Energy Transitions in Regulated Markets*

Gautam Gowrisankaran† Ashley Langer‡ Mar Reguant§

February 1, 2024

Abstract

Natural gas has replaced coal as the dominant fuel for U.S. electricity generation. However, U.S. states that regulate electric utilities have retired coal more slowly than others. We build a structural model of rate-of-return regulation during an energy transition where utilities face tradeoffs between lowering costs and maintaining coal capacity. We find that the current regulatory structure retires only 45% as much coal capacity as a cost minimizer. A regulated utility facing a carbon tax does not lower carbon emissions immediately but retires coal similarly to the social planner. Alternative regulations with faster transitions clash with affordability and reliability goals.

JEL Codes: L51, Q48, L94, Q50
Keywords: rate-of-return regulation, electricity, prudence, used and useful, coal generation, stranded assets.

*We thank Steve Cicala, Lucas Davis, Meredith Fowlie, Catie Hausman, Paul Joskow, Erin Mansur, Nancy Rose, 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. Gowrisankaran and Langer acknowledge financial support from the National Science Foundation through Grant SES-2149335.

†Columbia University, CEPR, and NBER
‡University of Arizona and NBER
§Northwestern University, BSE, CEPR, and NBER

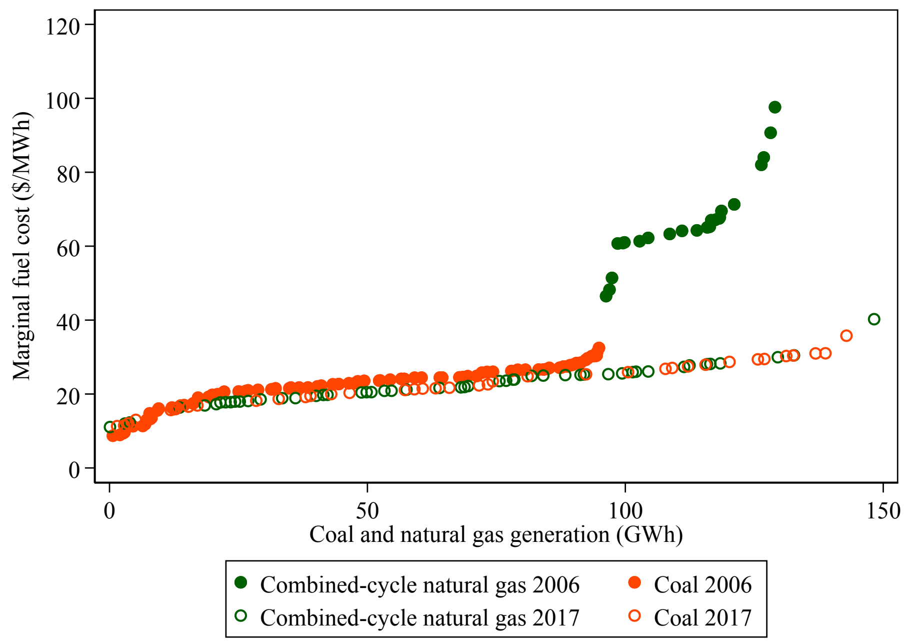
1 Introduction

Electricity generation is a critically important component of the economy and modern life. However, it also generates substantial negative externalities. In particular, electricity generation contributed 31% of U.S. CO_2 emissions in 2019 (Energy Information Administration, 2020). Beyond its climate change impacts, electricity generation 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).

Amid growing concerns about the damages from electricity emissions, two major transitions in electricity generation are underway. The first transition, occurring over the past two decades, is marked by the significant reduction in the cost of generating electricity with natural gas, thanks to combined-cycle technology and hydraulic fracturing (fracking). Figure 1 illustrates the marginal electricity fuel costs in 2006 and 2017 for combined-cycle natural gas and coal generators in our sample. The figure sorts generators by dispatch order—i.e., in order of increasing fuel cost—with capacities displayed cumulatively. In 2006, combined-cycle natural gas generators (in green) had uniformly higher fuel costs than coal. Yet, by 2017, 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). The second transition stems from a substantial decrease in the costs of solar panels and batteries, both having 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 considered natural monopolies and faced rate-of-return (RoR) regulation. The 1990s saw extensive restructuring in the U.S. and Europe, where electricity generation—and sometimes retailing—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

Figure 1: Marginal Fuel Costs Over Time



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

where states have substantially different levels of regulatory control. The impact of falling natural gas generation costs varied across regulated and restructured markets. For instance, between 2006 and 2018, 26.0% of coal capacity exited in restructured states, whereas only 17.2% exited in regulated states.[1] This suggests that it is important to study whether electricity regulation increases social costs by slowing transitions to new energy sources.

RoR electricity regulation has the twin goals of "reliability" and "affordability" (Energy, Climate, and Grid Security Subcommittee, 2023). To meet these goals, the regulator creates a structure under which utilities are incentivized to meet electricity load (demand) while encouraging low-cost generation and limiting underused capital (Joskow, 1974). This paper develops and estimates a model of electricity regulation. In the model, the utility optimizes against the regulatory structure by choosing investment and retirement capacity levels in the long run and generation quantities by fuel-technology and imports in each hour. We estimate our model using publicly available data on utilities' electricity generation, load,

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

revenues, and capacity. We use our model to evaluate how both the current and alternative regulatory structures affect generation and capacity decisions and the resulting costs and pollution.

RoR 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 RoR 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 RoR regulation leads to over-investment in capital because it limits the utility's return in proportion to the capital base, and the utility endogenously responds by increasing capital (Averch and Johnson, 1962). Regulators therefore have combined RoR regulation with the requirement that only "prudent" capital investments be included in the rate base. For older, existing technologies needing maintenance, repair, or upgrade, 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. The regulator's task has become even more complicated over the past 25 years due to changing technologies, fuel prices, and environmental concerns. These changes exacerbate the regulator's imperfect information about the current and future costs of alternative utility investment and operation decisions.

Our model captures these key features of RoR regulation. 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 considers capacity usage in assessing the extent to which capital is included in the rate base.

The utility optimizes against this regulatory structure in its investment and operations decisions. Each three-year period, the utility chooses coal capacity retirement and combined-cycle natural gas capacity investment, facing quadratic costs of adjustment. Investments increase its rate base and therefore its variable profits, conditional on the rate of return. When choosing which fuel-technologies to operate and how much to import to meet load, the 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.[2] Second, particularly after the decline in natural gas generation costs, it may use expensive coal generators to ensure that they are deemed used and useful.

Our model relies on both regulatory and cost parameters. The regulatory parameters include the determinants of the maximum rate of return and each fuel-technology's contribution to the rate base, which for coal capacity depends on its usage. Our cost parameters include operations and maintenance, ramping, and retirement/investment costs. We estimate the regulatory and operations parameters with a nested fixed-point indirect inference approach that seeks to match important data correlations. Specifically, we run regressions on our actual data that capture key features such as utilities revenues, ramping behavior, and usage, and find the structural parameters that yield the most similar regression coefficients in simulated data generated by the model. We also estimate the investment and retirement costs with a GMM nested fixed-point approach. We follow the Gowrisankaran and Schmidt-Dengler (2023) algorithm that facilitates the computation of models with continuous outcomes, in our case investment and retirement decisions.

We use our structural parameter estimates to analyze the impact of counterfactual policies on operations decisions and long-run outcomes in the presence of an energy transition. We consider four categories of counterfactuals: (1) adjustments to the existing regulatory parameters, (2) cost minimization, (3) a social planner which also internalizes carbon externalities at $190/ton (Environmental Protection Agency, 2023b), and (4) carbon taxes in the presence of RoR regulation.

We find that utilities faced with an energy transition in 2006 would retire 45% of coal capacity on average over a 30-year horizon and gradually reduce coal usage but increase combined-cycle natural gas capacity by 427% over the same period. Comparing this baseline outcome to counterfactual simulations that change the penalty for high variable costs, utilities generate more with combined-cycle natural gas capacity as the penalty for high variable costs increases, yielding a faster energy transition. Yet both coal and combined-cycle natural gas

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

capacity levels also increase with higher variable cost penalties, exacerbating the Averch and Johnson (1962) effect of over-investment from RoR regulation and decreasing affordability.

Turning to the potential costs of RoR regulation, cost minimizing utilities transition away from coal much more quickly than in the baseline: they immediately reduce coal generation by 53% and virtually eliminate coal capacity over 30 years. Similarly, a social planner that internalized carbon externalities would eliminate coal capacity over the same time-frame, but would also immediately reduce coal generation by 95%. Increasing the coal usage incentive or changing the penalty for high total variable cost within the existing RoR regulatory structure does not come close to replicating the speed of this energy transition. If regulated utilities instead faced a carbon tax, they would pass through 90% in the short-run, but transition away from coal toward combined-cycle natural gas in the long run. However, variable profits with the social planner, cost minimizer, or regulation with a carbon tax are all much lower. This may ultimately reduce reliability by undermining utilities ability to remain in business.

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 RoR regulation (e.g., Averch and Johnson, 1962; Baumol and Klevorick, 1970; Klevorick, 1971, 1973; Joskow, 1974; Gilbert and Newbery, 1994; Joskow, 2007). A different strand of the literation has focused on setting incentives correctly in the presence of asymmetric information (Laffont and Tirole, 1986). We extend the models in these literatures by investigating the role of regulation in the face of an energy transition.

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); Dunkle Werner and Jarvis (2022); Aspuru (2023); and Jha (2023). The closest paper in this literature to ours is Abito (2020), which structurally estimates a Laffont and Tirole style model of regulation under asymmetric information where an electric utility makes operations decisions trading off effort against costs. 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 RoR 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 et al. (2022); 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.

Section 2 presents background on the industry and introduces the data. Section 3 presents the model. Section 4 presents reduced-form evidence that supports key assumptions in our model. Section 5 presents our estimation approach, and Section 6 discusses our estimation and counterfactual results. Section 7 concludes.

2 Industry Background and Data

2.1 Industry Background

RoR regulation often exists in industries that are considered "natural monopolies" (Viscusi et al., 2018). The goal of this regulatory approach is to ensure reliability—in electricity, to literally keep the lights on—while maintaining affordability. In the U.S., state regulatory agencies, generally called Public Utility Commissions (PUCs), implement electricity RoR regulation with these goals in mind.[3]

Specifically, in our context, PUCs make three different types of decisions.[4] 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 RoR. Second, they determine which of utilities' reported non-capital costs are reimbursable. Third, they decide on the allowable

[3] 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).

[4] 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).

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 the regulator objectives of reliability and affordability. PUCs then 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) explains on page 52: "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. Regulated utilities own most of the generation capacity within their states (Shwisberg et al., 2020), while simultaneously trading 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.

However, 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).

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 2017.

Our primary annual data derive from the Energy Information Administration (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. 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-technologies: coal (COAL), combined-cycle natural gas (CCNG), and other (non combined-cycle) natural gas turbines (NGT), many of which are used to meet peak load. 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 heat rates, which indicate fuel energy input per unit of generation. We allow for different heat rates by utility and fuel-technology, but restrict each utility's heat rate to be the same within a fuel-technology for all existing and new capacity.

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. 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.[5]

We limit our data geographically because regulated utilities in the Eastern Interconnection all have relatively nearby Independent System Operators (ISOs). We use wholesale electricity price data from the nearest state in an ISO to define import prices. We also merge in coal and natural gas fuel prices at the state-year level derived from EIA Form 423.

Our last major data source is the Federal Energy Regulatory Commission (FERC) Form 714 data. These data include hourly load served by each utility. We then use the combined data to define hourly imports into the utility as load net of our three primary generation fuel-technology types. Thus, imports will include both nuclear and renewable (which is relatively moderate during this time-frame) generation.

On-line Appendix A2 discusses details of our data construction and includes summary statistics of our analysis data at the utility-year and utility-hour levels, respectively. Our final analysis data consist of nearly 2.5 million utility-hour observations across 283 utility-years for 26 unique utilities.

3 Model of Electricity Regulation

We present 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 RoR regulation to provide incentives to a natural monopoly electric utility. Following Viscusi et al. (2018), the regulator has two primary objectives:

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

it wants enough generation and imports to meet load in every hour ("reliability"), and it wants to keep consumer rates low ("affordability").[6] 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 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 evolving technologies and market conditions. Specifically, fluctuations in fuel prices determine which generation technologies are most cost effective, technological advancements create new generation options, and shifting perceptions of environmental harm affect 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 usage is one objective way to measure the prudence of an investment (Gilbert and Newbery, 1994). These

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

two incentives create a tension for the utility: it wants to use low-cost fuel-technologies but also wants to use expensive legacy capital.

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

The utility needs to meet load, ℓ_{yh}, in each hour, $h = 1, \ldots H$, 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 outside of its service area. We let $f \in \{COAL, CCNG, NGT\}$, represent coal, combined-cycle natural gas, and natural gas turbines respectively. We denote the quantity of generation from each fuel-technology q_{yh}^f and import choice q_{yh}^m, which we combine into $\vec{q}_y$.

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, $TVC(\vec{q}_y)$:

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

where $CostBasis$ is a baseline, observable cost metric that varies across utilities.[7]

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-

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

technologies, converted to dollar terms:

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

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 usage rate of coal, where $\overline{Q}_y^{COAL} \equiv \frac{1}{H} \sum_{h=1}^{H} q_{yh}^{COAL}$, 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 CCNG 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 CCNG generators both decreased costs and demonstrated their usefulness for baseload generation. NGT generators have higher average heat rates than CCNG ones, but since they often serve as "peakers," they would generally not need to prove usefulness via high usage rates. In contrast, coal 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 usage:

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

where μ_1 and μ_2 are parameters, with μ_2 determining the incentive to increase coal usage.

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 allows the utility to earn $Revenues_y$ sufficient for the utility to earn 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 operations decisions 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 CCNG capacity and whether to retire existing coal capacity. We make these choices since the vast majority of entry decisions are for CCNG capacity, and the vast majority of exit decisions are for coal capacity.[8] Finally, to limit the complexity of our model, we do not endogenize the choice of NGT capacity, which we view as less central to utility regulation.

In each period, the utility makes optimizing 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. 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

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

AR(1) process. The state Ω also includes the coal and CCNG capacities, both of which vary deterministically 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, NGT 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 with an annual discount factor of $\beta = 0.95$. We only allow the utility to make these decisions over the 30 years starting in 2006 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 models with continuous choices specify *i.i.d.* shocks for each candidate choice (e.g., Rust and Rothwell, 1995). 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 data variation.

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 its shock to the CCNG marginal cost of investment, ε_t^{CCNG}, and makes its CCNG 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:

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

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 CCNG 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:

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

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. The assumption that the state does not evolve when $t > 10$ allows us to solve the dynamic programming problem by backward induction, with the state-contingent value function at $t = 10$ being the discounted sum of future profits.

3.4 Operations Decisions

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 utility's profits depend upon its rate of return on its rate base, where its rate of return is a function of its total variable costs over the year:

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

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.[9] We assume that NGT generators 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. Following the literature (Bushnell et al., 2008; Gowrisankaran et al., 2016; Reguant, 2019), we let the utility's import costs be the integral under the inverse supply curve rather than the maximum import price times quantity imported.

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,

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

optimized profits are:

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

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

where the i subscript makes explicit which terms depend on the fixed state. As defined in Equation (2), the rate base, $B(\vec{q}_y, K^{COAL}, K^{CCNG})$, depends upon $\vec{q}_y$, through coal usage. Equation (8) makes explicit 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, for computation, we conceptualize this problem as a discretized finite-horizon Bellman equation. Without loss of generality, our model allows us to specify that the utility receives its only payoff, the regulatory profit, in the terminal hour. This payoff is a function of TVC and coal usage, $\overline{Q}^{COAL}$, but ramping costs imply that TVC depends on the hourly sequence of coal and CCNG generation. Thus, in any hour, h, the state for the Bellman equation includes the cumulative TVC and coal usage prior to this hour (which eventually pertain to profits), last hour's coal and CCNG generation (which affect ramping costs), and the hour of year h. 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 for the state-contingent value functions backwards to the first hour of the year, we then forward simulate—using the calculated state-contingent optimal policies—to recover the optimal action path. Specifically, we start in the first hour and record the optimal generation choices. We then use these choices to update the state for the next hour, which in turn allows us to record the state-contingent optimal actions for that hour. Iterating through the year, we obtain the utility's optimal operations decisions.

4 Reduced Form Evidence

Our model includes two key instruments that help the regulator meet its objectives. First, the maximum allowable rate of return offered declines in total variable costs. Second, capacity contributes more to effective capital when it is used sufficiently. This section analyzes the empirical support for these assumptions.

4.1 Determinants of the Allowable Rate of Return

We use utility-year level regressions to understand how the regulator adjusts the maximum rate of return, $\bar{s}_y$, as TVC_y changes. Since we do not directly observe either the rate of return or TVC_y, we proxy for them using observable information. Specifically, we proxy for TVC_y using fuel and import costs, which omits ramping and O&M costs, as these are not observable. Our dependent variable proxies for rate of return with a measure of variable profits—revenues net of fuel and import costs—divided by the sum of coal, CCNG, and NGT capacity in MWs. Our hypothesis is that increases in costs will lower the rate of return.

A central concern of this approach is that the return on capital needs to cover fixed costs—such as for capacity investment, transmission, and distribution—and these will vary across utilities. This implies that what the regulator considers a high cost for one utility may be a low cost for another, which then affects the maximum allowable rate of return.[10]

We use two approaches to address this concern here. First, we estimate specifications with utility fixed effects. Second, we divide variable costs by the utility's size, since costs per unit of utility size are likely to be similar across utilities. We measure size in two ways: first with the sum of coal, CCNG, and NGT capacities and second with a measure of the utility's peak load, the 95th percentile of hourly load in the utility-year.

Table 1 presents the results of these regressions. The three regressions with utility fixed-effects find strong, negative relationships between our proxies for the utility's rate of return and variable costs.[11] This provides support for our modeling assumption that generators

[10]Our structural model accounts for these differences by dividing TVC_y by the utility's $CostBasis$.

[11]The different normalizations mean that we cannot directly compare coefficients across specifications.

Table 1: Regressions of Rate of Return on Total Variable Cost per Effective Capital

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

Note: Each column presents regression results from a separate regression on our analysis data, with standard errors in parentheses. 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.

earn higher rates of return when they decrease variable costs. Without utility fixed effects, only one of the three regressions is statistically significant. We believe that the regressions with fixed effects are the most directly tied to the identification of our model.

4.2 Determinants of Effective Capital

We next motivate our assumption that coal usage affects the extent to which coal capacity contributes to effective capital, K^e_y. Our hypothesis is that regulated utilities face incentives to operate coal generators out of dispatch order to increase their effective capital. However, testing this hypothesis is complicated: it is difficult to understand the relative costs of using coal in any hour because ramping and O&M costs are not observed. Thus, we provide evidence on this hypothesis by comparing coal usage by firms in regulated states to restructured states. Firms owning coal in restructured states may face similar ramping and O&M costs but are not subject to used-and-useful considerations.[12]

Specifically, we run regressions at the utility-hour level of whether the utility is operating coal on whether the utility's fuel cost is above wholesale price and this indicator interacted

[12] We use our merged EIA/EPA dataset (before we merge in the FERC data) for these results. For both regulated and restructured markets, we define firms using the EIA's definition of a utility.

with whether the utility is in a restructured state, controlling for state and year fixed effects.[13] We also run these regressions for CCNG to understand the differential response of coal in restructured markets. The comparison of how usage responds to wholesale prices falling below fuel costs, differentially across fuel-technologies and regulatory status, identifies the used-and-useful incentive with a triple-difference-in-difference-style comparison.

Table 2: Out-of-Dispatch-Order Generation by Regulatory Status

	$\mathbb{1}$\{Fuel-Technology Operating\}	
	Coal	Combined Cycle Natural Gas
$\mathbb{1}\{\text{Fuel Cost} > \text{Price}\}$	-0.031	-0.201
	(0.031)	(0.031)
$\mathbb{1}\{\text{Fuel Cost} > \text{Price}\}\times$ Restructured	-0.122	0.005
	(0.050)	(0.029)
R^2	0.089	0.132
N	19,782,473	20,723,467

Note: Regressions are linear probability models that include state and year fixed effects. Data are at the utility-hour level for the Eastern Interconnection and include both regulated and restructured electric utilities. We cluster standard errors (in parentheses) at the state and year level.

Table 2 shows the results of these regressions for both coal and CCNG. The first column shows that coal in regulated states is not significantly less likely to operate when out of dispatch order, i.e., when its marginal fuel costs exceed the wholesale (import) price. However, coal generation in restructured states responds strongly to being out of dispatch order, with a statistically significant 12.2 percentage point reduction in the likelihood of operating relative to regulated states. The second column shows that CCNG generation in both regulated and restructured states react strongly to low wholesale prices. This provides evidence consistent with the hypothesis that regulated utilities gain value from generating with coal even when it is out of dispatch order.

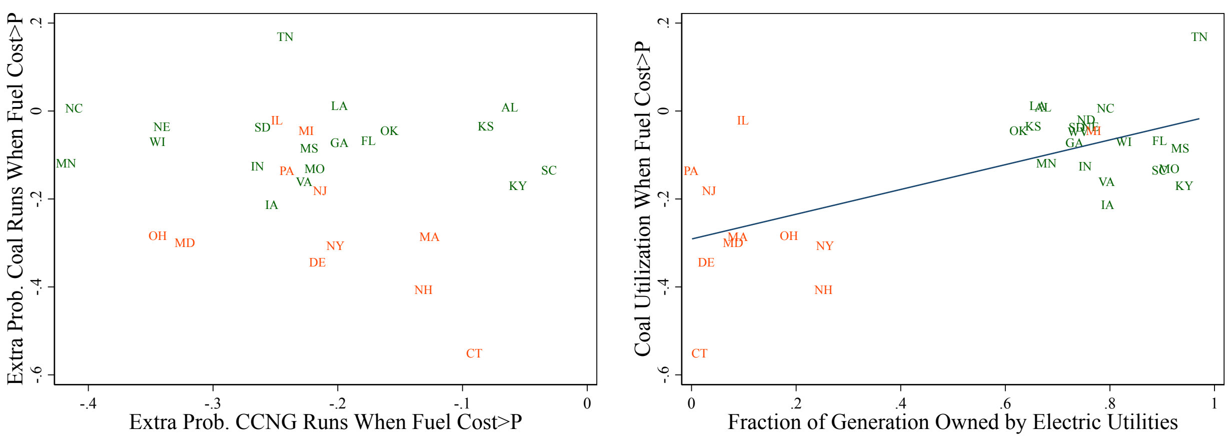
Figure 2 presents estimates from similar regressions to Table 2 on out-of-dispatch order generation, but allowing the coefficients to vary by state. Panel (a) shows the out-of-dispatch

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

Figure 2: Generation When Fuel Cost > Price in Regulated Versus Restructured Markets

(a) Out-of-Dispatch Order Generation for Coal and CCNG

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



Note: Panel (a) presents coefficients on coal and CCNG out-of-dispatch order generation by state. Panel (b) plots the fraction of generation owned by electric utilities against the same coal coefficients. In both panels, green states are regulated and red states are restructured.

order coefficients for coal (vertical axis) and CCNG (horizontal axis). We plot regulated states in green and restructured states in red. Out-of-dispatch order coal generation is clearly related to regulatory status while there is little pattern for CCNG. The seven states with the lowest differential coal usage—which are at the bottom of the graph—are all restructured states. This reinforces the idea that regulatory status significantly impacts coal usage.

Panel (b) of Figure 2 plots the share of generation owned by electric utilities in the state (as reported by Shwisberg et al., 2020) against the same coal coefficients as in Panel (a). Regulated states generally have utility ownership shares over 60%, whereas all restructured states but one have utility ownership shares under 30%. The best fit line shows that coal's responsiveness to low wholesale prices correlates 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 data and curves with utility revenue data to estimate the parameters governing the regulatory incentive structure and utilities' operations costs. Finally, we combine investment and retirement data with profits across states as predicted by our estimated operations decisions to recover investment and retirement cost parameters. We estimate both operations and investment/retirement parameters with full solution approaches and thus compute counterfactual outcomes with the same techniques.

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 CCNG capacity. Early in our sample, when natural gas prices were high, this utility would have generated with coal first, and only used natural gas in hours with high load. 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. By comparing the operations and investment/retirement decisions across utilities, we are able to identify the structural parameters.

We discuss our estimation approaches and the specific identification arguments for the operations and investment/retirement parameters in turn. On-line Appendix A3 discusses details of estimation including estimation of the import supply curves.

5.1 Operations Parameters

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 usage. Conditional on the state and parameters, these outcomes then determine the utility's annual profits.

We use utilities' hourly generation choices within a year to estimate the regulatory parameters: the penalty for high TVC—γ, the conversion from MW of CCNG capital to dollars of rate base—α, the other fuel-technologies' rate base contributions relative to CCNG—α^{NGT}

and α^{COAL}, and the used-and-useful terms—μ_1 and μ_2. We also use these choices to estimate ramping costs—ρ^{COAL} and ρ^{CCNG}, and O&M costs—om^{COAL}, om^{CCNG}, and om^{NGT}.

We estimate the structural parameters via a nested fixed-point indirect inference approach (Gourieroux et al., 1993). This involves a non-linear search to find the parameters that most closely match coefficients from regressions run on model-simulated data to those run on actual data. The solution of our model depends on $CostBasis$, which captures differences in fixed characteristics across utilities, such as size, that will influence the regulator's perception of costs. We define $CostBasis$ as fuel costs and a measure of import costs in the first year the utility appears in the data,[14] because the other components of TVC—including O&M and ramping costs—are not directly observable.

Solving for optimal operations decisions results in simulated hourly and annual data on which we run our indirect inference regressions. Indirect inference is a form of generalized method of moments (GMM), which specifies correlations in the data we would most like to match. Indirect inference regressions do not require a causal interpretation, but rather are meant to capture equilibrium behavior. We therefore choose indirect inference regressions that we believe best reflect the important equilibrium features of the data.

An alternative to indirect inference would be a GMM estimator that matched outcomes such as the rate base or the rate of return between the simulated model and the data. In our case, we do not observe these key elements, but rather observe outcomes that indirectly relate to them, notably revenues, generation, and fuel and import costs. The indirect inference approach allows us to capture the equilibrium co-movements between different observable variables.

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. While identification of the structural parameters derives from all of the indirect inference regressions together, we motivate particular regressions as aiding identification of particular parameters. Fundamentally, much of the variation that identifies

[14] In most cases, this will be 2006, before natural gas prices declined, and hence during a period when base load was predominantly met with coal generation.

these parameters will stem from the sharp decline in natural gas fuel prices, as discussed in the introduction of Section 5.

At the hourly level, we regress generation 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 regress 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 regress the log of the share of hourly generation from coal and CCNG that is met by coal on quintiles of annual coal usage,[15] 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 identify coal usage's contribution to effective capital. 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 usage 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 load is between 75% and 125% of total CCNG capacity, to isolate hours where the incentive to increase coal's contribution to the rate base are most likely to affect generation.

At the annual level, we regress an observable measure of variable profits—revenues net of fuel and a measure of import costs—on a constant. We further regress this same measure on capacity by fuel-technology and coal capacity interacted with coal usage. 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}. These regressions also combine with the hourly regressions to help identify the coal usage incentives, μ_1 and μ_2. Finally, we regress a measure of the rate of return on fuel and import costs and utility

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

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 capital and usage into the rate base. For both of these sets, we identify them jointly with multiple indirect inference regressions.

5.2 Investment and Retirement Cost Parameters

As 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 estimate the time-invariant terms, $\delta_0^f, \delta_1^f, \delta_2^f$, and standard deviations of unobservable components of investment and retirement costs, σ^f, with a GMM nested fixed-point estimator.[16] We use moments on the difference between the data and the model for investment or retirement. Specifically, for coal, we include the retirement amount and its square conditional on a non-zero amount, and indicators for non-zero retirement and whether they exceed certain thresholds. We also interact each of these terms with the utility's starting capital. Finally, we include the variance of the retirement amount. For CCNG, we include the analogous moments, but for investment.

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 two important features of our implementation. First, we solve the value function iteration and calculate the distributions of investment and retirement decisions

[16] 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.

using recent advances in the literature on continuous choice estimation (Gowrisankaran and Schmidt-Dengler, 2023). Following this 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 are not chosen for a given parameter vector and state. The algorithm provides a computationally quick way of ruling out the levels that are not chosen and of then solving for the probability of each remaining level based on cutoffs between them. We use these probabilities to solve for the continuation values in the Bellman recursion.[17] 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 grid of 10 investment/retirement levels.[18] 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.[19] 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 observed investment/retirement levels, yielding a zero likelihood.

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 a grid of coal capacity, CCNG capacity, and natural gas fuel prices that enters our investment and retirement Bellman, using the estimated operations parameters.

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

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

[18]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.

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

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 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 begins by presenting our estimation results. Following that, we describe our counterfactual simulations. We first analyze short-run counterfactuals that evaluate the impact of alternate regulatory policies on operations decisions holding utility capacity constant. Then, we explore long-run counterfactuals that simulate an energy transition over a 30-year horizon following a fall in natural gas prices.

6.1 Results

Table 3 presents estimates and standard errors for the structural parameters estimated using operations decisions.[20] Focusing first on how much a change in a utility's capacity would affect variable profits—which is primarily determined by both γ and α[21]—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%. We find that a 500 MW change in effective capital, which is roughly the mean CCNG generator capacity, would change variable profits by 6.5% on average across our sample.

Considering next 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

[20]Table A1 in the On-line Appendix A1 displays how simulations of our operations model compare to observed data and Table A2 in the same Appendix displays how the indirect inference coefficients estimated on the data match those estimated on the simulated data. Overall, we find that the model replicates patterns in the data reasonably well, replicating the pattern of usage across CCNG and coal usage quantiles. However, the model over-predicts revenues and NGT usage.

[21]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 3: Coefficient Estimates for Operations Model

Parameter	Notation	Estimate	Std. Error
Penalty for High TVC_t	γ	0.429	(0.08)
Rate Base per MW of Effective Capital (Millions $)	α	0.221	(0.06)
Coal Capacity Contribution to Effective Capital	α^{COAL}	1.117	(0.51)
Coal Usage Logit Base	μ_1	-0.589	(0.11)
Coal Usage Logit Slope	μ_2	5.641	(0.87)
NGT Contribution to Effective Capital	α^{NGT}	2.134	(1.00)
Ramping Cost for Coal (100$ / MW)	ρ^{COAL}	0.578	(0.11)
Ramping Cost for CCNG (100$ / MW)	ρ^{CCNG}	0.219	(0.31)
O&M Cost for Coal ($ / MWh)	om^{COAL}	16.350	(3.92)
O&M Cost for CCNG ($ / MWh)	om^{CCNG}	2.594	(0.10)
O&M Cost for NGT ($ / MWh)	om^{NGT}	19.767	(14.40)

Note: Structural parameter estimates from indirect inference nested fixed point estimation. All values are in 2006 dollars.

capital, is positive and statistically significant. In other words, coal capacity affects the rate base more as its usage increases. 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. Coal capacity owned by an (out-of-sample) utility that never generated electricity with coal would contribute to the rate base 40% as much as a similar amount of CCNG capacity. Alternatively, coal capacity owned by a similarly out-of-sample utility that generates with its full coal capacity in all hours would contribute to the rate base 111% as much as a similar amount of CCNG capacity. We find that a unit of NGT capacity would affect the rate base more than twice as much as a unit of CCNG capacity.

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 $5,780 while the figure is lower for CCNG ramps at $2,190. The coal estimates are between those of a full startup and a ramp of a unit that is already generating from Reguant (2014). They are lower than Borrero et al. (2023), but pertain to ramping across generators in a utility rather than for a specific generator as in that paper. We estimate O&M costs of $16.35/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 \$2.59/MWh and \$19.77/MWh respectively, though only the CCNG O&M cost is statistically significant. For CCNG, this number is extremely close to the reported variable O&M costs for single-shaft and multi-shaft CCNG turbines of \$2.67 and \$1.96, respectively (Energy Information Administration, 2022).

Table 4: Coefficient Estimates for Investment/Retirement Decisions

Parameter	Notation	Estimate	Std. Error
Fixed cost of coal retirement × 1e2	δ_0^{COAL}	-0.446	(9.79)
Linear coal cost per MW	δ_1^{COAL}	3.196	(0.44)
Quadratic coal cost per MW / 1e3	δ_2^{COAL}	0.117	(0.02)
Coal shock standard deviation per MW	σ^{COAL}	-0.430	(0.02)
Fixed cost of CCNG investment × 1e2	δ_0^{CCNG}	-0.509	(0.01)
Linear CCNG cost per MW	δ_1^{CCNG}	6.487	(0.08)
Quadratic CCNG cost per MW / 1e3	δ_2^{CCNG}	0.270	(0.05)
CCNG shock standard deviation per MW	σ^{CCNG}	-1.671	(0.06)

Note: Structural parameter estimates from GMM nested fixed point estimation. All values are in 2006 dollars.

Table 4 presents our investment/retirement parameter estimates. We find that the fixed costs of adjusting coal or gas capacity are small and slightly negative, although only statistically significantly for CCNG investment. The marginal costs of both coal retirement and gas investment are convex, with positive quadratic terms. The cost of CCNG investment offsets the negative fixed cost of investment as long as the investment amount exceeds 8 MW, a negligible size.

Evaluating the magnitude of standard investment and retirement decisions, a 250 MW coal retirement yields \$836 million in scrap value with the mean cost shock. The coal shock has a standard deviation of \$430,000 per MW. This means that the 250 MW of coal retirement would yield a scrap value between \$729 million and \$944 million with a one standard deviation positive or negative cost shock. In contrast, a 250 MW CCNG investment costs \$1.6 billion with the mean cost shock. For a 250MW CCNG investment, the costs lie between \$1.17 and \$2.01 billion with one standard deviation less or more than the mean CCNG investment cost.

Our estimates of the mean capital cost of CCNG investment are approximately 3-6 times higher than the capital costs reported in Energy Information Administration (2022). We would expect our estimates to be higher since they are based on revealed preferences and thus include substantially more than just capital costs. For instance, 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). Moreover, investments in our model generally occur when the utility receives a favorable draw of the CCNG cost shock, resulting in lower realized costs than the mean. Similarly, for coal retirement, we estimate large scrap values, but these estimates include avoided investments in coal generators that would have been necessary to keep these generators running (e.g., mercury abatement technologies as in Gowrisankaran et al., 2023).

6.2 Counterfactuals

We use our estimates to conduct a series of counterfactuals that illuminate the role of regulation in energy transitions. We first examine how alternative incentives would affect utilities' operations decisions for each counterfactual. We then evaluate the impact of a technological innovation that generates an energy transition by simulating utilities' investment and retirement decisions when they start with the 2006 environment but are suddenly faced with the much lower average 2018–20 natural gas fuel prices.[22]

Table 5 presents the results of our operations counterfactuals. The first row presents outcomes from the baseline model. The next two rows consider outcomes from two versions of the first best solution: a social planner who perceives a cost of CO_2 emissions of \$190/ton and then a cost minimizer (who does not consider carbon externalities).[23] The social planner reduces coal usage from 62% to 3% of capacity—a 95% reduction—by substituting to other sources. Specifically, CCNG capacity usage more than doubles. The cost minimizer also uses substantially less coal than in the baseline, 53% less.[24] This demonstrates that even without a carbon tax, moving to cost minimization over our sample period would significantly reduce coal generation relative to regulation. Consistent with these generation decisions, average

[22]On-line Appendix A4 provides implementation details on the counterfactuals.

[23]We assume that electricity imported from restructured markets has the U.S. average carbon intensity.

[24]For the operations counterfactuals, cost minimization is equivalent to the current regulatory framework with $\mu_2 = 0$.

Table 5: Operations Counterfactuals

	Coal Usage (%)	CCNG Usage (%)	Total Var. Production Costs (Mil. $)	Carbon Costs (Mil. $)	Electricity Revenues ($/MWh)	Variable Profits (Mil. $)
Baseline	61.80	21.66	1,338	5,057	92.62	1,582
Social Planner	2.98	48.94	4,482	3,004	151.30	651
Cost Min., $\mu_2 = 0$	29.32	36.79	1,183	4,050	73.94	1,155
2× Usage Incentive, μ_2	47.44	29.62	1,266	4,575	92.29	1,650
Half TVC Penalty, γ	71.98	16.98	1,382	5,381	95.42	1,597
2× TVC Penalty, γ	51.59	27.01	1,291	4,735	93.40	1,633
Carbon Tax w/ RoR	63.81	31.14	6,661	5,106	238.87	792

Note: Table presents counterfactual simulations of operations decisions at estimated parameter values. The social planner minimizes costs including a $190/ton carbon cost. The cost minimizer has the same incentives but does not value carbon externalities. This is equivalent to current regulation with $\mu_2 = 0$. The next three counterfactuals change regulatory incentives as indicated. The final counterfactual preserves the ROR regulatory structure adding the $190/ton cost to TVC.

annual carbon costs per utility are $5.1 billion in the baseline but fall to $3.0 billion with a social planner and $4.1 billion under cost minimization.

Both the social planner and cost-minimizing outcomes involve far lower utility profits, as shown in the final column of Table 5. Utilities will potentially not cover other costs such as transmission and distribution with these lower profits. This implies that RoR regulation may not be able to directly implement these solutions. We, therefore, compare these outcomes to those of counterfactuals where we keep the current regulatory structure the same but alter parameters that govern utility incentives. This allows us to evaluate the effects of potentially feasible policy changes.

Specifically, rows 4-6 of Table 5 consider regulatory structures that modify incentives to use coal generation and the penalty for high total variable costs. The fourth row doubles the value of coal usage in increasing effective capital. The effect of this counterfactual depends on how it changes each utility's marginal incentive to use coal.[25] Figure A1 shows that

[25]We do not conduct a counterfactual that decreases μ_2 since cost minimization is equivalent to $\mu_2 = 0$.

this marginal incentive is decreasing in coal usage. Doubling this incentive would increase the marginal incentive to use coal at low usage rates but decrease it at higher usage rates. Empirically, we see that the average effect across utility-years in our sample is that doubling this incentive decreases coal usage by 23%. We also simulate changing the penalty for high TVC. Halving the penalty results in 16% more coal usage, while doubling the penalty decreases coal usage by 17% relative to the baseline. None of these three counterfactuals result in outcomes that are close to the social planner or cost minimization levels for operations decisions but all increase utility profits relative to the baseline.

Finally, we consider adding carbon taxes to the existing regulatory structure. In a competitive environment, Pigouvian carbon taxes will yield first-best outcomes. In a regulated environment, however, since utilities typically pass through costs to consumers, carbon taxes will be (at least partially) passed through and may not change generation decisions. In our model, higher costs yield lower rates of return, which means that pass-through will be incomplete. We see that electricity revenues per MWh more than double from \$92.62 in the baseline to \$238.87, which represents a 90% pass-through rate.[26] Despite this change, the carbon tax does not substantially change operations decisions and actually leads to *higher* carbon emissions, increasing coal usage by 3%. This is because the tax increases TVC so much that the marginal incentive to minimize costs is then outweighed by the incentive to use coal to increase effective capital.[27]

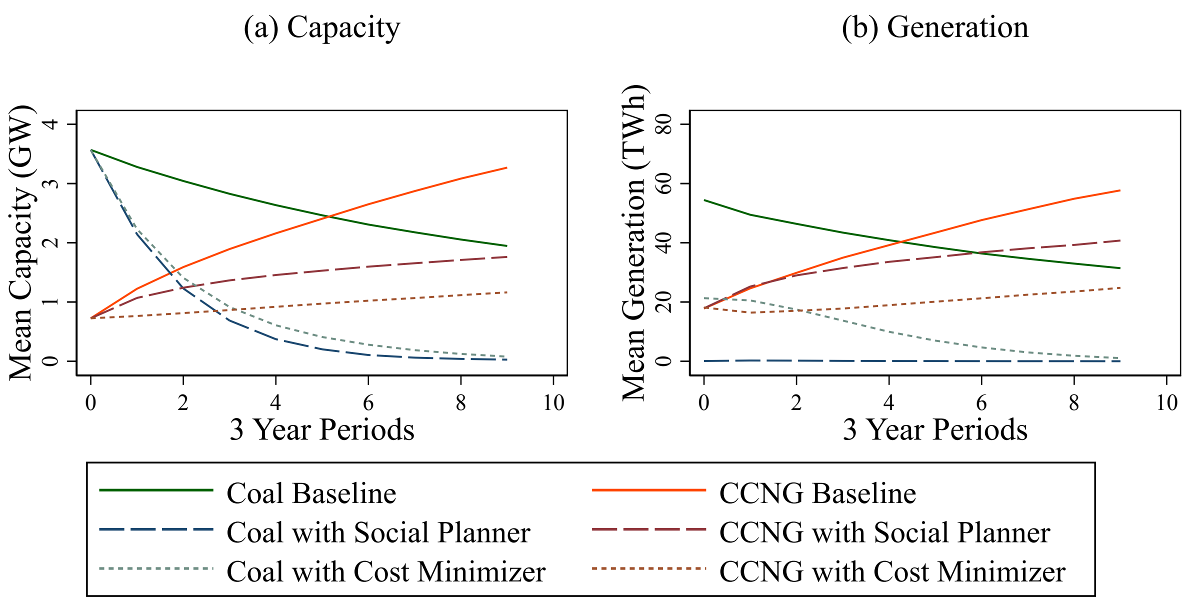
While these counterfactuals highlight the impact of changing regulatory incentives on operations decisions, changing these incentives would also have long-run ramifications for utilities' investment and retirement decisions. In particular, when a utility is facing an energy transition, it will respond with capital investment and retirement. We simulate the effect of an energy transition by calculating 30 years of utility decisions when the utilities in

[26]Carbon costs are \$5.1 billion while average utility revenues (not shown in Table 5) increase from \$2.92 to \$7.45 billion with a carbon tax.

[27]The \$190/ton carbon tax generates costs that are far from the variation in our data. With a \$20/ton tax we find that coal usage and carbon costs decline. With a large carbon tax, there is a question of whether the regulator will update $CostBasis$ to maintain reliability. We assume that $CostBasis$ does not change, so the utility's marginal incentive to reduce costs is attenuated. Since utilities are reimbursed for costs, the short-run disincentive for using more carbon-intensive coal plants is the TVC penalty.

2006 suddenly face the average 2018–20 natural gas fuel price. This allows us to understand how a utility that had adapted to high gas prices, but suddenly faced a future with expected lower gas prices, would respond.

Figure 3: Capacity and Generation for Social Planner and Cost Minimizer



Note: Figures present counterfactual simulations over 10 3-year periods starting with 2006 capacities but imposing 2018–20 natural gas fuel prices. The social planner minimizes costs including a $190/ton carbon cost. The cost minimizer has the same incentives but does not value carbon externalities.

Figure 3 presents the results of our investment and retirement counterfactuals for the social planner and cost minimizer. Panel (a) shows that while utilities in the baseline only retire 45% of coal capacity over 30 years, in both counterfactuals utilities virtually eliminate coal capacity over this horizon. Panel (a) further shows that CCNG investment is higher in the baseline than under both the social planner and the cost minimizer. This is because RoR regulation incentivizes over-investing in capital (Averch and Johnson, 1962). Panel (b) shows that the social planner effectively stops using coal in the first period, while the cost minimizer only reduces coal generation by 61% relative to the baseline. The cost minimizer only approaches the planner level of coal generation at the end of the 30-year horizon. Thus, during the energy transition we study, the primary benefit of carbon taxes relative to market incentives would have been in reducing coal generation rather than encouraging coal retirement.

Our main counterfactuals keep import curves fixed, consistent with a single utility facing alternate incentives. This allows the utility to increase imports when coal generation is costly, for instance when it faces the social planner's incentives. An alternative assumption is that the utility holds constant its import *quantities* at the 2006 baseline level.[28] Figure A2 in On-line Appendix A1 shows the decisions of the social planner and cost-minimizing utility in this environment. Comparing panel (a) of this figure to that of Figure 3, coal retirement decisions look quite similar. However, the social planner invests in CCNG more quickly since it cannot rely on imports to reduce carbon emissions in the short run. Turning to generation, panel (b) of Figure A2 shows that the social planner continues to generate with coal in the short-run since it cannot rely on imports to lower costs and carbon emissions. In the long run, however, investments in CCNG capacity allow the social planner to move completely away from coal generation, as when we allow import quantities to vary.

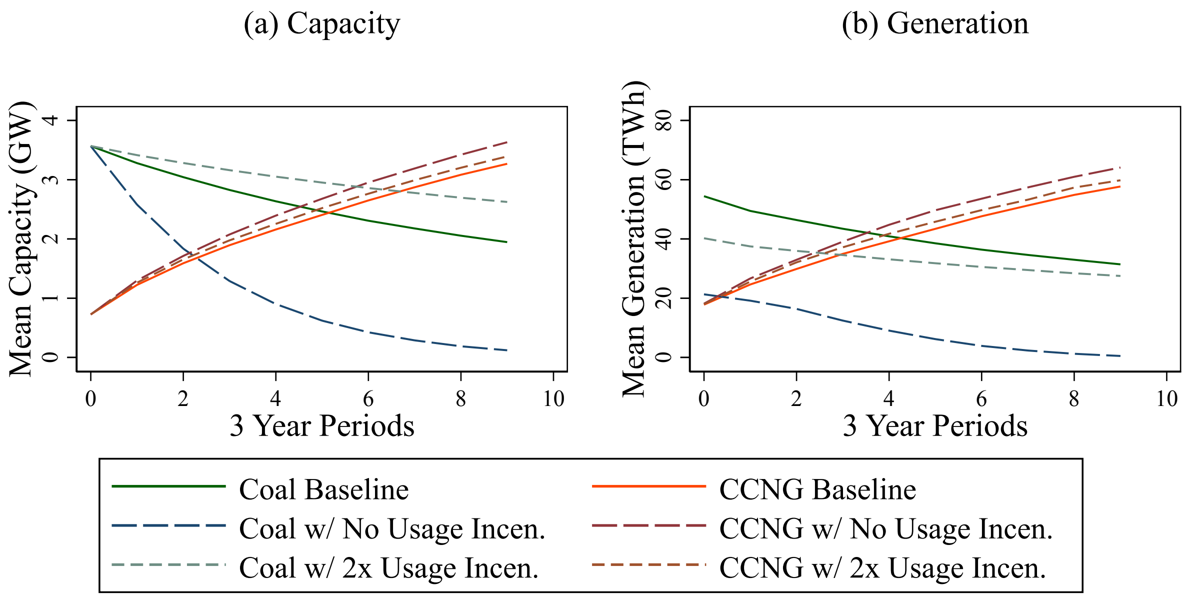
Turning to coal usage incentives, since coal does not contribute as much to effective capital, setting $\mu_2 = 0$ yields a decrease in coal capacity of 97% and coal generation of 98% by the end of the 30-year horizon. However, doubling the coal usage incentive leads to slower retirement of coal capacity, but also lower generation with that capacity, the latter for the same reasons as in the operations counterfactual. In contrast, changing coal usage incentives has very little effect on CCNG investment or usage.

When we evaluate the impact of changing the penalty for high TVC, we see a very different picture. Specifically, we find a large effect of changing this penalty on CCNG capacity and generation, since increasing the TVC penalty will cause the utility to invest in (and use) the newly lower marginal cost CCNG more. We find little effect on coal retirement decisions, but increasing the penalty lowers coal generation by 67% by the end of the horizon. This is because utilities will want to reduce generation with coal in order to lower TVC, but coal capacity still contributes to effective capital.

Finally, Figure 6 compares the baseline to carbon taxes assessed on the regulated utility

[28]Import prices will be different because they will reflect the different natural gas prices.

Figure 4: Capacity and Generation for Different Coal Usage Incentives



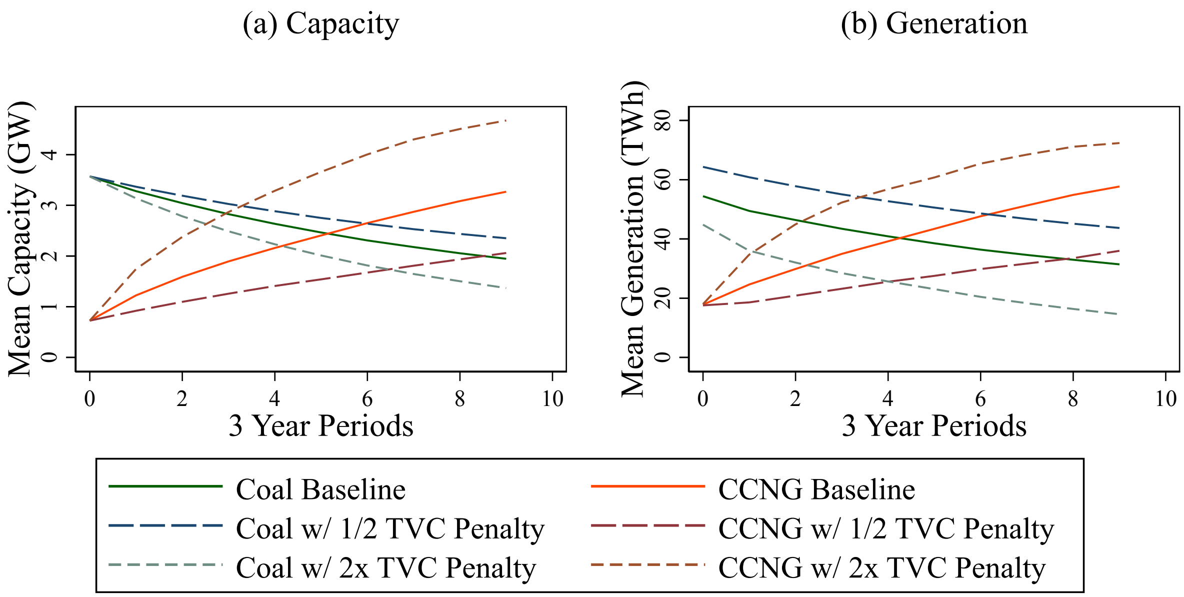
Note: Figures present counterfactual simulations over 10 3-year periods starting with 2006 capacities but imposing 2018–20 natural gas fuel prices. The counterfactuals change the coal usage incentive, μ_2, as indicated.

and the social planner. We find different results in the short run and the long run. While short-run effects are quite small, and even increase coal generation as in Table 5, we find that coal capacity and generation decline substantially in the long run—both by 96%—over the horizon as taxes incentivize coal retirement and CCNG investment. Figure A3 in On-line Appendix A4 shows that this results in roughly 25% below baseline carbon emissions for the social planner, cost minimizer, and regulated utility facing a carbon tax.

7 Conclusion

This paper develops and estimates a model of rate-of-return regulation and analyzes how regulation performs when confronted with an energy transition. The regulator creates an incentive structure that seeks to make electricity reliable and affordable. The utility optimizes against this structure, facing a tension between keeping variable costs low and proving that coal capacity is prudent by keeping its usage high.

Figure 5: Capacity and Generation for Different TVC Penalties

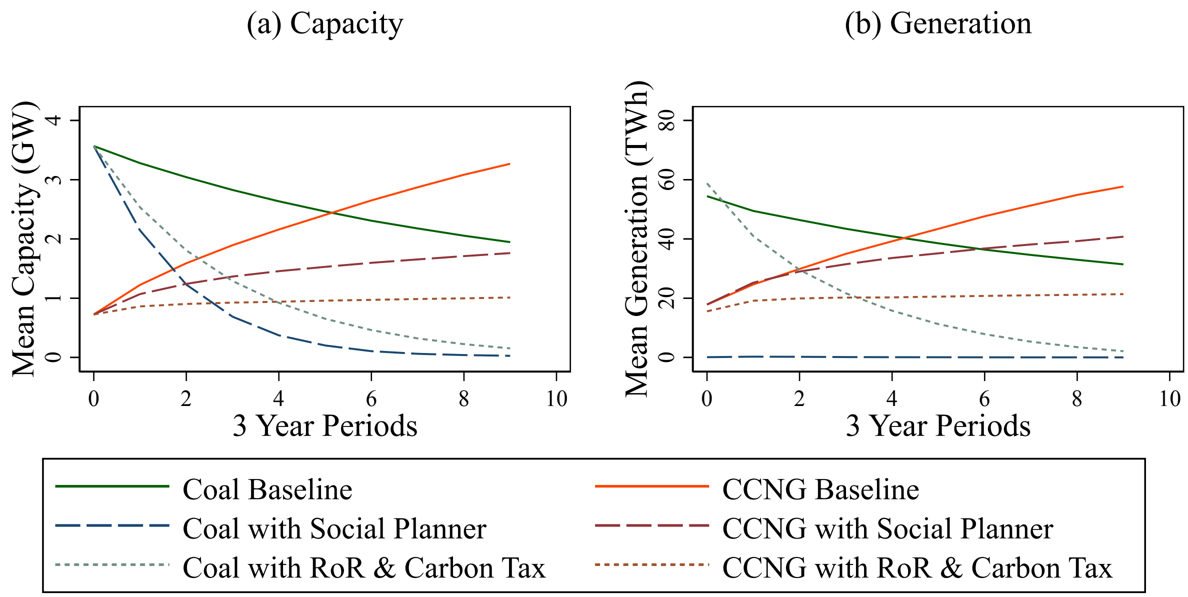


Note: Figures present counterfactual simulations over 10 3-year periods starting with 2006 capacities but imposing 2018–20 natural gas fuel prices. The counterfactuals change the TVC penalty, γ, as indicated.

We find that, in the face of an energy transition, the current regulatory structure is very different from either a social planner or cost minimizer, retiring only 45% of coal capacity in the 30 years after natural gas prices fall. In contrast, the cost minimizer virtually eliminates coal capacity over this horizon, while the social planner—which also internalizes carbon externalities of \$190/ton—essentially stops even *using* coal immediately.

We further find that adjusting the regulatory structure on the margin, by changing the penalty for high variable costs or the incentive to use coal, does not approach the cost-minimizing solution. For instance, increasing the penalty for high variable costs increases combined-cycle natural gas generation, but at the expense of substantial extra coal and combined-cycle capacity. Entirely eliminating the coal usage incentive approaches the cost minimizing coal capacity and generation in the long run, but with much more combined-cycle investment. Adding the \$190/ton carbon tax to existing RoR regulation leads to 90% of the tax just being passed through to customers in the short-run, though it does lead to the elimination of coal capacity over our horizon.

Figure 6: Capacity and Generation for Social Planner, and Carbon Tax with Regulation



Note: Figures present counterfactual simulations over 10 3-year periods starting with 2006 capacities but imposing 2018–20 natural gas fuel prices. The social planner minimizes costs including a \$190/ton carbon cost. The carbon tax counterfactual leaves the regulatory structure unchanged but adds the carbon cost to TVC.

While there are drawbacks to the existing RoR regulation, all three alternatives that lead to the elimination of coal capacity over 30 years—social planner, cost minimization, and a carbon tax—also substantially reduce utility variable profits. These approaches may therefore necessitate lump-sum transfers to maintain electricity reliability by ensuring resources for capacity investment, transmission, and distribution. Consistent with this result, the 2022 Inflation Reduction Act included substantial transfers for clean energy investment rather than carbon taxes.

Finally, our results speak to the importance of regulatory incentives in the *next* energy transition to electricity storage and renewables. Regulatory approaches that resulted in over-investment in combined-cycle capacity may hinder this ongoing transition by requiring ratepayers to fund stranded assets, thereby reducing electricity affordability. Our results further suggest that usage incentives for combined-cycle capacity are likely to further hinder this transition.

References

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

Abito, J. M., C. R. Knittel, K. Metaxoglou, and A. Trindade (2022). The role of output reallocation and investment in coordinating environmental markets. *International Journal of Industrial Organization 83*, 102843.

Aspuru, P. (2023). Delaying the coal twilight: Local mines, regulators, and the energy transition. *Working Paper*. https://pelloaspuru.github.io/papers/main.pdf.

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).

Cross-Call, D., R. Gold, C. Goldenberg, L. Guccione, and M. OBoyle (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, Climate, and Grid Security Subcommittee (2023). Oversight of FERC: Adhering to a Mission of Affordable and Reliable Energy for America. Available at https://energycommerce.house.gov/events/.

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

Energy Information Administration (2022). Cost and Performance Characteristics of New Generating Technologies, Annual Energy Outlook 2022. Available at https://www.eia.gov/outlooks/aeo/assumptions/pdf/table_8.2.pdf.

Energy Information Administration (2023). Carbon Dioxide Emissions Coefficients. Available at https://www.eia.gov/environment/emissions/co2_vol_mass.php.

Environmental Protection Agency (2023a). Greenhouse Gas Equivalencies Calculator. Available at https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator.

Environmental Protection Agency (2023b). Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances. Available at https://www.epa.gov/environmental-economics/scghg.

Fisher, J., A. Armendariz, M. Miller, B. Pierpont, C. Roberts, J. Smith, and G. Wannier (2019). Playing with other peoples 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.

Jha, A. (2023). Dynamic regulatory distortions: coal procurement at us power plants. *Available at SSRN 3330740*.

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. (2014). Complementary bidding mechanisms and startup costs in electricity markets. *Review of Economic Studies 81*(4), 1708–1742.

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.

Rust, J. and G. Rothwell (1995). Optimal response to a shift in regulatory regime: The case of the us nuclear power industry. *Journal of Applied Econometrics 10*(S1), S75–S118.

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. Oboyle, 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

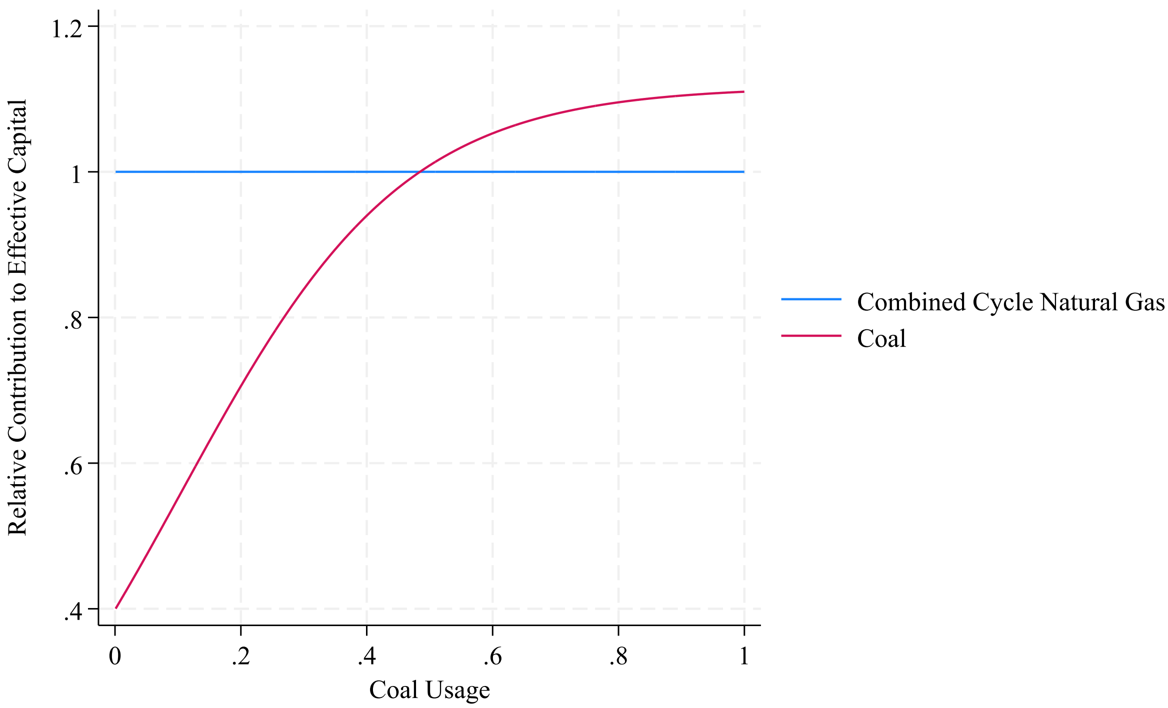
A1 Additional Tables and Figures Referenced in Main Paper

Table A1: Operations Model Fit

	Data	Model
Annual Electricity Production (TWh):		
Coal	16.14	19.43
CCNG	6.93	3.94
Imports	13.04	11.46
Mean Usage Share (%):		
Coal	52.40	61.80
CCNG	35.89	21.66
Annual Costs (Millions of Dollars):		
Coal Fuel	397.83	477.84
CCNG Fuel	230.90	103.69
NGT Fuel	36.32	147.87
Coal O&M	263.93	317.60
CCNG O&M	17.99	10.23
NGT O&M	12.14	37.64
Coal Ramping	15.48	13.76
CCNG Ramping	4.06	3.12
Imports	665.30	226.66
Total Variable Production Costs	1,644	1,338
Electricity Revenues (Dollars/MWh):	65.41	92.62

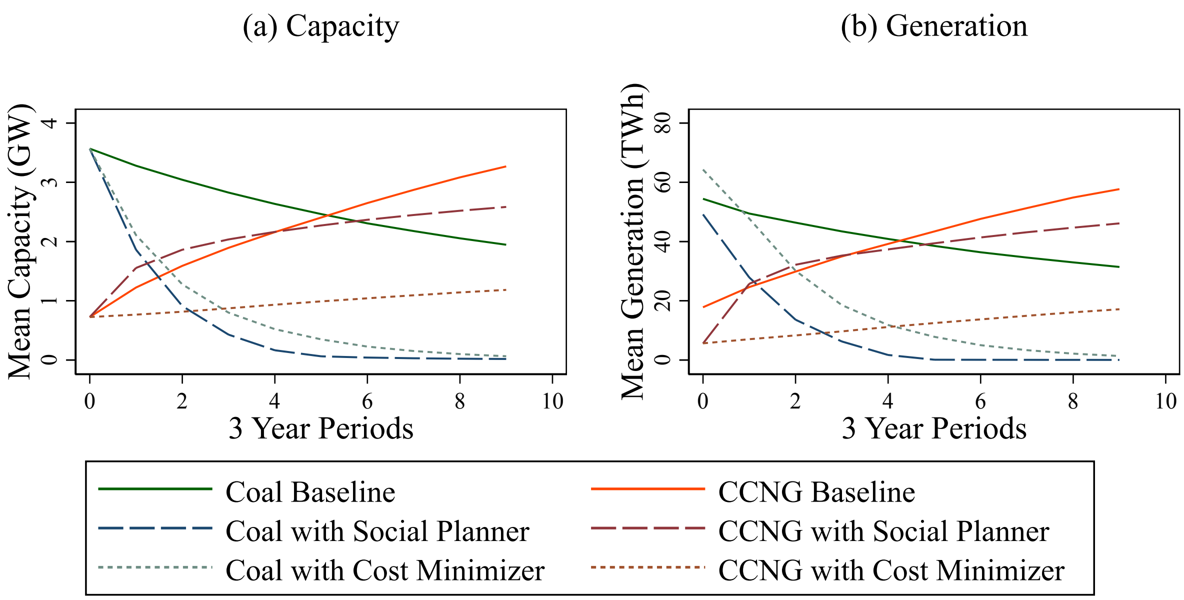
Note: Table presents key outcomes from the data and the model simulated at the estimated parameter values for the analysis sample. In the "data" column, we use observed operations decisions but calculate O&M and ramping costs using estimated parameters and import costs using estimated import supply curves.

Figure A1: Coal Contribution to Effective Capital Relative to CCNG



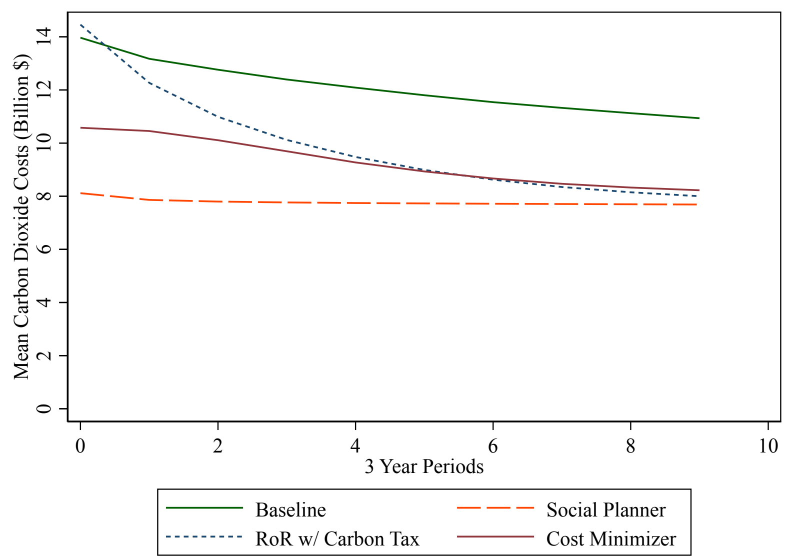
Note: Figure calculated from estimated model coefficients.

Figure A2: Capacity and Generation for Baseline, Social Planner, and Cost Minimizer with Fixed Imports



Note: Figures present counterfactual simulations over 10 3-year periods starting with 2006 capacities but imposing 2018–20 natural gas fuel prices. The social planner minimizes costs including a \$190/ton carbon cost. The cost minimizer has the same incentives but does not value carbon externalities. In both of these cases, we hold hourly imports for each utility fixed at their simulated quantities for the first year the utility appears in the analysis sample.

Figure A3: CO_2 Carbon Costs for Baseline, Planner, Cost Minimizer, and Carbon Tax



Note: The Figure presents counterfactual simulations over 10 3-year periods starting with 2006 capacities but imposing 2018–20 natural gas fuel prices. The social planner minimizes costs including a \$190/ton carbon cost. The cost minimizer has the same incentives but does not value carbon externalities. The carbon tax counterfactual leaves the regulatory structure unchanged but adds the carbon cost to TVC.

Table A2: Indirect Inference Coefficient Matching

Dependent Variable	Regressor	Actual Data	Simulated Data
Usage Variable:			
Coal	Constant	0.524 (0.000)	0.524 (0.001)
CCNG	Constant	0.359 (0.001)	0.166 (0.001)
NGT	Constant	0.087 (0.001)	0.168 (0.001)
Variable Profit Proxy			
	Constant	861.102 (99.895)	1963.903 (122.898)
Rate of Return Proxy			
	Fuel and Import Costs	-26.000 (7.000)	-11.000 (2.000)
Variable Profit Proxy			
	Coal Capacity (MW)	-0.358 (0.060)	0.270 (0.017)
	Coal Capacity x Usage	0.603 (0.110)	0.054 (0.026)
	CCNG Capacity (MW)	0.254 (0.021)	0.269 (0.004)
	NGT Capacity (MW)	0.086 (0.076)	0.541 (0.016)
Log Coal Share			
	First Quintile Coal	0.461 (0.077)	-0.018 (0.070)
	Second Quintile Coal	1.072 (0.077)	1.129 (0.073)
	Third Quintile Coal	1.452 (0.076)	0.884 (0.065)
	Fourth Quintile Coal	1.263 (0.078)	1.852 (0.049)
Log CCNG Share			
	First Quintile CCNG	-2.369 (0.003)	0.000 (0.004)
	Second Quintile CCNG	-1.298 (0.004)	-2.867 (0.002)
	Third Quintile CCNG	-0.708 (0.004)	-2.796 (0.002)
	Fourth Quintile CCNG	-0.294 (0.028)	-1.614 (.)
Ramping:			
Coal Usage	Lagged Coal Usage	0.972 (0.000)	0.979 (0.002)
CCNG Usage	Lagged CCNG Usage	0.968 (0.001)	0.965 (0.000)

Note: Table presents the indirect inference regression coefficients estimated on the actual data and data simulated by the model. Standard errors are in parentheses.

A2 Data

A2.1 Construction of Analysis Data

Our analysis data include information principally from the EIA, the EPA, and FERC. To construct our analysis data, we need to merge together information from these three sources, at the utility-year and utility-hour levels and in some cases also separately by fuel-technologies. The EIA data contain a plant ID and the EPA CEMS data contain a facility ID. We merge these two data sets together using these fields. The EIA Form 861 data contain a utility ID, which we use to collapse the data across generators with the same fuel-technology within the same utility. The FERC Form 714 data include an "EIA code" field, equivalent to EIA's utility ID field, which further allows us to merge these data to the combined EIA/EPA data.

The CEMS data also include information for the state within which each plant is located. We used this information to convert each hour in these data to eastern standard time. In some cases, this required us to approximate the time zone by state; e.g., we assumed that Kentucky is in the eastern time zone and Tennessee is in the central time zone. The FERC data include the time zone at which each utility reports hourly load. We used this reported information to convert each hour in these data to eastern standard time. In some cases, this required us to interpret utilities' responses to the time zone question, e.g., that "CEN" refers to the central time zone. We also converted hours in the FERC data from daylight savings time to standard time.

We deflate all revenues and prices to January, 2006 dollars. We used the CPI net of food and energy as our measure of inflation.

We retain in our sample only those utilities with at least five years of revenue and load data. We also drop three utilities which reported excessive exports or very low capacity.

We collected nodal prices from ISOs and then constructed an average hourly wholesale electricity price for each U.S. state and hour. As we describe in Section A3.1, to estimate import supply curves, we pair these wholesale electricity price data with average daily temperature at the state level from PRISM.[29]

[29]We downloaded these data from Prof. Wolfram Schlenker's website, `http://www.columbia.edu/`

Finally, for our coal and gas fuel price measures, we aggregate the EIA Form 423 information on annual contracted fuel prices by plant to the state-year level by taking the mean, weighting by annual generation at each generator. Using these data at the state-year level captures the opportunity cost of fuel faced by utilities.

A2.2 Summary Statistics on Data

Table A3 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. 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 just over \$20/MW over our sample. 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 41% over our time period. Finally, our data record information on 26 unique utilities. The average annual revenues of these utilities is approximately \$2 billion per year (in January, 2006 dollars), a figure that is consistent across years.

Table A4 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,[30] 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.

~ws2162/links.html.

[30] 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.

Table A3: Summary Statistics from Data at Utility/Year Level

	Overall	2006	2017
Capacity (GW):			
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 ($/MMBtu):			
Coal	2.45	2.02	2.37
	(0.79)	(0.65)	(0.58)
Natural Gas	5.35	7.97	3.12
	(2.27)	(1.02)	(0.42)
Fuel Cost ($/MWh):			
Coal	25.22	20.84	23.99
	(8.11)	(6.19)	(7.33)
CCNG	39.45	64.46	23.15
	(19.14)	(12.65)	(3.75)
NGT	73.05	103.68	88.53
	(56.55)	(25.35)	(186.79)
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 entire 2006–17 period. We report fuel costs conditional on a utility having positive capacity for that fuel-technology. Standard deviations are in parentheses.

Table A4: Summary Statistics from Data at Utility/Hour Level

	Overall	2006	2017
Load Served (GWh):			
	4.16	4.10	4.58
	(5.03)	(5.05)	(5.14)
Production (GWh):			
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 (GWh):			
	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 entire 2006–17 period. Standard deviations are in parentheses.

A3 Details of Estimation

This appendix section details the assumptions underpinning the estimation of our model. We begin with the estimation of import supply curves, then provide information on the estimation of regulatory and operating cost parameters using utilities' operating decisions. Finally, we discuss the estimation of investment and retirement cost parameters.

A3.1 Import Supply Curves

We estimate import supply curves for each utility in each hour in an initial step before these curves enter into the estimation of regulatory and operating cost parameters. 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.

Building on this literature—and important in our context, because the supply curves in exporting regions will change as fuel prices change—we allow the intercept and slope of the import supply curve to vary with the natural gas fuel price:

$$q_{uyh}^{m} = (\psi_{u0} + \psi_{u1} p^{NG}) p_{yh}^{m} + \psi_{u2} p^{NG} + \psi_{u3} X_{uyh} + \varepsilon_{uyh}^{m}. \tag{A1}$$

We allow all parameters to vary across utilities and, 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. 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 (A1) 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}.[31] 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 (A1) 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 (A1), i.e. $\hat{\psi}_{u2}p^{NG} + \hat{\psi}_{u3}X_{uyh} + \hat{\varepsilon}^{m}_{uyh}$ where the hats indicate estimated values.

In a few utility-years, we estimated import supply curves where the slope—of import quantity with respect to import price—was negative and very flat. With those slopes, utilities' profits became implausibly large with exports, which could result in utilities who export unreasonable amounts. To avoid this issue, we limited the slope to be at most -100 when we estimated a negative slope.

A3.2 Operations Decisions

We estimate regulatory and operating cost parameters from utilities' operating decisions by solving for the utility's optimal actions given a state and candidate parameter vector and then running indirect inference regressions on those actions. We then find the parameter vector that matches these indirect inference coefficients to those obtained when the same regressions are run on the data. We present the details of how we solve for utilities' optimal actions before turning to the details of the indirect inference regressions.

To solve for utilities optimal operations decisions, we construct a sample of 8 weeks across

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

the year. This sample includes four two-week spans starting at midnight on February 8th, May 8th, August 8th, and November 8th of each year. We assume that utilities pay ramping costs between hours within these two-week spans but not between the spans.

As discussed in the main text, the utility's Bellman equation in a given hour depends upon four states: (1) cumulative TVC up to that hour, (2) cumulative coal usage up to that hour, (3) lagged coal generation, and (4) lagged CCNG generation. We discretize each of these states into ten bins so that we have $10^4 = 10,000$ states for each utility-hour and interpolate across discretized states. For the cumulative TVC state, we keep track of the average variable cost—so that the state has a similar scale for earlier and later hours of the year—and divide the bins evenly between a minimum marginal cost (defined as 50% of the utility's lowest marginal cost for available fuel-technologies in the year) and a maximum marginal cost (defined as the maximum of \$200/MWh or 150% of the utility's highest marginal cost for available fuel-technologies in the year). For cumulative coal usage, we choose evenly divided bins between 0 and 1. For the lagged generation states, we choose evenly divided bins between zero and the capacity of the respective fuel-technology.

In each hour, the utility chooses its coal and CCNG generation levels, both of which affect the future state. We allow the utility to choose between 10 potential values of each of coal and CCNG generation, for 100 possible generation choices. We define the minimum of these equally-spaced bins as either 500 MWh below the lagged generation for that fuel-technology or zero, whichever is bigger. We define the maximum of the bins as either 500 MWh above the lagged generation for the fuel-technology or the fuel-technology's installed capacity for the utility, whichever is smaller. Thus, we do not allow utilities to ramp or deramp more than 500 MWs per hour, for both fuel-technologies.

For each hourly choice of coal and CCNG generation, the utility meets the remaining load with some combination of NGT or imports. Since these fuel choices do not enter into the utility's end-of-year payoff except through TVC, the utility is incentivized to make the cost-minimizing choice across these options. For each potential choice of coal and CCNG, we find the quantity of imports that sets the price of imports equal to the marginal cost of NGT. We then check whether this choice is feasible, or implies an NGT choice less than 0 or more than NGT capacity. In the first case, we use the computed quantity of imports. In the latter cases, we choose the boundary condition of NGT of 0 or capacity, as this will minimize costs. We then compute variable costs for the hour with this combination of generation and import choices and find the expected continuation value given this choice.

As discussed in the main text, we assume that the utility receives its regulatory profit in the terminal hour. We implement this by scaling annual outcomes (e.g. *Revenues* and fuel and import costs) from the 8 week sample to the annual level by multiplying by the hours in a year divided by the hours in the sample, 8760/1344 for non-leap years and 8784/1344 for leap years. The maximum rate of return that the regulator provides the utility increases as TVC decreases. We limit this rate of return to what it would be if TVC was 10% of $CostBasis$ in order to avoid some utilities choosing to export so much that they reach a negative, and hence unrealistic, TVC.

After solving for the utility's optimal operations decisions, we use these simulated data in our indirect inference regressions. For the regressions using hourly data, we run these regressions on the same 8 weeks of data on which we solve for the utilities' optimal operations decisions. For the regressions run on annual data, we run the regressions on the true annual data and the model-simulated data scaled to the annual level.

We run a total of 10 regressions on both the observed data and the model-simulated data and match 29 coefficients from these regressions. These regressions include:

1. **Scale of Generation:** We run regressions of the hourly utilization (generation divied by capacity) for each fuel-technology on a constant. This yields three regressions, one for each of coal, CCNG, and NGT. We cluster the standard errors of these regressions at the utility level and match the three coefficients on the constants.

2. **Scale of Variable Profit:** We run one regression at the utility-year level of revenues net of fuel and import costs on a constant. We cluster the standard errors of this regression at the utility level and match the coefficient on the constant.

3. **Determinants of Rate of Return:** We run one regression of a proxy for the utility's rate of return on a proxy for total variable costs. Specifically, our dependent variable

is the utility's revenues net of fuel and import costs divided by the utility's total coal, CCNG, and NGT capacity in the year. Our independent variable is fuel and import costs. We include utility fixed effects in this regression, but we only match the coefficient on our TVC proxy, not the fixed effect estimates.

4. **Determinants of Variable Profit:** We run one regression at the utility-year level of a proxy for variable profits (revenues net of fuel and import costs) on the utility's coal capacity, coal capacity multiplied by coal usage rate, CCNG capacity, and NGT capacity. We cluster the standard errors at the utility level and match the three coefficients on capacity and the interaction term.

5. **Usage of Coal and CCNG:** We run two regressions at the hourly level where the dependent variables are the log of coal (or CCNG) generation divided by the sum of coal and CCNG generation in the hour. For the coal regression, the primary dependent variables are quintiles of annual coal utilization across all utility-years where a utility has positive coal capacity and these quintiles interacted with the difference in marginal cost between coal and CCNG. We include analogous regressors in the CCNG generation share regression. We also include utility fixed effects in both regressions. We run these regressions only on hours of the year where the load is between 75% and 125% of the utility's CCNG capacity (for the coal regression) or coal capacity (for the CCNG capacity) in that year. We cluster the standard errors at the utility level. We match the nine coefficients in each regression on the usage shares and their interactions with fuel prices.

6. **Extent of Ramping:** We run two regressions at the hourly level of current coal or CCNG generation on lagged generation with the same fuel-technology. In these regressions we control for the fuel price of coal, the fuel price of CCNG, the current electricity price, current load, and six leads for each of load, the import supply curve intercept, and the electricity price. We include utility, month-of-year, and hour-of-year fixed effects and cluster standard errors at the utility and hour-of-year level. We only match the one coefficient from each regression (two coefficients total) on lagged

generation.

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 inverse of the variance-covariance matrix of the regressions on the actual data above. We assume that there is no covariance across regressions.

A3.3 Investment and Retirement Decisions

We estimate investment and retirement decisions with a nested fixed point GMM estimator that requires solving for the dynamically optimal investment/retirement decisions across states. We compute the optimal investment/retirement decisions with a Bellman equation. After the final decision period, when $t > 10$, the state no longer evolves and the utility no longer makes investment/retirement decisions. Hence, we solve for the value at this state as the discounted flow of π^*, evaluated at the terminal state. We then solve the remaining 10 period problem with backward recursion, starting with the CCNG investment decision for all states at $t = 10$, then the coal retirement decision for all states at $t = 10$, the CCNG investment decision for all states at $t = 9$, etc.

As with the operations decision detailed in Section A3.2, we discretize the state space and compute continuation values by interpolating across discretized states. Here, we use 10 evenly divided bins for each of the time-varying states of coal capacity, CCNG capacity, and natural gas fuel price. Since we consider only retirements for coal, we let the coal capacity bins range from 0 to the observed coal capacity at the beginning of the sample. Since we consider only investment for CCNG, we let CCNG capacity bins range from the observed capacity CCNG capacity at the beginning of our sample to 110% of peak load, defined as the 95th percentile of hourly load. Finally, we let the natural gas fuel price bins range from 75% of the lowest three-year mean fuel price (as described below) to the maximum three-year mean price. Given that there are 10 evaluation time periods and two decisions (investment and retirement) in each period, we solve for continuation values at $10^4 \times 2 = 20{,}000$ states per utility.

Although the value function varies across period t, the period profits, π^* do not, and vary only across utility and the three time-varying states. For each utility i, we calculate π_i^* using the utility's mean NGT capacity and coal fuel price over the sample period. For the utility's fixed states that vary by hour—which are load and the import supply curve residual—we use hours from the first year that utility i is observed in our data, generally, 2006. We use data from one year here rather than using the mean across years to preserve the level of fluctuations that occurs between hours and accurately capture ramping and other costs.

For each of the 20,000 states, we solve for the continuation value using the algorithm developed by Gowrisankaran and Schmidt-Dengler (2023). This algorithm discretizes the continuous (in our case, investment/retirement) decision and requires that we specify the number and values of the discrete levels. Based on our examination of changes in the data, we specify 10 bins for coal capacity change: 0, 200, 400, 600, 800, 1000, 1500, 2000, 3500, and 5000 MWs of capacity retirement and 10 bins for CCNG capacity change: 0, 100, 200, 400, 600, 800, 1000, 1500, 2000, and 3000 MWs of capacity investment. We exclude coal retirement bins that would imply negative coal capacity.

We estimate our GMM objective function with 18 moments. Each of the moments indicates the difference between the estimated model value of a statistic and the value in the data. We use 9 moments each for coal and CCNG decisions. For coal, we include (1) an indicator for positive retirement, this indicator interacted with (2) the quantity of capacity retired and (3) quantity squared, and (4) an indicator for retirement of more than 500 MW. We interact these four moments with coal capacity (5–8) and include (9) the retirement quantity variance. For CCNG, we include the analogous moments, but for investment (10-18).

We estimate an asymptotically optimal GMM weighting matrix by bootstrapping the model moment values across observations and using the inverse estimated variance-covariance matrix as the weighting matrix. We calculate standard errors using the standard GMM formulas. Because of computational complexity, our standard errors for the investment/retirement decisions do not account for the fact that π^* is estimated.

Finally, we estimate the natural gas fuel transitions using a panel of Henry Hub natural gas spot prices as reported by `https://www.eia.gov/dnav/ng/hist/rngwhhdM.htm`. We use data from 2003–20, and let each observation denote the three-year mean price. We then estimate gas price transitions with a simple autoregressive specification of price on lagged price. We take the slope and residual from the regression and discretize quantiles of the prediction to obtain transition probabilities of the natural gas fuel price state from period to period.

A4 Implementation of Counterfactuals

This appendix discusses our implementation of the counterfactuals for both the operations decisions and the long-run decisions that simulate an energy transition. These counterfactuals compare the current regulatory structure to (1) the cost minimization solution and (2) the social planner solution. They also analyze alterations to the current regulatory framework, specifically (3) altering the penalties for high TVC and (4) altering usage incentives. Finally, (5) we simulate counterfactual carbon taxes within the context of RoR regulation.

To simulate operations decisions, we start with each utility-year in our analysis sample and simulate how operations would change under these counterfactual environments. To simulate the long-run decisions under counterfactual environments, we calculate a grid of state-contingent profits π_i^* for each utility i observed in our sample under these environments. Our simulation process then follows the computation described in the estimation of the investment and retirement parameters, in On-line Appendix A3.3, but with different profit grids from those used in the baseline estimation.

We report outcomes for each counterfactual that include information on the generation decisions, carbon externalities, and—in the case of the long-run counterfactuals—coal and CCNG capacity. To calculate the carbon externalities, we multiply the EPA's current carbon cost (Environmental Protection Agency, 2023b) by the carbon intensity of each fuel source in CO_2 tons per MWh. We calculate the carbon intensity of each utility and fuel-technology by multiplying its heat rate, measured in MMBtu per MWh, by its emissions tons per heat input, measured in CO_2 tons per MMBtu. We recover the heat rate by utility and fuel-technology from our analysis data and the emissions per heat input for coal and natural gas from Energy Information Administration (2023). Our reported counterfactual carbon cost measures further account for the carbon intensity of imports. We assume that the carbon intensity for imports is the 2019 national mean carbon intensity of generation, which we calculate from Environmental Protection Agency (2023a). Because we fix the carbon intensity of imports across counterfactual policies, the carbon impact of these policies most accurately indicate the impact of a policy change affecting a *single* utility.

Finally, we discuss our implementation of each of the five types of counterfactuals.

1. **Cost Minimization.** For the operations model, cost minimization is equivalent to the current regulatory problem with μ_2 set to 0. This is because, without usage incentives, the regulated utility is incentivized to minimize operations costs. In the long run, however, investment/retirement decisions will differ between the cost minimization solution and the current regulatory framework with $\mu_2 = 0$. To solve the cost minimization solution for the energy transition, we maximize a value function where the period objective is the negative of total cost, rather than profits.

2. **Social planner.** The social planner in our model seeks to minimize the expected discounted costs of electricity production plus the CO_2 externality from this production. Thus, we compute the social planner solution in the same way as the cost minimization solution, except that we subtract from the criterion function the social cost of carbon from generation with each of the fuels and from imports.

3. **Altering the penalties for high TVC.** We consider counterfactuals that halve or double γ, which indicates the penalty that a high TVC gives the utility in terms of a lower maximum rate of return $\overline{s}$. An increase in γ implies both a steeper drop in profits from higher TVC and a drop in profits overall. For these counterfactuals, we would like to study the impact of changing the slope of profits with respect to TVC rather than the level. Thus, for these counterfactuals, we also adjust α—which indicates the rate base per MW of effective capital—to keep mean variable profits across utility-year observations at the baseline operating decisions the same as in the baseline.

4. **Altering usage incentives.** We consider counterfactuals that eliminate or double the logit slope μ_2 on coal capacity's extra contribution to the rate base with additional usage. As noted above, $\mu_2 = 0$ is equivalent to cost minimization in operations decisions. However, it is different in its long-run implications and hence we report the impact of an energy transition separately for cost minimization and for the regulatory framework with $\mu_2 = 0$.

5. **Carbon tax within existing regulatory structure.** We assume here that utilities are charged a carbon tax of \$190/ton on both their generation and imports. Because this carbon tax then enters TVC, utilities can partly increase consumer rates in response. The penalty for high TVC limits their ability to fully pass through this tax, and creates long-run incentives to move to low-carbon fuel sources. Our carbon tax counterfactual results account for these mechanisms.