostBasis$ to capture differences in fixed characteristics across utilities such as size that will influence the regulator's perception of costs.

Understanding that its choices in each hour will affect state transitions, the utility trades off higher coal utilization against lower variable costs, both of which will increase its profits in the terminal hour. Ramping costs encourage the utility to limit large hourly adjustments to coal and CCNG generation unless there are substantial profit gains from these adjustments.

Once we have solved for optimal state-contingent policies, we use a forward simulation process to recover optimal decisions and resulting outcomes over the year. Specifically, we start in the first hour and record the optimal generation choices. We then use these choices to update the state for the following hour, which in turn allows us to record the state-contingent optimal actions for that hour. Iterating through the year, we then obtain simulated hourly and annual data that we use in our indirect inference regressions. Online Appendix A3.1 provides more detail on the computation of the operations decisions.

This process results in simulated data on which we run a set of indirect inference regressions. Our estimation approach matches the coefficients from these regressions to coefficients produced by the same regressions run on real data. Indirect inference is a form of generalized method of moments, where we specify correlations in the data we would most like to match. We chose indirect inference regressions to capture key correlations in the data that are similar to those presented in Section 4. Unlike like reduced form evidence, indirect inference regressions do not require a causal interpretation of the regression coefficients, but rather capture equilibrium relationships from the data that we would like to hold in the model. We therefore choose indirect inference regressions that we believe best reflect important equilibrium features of the data.

The indirect inference approach offers particular advantages in our case. Our model involves regulatory objects—notably the rate base, B_y, maximum rate-of-return on capital, $\overline{s}$, and effective coal capital, $UU(\cdot) \times K_y^{COAL}$—that are not observable in the data. Unlike these terms, revenues, fuel costs, and import quantities and prices—key components that influence profits—are observable in the data. Our indirect inference regressions help us to identify the structural parameters by matching conditional correlations between these key components of profits on outcomes, and also by matching other conditional correlations resulting from generation choices, which reflect marginal incentives. While identification of the structural

parameters comes from the combination of all of the indirect inference regressions, we choose particular regressions to aid in identifying particular parameters. For instance, we primarily identify coal utilization's contribution to effective capital with two types of regressions: (1) the relationship between utilities' choices of coal utilization and realized profits and (2) how coal generation responds to its marginal impact on effective capital. Fundamentally, much of the variation that identifies these parameters will stem from the sharp decline in natural gas fuel prices, as discussed in the introduction of Section 5.

To capture these correlations, we run indirect inference regressions at both the utility-hour and utility-year level. We summarize each of these sets of regressions here and include a more complete discussion in On-line Appendix A3.2.

The hourly-level regressions all capture moments of the hourly generation decisions. We include regressions of hourly utilization by fuel-technology on a constant to match the scale of generation of each fuel. Because utilities have an incentive to reduce costs, these scales are particularly useful for identifying O&M costs. We also include regressions of current generation on lagged generation, controlling for current and future predictors of demand, separately for coal and CCNG. These regressions help identify ramping costs, because the higher the ramping costs the less the utility will change generation from hour to hour.

We also include an hourly regression of the log of the share of total generation from coal and CCNG that is met by coal on quintiles of annual coal utilization,[18] the coal fuel price minus the natural gas fuel price, their interactions, and utility fixed effects. We run an analogous regression for CCNG. These regressions help us to understand the utility's incentive to run coal out of dispatch order, which is captured by the used and useful parameters in the model. We would expect that to the extent that used and useful incentives bind, coefficients on coal quintiles—unlike for CCNG—should exhibit an inverse U-shape, with the marginal incentive to use coal in an hour being highest when the return to coal utilization via the used-and-useful incentive is the highest. While this relationship may reflect the level of used-and-useful incentives for coal, it may also reflect the fact that the use of a fuel in a given hour reflects the fuel's overall value. We therefore limit our regressions to hours where total

[18] We define the quintiles of utilization across all utility-years where the utility has positive coal capacity.

load is between 75% and 125% of total CCNG capacity, to isolate hours where the incentives to increase coal's contribution to the rate base are most likely to affect generation.

Using data at the utility-year level, we regress a measure of profits based on key observable components—average level of revenues net of the sum over hours of fuel costs and import price times import quantity—on a constant. We further regress this same measure on capacity by fuel-technology and coal capacity interacted with coal utilization. These regressions help recover the conversion between MW of CCNG capacity and dollars of revenue, α, as well as coal and NGT's relative contributions to the rate base, α^{COAL} and α^{NGT}.[19] These regressions also combine with the hourly regressions to help identify the used and useful parameters. Finally, we regress a measure of the rate of return on fuel and import costs and utility fixed effects to help identify γ, which indicates how the maximum allowable rate of return responds to changes in TVC.

Note that there are two sets of parameters that are jointly identified. Both α and γ determine how generation capital translates into dollars of allowable return. Similarly, α^{COAL}, μ_1, and μ_2 combine to translate coal utilization into the rate base. For both of these sets, we identify them as a set with sets of indirect inference regressions, and adjust them accordingly in counterfactuals.

5.3 Investment and Retirement Cost Parameters

As we explained in Section 3.3, the utility makes CCNG investment and coal retirement decisions to maximize its expected discounted sum of operating profits net of investment and retirement costs, with Bellman equations (6) and (7). We now describe our estimation of the parameters of the investment and retirement costs, which include both time-invariant terms and the standard deviations of the unobservable components of costs.

We estimate these parameters with a GMM nested fixed-point estimator, similar to the operations parameters.[20] We use moments that capture the difference between the data and

[19] Recall from Section 3 that we normalize $\alpha^{CCNG} = 1$ so that α is the contribution of CCNG.

[20] An alternative approach would be to use a conditional choice probability (CCP) estimator (Hotz and Miller, 1993). In our context, the nested fixed-point approach is computationally very rapid, reducing the potential benefits of a CCP approach.

the model regarding functions of investment or retirement, for CCNG and coal respectively. Specifically, we include the investment/retirement amount and its square conditional on a non-zero amount, and indicators for non-zero investment/retirement and whether they exceed certain thresholds. We also interact each of these terms with the utility's starting capital. Finally, we include standard deviations of the investment/retirement amounts over the analysis data.

We estimate the structural parameters with a search over candidate parameters values that minimize the moment condition. For each candidate parameter vector, we solve the investment/retirement dynamic programming problem and find the value of the moments. Our moment function uses an asymptotically efficient weighting matrix, which we construct by bootstrapping the data to solve for the variance of the moments and then taking the inverse of the variance.

We now describe three important features of our implementation. First, we solve the dynamic programming problem by backward induction with a discretized state space. While the utility earns profits over an infinite horizon, the lack of decisions for $t > 10$ allows us to solve backwards for optimal decisions starting at $t = 10$. For each utility, we choose a grid of 10 dimensions for each of the three time varying states: coal capacity, CCNG capacity, and natural gas fuel prices. Given that there are ten decision periods and investment and retirement decisions in each period, we solve for continuation values at 20,000 states per utility.

Second, in an initial step, we calculate each utility's operating profits—which are key inputs in the investment and retirement Bellman equation— across states. We implement this by solving profits across the 1,000 state grid of coal capacity, CCNG capacity, and natural gas fuel prices that enters our investment and retirement Bellman, using the estimated operations parameters.[21] In counterfactuals such as those that evaluate a social planner or cost minimizing solution, we re-calculate this profit grid, changing the operations objective function, and then use these counterfactual profit grids together with counterfactual investment

[21] Because hourly load and import supply curves vary across years for each utility, we calculate profits using their values from the first year the utility is in the sample.

and retirement objective functions to simulate outcomes.

Finally, we solve the value function iteration and calculate the distributions of investment and retirement decisions using recent advances in the literature on continuous choice estimation (Gowrisankaran and Schmidt-Dengler, 2023). Following their method, we define a finite grid of investment/retirement levels and then find the cutoff ε values between the different levels. A complication that Gowrisankaran and Schmidt-Dengler address is that some levels may never be chosen. The algorithm provides a computationally quick way of ruling out the levels that are never chosen and of then solving for the probability of each level based on the cutoffs between the remaining levels. We use these probabilities to solve for the continuation values in the Bellman recursion.[22] Because we solve for the probability of each level rather than simulating a distribution of levels, the moment condition is continuous in the structural parameter vector, which reduces the implementation costs of a nested fixed-point approach.

In our case, we employ a finite set of 10 investment/retirement levels. For a given parameter vector, the computed model provides the probability that the firm would choose each grid point for any investment/retirement decision in the data.[23] This probability then enters into the moment condition for this parameter vector. We estimate our parameters with GMM instead of maximum likelihood because, for many parameters, we found a zero model probability for some investment/retirement levels, yielding a zero likelihood. We experimented with increasing the number of investment/retirement levels in our GMM estimator but found that our results were not sensitive to this change.

Identification of the investment and retirement cost parameters comes from the extent to which utilities choose to retire coal or invest in CCNG given differences in expected profits across these states. For instance, the extent to which utilities choose to delay entry of CCNG—even when profits are potentially higher with additional natural gas capacity—identifies the average investment costs for CCNG. The extent to which there is heterogeneity in utilities' investment and retirement decisions given similar differences in expected profits conditional on an action identifies the standard deviation of the investment and retirement

[22]The additional difficulty in calculating continuation values, which this algorithm addresses, is that they include the expected value of ε given the choice.

[23]We find the probability for each decision by interpolating the computed policy function.

cost shocks. As with the operations parameters, declines in natural gas fuel prices together with heterogeneity across utilities in their capital mixes provide variation in profit differences across states that identify the investment/retirement parameters.

6 Results and Counterfactuals

This section presents our estimation results and counterfactuals. Section 6.1 presents our operations estimates and then our investment/retirement estimates. Appendix A4 discusses our import supply curve estimates. Section 6.2 discusses our counterfactuals. We focus first on short-run counterfactuals that evaluate the impact of alternate regulatory policies on operations decisions holding utility capacity constant. We then examine counterfactuals of how utility capacity would adjust over a 30-year period following an energy transition where natural gas prices fall.

6.1 Results

Table 5 presents estimates and standard errors for the operations parameters.[24] Focusing first on the parameters underlying how coal usage affects the rate base, we find that μ_2, the slope term on usage in the function that determines coal's contribution to effective capital, is positive and statistically significant. In other words, coal capacity affects the rate base more the more electricity it generates. Figure A1 in On-line Appendix A1 displays the impact of coal capacity on the rate base by usage level relative to the impact of CCNG capacity. An (out-of-sample) coal unit that is never used would increase the rate base by 24% as much as a CCNG unit with the same capacity while a coal unit that is always generating at full capacity would increase the rate base by 95% as much. We find that a unit of natural gas turbine capacity would affect the rate base 16% more than would a unit of CCNG capacity. Thus, with full coal utilization, all three types of capacity affect the rate base similarly.

[24]Table A1 in the On-line Appendix A1 displays how simulations of our operations model compare to observed data. Overall, we find that the model replicates patterns in the data well, although it slightly overpredicts coal usage and underpredicts CCNG usage.

Table 5: Coefficient Estimates for Operations Model

Parameter	Notation	Estimate	Std. Error
Penalty for High TVC_y	γ	0.817	(1.53)
Rate Base per MW of Effective Capital (Millions $)	α	0.193	(0.10)
Coal Capacity Contribution to Effective Capital	α^{COAL}	0.968	(0.20)
Coal Utilization Logit Base	μ_1	-1.092	(1.98)
Coal Utilization Logit Slope	μ_2	5.650	(0.70)
NGT Contribution to Effective Capital	α^{NGT}	1.157	(0.14)
Ramping Cost for Coal ($ / MW)	ρ^{COAL}	30.800	(62.609)
Ramping Cost for CCNG ($ / MW)	ρ^{NGCC}	6.764	(48.677)
O&M Cost for Coal ($ / MW)	om^{COAL}	11.953	(2.19)
O&M Cost for CCNG ($ / MW)	om^{NGCC}	1.518	(5.59)
O&M Cost for NGT ($ / MW)	om^{NGT}	25.079	(178.73)

Note:

We next consider how much a change in a utility's capacity would affect variable profits, which is primarily determined by both α and γ.[25] A 500 MW change in effective capital, which is roughly the mean CCNG generator capacity, would change variable profits by 6.7%, on average across our sample. Unlike effective capital, TVC has a nonlinear impact on variable profits. We find that, across sample observations, a 10% increase in TVC would decrease variable profits by 4%, while a 10% decrease in TVC would increase variable profits by 4.6%.

Turning to other operations costs, a 100 MW coal ramp in one hour—which corresponds to increasing output by 15% for a coal generator with mean capacity—would cost the utility $3,080 while the figure is lower for CCNG ramps at $676. The coal number is lower than Borrero et al. (2023), but is for ramping across all generators in a utility rather than for a specific generator as in that paper. We estimate O&M costs of $16.30/MWh for coal capacity. This figure is similar to Linn and McCormack (2019) and Borrero et al. (2023). Our O&M costs for CCNG and NGT are $2.50/MWh and $19.42/MWh respectively, though neither term is statistically significant.

Table 6 presents our investment/retirement parameter estimates. We find that the fixed

[25]Recall that we do not have data on the rate base and hence α is not separately identified from γ. Because ramping and O&M costs affect TVC, they also enter into the rate of return and affect variable profits.

Table 6: Coefficient Estimates for Investment/Retirement Decisions

Parameter	Notation	Value	Std. Dev.
Fixed cost of coal retirement × 1e2	δ_0^{COAL}	-0.468	(7.54)
Linear coal cost per MW	δ_1^{COAL}	2.155	(3.13)
Quadratic coal cost per MW / 1e3	δ_2^{COAL}	0.137	(0.05)
Coal shock standard deviation per MW	σ^{COAL}	0.469	(0.42)
Fixed cost of NGCC investment × 1e2	δ_0^{NGCC}	-1.006	(0.25)
Linear NGCC cost per MW	δ_1^{NGCC}	8.712	(5.72)
Quadratic NGCC cost per MW / 1e3	δ_2^{NGCC}	0.651	(0.09)
NGCC shock standard deviation per MW	σ^{NGCC}	3.644	(5.61)

Note: All values in millions of 2006 dollars.

costs of adjusting coal or gas capacity is small. Not accounting for the standard deviation in costs, coal retirement earns a scrap value while gas investment is costly, both for small amounts. The costs of both coal retirement and gas investment are convex, with positive quadratic terms.

Evaluating magnitudes, a 250 MW coal retirement yields $577 million in scrap value with the mean cost shock. In contrast, a 250 MW CCNG investment costs $2.3 billion with the mean cost shock. Since these estimates are based on revealed preferences, they include substantially more than the capital costs of adjusting capacity. They also include permitting costs, the costs of the PUC approval processes, and any additional regulatory costs (or avoided regulatory costs in the case of coal capacity retirement). The coal shock has a standard deviation of $469,000 per MW.

6.2 Counterfactuals

We conduct a number of key counterfactuals for both our operations model and our investment/retirement model. First, we seek to understand how rate-or-return regulation affects outcomes relative to efficiency by comparing the baseline model to outcomes with a social planner who minimizes costs inclusive of carbon externalities, and a cost-minimizing social

planner who minimizes generation costs, but does not consider carbon externalities.[26] We then investigate how changing the extent of different regulatory incentives (i.e. used and useful and the penalty for high TVC) while maintaining the same rate-of-return structure affects key outcomes. On-line Appendix A5 provides the details of how we implement these counterfactuals.

Table 7 shows how our first set of counterfactuals affect utilities' operations decisions during our analysis period. We see that the social planner reduces coal generation by 95% (from 62% of capacity usage to 3% usage) because of its combination of high costs and substantial carbon externalities relative to CCNG. Both CCNG generation and imports respond to meet load.[27] This adjustment decreases carbon externalities by 40% but increases variable production costs by 19% and decreases utilities average variable profits by 73%. Because variable profits need to cover costs such as transmission and debt service that are not included in our model, variable profit declines of this magnitude are likely not feasible for the regulator to keep the utility in business and keep the lights on.

The final column of Table 7 shows what happens when we double the slope of the usage incentive, μ_2. We see that in both cost minimization (when there is no usage incentive) and with twice as large a usage incentive, coal usage declines. This is because when the usage incentive is steeper the utility can ensure a higher contribution of coal capital to the rate base with lower usage, whereas when there is no usage incentive then the utility tends not to run high-cost coal. In this way, the observed usage incentive substantially distorts coal usage incentives.

Table 8 presents results of counterfactual operations model outcomes where we change the penalty for high TVC, γ. In these counterfactuals, we change γ, but simultaneously change α, the conversion of MWs of capacity into dollars of rate base, to keep average utility revenues constant at the observed operations choices. We see that halving the TVC penalty increases coal usage by 15%, while doubling it decreases coal usage by 15%. As in the other counterfactuals, these changes in coal generation are offset by changes in CCNG generation

[26] We assume a social cost of carbon of \$190/ton.

[27] We assume that electricity imported from restructured markets has the average carbon of the entire U.S.

Table 7: Operations Counterfactuals: Coal Utilization Incentives

	Baseline	Social Planner	Cost Min. / No U&U	Double U&U
Annual Electricity Production (TWh):				
Coal	19.42	1.01	11.41	15.58
CCNG	4.23	8.64	6.30	5.33
Imports	11.18	25.00	17.02	13.91
Mean Utilization Share (%):				
Coal	61.81	3.10	32.31	48.97
CCNG	23.10	49.99	37.22	30.12
Annual Costs (Millions of Dollars):				
Coal Fuel Costs	477.65	28.19	265.56	378.84
CCNG Fuel Costs	111.19	248.87	165.47	139.16
NGT Fuel Costs	147.77	159.77	151.14	148.45
Coal O&M Costs	231.90	12.00	136.24	186.06
CCNG O&M Costs	6.53	13.34	9.73	8.22
NGT O&M Costs	40.99	45.11	43.19	41.42
Coal Ramping Costs	8.19	2.07	7.74	8.90
CCNG Ramping Costs	1.15	0.88	1.00	1.14
Imports	215.68	972.04	375.16	283.73
Total Variable Production Costs	1,241.06	1,482.26	1,155.22	1,195.93
Carbon Costs	5,054.48	3,007.64	4,157.87	4,623.75
Variable Profits (Millions of Dollars):	1,251.70	336.98	891.90	1,341.31
Electricity Price (Dollars/MWh):	36.27	41.59	32.54	34.46

Note:

Table 8: Counterfactual Simulations of Operations Decisions

	Baseline	Half TVC Penalty, γ	Double TVC Penalty, γ
Annual Electricity Production (TWh):			
Coal	19.42	22.31	16.51
CCNG	4.23	3.41	5.15
Imports	11.18	9.11	13.17
Mean Utilization Share (%):			
Coal	61.81	72.21	51.21
CCNG	23.10	18.13	28.67
Annual Costs (Millions of Dollars):			
Coal Fuel Costs	477.65	551.94	401.65
CCNG Fuel Costs	111.19	92.74	133.83
NGT Fuel Costs	147.77	147.65	148.28
Coal O&M Costs	231.90	266.37	197.05
CCNG O&M Costs	6.53	5.27	7.95
NGT O&M Costs	40.99	40.91	41.33
Coal Ramping Costs	8.19	7.26	8.79
CCNG Ramping Costs	1.15	1.09	1.14
Imports	215.68	171.69	271.92
Total Variable Production Costs	1,241.06	1,284.93	1,211.93
Carbon Costs	5,054.48	5,381.63	4,723.26
Variable Profits (Millions of Dollars):	1,251.70	1,227.99	1,498.08
Electricity Price (Dollars/MWh):	36.27	38.04	34.98

Note:

and imports.

In order to understand counterfactual regulatory structures affect energy transitions, we simulate counterfactual investment and retirement outcomes over 30 years starting with each utility's capacity in 2006 but with the average natural gas price from the 2018-2020 period. This approach allows us to see how the average utility will adjust to new market conditions under different regulatory regiemes. We present these couterfactual result in figures in order to show the full transition path for each model.

We begin by analyzing how the social planner (who internalizes carbon externalities) and the cost minimizing utility would change their coal retirement and CCNG investment decisions relative to the baseline regulated utility. Figure 3 shows that the social planner retires coal substantially faster than the regulated utility, with coal capacity 96.7% lower after 30 years. Interestingly, the cost minimizing utility also substantially reduces coal capacity, by 78.8% relative to the regulated utility. Neither the social planner nor the cost minimizing utility face the Averch-Johnson incentives to over-invest in capital, so they invest substantially less than the regulated utility in new CCNG capacity, with the social planner investing in 45.7% less CCNG capacity and the cost minimizing utility investing in 62.8% less. The social planner invests in more CCNG in order to meet load predominantly with CCNG rather than with coal and imports. This result is clear from Figure 4, which shows that the social planner switches away from coal generation immediately, while the cost minimizing utility reduces coal generation immediately, but then continues to reduce coal generation as more natural gas capacity is added to the capital stock.

Pollution outcomes:

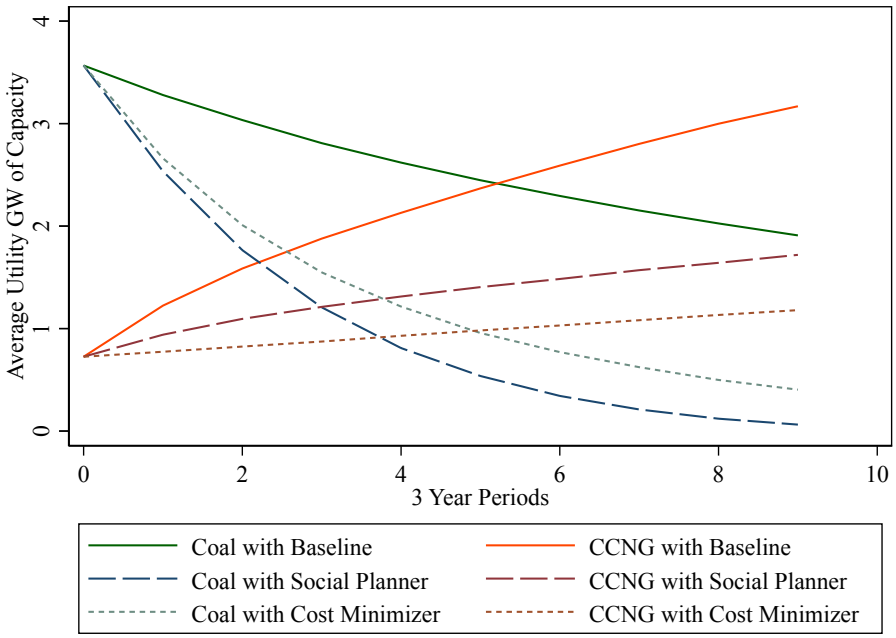
- Figure 5 shows that CO2 costs follow generation, with the social planner decreasing CO2 costs immediately and the cost minimizer decreasing them more over time.

Changing usage incentive:

- Figure 6 shows that changing usage incentives substantially affects coal retirement decisions, but does not have a substantial effect on CCNG investment.

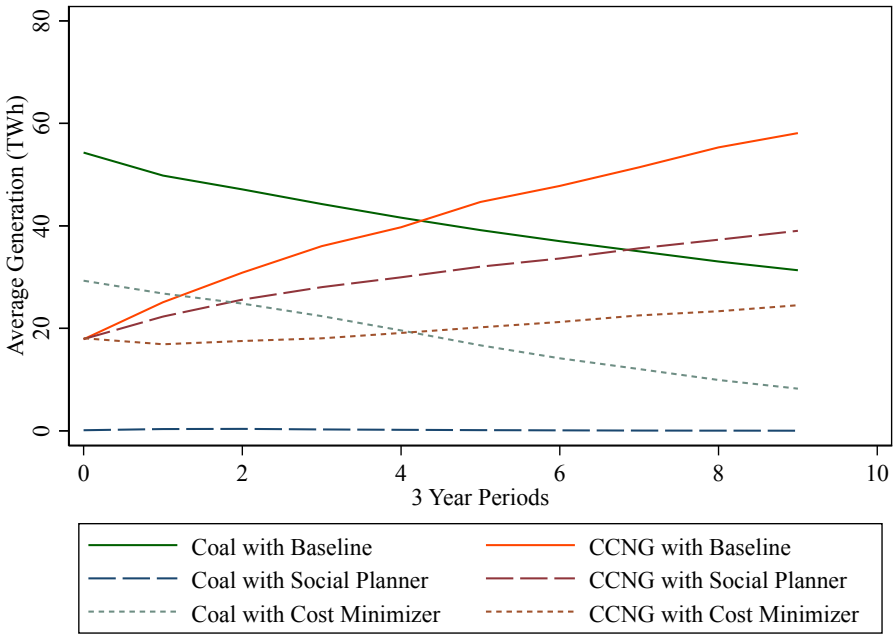
Changing the penalty for high total variable costs:

Figure 3: Capacity Changes: Baseline, Social Planner, Cost Minimizer



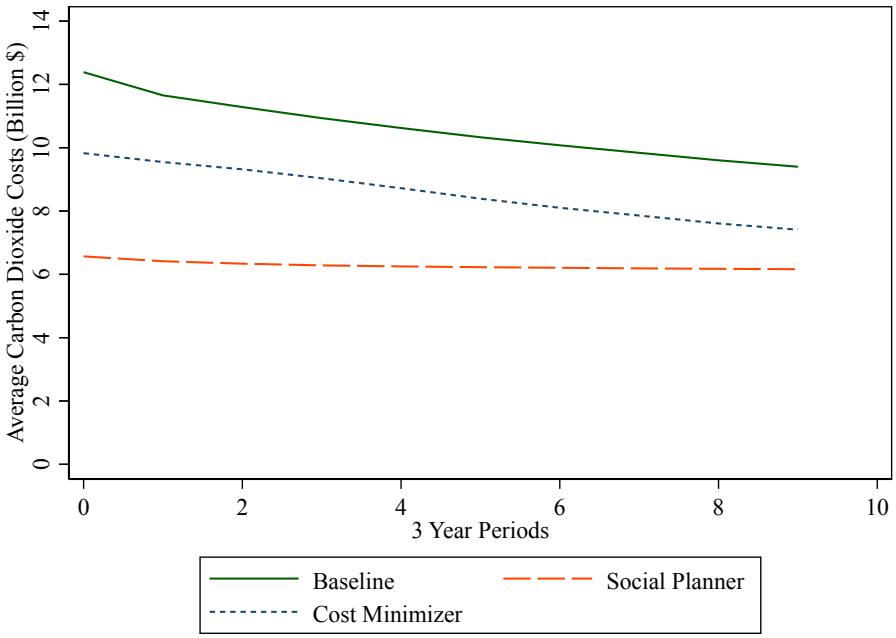
Note:

Figure 4: Generation Changes: Baseline, Social Planner, Cost Minimizer



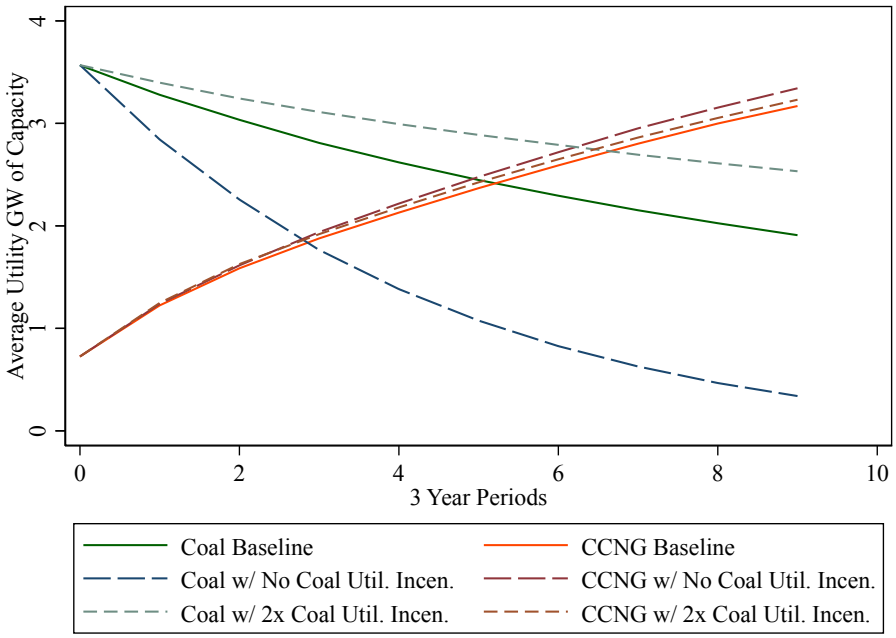
Note:

Figure 5: CO_2 Costs: Baseline, Social Planner, Cost Minimizer



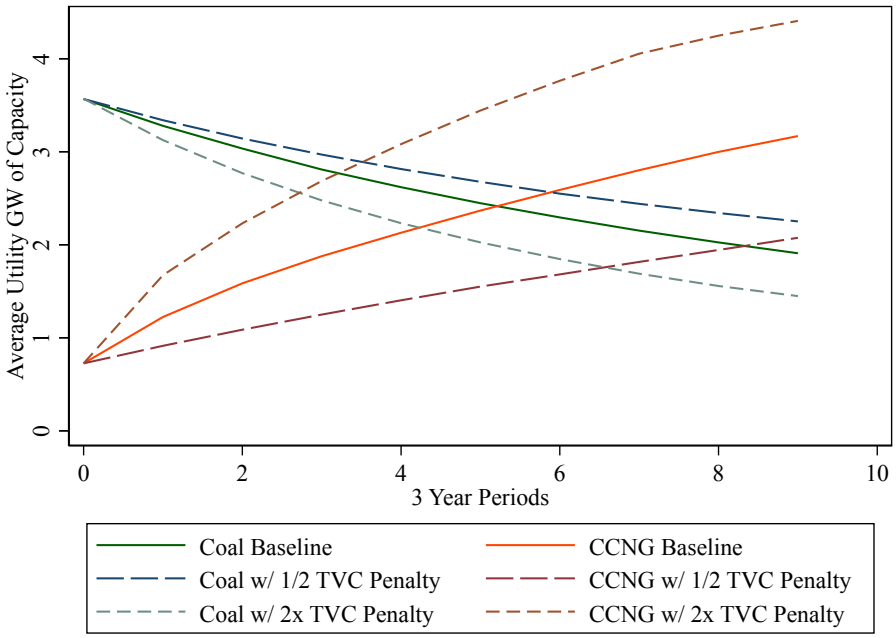
Note:

Figure 6: Changing Utilization Incentive (μ_2): Capacity



Note:

Figure 7: Changing Penalty for High TVC (γ): Capacity



Note:

- Figure 7 shows that changing the penalty for high TVC leads to substantial changes in CCNG investment, with smaller effects for coal retirement.
- In particular, with high TVC penalty, the utility will want to immediately invet in CCNG in order to have the option to generate with CCNG. This exacerbates the Averch-Johnson incentive to over-invest in capital.

7 Conclusion

IN PROGRESS

References

Abito, J. M. (2020). Measuring the welfare gains from optimal incentive regulation. *The Review of Economic Studies 87*(5), 2019–2048.

Averch, H. and L. L. Johnson (1962). Behavior of the firm under regulatory constraint. *The American Economic Review 52*(5), 1052–1069.

Baumol, W. J. and A. K. Klevorick (1970). Input choices and rate-of-return regulation: An overview of the discussion. *The Bell Journal of Economics and Management Science*, 162–190.

Biewald, B., D. Glick, J. Hall, C. Odom, C. Roberto, and R. Wilson (2020). Investing in failure. *Synapse Energy Economics, Inc*. Available at `https://www.synapse-energy.com/sites/default/files/Investing-in-Failure-20-005.pdf`.

Borrero, M., G. Gowrisankaran, and A. Langer (2023). Ramping costs and coal generator exit. *Working Paper*. `https://tinyurl.com/BorreroGL`.

Bottorff, C., N. Ver Beek, and L. C. Stokes (2022). The dirty truth about utility climate pledges. *Sierra Club*. Available at `https://www.sierraclub.org/sites/www.sierraclub.org/files/2022-09/sierra_club_the_dirty_truth_report_v2_2022_0.pdf`.

Bushnell, J. B., E. T. Mansur, and C. Saravia (2008). Vertical arrangements, market structure, and competition: An analysis of restructured us electricity markets. *American Economic Review 98*(1), 237–66.

Butters, R. A., J. Dorsey, and G. Gowrisankaran (2023). Soaking up the sun: Battery investment, renewable energy, and market equilibrium.

Cicala, S. (2015). When does regulation distort costs? Lessons from fuel procurement in us electricity generation. *American Economic Review 105*(1), 411–44.

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

Cole, W. J. and A. Frazier (2019). Cost projections for utility-scale battery storage. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Collard-Wexler, A. (2013). Demand fluctuations in the ready-mix concrete industry. *Econometrica 81*(3), 1003–1037.

Cross-Call, D., R. Gold, C. Goldenberg, L. Guccione, and M. O'Boyle (2018). Navigating utility business model reform: A practicalguide to regulatory design. *Rocky Mountain Institute*. Available at `https://rmi.org/insight/navigating-utility-business-model-reform`.

Daniel, J. (2021). Has consumers energy found a loophole in its clean energy pledge? *Union of Concerned Scientists*. Available at `https://blog.ucsusa.org/joseph-daniel/has-consumers-energy-found-a-loophole-in-its-clean-energy-pledge/`.

Daniel, J., S. Sattler, A. Massie, and M. Jacobs (2020). Used, but how useful? how electric utilities exploit loopholes, forcing customers to bail out uneconomic coal-fired power plants. *Union of Concerned Scientists Report*. Available at `https://www.ucsusa.org/resources/used-how-useful`.

Davis, L. W. and C. Wolfram (2012). Deregulation, consolidation, and efficiency: Evidence from US nuclear power. *American Economic Journal: Applied Economics 4*(4), 194–225.

Dunkle Werner, K. and S. Jarvis (2022). Rate of return regulation revisited. *Energy Institute at Haas Energy Institute WP 329*.

Eisenberg, T. (2019). Regulatory distortions and capacity investment: The case of China's coal power industry.

Elliott, J. T. (2022). Investment, emissions, and reliability in electricity markets.

Energy Information Agency (2020). Frequently asked questions. Available at https://www.eia.gov/tools/faqs/faq.php?id=77&t=11.

Fisher, J., A. Armendariz, M. Miller, B. Pierpont, C. Roberts, J. Smith, and G. Wannier (2019). Playing with other people's money: How non-economic coal operations distort energy markets. *Sierra Club*. Available at https://www.sierraclub.org/sites/default/files/Other%20Peoples%20Money%20Non-Economic%20Dispatch%20Paper%20Oct%202019.pdf.

Fowlie, M. (2010). Emissions trading, electricity restructuring, and investment in pollution abatement. *American Economic Review 100*(3), 837–69.

Fowlie, M., M. Reguant, and S. P. Ryan (2016). Market-based emissions regulation and industry dynamics. *Journal of Political Economy 124*(1), 249–302.

Gilbert, R. J. and D. M. Newbery (1994). The dynamic efficiency of regulatory constitutions. *The Rand journal of economics*, 538–554.

Goldie-Scot, L. (2019, December). A behind the scenes take on lithium-ion battery prices. *BloombergNEF*. Available at https://about.bnef.com/blog/behind-scenes-take-lithium-ion-battery-prices/.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of applied econometrics 8*(S1), S85–S118.

Gowrisankaran, G., A. Langer, and W. Zhang (2023). Policy Uncertainty and the in the Market for Coal Electricity: The Case of Air Toxics Standards. Available at https://drive.google.com/file/d/1WEpoPI2R052mM76agtDdgrfMBfoOqH3i/view?usp=sharing.

Gowrisankaran, G., S. S. Reynolds, and M. Samano (2016). Intermittency and the value of renewable energy. *Journal of Political Economy 124*(4), 1187–1234.

Gowrisankaran, G. and P. Schmidt-Dengler (2023). A computable dynamic oligopoly model of capacity investment. Working Paper.

Hausman, C. (2019). Shock value: Bill smoothing and energy price pass-through. *The Journal of Industrial Economics 67*(2), 242–278.

Holland, S. P., E. T. Mansur, N. Z. Muller, and A. J. Yates (2020). Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation. *American Economic Journal: Economic Policy 12*(4), 244–74.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

International Renewable Energy Agency (2020). Renewable power generation costs in 2019. Available at `https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Jun/IRENA_Costs_2019_EN.pdf?la=en&hash=BFAAB4DD2A14EDA7329946F9C3BDA9CD806C1A8A`.

Joskow, P. L. (1974). Inflation and environmental concern: Structural change in the process of public utility price regulation. *The Journal of Law and Economics 17*(2), 291–327.

Joskow, P. L. (2007). Regulation of natural monopoly. *Handbook of law and economics 2*, 1227–1348.

Joskow, P. L. (2014). Incentive regulation in theory and practice: electricity distribution and transmission networks. *Economic regulation and its reform: What have we learned?*, 291–344.

Klevorick, A. K. (1971). The" optimal" fair rate of return. *The Bell Journal of Economics and Management Science*, 122–153.

Klevorick, A. K. (1973). The behavior of a firm subject to stochastic regulatory review. *The Bell Journal of Economics and Management Science*, 57–88.

Laffont, J.-J. and J. Tirole (1986). Using cost observation to regulate firms. *Journal of political Economy 94*(3, Part 1), 614–641.

Lazar, J. (2016). Electricity regulation in the us: A guide. second edition. *The Regulatory Assistance Project*. Available at http://www.raponline.org/knowledge-center/electricityregulation-in-the-us-a-guide-2.

Lim, C. S. and A. Yurukoglu (2018). Dynamic natural monopoly regulation: Time inconsistency, moral hazard, and political environments. *Journal of Political Economy 126*(1), 263–312.

Linn, J. and K. McCormack (2019). The roles of energy markets and environmental regulation in reducing coal-fired plant profits and electricity sector emissions. *The RAND Journal of Economics 50*(4), 733–767.

MacKay, A. and I. Mercadal (2019). Shades of integration: The restructuring of the US electricity markets. *Harvard Business School, Working Paper 18*.

Myatt, J. (2017). Market power and long-run technology choice in the US electricity industry.

Potomac Economics (2020). A review of the commitment and dispatch of coal generators in MISO. Available at https://www.potomaceconomics.com/wp-content/uploads/2020/09/Coal-Dispatch-Study_9-30-20.pdf.

Reguant, M. (2019). The efficiency and sectoral distributional impacts of large-scale renewable energy policies. *Journal of the Association of Environmental and Resource Economists 6*(S1), S129–S168.

Ryan, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica 80*(3), 1019–1061.

Shwisberg, L., M. Dyson, G. Glazer, C. Linvill, and M. Anderson (2020). How to build clean energy portfolios: A practical guide to next-generation procurement practices. *Rocky Mountain Institute*. Available at http://www.rmi.org/insight/how-to-build-clean-energy-portfolios.

U.S. Energy Information Administration (2018, October). Today in energy. Available at https://www.eia.gov/todayinenergy/detail.php?id=37392.

Viscusi, W. K., J. E. Harrington Jr, and D. E. Sappington (2018). *Economics of Regulation and Antitrust*. MIT press.

Wilson, J. D., M. O'boyle, R. Lehr, and M. Detsky (2020). Making the most of the power plant market: Best practices for all-source electric generation procurement. *Energy Innovation Policy and Technology and Southern Alliance for Clean Energy*. Available at https://energyinnovation.org/wp-content/uploads/2020/04/All-Source-Utility-Electricity-Generation-Procurement-Best-Practices.pdf.

On-Line Appendix

A1 Additional Tables

A2 Data

- Short discussion of how we cannot observe TVC in the data. We therefore let CostBasis be equal to fuel costs plus import price times import quantity, which ignores the fact that utilities are price discriminating monopsonists. This is a close approximation to what the regulator will observe since the regulator cannot observe import supply slopes.

A3 Details of Estimation

A3.1 Operations Model Computation

Operations Bellman:

- 4 dimensional grid for bellman in each hour. Choice bins depend on lagged generation. Min and max of each grid.
- Approximate with 8 weeks. What are they? Assuming no ramping costs if between weeks.
- Cost basis: generation and import costs as observed in the data. But, use all costs against this.
- Leap year: what we do.
- import costs are the area under the supply curve footnote: price discriminating purchaser of imports who buys from different sources with many small contracts
- Limit return to what it would be if TVC is 10% the base year TVC. This also removes negative TVC where profits aren't defined.

Table A1: Operations Model Fit

	Data	Baseline
Annual Electricity Production (TWh):		
Coal	16.14	19.42
CCNG	6.93	3.98
Imports	7.44	11.43
Mean Utilization Share (%):		
Coal	52.40	61.79
CCNG	35.89	21.83
Annual Costs (Millions of Dollars):		
Coal Fuel Costs	397.83	477.76
CCNG Fuel Costs	230.90	104.56
NGT Fuel Costs	36.32	147.87
Coal O&M Costs	263.20	316.66
CCNG O&M Costs	17.34	9.94
NGT O&M Costs	11.93	36.98
Coal Ramping Costs	15.42	13.72
CCNG Ramping Costs	4.05	3.12
Imports	155.18	232.82
Total Variable Production Costs	1,132.18	1,343.44
0.00	5,056.90	
Variable Profits (Millions of Dollars):	844.19	1,594.96
Electricity Revenues (Dollars/MWh):	40.70	39.72

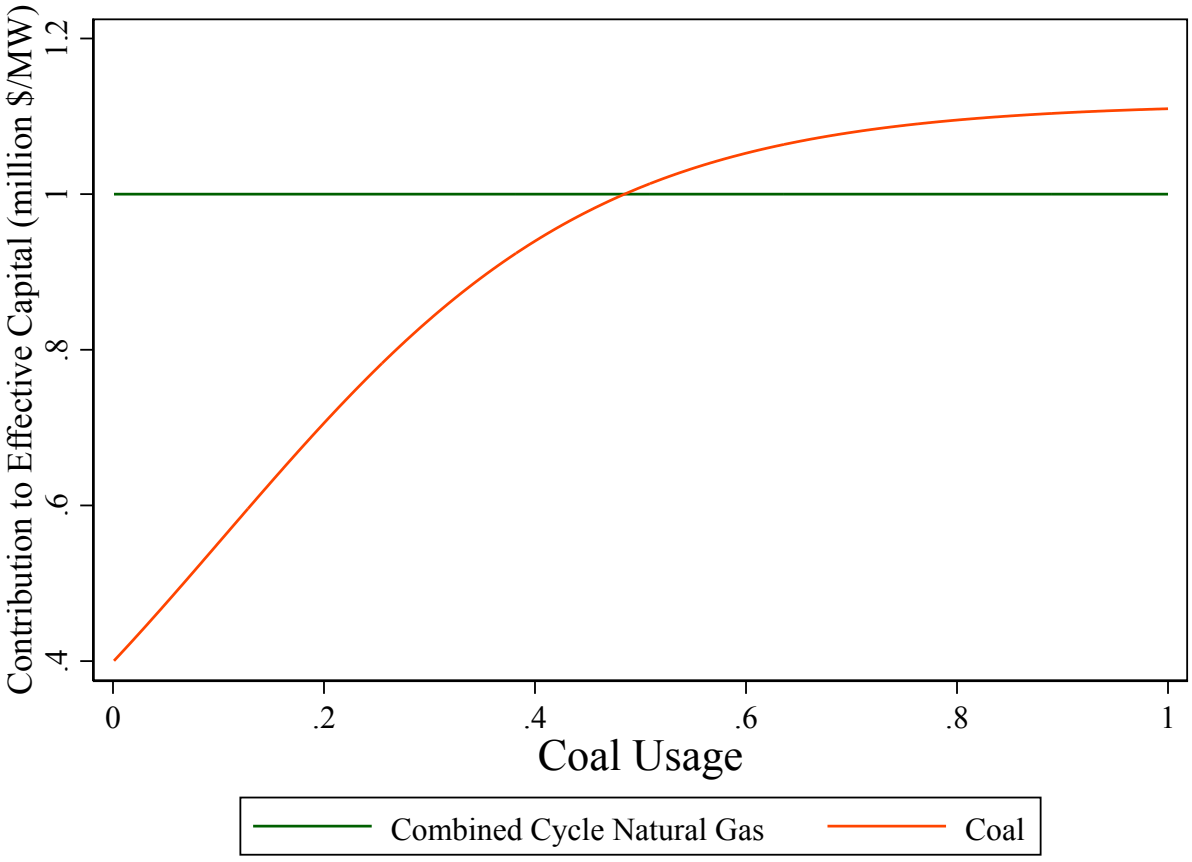
Note:

Table A2: Indirect Inference Coefficient Matching

Reg #	Dependent Variable	Regressor	Actual Data	Simulated Data
1-3	Utilization Levels:			
	Coal	Constant	0.521 (0.032)	0.560 (0.079)
	CCNG	Constant	0.357 (0.040)	0.220 (0.032)
	NGT	Constant	0.082 (0.023)	0.01 (0.015)
4	Utility Revenue Level:			
		Constant	874.789 (350.86)	2170.8 (423.43)
5	$\overline{s}$ proxy: $(Revenue_y - TVC_y)/\sum_f K_f$			
		Total Variable Cost	-0.0282 (7.26e-3)	-0.0313 (1.68e-3)
6	Variable Profits: $Revenue_y - TVC_y$			
		Coal Capacity (MW)	-0.381 (0.268)	0.378 (0.215)
		Coal Capacity × Utilization	0.651 (0.279)	-0.0549 (0.231)
		CCNG Capacity	0.259 (0.024)	0.266 (0.07)
		NGT Capacity	0.097 (0.210)	0.588 (0.69)
7	Log Coal Share:			
		First Quantile Coal	0.083 (0.422)	-0.375 (0.12)
		Second Quantile Coal	1.587 (0.447)	0.595 (0.331)
		Third Quantile Coal	1.610 (0.397)	0.549 (0.147)
		Fourth Quantile Coal	1.223 (0.121)	1.168 (0.078)
8	Log CCNG Share:			
		First Quantile CCNG	-2.361 (0.355)	-2.649
		Second Quantile CCNG	-1.069 (0.243)	-2.158 (0.147)
		Third Quantile CCNG	-0.3875 (0.185)	-1.349 (0.098)
		Fourth Quantile CCNG	0.090 (0.217)	-0.705 (0.069)
9-10	Ramping:			
	Coal Utilization	Lagged Coal Utilization	0.978 (0.003)	0.990 (0.002)
	CCNG Utilization	Lagged CCNG Utilization	0.971 (0.006)	0.921 (0.003)

Note:

Figure A1: Coal Contribution to Effective Capital



Note:

- Grid over CCNG and COAL because they're dynamic. For NGNCC, closed form if interior solution, but also check endpoints.

- If import slopes are slightly negative, the import supply curve is too steep, because slope is q as a function of p, so inverse supply is $1/slope$

A3.2 Indirect Inference

Total of 29 moments from 10? regressions. Need to change TVC here because it's actually $TVC_fuel_imports$ in the data.

Four of these are regressions on a constant term, meant to capture the scale of parameters. Specifically, we run a regression for each fuel (coal, combined-cycle natural gas, and non-combined-cycle natural gas) of hourly utilization (generation divided by capacity) on a constant. We use utilization instead of generation in levels because of the vast differences in the scale of utilities for different fuels.

The fifth indirect inference regression captures how allowable rates of return vary with utilities' total variable costs. While we do not observe the maximum allowable rate of return, $\overline{s}$, a transformation of Equation (2) gives us that $\overline{s}_y K^e = Revenues_y - TVC_y$, and we do observe revenues, total variable costs, and capacities. The concern is then that effective capital, as defined in Equation (2), depends upon both used-and-useful considerations and the value of MW of capacity as expressed in dollars. We, therefore, run a regression of revenues net of total variable costs divided by total MW of capacity on total variable costs and utility fixed effects. This is effectively a regression of the rate of return, with MW of capacity proxying for effective capital.

While this regression controls for utility fixed effects to account for differences in base year total variable costs, TVC_0, we do not match these coefficients with indirect inference and instead focus only on matching the coefficient on total variable costs. We view these fixed effects as capturing features that are outside of our model, such as, for instance, high transmission costs for utilities where customers are spread out.

Our sixth regression aims to understand how capacity is valued directly, via the α pa-

rameters. This regression complements our fifth regression, which uses a proxy for effective capital. In particular, our sixth regression is of total variable profits, $Revenue_y - TVC_y$ on coal capacity, coal capacity interacted with coal utilization, combined-cycle natural gas capacity, and non-combined-cycle natural gas capacity.[28] One key point to note here is that α, the conversion of MW of combined-cycle natural gas capacity into dollars of effective capital, does not affect the operation decisions conditional on fuel-technology capacities, which are fixed within a year. However, α does affect overall revenues (and therefore profits), so the identification of α largely comes from the scale regressions.

The seventh and eighth regressions are intended to capture utilities' incentives to have the regulator regard coal capacity as used and useful. We therefore look at hours where a utility is likely to be considering a combination of both coal and combined-cycle natural gas generation.

Our eighth regression is similar, but examines the extra use of gas generation in a similar window to the above. Here, we run a similar regression for the log of the combined-cycle natural gas generation share, but for hours where load is between 75% and 125% of coal capacity. From these two regressions, we match the coefficients on the first through fourth quintile of utilization. This regression is meant as a falsification, since we do not expect gas capacity to be subject to the used-and-useful standard.

Finally, the ninth and tenth regressions relate to the ramping costs for both coal and combined-cycle natural gas capacity. The intuition behind these regressions follows from Gowrisankaran et al. (2023), where ramping costs are identified by comparing hours where continuation values given generation decisions are very similar, but past generation values may differ. Specifically, we run two regressions, one for coal and one for combined-cycle natural gas, where the dependent variable is hourly utilization (the share of capacity that is generating) and the independent variables are the hourly generation from the previous

[28] Relative to the fifth indirect inference regression, this specification does not control for variation in TVC. Therefore, the coefficients can be interpreted as incorporating both the α and the TVC terms. We explicitly do not account for utility fixed effects in this regression because the variation in total capacity for each fuel-technology is somewhat limited within utilities. While this means that variation in the coefficients in this regression comes largely from the cross-section, it also accounts for the fact that utilities with larger capacities need larger payments to finance that capacity.

hour, coal and gas fuel prices, and six hourly future values of load, the import intercept, and the electricity price. Our indirect inference approach only matches the coefficient on lagged generation in each regression to the results generated from the model.

To implement our indirect inference estimation approach, we begin by estimating these ten regressions in our data and recovering the coefficients and their standard errors. We then choose a candidate parameter vector and solve the operations model to generate data for each utility and year. We use these simulated data to run the same regressions and recover the coefficients and standard errors of interest. Finally, we search over parameter values using a non-linear search for the vector of values that best matches the coefficients and standard errors across the two sets.

This indirect inference approach also requires us to choose a weighting matrix to determine how differences across moments will be summed. We use a weighting matrix based on the variance-covariance matrix of the regressions on the actual data above, where all regressions are clustered at the utility level, and the hourly regressions are also clustered at the month-of-year and hour-of-day. We assume that there is no covariance across regressions.

A3.3 Investment and Retirement Cost Estimation

- Likelihood by interpolation.
- Finite penalty for 0 probability events?
- Number of choices and level of choices for PSD stuff.
  - Give base bin numbers from code.
  - Also note that we exclude bins that would give negative coal capacity.

A4 Import Supply Curve Results

A5 Counterfactual Computations

Cost Minimization

- Set μ_2 to zero so that the utility just minimizes cost
- Negative cost from the operations model becomes the "profit" in the investment model code

Social Planner

- Same as cost minimimzation except that costs include carbon external cost

Used and Useful

- Change μ_2 and then maybe? change α^{COAL}

Penalty for High Variable Costs

- Change γ and then change α to keep average revenues the same