International Journal of Industrial Organization 83 (2022) 102843

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

The role of output reallocation and investment in coordinating environmental markets[☆]

Jose Miguel Abito[a,*], Christopher R. Knittel[b,e], Konstantinos Metaxoglou[c], André Trindade[d]

[a] *Ohio State University, United States*
[b] *Shultz Professor of Applied Economics, MIT and NBER, United States*
[c] *Carleton University, Canada*
[d] *FGV EPGE, Brazil and Amazon, United States*
[e] *MIT Sloan School of Management and NBER, United States*

ARTICLE INFO

Article history:
Received 19 May 2020
Revised 2 November 2021
Accepted 11 April 2022
Available online 26 April 2022

JEL classification:
L1
L5
L9
Q4
Q5

Keywords:
Emissions markets
Electricity markets
Uncoordinated regulation
Optimal investment
Empirical games

ABSTRACT

We examine the inefficiency of uncoordinated environmental regulation of CO_2 emissions from electricity generation for a large regional U.S. wholesale electricity market that spans multiple states. We estimate a dynamic structural model of production and investment to compare the social welfare of two counterfactual regulatory scenarios. In both scenarios, emissions targets set by the regulator are met via endogenously determined CO_2 prices in markets aiming to correct the externality. In the first scenario, the CO_2 prices are state-specific. In the second scenario, there is a regional CO_2 price. According to our social welfare estimates, the inefficiency of uncoordinated regulation is mitigated substantially because of the firms' participation in an integrated product market in two main ways. First, firms reallocate output from states with high CO_2 prices to states with low prices. Second, this reallocation spurs investment in cleaner capacity that is exempt from CO_2 regulation. Our finding regarding the mitigation and, potentially, elimination of the inefficiency is robust to alternative models of optimal investment behavior.

1. Introduction

Product markets typically fail in the presence of externalities. Since the work of Coase, economists have known that one solution to an externality problem is to establish property rights and let agents negotiate. In the presence of hetero-

☆ Abito received funding for this project from the Kleinman Center for Energy Policy and from the Dean's Research Fund, at the Wharotn School, University of Pennsylvania. We thank seminar participants in numerous conferences during 2016–2019 for many useful comments. We also thank Erin Mansur for providing some of the data used in the paper, Mushin Abdurrahman from PJM for sharing details about PJM's studies of the Clean Power Plan, as well as Aviv Nevo and Uli Doraszelski for valuable feedback in earlier versions of the paper. The comments of two anonymous referees helped us to significantly improve the current draft. Any remaining errors are ours.

* Corresponding author.
E-mail addresses: abito.1@osu.edu, abito@wharton.upenn.edu (J.M. Abito), knittel@mit.edu (C.R. Knittel), konstantinos.metaxoglou@carleton.ca (K. Metaxoglou), andre.trindade@fgv.br (A. Trindade).

https://doi.org/10.1016/j.ijindorg.2022.102843
0167-7187/

geneous agents, market-based mechanisms (henceforth, MBMs) have also been shown to provide cost-effective solutions. These MBMs are designed around a price that corrects the externality efficiently.

In the case of pollution, several MBMs, including markets for emissions permits, have been proposed by policy makers. A market for permits is an efficient way to correct the negative externality accounting for heterogeneous marginal abatement costs. In the textbook setting, a single price-oriented MBM that corrects the externality, such as a market for permits, maximizes the gains from trade.[1] In practice, organizing a single MBM often requires coordination across multiple jurisdictions, which is generally difficult due to differences in their preferences and priorities. At best, jurisdictions implement separate and often uncoordinated MBMs. For example, 27 jurisdictions had implemented or were scheduled to implement a form of a trading system for carbon emissions in 2019. Most of these trading systems were organized at the provincial or city level, and they were not linked to each other.[2] In the U.S., attempts to address greenhouse gas emissions at the federal level have largely failed. Unless new legislation is passed, future attempts will most likely be at the state level, which will result in uncoordinated MBMs.[3]

The goal of the paper is to measure empirically the outcomes of uncoordinated MBMs and compare them to those of a single MBM regulating emissions in order to quantify potential inefficiencies using social welfare as our metric. Our key insight is that firms invest more to increase their production capacities with uncoordinated MBMs, which mitigates, and potentially eliminates, the inefficiencies resulting from the lack of coordination of regulatory actions. In particular, given that the lack of coordination leads to spatial dispersion in prices aiming to correct the externality, there are locations where it is cheaper to emit. All else equal, profit-maximizing firms reallocate production, along with emissions, towards these locations until it is no longer profitable to do so or capacity constraints in the low-priced locations prevent them from doing so. These constraints spur investment in new capacity that is cleaner and is exempt from CO_2 regulation. Because of the dispersion in CO_2 prices, the incentive to invest in new capacity is stronger in the case of uncoordinated MBMs.

The regulation of CO_2 emissions from electric power plants is our empirical setting. The electricity sector is a major source of CO_2 emissions, just behind transportation, and has been the direct target of efforts to curb emissions. We focus on power plants in the Pennsylvania–New Jersey–Maryland (PJM) Interconnection, which is the world's largest wholesale electricity market spanning multiple U.S. states. PJM is an ideal setting to study inefficiencies with uncoordinated regulation of markets for CO_2 emissions because there is considerable heterogeneity in the plants' fuel mix and emissions resulting in substantial heterogeneity in the stringency of emissions regulation across states. In the presence of such heterogeneity, a regional CO_2 market allows gains from trade.

Using rich plant-level data, we set up and estimate a dynamic structural model of production and investment preserving the heterogeneity both within and across firms. We simulate the model and estimate social welfare for two counterfactual scenarios of CO_2 regulation. In both scenarios, emissions targets set by the regulator are met via CO_2 prices, namely, the shadow prices of the emissions constraints. We use the CO_2 emissions targets from the Clean Power Plan (CPP) proposed by the Obama Administration in August 2015. Although the CPP was repealed by the Trump Administration, which proposed the Affordable Clean Energy (ACE) Rule in August 2018 to replace the CPP, it is the most recent example of comprehensive CO_2 emissions regulations by the U.S. Environmental Protection Agency under the Clean Air Act. In the first scenario, there are separate (state-specific) CO_2 markets that correct the negative externality. In the second scenario, there is a single regional (PJM-wide) CO_2 market. In the first scenario, each state is required to keep CO_2 emissions below its specified target. In the second scenario, regional compliance requires the sum of CO_2 emissions across states to be below the sum of the states' targets. Hence, in the first scenario, the market participants pay CO_2 prices that depend on the plants' location. In contrast, the market participants in the second scenario pay a single CO_2 price regardless of the plants' location.

Given the heterogeneity in plant characteristics and regulatory stringency across states, state-specific CO_2 prices will most likely differ. In this case, the marginal abatement costs of plants are not equal. Based on insights from the trade literature (e.g. Samuelson, 1948; Mundell, 1957), if CO_2 emissions are treated as a factor of production, inefficiencies from the lack a single CO_2 market may not be large if firms participate in an integrated product market and spatial reallocation of production (hence, emissions) is possible. In our setting, although power plants are located in different states and are subject to different CO_2 prices, they supply electricity to an integrated product market (the PJM wholesale market). All else equal, profit-maximizing firms move production from plants in states with high CO_2 prices to states with low CO_2 prices. In fact, the larger the spatial heterogeneity in CO_2 prices, the stronger the incentive to reallocate production. In the absence of any frictions in output reallocation and CO_2 price adjustment, a firm continues to reallocate output until there is no difference in CO_2 prices. However, frictions that impede reallocation and sustain the inefficiency due to separate CO_2 markets are possible. One class of important frictions that we focus on are capacity constraints.[4]

[1] If the marginal damage from the externality differs across jurisdictions, such as in the case of NO_x emissions (Muller and Mendelsohn, 2009; Fowlie and Muller, 2019), having separate but coordinated MBMs is ideal from a welfare standpoint. However, the benefit of having different NO_x prices seems to be of second order relative to having correct expectations on what abatement costs will be in the future (Fowlie and Muller, 2019; Holland and Yates, 2015). For uniformly mixed pollutants like CO_2, the damage depends on the total emissions entering the atmosphere and not the location of the source.

[2] Among these jurisdictions, one is supranational (European Union), four are national (China, Colombia, New Zealand, Switzerland), fifteen are provinces and states, and seven are cities. See https://bit.ly/3j38Qf0 for an updated interactive map of emissions trading systems in force, scheduled or under consideration at the national and subnational levels.

[3] Absent new legislation, efforts to regulate CO_2 will fall under the purview of the Clean Air Act (Goulder and Stavins, 2010). The Act authorizes the Environmental Protection Agency (EPA) to set state-level targets and solicit state implementation plans to achieve these targets. While the EPA can encourage coordination among states, it does not have the power to force them to do so.

To understand the role of capacity constraints, we first examine the difference in social welfare between a single and separate CO_2 markets assuming investment in capacity is exogenous. We refer to this difference in social welfare as *static* inefficiency. Next, we investigate whether optimal investment is different in the two regulatory scenarios. The comparison of social welfare in this case takes into account potential differences in the incentives to invest, and, therefore, differences in capacity levels. The difference in social welfare between the two scenarios that accounts for optimal investment in capacity is a measure of what we refer to as *dynamic* inefficiency.

Our results show that the static inefficiency, namely the reduction in social welfare due to separate CO_2 markets is small, and it is completely eliminated for sufficiently large capacity levels. In this latter case, the capacity constraints no longer bind and, hence, frictionless production reallocation is possible. As a result, the social welfare with separate CO_2 markets is qualitatively similar to the social welfare with a single CO_2 market.

Because the size of the inefficiency depends on the level of capacity, we examine the firms' level of investment in capacity for each regulatory scenario. When we allow for optimal investment, which is exempt from CO_2 regulation, the capacity with separate CO_2 markets exceeds its counterpart with a single CO_2 market. Given that investment in new capacity facilitates reallocation of production when capacity constraints bind, there is an additional benefit from investment in new capacity, which is increasing in the dispersion of CO_2 prices. This additional incentive to invest is not present in a single CO_2 market because there is no dispersion in CO_2 prices. In our simulations, the additional investment with separate CO_2 markets is sufficiently large to cause the generation cost to fall below its level with a single CO_2 market, eliminating the dynamic inefficiency. Moreover, and depending on how we model optimal investment behavior, the decline in the generation cost with separate CO_2 markets may exceed the increase in investment cost. As a result, the cost—generation plus investment— may actually be lower with separate CO_2 markets.

Related literature Our paper is related to the literature on incomplete regulation. Incomplete regulation occurs when there is no uniform adoption of regulations across jurisdictions, hence, exempting a subset of polluting sources from regulation. As such, separate and uncoordinated markets to correct the externality can be a direct consequence of incomplete regulation. The literature has focused on the important problem of emissions leakage whereby firms relocate production (and emissions) to unregulated jurisdictions, which reduces the efficacy of the regulation (see, e.g., Fowlie, 2009; Fowlie et al., 2016). A similar form of leakage occurs when firms face overlapping state and federal regulations in only a subset of states and state regulations are stricter than federal ones (Goulder et al., 2012; Goulder and Stavins, 2010).[5]

We contribute to the literature by illustrating the importance of thinking about adjustments that may not happen in the short-run and showing how ignoring these adjustments may overstate problems with incomplete regulation.[6] Although the short-run inefficiency with uncoordinated regulation can be substantial, what drives the inefficiency—different state-specific CO_2 prices in our case—can actually encourage more investment. More investment in turn mitigates the inefficiency due to lack of coordination. In fact, if there are distortions that weaken the incentives to invest (e.g. strategic capacity withholding and lax environmental regulations), the long-run social welfare may end up being higher.

The paper is also germane to the literature that investigates the interaction between environmental regulation and other forms of regulation and market structure.[7] Our work is most related to Ryan (2012) and Fowlie et al. (2016) who build a Markov Perfect Nash Equilibrium (MPNE) framework and use a two-stage estimation method based on Bajari et al. (2007) to study the effects of environmental regulation in an oligopoly setting.

In terms of estimation, we follow Ryan (2012) and Fowlie et al. (2016), and adapt their methods to the electricity industry and our institutional setting. One important departure from their work is that we only estimate investment costs because we compute production costs directly from data on plant-level heat rates, emission rates for various pollutants, and other operations-and-maintenance related costs (e.g., Mansur, 2007; Bushnell et al., 2008; Gowrisankaran et al., 2016). Furthermore, our approach for the counterfactual simulations differs significantly from Ryan (2012) and Fowlie et al. (2016) in two ways. First, the computation of the stage-game equilibrium is more involved because we find the set of prices that simultaneously clear a large number of (CO_2 and the electricity) markets. Second, given the stage-game equilibrium payoffs

[4] An important friction that we abstract from are transmission constraints. Modeling transmission constraints in the context of a dynamic model is computationally difficult. In fact, even papers employing static analyses (e.g., Bushnell et al., 2008) do not model nodal pricing. Nevertheless, given our assumption that new investment is not included in the emissions target but instead must have the best available technology (see Section 2.2), the location of new capacity does not matter in determining equilibrium CO_2 prices. Thus, in principle, new investment can be located where the nodal price is close to the average (across nodes) price without affecting our results.

[5] Bushnell et al. (2017b) study differences in regulatory environment across states resulting from lack of coordination and strategic policy choice. They study a state-level policy choice in the context of the CPP: whether to implement a mass- or a rate-based target. They show that states can strategically choose between these two policies in a way that leads to lower welfare and increased emissions (due to leakage) highlighting the importance of coordinating regulation. In contrast, we take a step back from the specific design of the policy, and focus on the question of single versus separate markets.

[6] There are actually two adjustments in the paper: firms adjust their capacity to facilitate further reallocation when capacity constraints bind and CO_2 prices adjust across markets as production (and emissions) is moved from one location to another. By assuming implementation of CO_2 regulations via markets, we are, in a way, assuming that states without any regulation—as in the typical case of incomplete regulation—will eventually impose one as emissions are dumped into the state. One example of this type of regulatory adjustment is how California's more stringent fuel efficiency standards eventually led to the adoption of similar standards by other states.

[7] Recent papers in this literature include Fowlie (2010) on the interaction of the NO_x Budget Program with rate-of-return regulation, Cicala (2015) and Abito (2020) on the interaction between the Acid Rain Program and agency problems, Davis and Muehlegger (2010) on U.S. natural gas distribution, Hausman and Muehlenbachs (2019) on methane leaks, Ryan (2012) on industry concentration and the Clean Air Act Ammendments, and finally Fowlie et al. (2016) on the interaction of market power, industry dynamics and market-based mechanisms to limit CO_2 emissions.

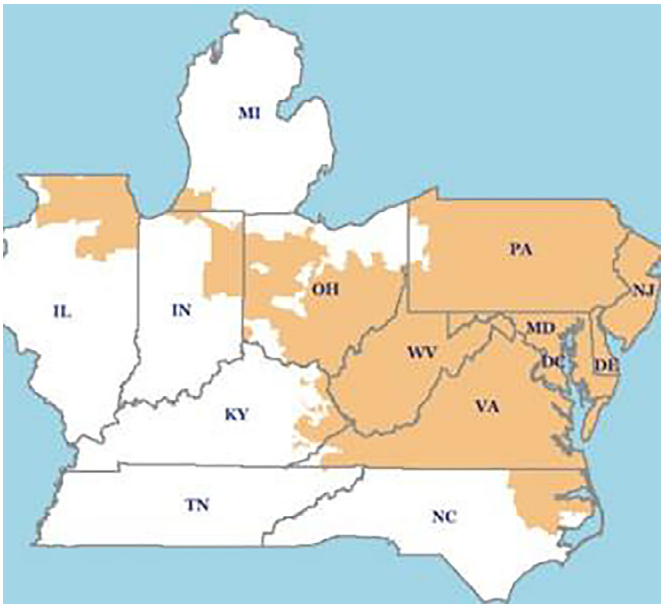


Fig. 1. Area covered by the Pennsylvania–Jersey–Maryland (PJM) Interconnection. *Note:* The figure is from the following link. For additional details, see the discussion in Section 2.1.

for each point in the state space, we solve for the MPNE by combining an Upwind Gauss-Seidel approach (Judd, 1998) with a state-wise Nash Equilibrium (NE) approach. This hybrid approach is less prone to convergence issues because finding the MPNE reduces to sequentially solving the NE of normal-form games in the state space.[8]

Finally, our paper contributes to the empirical literature on electricity markets. Most of the literature has focused on firms exercising market power through strategic bidding and withholding of existing capacity—see Green and Newbery (1992) and Wolfram (1998) for early contributions, and more recently, Borenstein et al. (2002); Hortacsu and Puller (2008); Mansur (2007), and Bushnell et al. (2008). In contrast to these papers, we model strategic investment in new capacity, which has only received limited attention (e.g. Bushnell and Ishii, 2007).

Roadmap The remainder of the paper is organized as follows. Section 2 gives some background for our empirical setting. We present our model in Section 3, followed by a discussion of estimation and results in Section 4. Section 5 contains our counterfactual analysis. We finally conclude in Section 6. We relegate some additional material to the online Appendix.

2. Background

2.1. The PJM electricity market

The Pennsylvania–New Jersey–Maryland (PJM) Interconnection operates the world's largest wholesale electricity market as the regional transmission organization for all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia and the District of Columbia (Fig. 1). PJM coordinates the buying, selling and delivery of wholesale electricity through its Energy Market which began operations in 1997. As the market operator, PJM balances the needs of buyers, sellers, and other market participants, and monitors market activities. As of December 31, 2012, PJM had installed generation capacity of about 182,000 megawatts (MW) and a peak load close to 154,000 MW.[9] Between 2003 and 2012, the value of transactions in PJM's real-time energy market grew from approximately $13 billion to $26 billion (Table A1). Total billings in 2012 were close to $29 billion.

Table 1 shows installed capacity by source using data from the PJM State-of-the-Market (SOM) reports for 2005–2012.[10] Total capacity increased from 163,500 MW in 2005 to 182,000 in 2012 primarily because of the expansion in PJM's footprint. During the same time, coal-fired capacity increased from 67,000 MW to 76,000 MW, while gas-fired capacity increased from 44,000 MW to 52,000 MW, respectively. Averaged across years, the two fuels combined account for 70% of the total capacity,

[8] Chen et al. (2009) compute the MPNE in their game with network effects by solving a two-stage subgame of compatibility and pricing at each given state. Doraszelski and Escobar (2010) characterize an MPNE as the NE of normal form games (one for each state) to apply an analogous purification argument as Harsanyi (1973). Abito and Chen (2021) use a similar trick to construct bounds in the context of supergames.

[9] See Table 1-1 in Volume 1 of the State-of-the-Market report for 2013.

[10] See https://bit.ly/3yifOTJ.

Table 1
Electricity generating capacity by source.

Year	Coal	Gas	Nuclear	Oil	Hydro	Solid waste	Wind	Total
(a) MW (thousands)								
2005	67.8	45.0	31.2	11.8	7.0	0.5		163.5
2006	66.5	47.0	30.0	10.7	7.1	0.6		162.1
2007	66.2	47.6	30.9	10.6	7.4	0.7	0.2	163.5
2008	66.9	48.1	30.4	10.7	7.4	0.7	0.3	164.3
2009	68.1	48.9	30.8	10.7	7.9	0.7	0.7	167.3
2010	67.9	48.5	30.5	10.2	8.0	0.7	0.7	166.5
2011	75.1	50.6	32.6	11.3	8.0	0.7	0.7	178.8
2012	76.1	52.0	32.9	11.5	7.8	0.7	0.7	182.0
(b) MW (%)								
2005	41.5	27.5	19.1	7.2	4.3	0.3		100
2006	41.0	29.0	18.5	6.6	4.4	0.4		100
2007	40.5	29.1	18.9	6.5	4.5	0.4	0.1	100
2008	40.7	29.3	18.5	6.5	4.5	0.4	0.2	100
2009	40.7	29.2	18.4	6.4	4.7	0.4	0.4	100
2010	40.8	29.1	18.3	6.1	4.8	0.4	0.4	100
2011	42.0	28.3	18.2	6.3	4.5	0.4	0.4	100
2012	41.8	28.6	18.1	6.3	4.3	0.4	0.4	100

Note: The table is constructed using information from the PJM State of the Market reports available here. For additional details, see Section 2.1.

with coal accounting for 40% and gas accounting for the remaining 30%. Nuclear's share of total capacity is 18.5% while that for oil is 6.5%. The remaining sources—hydro, wind, and solid waste— account for the remaining 5% of total capacity.

2.2. Emissions regulation

The closest that the U.S. has come to regulating CO_2 emissions from existing fossil fuel-fired power plants at the national level was through the Clean Power Plan (CPP) formulated during the Obama Administration. Fossil fuel-fired plants, which are mostly coal- and gas-fired, are one of the largest single source of CO_2 emissions, accounting for about a third of U.S. total greenhouse gas emissions. The CPP called for a 32% reduction in CO_2 emissions from the power sector by 2030 relative to its 2005 levels. Although the Trump Administration repealed the Obama-era rules in June 2019, the CPP still provides a useful example of CO_2 emissions regulation under the Clean Air Act (CAA).

Using the authority given by the CAA, the U.S. Environmental Protection Agency (EPA) finalized two sets of rules aimed to address CO_2 emissions from fossil-fired power plants (EPA, 2015). In this paper, we collectively refer to both sets of rules as the CPP, though technically the CPP refers to the set of emissions targets applied to existing plants (Section 111(d) of the Clean Air Act) while the rules that are applicable to new sources are part of the Carbon Pollution Standard for New Plants (Section 111(b)).

Section 111(b) gives the EPA authority to set standards or emissions limitations on new, modified, or reconstructed plants.[11] Even though the EPA cannot require a specific technology that firms should adopt under Section 111(b), the emission limits set by the EPA in the case of the CPP essentially precluded technologies that would not meet the limit. For example, the final CPP rule specified a limit of 1000 lbs. of CO_2 per MWh for gas-fired plants, which was feasible only for the latest combined-cycle technology. For coal-fired plants, the limit was 1400 lbs. of CO_2 per MWh, which was achievable only with carbon capture and storage technology, a technology that was costly and not widely available.

Under Section 111(d), the CPP established interim and final rate-based (lbs./MWh) and mass-based (short tons) state goals regarding CO_2 emissions. The interim goals were for the period 2022–2029, while the final goals were for 2030. The EPA gave the states the flexibility to develop and implement plans to ensure that power plants in their state—either individually, together, or in combination with other measures—were capable to achieve the interim and final goals.[12]

Table 2 shows the CPP mass-based targets for the eleven PJM states used in our empirical analysis, noting that the targets have been adjusted to account for the fact that only some of the plants located in Illinois, Indiana, Kentucky, and North Carolina, participate in the PJM wholesale market. The first observation regarding this table is the gradual reduction in total emissions (short tons) for all states between the first year and the final year of the CPP. The second observation is the notable heterogeneity in targets across states. For example, in the first year of the CPP, the target for Maryland is 18.2

[11] Units that are built, modified or reconstructed after the prevailing Section 111(d) targets were set are, by statute, classified as "new" if the same targets are in place. For example, in the CPP, the targets were expected to remain at least until 2030. These sources are reclassified as existing only when targets are revised.

[12] To set these targets, the EPA determined the best system of emission reductions (BSER) that had been demonstrated for a particular pollutant and particular group of sources by examining technologies and measures previously used. The BSER consisted of three building blocks: (i) reducing the carbon intensity of electricity generation by improving the heat rate of existing coal-fired power plants, (ii) substituting existing gas-fired generation for coal-fired generation, and (iii) substituting generation from new renewable sources for existing coal-fired generation.

Table 2
Clean power plan mass-based targets (million short tons).

State	2022	2023	2024	2025	2026	2027	2028	2029	2030
DE	5.524	5.355	5.166	5.072	4.971	4.846	4.806	4.762	4.712
IL	32.087	30.907	29.371	28.737	28.050	27.224	26.686	26.102	25.458
IN	30.510	29.389	27.931	27.328	26.676	25.892	25.382	24.829	24.218
KY	14.327	13.793	13.091	12.805	12.494	12.122	11.871	11.598	11.297
MD	18.197	17.518	16.626	16.263	15.869	15.396	15.076	14.730	14.348
NC	1.333	1.286	1.227	1.201	1.174	1.140	1.121	1.101	1.078
NJ	16.678	16.222	15.778	15.519	15.241	14.892	14.858	14.819	14.766
OH	92.147	88.825	84.565	82.775	80.838	78.501	77.061	75.499	73.770
PA	110.196	106.388	101.664	99.598	97.364	94.653	93.188	91.596	89.822
VA	32.341	31.334	30.195	29.638	29.038	28.297	28.040	27.757	27.433
WV	65.266	62.818	59.587	58.277	56.857	55.154	53.986	52.720	51.325

Note: The mass-based targets reported in this table are based on the supporting data file for CPP compliance from PJM (2016) and are based on electric generating units in the PJM footprint for each state noting that PJM covers only parts of Illinois, Indiana, Kentucky, and North Carolina. The rate-based targets reported in panel B are from the Appendix 5-State Goals sheet in CPP State Goal Visualizer spreadsheet. A spreadsheet with the calculation of the mass-based targets was provided to the authors by PJM. For additional details, see the discussion in Section 2.2.

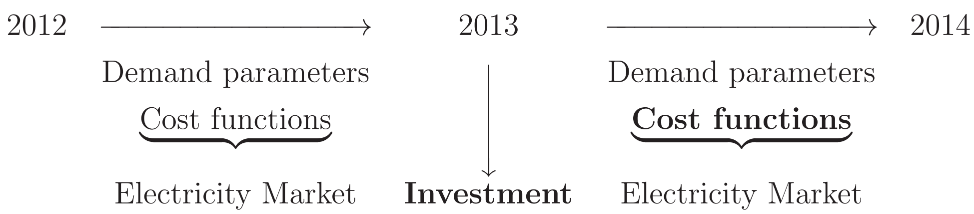


Fig. 2. Overview of the model timing. *Note:* The bold text emphasizes the fact that investment in 2013 affects the cost functions in 2014. For additional details, see the discussion in Section 3.

million short tons, while its counterparts for Ohio and Pennsylvania are 92.1 and 110.2, respectively. This difference in CO_2 emissions reflects the difference in generation from coal, gas, and oil, for the three states in 2012. This baseline generation, in the language of the CPP, is a key component in the calculation of the targets (Table A2). The stringency of the target varies substantially across states and this variation is the source of gains from trade from coordinating separate (state-specific) CO_2 markets.

The different rules for existing and new plants provide two useful modeling shortcuts. First, because only emissions from existing plants are counted against the state-level CO_2 targets, the location of a new plant is irrelevant with respect to the CO_2 price. Location choice for new capacity is an interesting but a complicated problem, especially in our case, where multiple CO_2 markets and the electricity market all have to clear simultaneously. Second, because firms must essentially invest in plants that have the best available technology (BAT), we assume that such investment is in plants that are inframarginal, which helps us reduce the size of the state space.

3. Model

In this section, we present our model of the PJM wholesale electricity market starting with an overview of the timing of the model in Fig. 2. We model the market interaction as a dynamic stochastic game. Firms first decide on whether to invest in new plants. Subsequently, given their existing portfolio of plants, they compete to supply electricity. Each firm owns a portfolio of plants that can differ in the fuel used to generate electricity, capacity, efficiency, emission rates, and location. Investment and supply decisions determine the portfolio of plants and the share of electricity output for each fuel type, which, in turn, determine the level and location of CO_2 emissions. To maintain model tractability, and for data availability reasons, we assume that each power plant has a single electric generating unit despite the fact that power plants may have different units.

We distinguish between two groups of firms in our model. There is a group of strategic firms, which is much smaller than the total number of firms. We treat the rest of the firms as a fringe. We assume that only strategic firms can invest in new plants. The fringe is exogenously endowed with a portfolio of plants that remains fixed throughout the analysis. In what follows, we first describe how we model electricity demand and firms' supply decisions conditional on the portfolio of plants. We then discuss how plant portfolios endogenously evolve because of investment. We close the section with a discussion of the equilibrium notion.

Table 3
List of strategic firms.

Abbreviation	Full name
AEP	American Electric Power
AES	Applied Energy Services
DOM	Dominion
DUKE	Duke
EXE	Exelon
FE	First Energy
GEN	Genon
NRG	NRG
PPL	Pennsylvania Power and Light
PSEG	Public Service Enterprise Group

Note: The table provides the list of the 10 strategic firms in our model. For additional details, see the discussion in Section 3.1.

3.1. Electricity demand

To model demand, we adapt the approach in Bushnell et al. (2008) (henceforth, BMS) using monthly data and a more parsimonious specification. The need for parsimony stems from the fact that we only have 120 monthly observations for 2003–2012, whereas BMS uses roughly 3000 hourly observations. We use fringe supply to refer to the supply subtracted from the vertical inelastic market demand to obtain the residual demand for the strategic firms, which we assume to be the firms listed in Table 3. This fringe supply consists of net imports and the supply of fringe firms. Using t to denote a month in our sample (e.g., Jan-2005) throughout this section, we estimate the following fringe supply function:

$$q_t^{fringe} = \sum_{m=1}^{12} \alpha_m d_{mt} + \sum_{y=2}^{10} \alpha_y d_{yt} + \beta ln(p_t^w) + \mu_1 CDD_t + \mu_2 CDD_t^2 + \mu_3 HDD_t + \mu_4 HDD_t^2 + \varepsilon_t, \tag{1}$$

where d_{mt} and d_{yt} are the fixed effects for month m and year y, respectively. Additionally, p_t^w is the average monthly real-time system-wide locational marginal price in the PJM wholesale electricity market. We proxy for electricity prices in the states surrounding PJM using average cooling (CDD_t) and heating (HDD_t) degree days and their squares accounting for the fact that the PJM footprint expanded during the period in our sample. Finally, ε_t is the econometric error term. Using "hats" to denote estimates and some compact notation we write:

$$\widehat{q}_t^{fringe} = \widehat{\lambda}_t + \widehat{\beta} ln(p_t^w) \tag{2}$$

$$\widehat{\lambda}_t \equiv \sum_{m=1}^{12} \widehat{\alpha}_m d_{mt} + \sum_{y=2}^{10} \widehat{\alpha}_y d_{yt} + \widehat{\mu}_1 CDD_t + \widehat{\mu}_2 CDD_t^2 + \widehat{\mu}_3 HDD_t + \widehat{\mu}_4 HDD_t^2. \tag{3}$$

An estimate of the residual demand $\widehat{Q}_t^S$ for the strategic players is then given by the difference between the total demand Q_t and the estimated fringe supply $\widehat{q}_t^{fringe}$:

$$\widehat{Q}_t^S = Q_t - \widehat{q}_t^{fringe} = \widehat{\alpha}_t - \widehat{\beta} ln(p_t^w), \quad \widehat{\alpha}_t \equiv Q_t - \widehat{\lambda}_t. \tag{4}$$

Seasonal and intraday variation We accommodate seasonal (monthly) variation and distinguish between off-peak and peak periods (intraday variation) in the demand for electricity. Both sources of variation are important for a realistic representation of electricity wholesale markets and are introduced in a parsimonious fashion through the intercept of the residual demand curve. Moreover, we define markets as month and off-peak/peak combinations, and solve for the equilibrium of the stage game for $12 \times 2 = 24$ markets in a given year.[13]

3.2. Firms

3.2.1. Generation cost

Following BMS and Mansur (2007), the marginal cost of generating electricity (\$/MWh) for plant i at time t, which denotes a month (e.g., Jan-2005), is given by:

$$c_{it} = VOM_{it} + HR_{it} \times \left(P_{it}^f + P_t^s r_{it}^s + P_t^n r_{it}^n\right), \tag{5}$$

[13] We follow PJM's peak/off-preak definition. Off-peak is a period of time when consumers typically use less electricity: normally, weekends, holidays or times of the day when many businesses are not operating. PJM typically considers New Year's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving Day and Christmas Day, as well as weekend hours and weekdays from 11 p.m. to 7 a.m. as off-peak. See https://bit.ly/3fc6JEx.

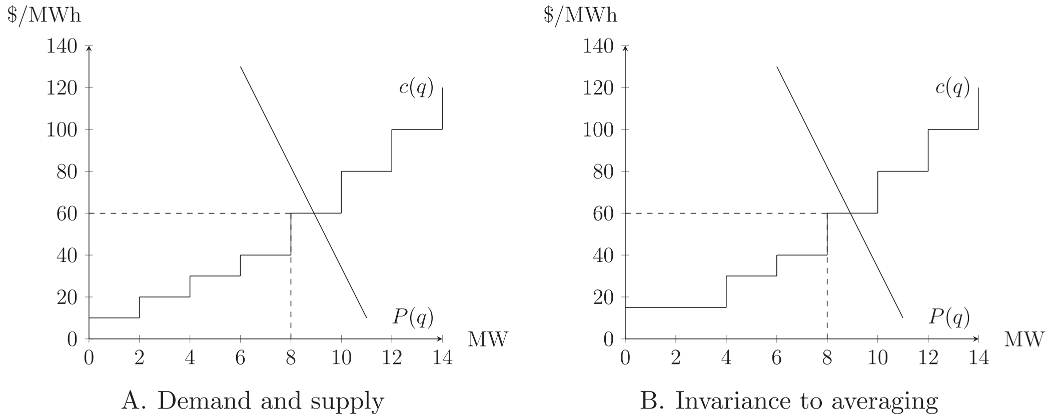


Fig. 3. Equilibrium invariance to averaging. *Note:* Let the first two steps in panel A represent new capacity. Panel A illustrates the market equilibrium when we retain all the information each time we add a new plant. Panel B instead only keeps track of the cumulative size of added capacity and updates a weighted-average cost of these additions. If new capacity is infra-marginal, equilibrium quantities, prices and profits are invariant to averaging of these plants. For additional details, see the discussion in Section 3.2.2.

where VOM_{it} is the variable non-fuel operations-and-maintenance cost ($/MWh), and HR_{it} is the heat rate (MMBtu/MWh) that captures efficiency in turning heat input from fuel to electricity. Additionally, r^s_{it} and r^n_{it} are the fuel-specific SO_2 and NO_x emission rates (lbs./MMBtu), when applicable. Finally, P^f_{it} is the fuel price ($/MMBtu) while P^s_t and P^n_t are the SO_2 and seasonal NO_x permit prices ($/lb.). The VOM costs, the heat rates, and the emission rates, exhibit variation by plant and year. The fuel prices exhibit variation by plant and month. The permit prices exhibit variation by month.

A firm's marginal cost function is a step function with each step representing a plant and it is constructed by ordering its plants in terms of their marginal costs. Because we observe all of the components in (5), we construct each firm's marginal cost curves directly from the data.

3.2.2. Evolution of plant portfolios

Investment affects the shape of the marginal cost function by changing the firm's portfolio of plants. In the beginning of each year, firms choose to invest in coal- or gas-fired capacity. To determine the heat rate of new plants, we rely on Section 111(b) of the Clean Air Act discussed in Section 2.2, which essentially requires that new capacity is of the best available technology (BAT). Hence, we assume that firms invest in plants that have the best (lowest) heat rate in the year the investment takes place.

Aside from simplifying the choice of the plant type firms invest in, the BAT assumption also helps us reduce the dimensionality problem of our model that arises for two reasons. First, we need to take stock of the type of plant firms invest in at each point in time. Second, when evaluating different investment strategies, firms have to be able to compute future profit flows under different investment scenarios involving different paths for their plant portfolio.

Fig. 3 illustrates how the BAT assumption helps us to address the dimensionality problem. Because new plants must have the best heat rate, we assume that they are infra-marginal, at least, in the time horizon. The two lowest steps of the supply curve in panel A of Fig. 3 represent investment in new capacity, while the remaining portion of the supply curve corresponds to existing capacity. Panel A illustrates the market equilibrium when we keep track of all the information about new capacity that the firm invests in. Panel B instead combines the two lowest steps into one. As the example illustrates, it is sufficient to keep track of an average of all the new capacity that the firm invests in because averaging across the individual plants does not affect the equilibrium quantities, prices, and profits. Thus, if new capacity is assumed to be infra-marginal, tracking the firm-level cumulative BAT capacity and the associated weighted average heat rate is sufficient for our empirical analysis.

In more detail, using $f \in \mathscr{F} = \{coal, gas\}$ to denote the fuel type, let i^f_{jt} be the investment by firm j in coal- or gas-fired capacity at time t that denotes a year throughout this section. In addition, let $\underline{K}_{jt}$ be the cumulative BAT capacity at time t, which, following ivestment, gives rise to cumulative BAT capacity at time $t+1$:

$$\underline{K}_{jt+1} = \underline{K}_{jt} + i^{coal}_{jt} + i^{gas}_{jt}. \tag{6}$$

Because the heat and emission rates for coal- and gas-fired capacity are different, we keep track of the share of gas-fired cumulative BAT capacity that evolves according to the equation below:

$$\underline{S}_{jt+1} = \frac{\underline{S}_{jt}\underline{K}_{jt} + i^{gas}_{jt}}{\underline{K}_{jt+1}}. \tag{7}$$

For heat rates, as well as the remaining components of the fuel-specific marginal costs, we track a weighted average at time t. For example, in the case of the heat rate for gas-fired cumulative BAT capacity, we track the following weighted average:

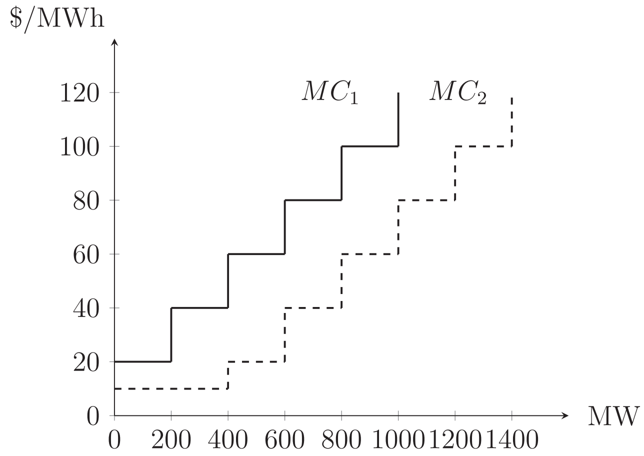


Fig. 4. Updating the marginal cost curve. *Note:* The step function MC_1 (solid line) indicates the marginal cost curve prior to investment and is constructed by ordering available sources to serve demand in terms of their marginal costs. The sources with the lowest (highest) costs are ordered first (last). The step function MC_2 (dashed line) indicates the marginal cost curve following a hypothetical investment of 400 MW in best available technology with a cost of $10/MWh. The vertical distance between the two curves at their origin shows the improvement in marginal costs due to the assumed investment. For additional details, see the discussion in Section 3.2.2.

$$\underline{HR}^{gas}_{jt+1} = \frac{\underline{S}_{jt}\underline{K}_{jt}}{\underline{S}_{jt}\underline{K}_{jt} + i^{gas}_{jt}}\underline{HR}^{gas}_{jt} + \frac{i^{gas}_{jt}}{\underline{S}_{jt}\underline{K}_{jt} + i^{gas}_{jt}} hr^{gas}_{jt}, \tag{8}$$

where hr^{gas}_{jt} is the heat rate for investment in gas-fired capacity.

The marginal cost of the cumulative BAT capacity for firm j at time t is:

$$\underline{c}_{jt} = (1 - \underline{S}_{jt})\underline{c}^{coal}_{jt} + \underline{S}_{jt}\underline{c}^{gas}_{jt}, \tag{9}$$

where $\underline{c}_{jt}$ is computed using (5) for the set of plants operated by firm j. A firm's new marginal cost curve following investment is obtained through a shift of the marginal cost curve prior to the investment as illustrated with the simple example in Fig. 4. *Renewable sources* When making investment decisions, firms take into account the expected evolution of generation capacity from renewable sources. Our model accommodates changes in capacity due to renewable sources in a flexible way through exogenous shifts in the BAT capacity over time. Our model, however, does not allow for investment in renewable sources to respond strategically to changes in the coal- and gas-fired capacity. This assumption is supported by the binding Renewable Portfolio Standards (RPSs) we observe in the data, at least, in the medium run. An RPS mandates that a specific fraction of electricity is generated by renewable sources. With a binding RPS, investment in renewable sources is driven by regulation rather than profit maximization. We collect information on the RPS future mandates for the different states that comprise the PJM market and use these in our simulations.

3.3. Equilibrium

3.3.1. Electricity market equilibrium

To model firms' supply decisions in the wholesale electricity market, we build on the results in Wolak (2000) and BMS. Wolak and BMS show that electricity markets in the presence of forward contracts generate outcomes that are much closer to perfect competition than to an oligopoly (Cournot) game. For PJM, where forward contracts are present, we confirmed the results from BMS by comparing equilibrium prices under perfect competition and Cournot.[14] Perfect competition implied equilibrium prices that were reasonable and consistent with predictions from futures markets, while a Cournot game implied equilibrium prices that were (unrealistically) higher. In our case, forward commitments are not as straightforward to deal with as in BMS because our model is dynamic. One approach to deal with forward commitments, is to assume that they are exogenous and determine their evolution outside the model, or to simply assume they are fixed. An alternative approach is to treat them as endogenous and introduce them in the model. While interesting, modeling the endogenous evolution of forward commitments is beyond the scope of the paper.

Given our assumption of a perfectly competitive wholesale electricity market, the equilibrium electricity price is determined by the intersection of the market demand and supply curves. Market supply is determined by ordering all available

[14] Our assumption of a perfectly competitive market is also supported by the conclusions in the State-of-the-Market (SOM) reports prepared by the PJM Market Monitoring Unit for 2003–2012. The SOM reports analyze competition within, and efficiency of the PJM markets using various metrics, such as market concentration, the residual supply index, and price-cost markups.

capacity in terms of its marginal costs, similar to Fig. 4. This merit order dictates the sequence in which the various plants are dispatched as the demand for electricity increases. The equilibrium wholesale price is the marginal cost of the most expensive plant called to serve demand. As we discussed earlier, given that we observe all of the cost components in (5), we construct the merit order directly from the data.

In our model, investment decisions are strategic. Hence, firms decide on investment considering its impact on other firms, and vice-versa. The assumption of a perfectly competitive wholesale market combined with strategic investment, under the existence of forward commitments, is consistent with theory. For example, Adilov (2012) models firms' investment in capacity in order to study the effects of forward markets on competition and efficiency extending the standard Allaz and Villa (1993) framework. The forward market takes place after the investment decisions are made but before the spot market. Importantly, endogenous capacity choices affect strategic behavior in the forward and spot markets. Outside the setting of electricity markets, Dixon (1985) analyzes a model where the market is competitive but investment is strategic. He finds that, in equilibrium, firms underinvest to drive prices above potential marginal cost, namely, the marginal cost implied by the socially optimal level of investment.

3.3.2. Markov perfect Nash equilibrium

Let $\boldsymbol{a}_{jt} \equiv \{\boldsymbol{q}_{jt}, i_{jt}^{coal}, i_{jt}^{gas}\}$ denote the actions of firm j at time t, and let $\boldsymbol{a}_t$ denote all firms' actions. We use t to denote a year and $\boldsymbol{q}_{jt}$ to denote the 24×1 vector whose elements are the off-peak/peak monthly outputs in the wholesale electricity market in a given year. We use i_{jt}^{coal} and i_{jt}^{gas} to denote annual investment in coal- and gas-fired capacity. The state vector is given by:

$$\mathbf{s}_t \equiv \left(\boldsymbol{\alpha}_t, \mathbf{p}_t^f, \left\{\underline{K}_{jt}, \underline{S}_{jt}, \underline{HR}_{jt}^{coal}, \underline{HR}_{jt}^{gas}\right\}_{j=1}^{N}\right). \tag{10}$$

The endogenous part of the state vector, $\left\{\underline{K}_{jt}, \underline{S}_{jt}, \underline{HR}_{jt}^{coal}, \underline{HR}_{jt}^{gas}\right\}$, is related to annual BAT capacity investment and its evolution was discussed in the previous section. In terms of the exogenous state variables, $\boldsymbol{\alpha}_t$ is a vector of dimension 24×1 whose elements are the off-peak/peak monthly residual-demand intercepts, and $\mathbf{p}_t^f \equiv \{p_{it}^f\}_{i=1}^{\mathcal{N}}$ is a vector of annual averages of $\mathcal{N}$ monthly plant-specific coal and gas prices.[15]

The annual profit function for firm j is as follows:

$$\pi_{jt}(\boldsymbol{a}_t, \mathbf{s}_t, \nu_{jt}) = \mathbf{1}\overline{\boldsymbol{\pi}}_{jt}(\boldsymbol{a}_t, \mathbf{s}_t) - \Gamma_{jt}(\boldsymbol{a}_t, \nu_{jt}), \tag{11}$$

where

$$\overline{\boldsymbol{\pi}}_{jt}(\boldsymbol{a}_t, \mathbf{s}_t) \equiv p_{jt}^r \times q_{jt}^r + p_t^w \times (q_{jt} - q_{jt}^r) - C(q_{jt}, \mathbf{s}_t) \tag{12}$$

is the 24×1 vector of profits from the wholesale electricity market for each off-peak/peak and month combination excluding the annual investment cost $\Gamma_{jt}(\boldsymbol{a}_t, \nu_{jt})$, and $\mathbf{1}$ is a row vector of ones of dimension 1×24. Here, p_{jt}^r is the price the firm receives from retail sales commitments q_{jt}^r, which are assumed to be sunk at the time production decisions are made for the wholesale market, and p_t^w is the equilibrium wholesale electricity price. Finally, $C(q_{jt}, \mathbf{s}_t)$ is the total cost of producing q_{jt} given $\mathbf{s}_t$.

The annual investment cost is given by:

$$\Gamma_{jt}(\boldsymbol{a}_t, \nu_{jt}) = \sum_{f \in \mathcal{F}} (\gamma^f + \nu_{jt}^f) i_{jt}^f, \tag{13}$$

where ν_{jt} is a private shock that is independently distributed across firms and years and it is drawn from a common distribution. The private shock ν_{jt} is the only source of uncertainty in the model because we assume that firms have perfect foresight of other state variables. Hence, there is option value to waiting for favorable investment cost draws before investing. Finally, γ^f is an investment parameter that we estimate.[16]

Firms' strategies depend only on the current state (including the private investment shock) as in Ericson and Pakes (1995). That is, for firm j, strategy σ_j maps the state and private shock into actions. The strategy profile $\boldsymbol{\sigma}$ is a Markov Perfect Nash Equilibrium (MPNE) if each firm j's strategy σ_j generates the highest value among all alternative Markov strategies $\sigma_j^!$ given the rivals' profile $\boldsymbol{\sigma}_{-j}$:

$$V_j(\mathbf{s}; \boldsymbol{\sigma}) \geq V_j(\mathbf{s}; \sigma_j^!, \boldsymbol{\sigma}_{-j}), \tag{14}$$

[15] The vector of monthly SO_2 and seasonal NO_x permit prices is set equal to zero, consistent with the situation in the electric power industry during the time that is relevant for our analysis. Therefore, they are not included in the state vector. Likewise, the remaining components of the BAT cost, such as the VOM cost, are held constant at their current values and, hence, need not be considered in the state vector.

[16] The investment-cost specification in (13) allows only for positive adjustments to capacity. A version of (13) with scrap value would be $\Gamma_{jt}(\boldsymbol{a}_t, \nu_{jt}) = \sum_f 1_{[i_{jt}^f > 0]}(\gamma_1^f + \nu_{1jt}^f) i_{jt}^f + 1_{[i_{jt}^f < 0]}(\gamma_2^f + \nu_{2jt}^f) i_{jt}^f$ as in Ryan (2012). Thus, unlike Ryan (2012) or Fowlie et al. (2016), there is no scrap value from closing down a plant. Given that we do not have fixed costs in our model, the firm will just keep unused plants idle. One may worry that firms choose to divest existing plants as a form of exercising market power (Myatt, 2017). If this is the case, then our results are even stronger. The socially optimal level of (net) investment will likely be higher with separate CO_2 markets than with a single CO_2 market, and, hence, the stronger incentives to invest with separate CO_2 markets are, in fact, welfare-enhancing.

where $V_j(\mathbf{s};\boldsymbol{\sigma})$ is the ex ante—before observing the realization of the private shocks—value function for firm j given by

$$V_j(\mathbf{s};\boldsymbol{\sigma}) = \sum_{t=0}^{\infty} \beta^t E\Big[\pi_{jt}(\boldsymbol{a}_t, \mathbf{s}_t, \nu_{jt})|\mathbf{s}_0\Big], \tag{15}$$

where β is the common discount factor and t denotes a year.

Capacity auctions In our model, the benefits from investment come from the profits firms earn in the wholesale electricity market. However, PJM encourages investment in new capacity through capacity auctions. The motivation for the capacity auctions is adequacy of resources to ensure that the demand for electricity can be met at all times in the near future. Utilities and other electricity suppliers, collectively known as load serving entities (LSEs), are required to have the resources to meet their customers' demand plus a reserve. The LSEs can meet the resource requirement with generation capacity they own, with capacity they purchase from others using bilateral contracts, through demand response—in which customers reduce their usage in exchange for payment—or with capacity obtained through the capacity auctions themselves.

During the period relevant for our analysis (2003–2012), capacity payments have only accounted for 6% of the total wholesale price per MWh when energy payments accounted for 82%.[17] Although the impact of capacity auctions in our estimating sample is small, capacity auctions may increase in importance in future years. Thus, one may be concerned that failure to account for capacity payments may significantly affect firms' investment incentives and our results. While we do not explicitly model capacity payments, our model can allow such payments by reinterpreting the estimated investment cost. With capacity payments, Γ_{jt} becomes the investment cost net of the additional benefit from capacity payments. Moreover, our model can accommodate heterogeneity in capacity payments because of zonal (location-specific) pricing through the private shock ν_{jt}. If this additional benefit is the same with separate and single CO_2 markets, capacity payments will not affect our results. If the capacity markets allow market participants to exercise market power by limiting investment, separate CO_2 markets mitigate such an incentive to underinvest.

4. Estimation

The key components of the structural model that we need to estimate are the fringe supply equation, the electricity generation costs, and the investment cost parameter. We estimate the fringe supply equation using 2SLS, where the monthly quantity demanded serves as our instrument for price based on the idea that short-run wholesale electricity demand is completely inelastic. We compute plant-level generation costs directly from the data as described in Section 3.2.1.

We estimate the investment cost parameter using the two-stage methodology in Bajari et al. (2007). In the first stage, we estimate policy functions from the data using the observed values of the state variables. These policy functions are reduced-form because they provide estimates of parameters that are not primitives of the underlying economic model of investment. In the second stage, we search for the investment cost parameter that best rationalizes the firms' observed behavior and the transitions of the state variables. The advantage of this approach is that the investment cost parameter can be estimated without the need to solve for the equilibrium of the game.

Our application of the two-step methodology follows Fowlie et al. (2016) with some exceptions. First, the lack of exit events and divestment in the data precludes us from obtaining reliable estimates of an exit policy function and scrap values. Second, we do not estimate a band equation in our (S,s) model. Instead, we assume that firms invest or divest to their target capacity if their current capacity is outside a 10% band around the target. Finally, although (in theory) the estimated investment policy function allows for divestment, our model does not have either scrap values or recoverable fixed costs. Hence, we do not model the firms' divestment decisions. Additional details regarding the methodology are available in Section A.4.

4.1. Estimation results

Fringe supply We report estimates for the fringe supply equation in Table 4.[18] The price coefficient, which is of main interest for the subsequent analysis, is generally highly significant in all specifications considered. According to our preferred specification, in which the price enters in logs, the implied elasticity at the sample averages of the fringe supply (MWh) and the wholesale price of electricity ($/MWh) is 0.78.

[17] See Table 9 of the 2012 PJM State of the Market Report Volume I. Modeling firm behavior in the capacity market is beyond the scope of the paper. As background, effective June 2007, the PJM Capacity Credit Market (CCM), which had been the market design since 1999, was replaced with the Reliability Pricing Model (RPM) capacity Market. Under the CCM, LSEs could acquire capacity resources by relying on the PJM capacity market, by constructing generation, or by entering into bilateral agreements. Under RPM, there is a must-offer requirement for existing generation that qualifies as a capacity resource and a mandatory participation for LSEs with some exceptions. LSEs must pay the locational capacity price for their zone and zonal prices may differ depending on transmission constraints. LSEs can own capacity or purchase capacity bilaterally and can offer capacity into the RPM auctions when no longer needed to serve load. Capacity obligations are annual and Base Residual Auctions (BRAs) are held for delivery years that are three years in the future. There are also incremental auctions that may be held for each delivery year if there is a need to procure additional capacity resulting from a delay in a planned large transmission upgrade that was modeled in the BRA for the relevant delivery year. Bushnell et al. (2017a) provide an in-depth discussion of the capacity markets.

[18] We refer the reader to Section A.2 for some additional descriptive statistics.

Table 4
Fringe supply.

Variable	(1) Log	(2) Level	(3) Square root	(4) Cubic root
Price	5,155.3179***	104.3555**	1,581.7544***	4,525.0135***
	(1,630.7993)	(42.4884)	(545.1234)	(1,474.5318)
CDD	−10.7387	−12.5746	−12.9437	−12.5531
	(7.4133)	(9.1130)	(8.3463)	(8.0017)
CDD Sq.	0.0202**	0.0219**	0.0219**	0.0215**
	(0.0087)	(0.0096)	(0.0092)	(0.0090)
HDD	−1.7906	−0.7717	−1.0955	−1.2855
	(2.2626)	(2.5595)	(2.4155)	(2.3623)
HDD Sq.	−0.0002	−0.0011	−0.0010	−0.0008
	(0.0019)	(0.0027)	(0.0023)	(0.0021)
Constant	−2,759.2502	3,350.0882***	925.4219	−1,317.6314
	(2,041.7490)	(583.7189)	(1,076.1755)	(1,679.8604)
Observations	119	119	119	119
R-squared	0.7123	0.6428	0.6678	0.6808
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes

Note: The table presents 2SLS coefficients estimates for various functional form specifications of price using monthly data for 2003–2012 and applying the Prais–Winsten methodology to correct for serial correlation. In all 4 specifications, the dependent variable, fringe supply, is in levels, and we include year and 11 month fixed effects. We use CDD (HDD) to refer to cooling (heating) degree days. The results reported in the paper are based on the log specification reported in column (1). The standard errors in parentheses are also corrected for heteroskedasticity. The asterisks denote statistical significance as follows: 1% (***), 5%(**), 10%(*). For additional details, see the discussion in Sections 3.1 and 4.1.

Table 5
Investment cost of gas-fired capacity in 2010$/MW.

Parameter	Estimate	Std. error
γ^{gas}	1,462,750	33,970

Note: This estimate pertains to γ^{gas} in (13). The reported standard error is calculated resampling moment inequalities from our 2-stage estimator and ignores any 1st-stage estimation error. For additional details, see the discussion in Section 4.1.

Investment cost The estimate of investment cost, $\widehat{\gamma}^{gas}$, reported in Table 5 is in $/MW of gas-fired capacity. Given the lack of investment in coal-fired capacity implied by our model, it is not possible to estimate the cost for coal-fired capacity ($\widehat{\gamma}^{coal}$). Our estimate of around $1.46 million per MW for gas-fired capacity is comparable to the engineering costs in Spees et al. (2011), which are up to $1 million per MW. The reported standard error of approximately $34,000 per MW, which we calculate using 1000 bootstrap replications by resampling from the moment inequalities, does not account for the first-stage estimation error of our 2-stage estimator (Bajari et al., 2013).

4.2. Time paths and model fit

Time paths of exogenous state variables Fig. 5 shows the time paths for various exogenous state variables in the model for 2013–2062. We start by showing the path of the monthly off-peak/peak residual demand intercepts (panel A). We obtain paths of the off-peak and peak monthly interecepts as follows. First, we take the annual average of the monthly intercept estimates in 2012 and apply an increase of 1% per year. Second, we scale the annual average intercept with a factor that reflects the number of off-peak and peak hours in 2012 and subsequently apply a monthly shift. The monthly shift is equal to the coefficient of the appropriate monthly dummy in a time-series regression with 120 observations of the monthly system-wide electricity demand (load) in PJM on year and 12 month fixed effects. Both the scaling factor and the monthly shifts are held fixed in the 50-year window 2013–2062.

The coal heat rates for new investment are assumed to be fixed at their 2012 levels (10 MMBtu/MW), while their gas counterparts are assumed to be falling over time from 7.5 MMBtu/MWh to 6.1 MMBtu/MWh; see panel B. We obtain the gas heat rates of new investment by projecting the linear trend of the log gas BAT heat rates for 2003–2012 to 2013–2062. The VOM costs and CO_2 rates, which are relevant in our counterfactual simulations discussed in a subsequent section, are held constant at their 2012 levels.

In the case of coal prices, we extrapolate the EIA annual projections for 2013–2035 from the 2012 Annual Energy Outlook reference case to 2062 using the implied compound annual growth rate (panel C). For gas prices, we use monthly NYMEX

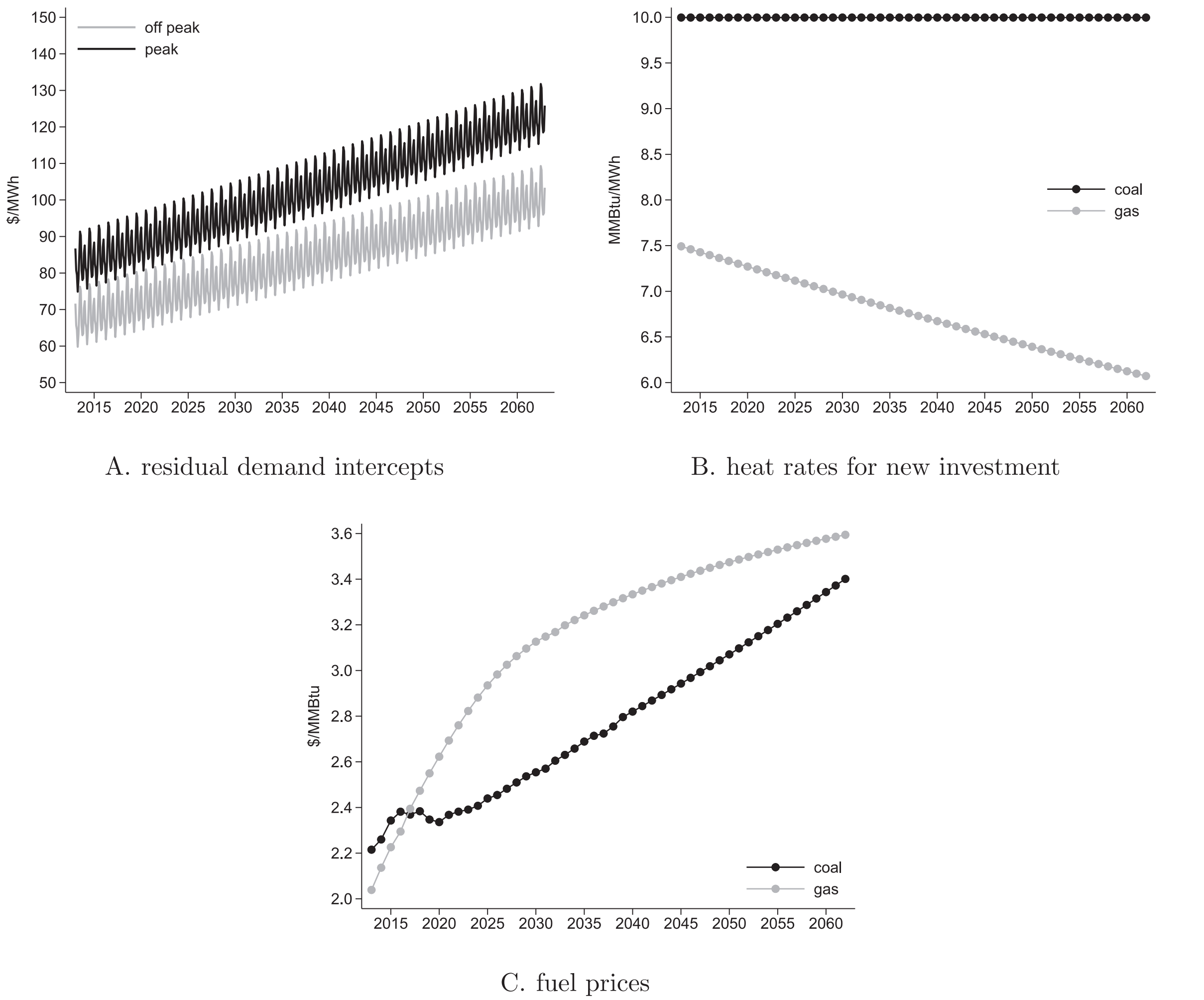


Fig. 5. Paths of exogenous variables, 2013–2062. *Note:* We show monthly off-peak and peak residual demand intercepts in panel A. The coal prices in panel C are in 2010$ and are based on EIA AEO projections. The gas prices in panel C are in 2010$ and are based on NYMEX Henry Hub futures prices for 2013–2028, which are extrapolated to 2062 using flat extrapolation. For additional details, see the discussion in Section 4.2.

Henry Hub futures prices for 2013–2028. We expand the series until 2062 using flat extrapolation of the 2028 levels. Given the collapse in SO_2 and seasonal NO_x permit prices in years around 2012, we set both of these prices equal to zero for 2013–2062.[19]

Time paths of endogenous variables Fig. 6 shows time paths for a variety of endogenous variables. The BAT capacity, which is exclusively gas-fired, exhibits an upward trend increasing from 1590 MW in 2014, the first year of investment, to 11,780 MW in 2062 (panel A). As a result, the share of output (electricity generation) that BAT capacity accounts for increases over time with roughly half of the increase taking place the first 15 years (panel B). Electricity generation (panel C) and prices (panel D) increase over time too. Following a period with a downward trend during 2013–2030, the share of gas in electricity generation increases from 23% to 31.5% (panel E). After about 20 years of growth, the share of coal in electricity generation plateaus at 38%. The share of sources other than coal and gas in electricity generation decreases from 47% in 2013 to 31% in 2062 (panel F)—recall that we assume no investment in these sources.

Table 6 shows the investments in gas-fired capacity by firm for 2013–2062. During the same period, there is no investment in coal-fired capacity. The table shows investment flow and not net investment. Investment may imply replacement of old units that become more costly to operate with new units. According to the model, there are 52 instances of investment for a total of 12,120 MW of gas-fired capacity. Three firms account for roughly 70% of the total investment. Exelon accounts for 3202 MW, followed by NRG (2756 MW) and AES (2625 MW). Exelon invests 15 times while AES and NRG both invest 12 times. A detailed timeline of investment by firm is available in Fig. 7.

[19] Our use of Henry Hub futures prices for gas and the assumption regarding zero permit prices are both consistent with the approach taken in PJM (2016) regarding projections of gas and permit prices.

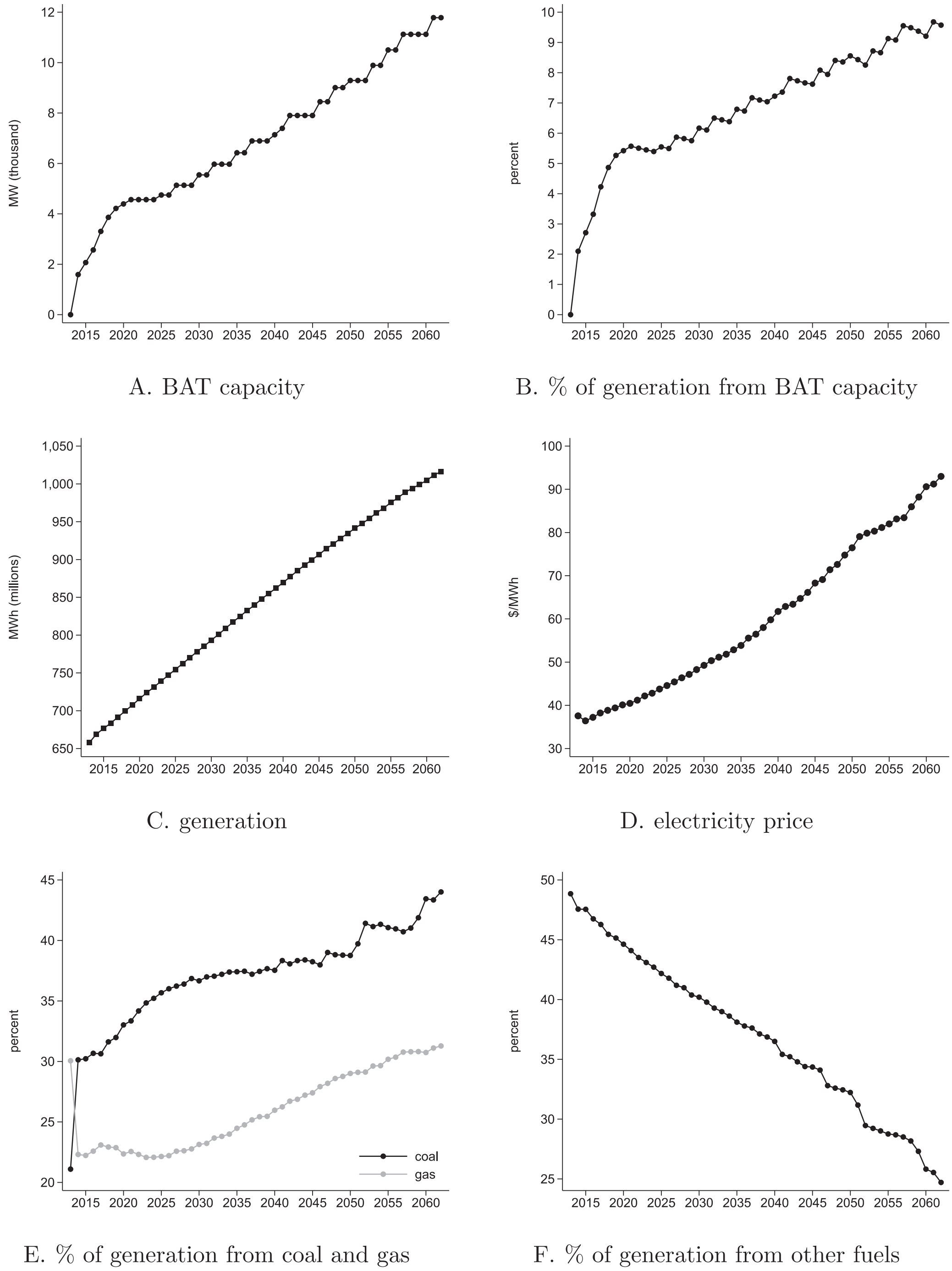


Fig. 6. Paths of endogenous variables, 2013–2062. *Note:* BAT refers to best available technology. We show annual total electricity generation in panel C. The electricity prices in panel D are annual averages of 24 monthly off-peak and peak prices and are in 2010$. We have 24 prices because we solve for the equilibrium of the electricity market for both off-peak and peak demand in each of the 12 months of the year. For additional details, see the discussion in Section 4.2.

Model fit In Fig. 8 we plot the monthly PJM electricity price predicted by our model along with the PJM electricity price from NYMEX futures between April 2016 and December 2019. The main message is that our model tracks reasonably well the NYMEX futures prices. The R^2 of a simple linear rgeression of the two series with 45 observations is around 0.4. The correlation between the two prices is 0.66. Figure A1 shows heat rates, fuel prices, generation and capacity before and after

Table 6
Investment in gas-fired capacity by strategic firm.

	Investment	
Firm	Size	Counts
AEP	0.000	0
AES	2.625	12
DOM	0.000	0
DUK	0.000	0
EXE	3.202	15
FE	1.753	7
GEN	0.916	3
NRG	2.756	12
PPL	0.868	3
PSEG	0.000	0
TOTAL	12.120	52

Note: The numbers reported are for 2013–2062 and pertain to the strategic firms in Table 3. The size of the investment is in thousand megawatts (MW). A firm is assumed to invest once a year. For example, according to our model, AES invested 12 times during 2013–2062 for a total investment of 2625 MW. For additional details, see the discussion in Section 4.2.

2012, the last year in our sample. In general, we see a transition that is smooth and a trend towards more gas in both generation and capacity.

5. Counterfactual simulations

We use the estimated model to compare the outcomes of two counterfactual regulatory scenarios regarding CO_2 emissions. In the first, there is a single (PJM-wide) market that determines the price of CO_2 emissions. In the second, there are separate (state-specific) markets that determine the price of CO_2 emissions. Our goal is to examine the inefficiency of separate CO_2 markets relative to a single CO_2 market. We begin with the static case, for which we compute outcomes holding investment in the best available technology (BAT) capacity fixed. Investment is exogenous and BAT capacity is the same in the two regulatory scenarios. We then shift our focus to the dynamic case, in which investment is the result of optimal behavior and BAT capacity is endogenous.

To implement the counterfactual regulatory scenarios, we assume that power plants in the PJM states are subject to the mass-based targets of the Clean Power Plan (CPP) in Table 2, which limit the annual quantity (short tons) of the power plants' CO_2 emissions. The interim targets for 2022—2029 are followed by a permanent target from 2030 onwards. With separate CO_2 markets, each state's power plant emissions have to be less than or equal to its annual target. With a single CO_2 market, power plant emissions only need to be less than or equal to the sum of the targets across the PJM states (see Fig. 9). Although we do not explicitly model a market for emissions permits, we take the shadow prices of the CO_2 emissions constraints as our CO_2 prices. In the case of the single CO_2 market, the CO_2 price corresponds to the shadow price of the PJM-wide emissions constraint. In the case of separate CO_2 markets, the CO_2 prices correspond to the shadow prices of the states' emissions constraint. Hence, the CO_2 prices are determined endogenously and ensure that the emissions targets are met.

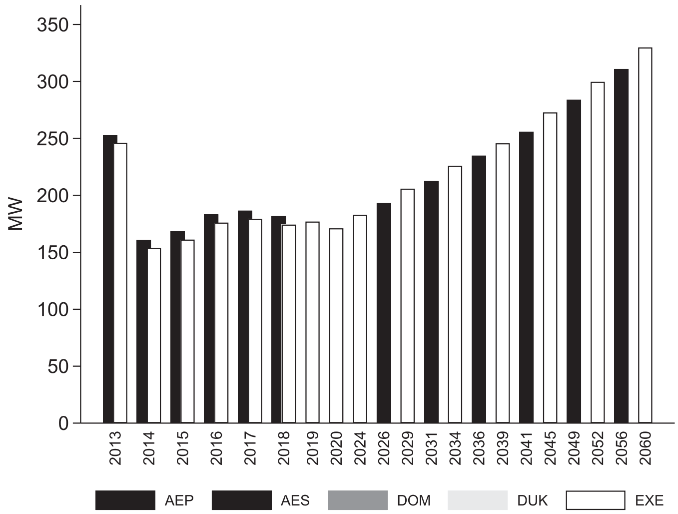
For both regulatory scenarios, the CO_2 price increases the electricity generation cost. This additional cost depends on the power plants' heat rates (MMBtu/MWh), emission rates (lbs./MMBtu), and location. Accounting for CO_2 emissions, the marginal cost for plant i in state s and time t that denotes an off-peak/peak month combination in a year is given by:

$$c^C_{ist} = c_{ist} + P^C_{st} \times r^C_{ist} \times \zeta, \tag{16}$$

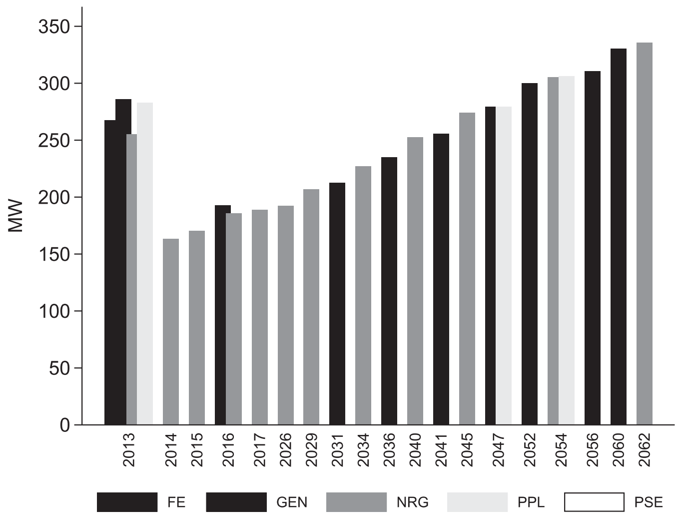
where c_{ist} is the plant's generation cost excluding the cost of emissions ($/MWh), P^C_{st} is the CO_2 price ($/ton), r^C_{ist} is the CO_2 emissions rate (lbs./MWh), and ζ is an appropriate scaling factor that accounts for units of measurement. In the case of a single CO_2 market, $P^C_{st} = P^C_t$, $\forall s \in S$, where S is the set of the eleven PJM states listed in Table 2. Therefore, we solve for 24 CO_2 prices in the case of a single CO_2 market, one for each month and off-peak/peak combination for every year in our simulations. In the case of separate CO_2 markets, we solve for $24 \times 11 = 264$ CO_2 prices for every year because there are eleven states participating in the PJM wholesale electricity market.

Given that CO_2 prices affect the plants' generation costs, they also affect the shape of the wholesale electricity market supply curve. Market demand and supply determine each plant's equilibrium electricity generation, which in turn determine emissions. Emissions in turn determine the extent to which the associated constraint binds and the resulting equilibrium CO_2 price. Therefore, in equilibrium, we solve for a set of prices that simultaneously clear the electricity and CO_2 markets.

We make a series of assumptions for computational feasibility of our simulations, which are consistent with the institutional details of our setting. First, only emissions from existing (built by 2012) capacity are subject to CO_2 prices. Emissions



A. 1st group of strategic firms



B. 2nd group of strategic firms

Fig. 7. Strategic firms' BAT investment in gas-fired capacity, 2013–2062. *Note:* BAT refers to best available technology. The figure shows only years for which there is investment. We divide the strategic firms in two groups and report their investment levels in two panels so that the figure is more legible. In the 1st group of firms, and consistent with the entries of Table 6, only Applied Energy Services (AES) and Exelon (EXE) invest. For additional details, see the discussion in Section 4.2.

from capacity built after 2012 are exempt from the CO_2 price, and post-2012 capacity must have the lowest heat and emissions rate during the investment year. Second, we assume that heat rate improvements are exogenous.[20] Third, generation from renewable sources increases exogenously according to annual growth rates in the CPP.[21] Finally, we assume an upper bound of $100 per ton for the CO_2 price and set the post-2030 CPP targets at their 2030 levels.[22]

[20] See the discussion on exogenous state variables in Section 4.2.

[21] This assumption is based on the June 2014 CPP proposed rule technical support documentation (TSD). The relevant TSD spreadsheet provides state-specific growth rates for renewable energy for 2020–2029. We assume that the average growth rate for 2020–2029 holds for the entire period of our simulations. Moreover, we assume that nuclear capacity does not change.

[22] Borenstein et al. (2019) argue that extreme price outcomes are likely in most cap-and-trade markets for greenhouse gas (GHG) emissions for two main reasons. The first is GHG emissions volatility. The second is the low price elasticity of GHG abatement over the price range generally deemed to be

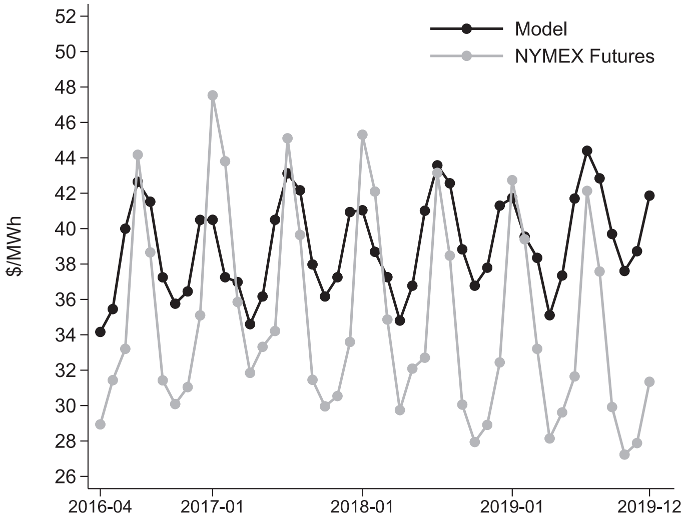


Fig. 8. Electricity prices implied by the model compared to NYMEX futures. *Note:* The figure shows our model's monthly predictions for PJM electricity prices along with their counterparts from NYMEX futures for April 2016–Demcember 2012. For additional details, see the discussion in Section 4.2.

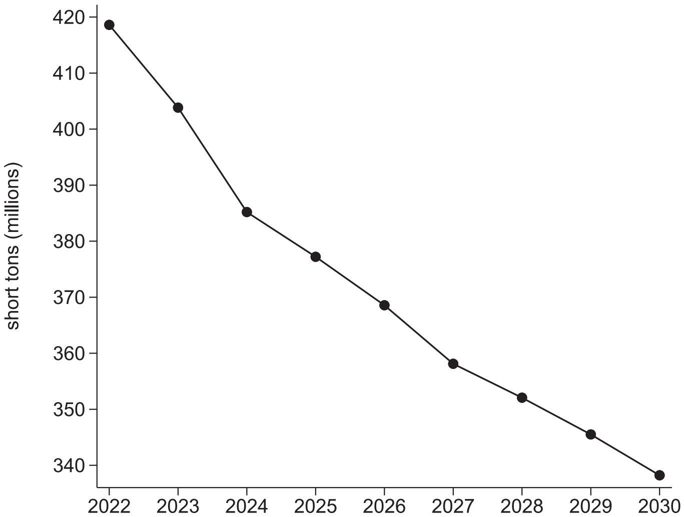


Fig. 9. Regional clean power plan mass-based targets. *Note:* The mass-based target in this figure is based on the supporting data file for the Clean Power Plan compliance from PJM (2016) and are based on electric generating units in the PJM footprint for each state noting that PJM covers only parts of Illinois, Indiana, Kentucky, and North Carolina. We plot the sum of state mass-based targets from panel A of Table 2. For additional details, see the discussion in Section 5.

In what follows, we first provide the computational details for the counterfactual simulations. We then discuss the static analysis with exogenous investment in BAT capacity. Finally, we present the dynamic analysis with optimal investment in BAT capacity.

5.1. Computational details

For the purpose of the discussion in this section, we focus on the vector of cumulative BAT capacities of each firm when describing our state space.[23] Based on our estimates, we assume that the strategic firms invest only in gas-fired capacity. Moreover, we allow for investment but no divestment and assume that capacity does not depreciate. These assumptions are

---

acceptable. Recognizing the problems created by uncertainty in emissions permit prices, hybrid mechanisms that combine caps on emissions and price collars (both floors and ceilings) have been proposed.

[23] The state vector includes the exogenous variables discussed in Section 4.2 and the weighted-average heat rate of each firm's cumulative BAT capacity. For the BAT heat rate dimension of the state space, we use 3 nodes corresponding to the minimum, average, and maximum heat rates for 2013–2030. We create a dense grid for the state along the BAT heat rate dimension using a cubic spline.

consistent with plants simply remaining idle if they are not utilized and long-lived capacity. They also imply that the BAT capacity either increases or stays at its current level. Finally, we assume that total BAT capacity across firms must be less than or equal to 60,000 MW, which, if fully utilized, represents about 60% of total output.[24]

In terms of the implementation, we start by discretizing the BAT capacity using a coarse grid of increments equal to 1000 MW. For each of the points in the grid, we compute the stage-game market equilibrium, which allows us to calculate firm profits, consumer surplus, and damages from CO_2 emissions as functions of BAT capacity; we describe how we solve for the stage-game equilibrium below. We then interpolate the values of profits, consumer surplus, and CO_2 damages over a finer grid of 100-MW increments. The finer grid aims to strike a balance between the approximation error in the value-function iterations and computational feasiblity.

The exact objective function in the optimal-investment problem depends on the simulated scenario. The investment problem is non-stationary because prices, demand, heat rates for new investment, and CO_2 targets change each year. To solve the model, we fix all exogenous state variables at their 2030 levels post 2030, and solve the associated stationary infinite-horizon problem. We follow this approach because we do not have emissions targets beyond 2030. Once we have the value functions for 2030, we proceed backwards, starting in 2029 and ending in 2013, noting that the exogenous variables change every year.

The equilibrium computation can be divided into two parts. In the first part, we compute the stage-game profits for each firm, point in the state space, and point in time (2013–2029 plus the infinite horizon beginning 2030), which require the simultaneous clearing of the CO_2 and electricity markets. In the second part, given stage-game profits, we solve for the Markov Perfect Nash Equilibrium, that is, a set of investment policy and value functions for each firm, point in the state space, and point in time. With a discretized state space, we can do the first part of the equilibrium computation separately and also exploit parallel computing to solve for each of the points in the state space. For the second part, standard value-function iteration approaches, including the parametric approximation method in Ryan (2012) and Fowlie et al. (2016) did not allow our solver to converge even with lax tolerance (e.g. 10^{-5}). Hence, we opted for an alternative approach described below.

5.1.1. Stage game market equilibrium

The equilibrium of the stage game requires the simultaneous clearing of the wholesale electricity and the CO_2 markets. The need to look for the simultaneous clearing of both markets is due to the complementary nature of the output and the CO_2 emissions. A change in CO_2 prices affects the relative cost of the different fuels. This in turn changes the relative position of each plant in the merit order of the aggregate electricity supply and, therefore, impacts the equilibrium in the wholesale electricity market. In the case of a single CO_2 market, we solve for 24 CO_2 prices. Each of these 24 prices pertains to a particular off-peak/peak month combination in a year. With separate (state-specific) CO_2 markets, we solve for $24 \times 11 = 264$ CO_2 prices because there are 11 states participating in PJM. For both regulatory scenarios, we solve for 24 electricity prices.

In more detail, let q_{ist} denote the electricity output of plant i located in state s at time t that denotes an off-peak/peak month combination in a year. In addition, HR_{ist} is the associated heat rate and r_{ist} is the rate of CO_2 emissions. The mass-based target of CO_2 emissions for state s at time t is $\overline{E}_{st}$, which we compute by multiplying the annual state mass-based target by the fraction of last year's net generation for the same off-peak/peak month combination. Finally, let S denote the set of the 11 states participating in the PJM wholesale electricity market.

With a single CO_2 market, the 24×1 vector of equilibrium CO_2 prices is the solution to the following minimization problem:

$$\mathbf{P}_t^C = \min\{\mathbf{P} : \sum_{s \in S} \sum_{i \in s} (q_{ist}(\mathbf{P}) \times HR_{ist} \times r_{ist}) \leq \sum_{s} \overline{E}_{st}\}. \tag{17}$$

With separate CO_2 markets, the solution is given by the following 264×1 vector of CO_2 prices:

$$\mathbf{P}_t^C = \min\{\mathbf{P} : \sum_{i \in s} (q_{is\tau}(\mathbf{P}) \times HR_{ist} \times r_{ist}) \leq \overline{E}_{st}\} \quad \forall s \in S. \tag{18}$$

We solve the minimization problem using the following 4-step algorithm:[25]

1. Start with zero CO_2 prices for all states and compute the electricity market equilibrium.
2. If at least one state has excess emissions, proceed to Step 3; otherwise, end.
3. Increase the CO_2 price of the state with the highest excess emissions by \$1.
4. Compute the electricity market equilibrium and check for excess emissions.

[24] We experimented with numerous alternative upper bounds for the BAT state variable and we noticed that there was no optimal investment beyond the assumed level of 60,000 MW for the models considered.

[25] With a single CO_2 market, we treat the entire PJM area as one state and the algorithm works in the same way. In this case, the dimension of the vector of CO_2 prices becomes 24×1 as we have already discussed.

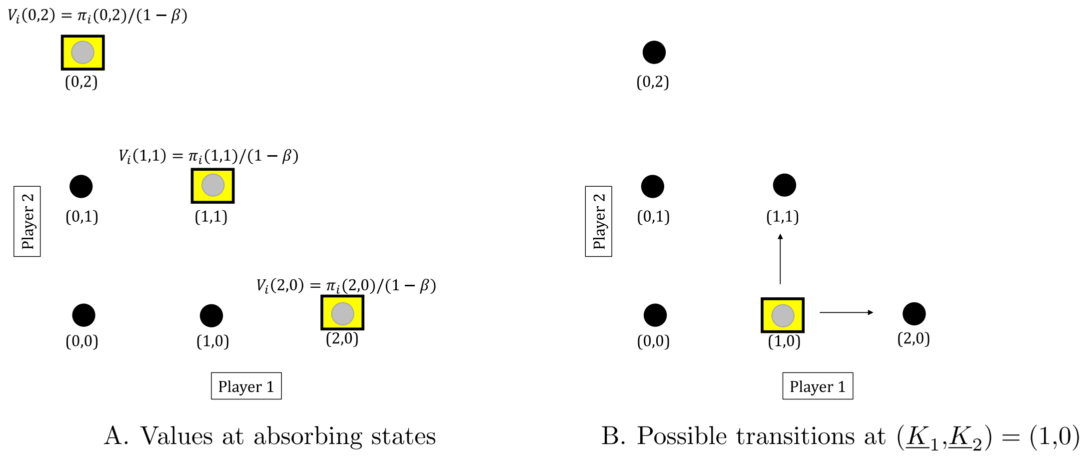


Fig. 10. Computing the MPNE using a simplified example. *Note:* To illustrate how we compute the Markov Perfect Nash Equilibrium (MPNE) for our counterfactual analysis, consider a simplified state space where the sum of firm 1 and 2's BAT capacity satisfies $\underline{K}_1 + \underline{K}_2 \leq 2$ and players can invest in unit increments ($i_1, i_2 \in \{0, 1\}$). We first start at the "edge" of the state space and then work backwards as shown in panel A. The points on the edge are absorbing states. The values at the points are of the form $\pi_i(\underline{K}_1, \underline{K}_2)/(1-\beta)$ where $\pi_i(\cdot)$ is the stage-game payoff at the given state and β is the common discount factor. Given these values, we can then consider the point (1,0). At (1,0), we can either transition to (1,0) (no investment), (2,0) (only player 1 invests) and (1,1) (only player 2 invests), as in panel B. The transition is determined by the pure-strategy Nash Equilibrium of the normal form game with payoff matrix given in panel C. Note that $V_i(2,1)$ is technically undefined because this is outside our state space. Moreover, existence and uniqueness of the pure-strategy NE is not guaranteed. Thus, we instead consider a sequential version of this normal form game where player 1 first decides on i_1 followed by player 2. This solves the existence and uniqueness problem, and also the indeterminacy of $V_i(2,1)$ (i.e. $V_i(2,1) = V_i(2,0)$). For additional details, see the discussion in Section 5.1.2.

5.1.2. Solving for the Markov perfect Nash equilibrium

To illustrate our computation of the Markov Perfect Nash Equilibrium (MPNE) for our counterfactual analysis, consider a simplified state space for a two-player investment game in which the sum of two players' BAT capacity satisfies $\underline{K}_1 + \underline{K}_2 \leq 2$ and investment is in unit increments ($i_1, i_2 \in \{0, 1\}$).

Our approach resembles the Upwind Gauss–Seidel alogirthm in dynamic programming (Judd, 1998). We first start at the "edge" of the state space and then move backwards as shown in panel A of Fig. 10. The points on the edge are absorbing states with values $\pi_i(\underline{K}_1, \underline{K}_2)/(1-\beta)$, where $\pi_i(\cdot)$ is the stage-game payoff for player i at the given state and β is the common discount factor. Given these values, we move backwards to the point (1,0).

As shown in panel B, at point (1,0), the players can transition to (1,0) (no investment), (2,0) (only player 1 invests) and (1,1) (only player 2 invests). Given that an MPNE is a set of policies and values such that these form a Nash Equilibrium (NE) at the subgame defined by the state, the transition from (1,0) is determined by the pure-strategy NE of the normal-form game with the payoff matrix given in panel C.[26] This is the statewise Nash approach implemented in Chen et al. (2009) (see also Doraszelski and Escobar, 2010; Abito and Chen, 2021). The payoffs $V_i(2,0)$ and $V_i(1,1)$ are already available from the previous iteration. The payoff $V_i(1,0)$ is just $\pi_i(1,0) + \beta V_i(1,0) = \pi_i(1,0)/(1-\beta)$. The payoff $V_i(2,1)$ is technically undefined because it lies outside the state space. To handle the indeterminacy and guarantee existence and uniqueness, we consider a sequential version of this normal-form game where player 1 first decides on i_1 followed by player 2.[27]

5.2. Static analysis with exogenous investment

Holding capacity fixed, the electricity generation cost with a single CO_2 market is expected to be lower than with separate CO_2 markets. A single CO_2 market equates marginal abatement costs across states leading to lower overall compliance costs. Insights from the trade literature, however, tell us that the integrated electricity market can mitigate inefficiencies associated with separate CO_2 markets if production from markets with high CO_2 prices can be reallocated to markets with low CO_2 prices, all else equal. In what follows, we show that mitigation—and even elimination—of the inefficiencies, is possible in our setup of the PJM wholesale electricity market.

We solve for the equilibrium of the electricity and CO_2 markets as a function of BAT capacity, which is set exogenously at a level between 1000 and 60,000 MW from 2013 onwards, and compare the outcomes of the two regulatory scenarios. In particular, we show the present discounted value (PDV) of social welfare for the two scenarios as a function of BAT capacity in panel A of Fig. 11. In panel B of the same figure, we show the PDV of electricity generation cost for the two scenarios as

[26] It is more straightforward to solve for the NE in a complete-information version of the normal-form game and, hence, we set the privately-observed cost shocks to be equal to zero in the counterfactuals.

[27] See Bresnahan and Reiss (1990); Berry (1992) for early examples in static entry game setting, and, more recently, Abbring and Campbell (2010) in an infinite-horizon setting.

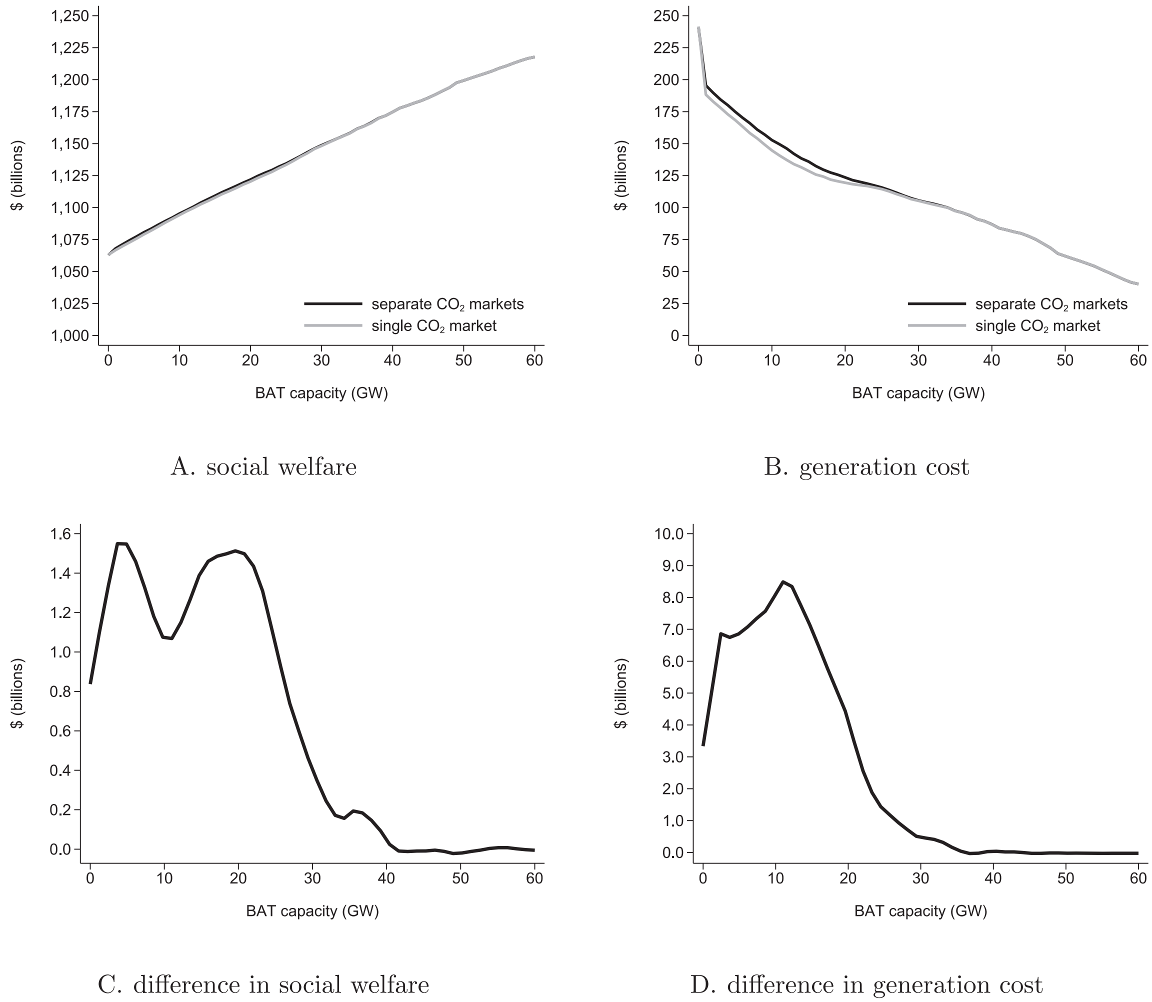


Fig. 11. Static inefficiency. *Note:* Panel A shows the present discounted value (PDV) of social welfare for the separate and single CO_2 markets scenarios as a function of best available technology (BAT) capacity. We use total surplus, defined as consumer surplus plus producer surplus plus revenues from the CO_2 markets minus damages from CO_2 emissions calculated using the social cost of carbon to measure social welfare. Panel B shows the present discounted value (PDV) of electricity generation cost for the separate and single CO_2 markets scenarios as a function of (BAT) capacity. The dollar amounts for both social welfare and generation cost are in 2010$. We keep the BAT capacity fixed from 2013 onwards for both scenarios of CO_2 emissions regulations in this analysis. Panels C and D plot the difference (separate-single) in the PDV of social welfare and electricity generation cost between the two scenarios as a function of BAT capacity. For additional details, see the discussion in Section 5.2.

a function of BAT capacity. We show the difference in social welfare and electricity generation between the two scenarios as a function of BAT capacity in panels C and D, respectively. We define social welfare as the sum of the strategic firms' profits and consumer surplus less damages from CO_2 emissions including those from BAT capacity. Consumer surplus refers to the surplus of buyers in the PJM wholesale electricity market and is calculated assuming a maximum willingness to pay of $200 per MWh. To compute damages, we use the annual social-cost-of-carbon (SCC) values for an assumed discount rate of 3% from Table A.1 in IWG (2013). The difference in social welfare is what we refer to as static inefficiency. All dollar amounts are in 2010$.

Social welfare is increasing in the BAT capacity ranging from $1060 billion to $1218 billion. The difference in social welfare between the two regulatory scenarios is small. For BAT capacity close to 22,000 MW, the static inefficiency reaches its maximum of about $1.6 billion, which is small compared to the values of social welfare mentioned above.[28] When BAT capacity exceeds 40,000 MW, there is a negligible difference in social welfare between the two scenarios and, hence, the inefficiency is eliminated. At very high levels of BAT capacity, frictionless production reallocation is feasible allowing CO_2 prices with single and separate markets to converge to zero and eliminate the inefficiency.

[28] For low levels of BAT capacity, CO_2 prices with single and separate CO_2 markets can sometimes hit the CO_2 price ceiling of $100. In these instances, there is no difference in the outcomes of interest between the two regulatory scenarios.

Table 7
Summary of outcomes for optimal investment scenarios.

	BAT capacity	Electricity price	Generation costs	Investment costs	CO_2 emissions	Social welfare
Scenario	MW	$/MWh	$ billion	$ billion	Million tons	$ billion
SOCPLAN	60,000	30	52	80	101	1129
NST-SIN	43,000	35	117	33	254	1100
NST-SEP	48,000	36	116	37	240	1102
FCI-SIN	5000	97	243	2	296	1036
FCI-SEP	11,000	92	220	6	284	1048
LFG-SIN	16,000	96	221	9	276	1068
LFG-SEP	21,000	88	189	22	243	1077

Note: BAT refers to best available technology. We report a quantity-weighted average price of electricity and a quantity-weighted average of CO_2 emissions, including those from BAT capacity, for 2013–2030. The present discounted dollar values (2010$) are calculated using a discount factor of 0.90 and assuming that the 2030 values are the steady state ones. A brief description of the scenario abbreviations is available in Table 8. For additional details, see the discussion in Section 5.3.

Table 8
Description of optimal investment scenarios.

Scenario	Description
SOCPLAN	Social planner
NST-SIN	Non-strategic investment, single CO_2 market
NST-SEP	Non-strategic investment, separate CO_2 markets
FCI-SIN	Fully coordinated investment, single CO_2 market
FCI-SEP	Fully coordinated investment, separate CO_2 markets
LFG-SIN	Leader-follower investment game, single CO_2 market
LFG-SEP	Leader-follower investment game, separate CO_2 markets

Note: The table provides a brief description of the alternative investment scenarios that pertain to different market structures and regulatory scenarios discussed in detail in Section 5.3.

The electricity generation cost is decreasing in BAT capacity with values between $40 and $242 billion. In the case of separate CO_2 markets, the generation cost is either higher than or equal to its counterpart with a single CO_2 market, and the difference in the generation cost between the two scenarios as a function of BAT capacity follows the same pattern as the difference in social welfare. We should, note, however, that the difference in generation costs is larger than the difference in social welfare. In particular, the generation cost with separate CO_2 markets can exceed its counterpart with a single CO_2 market by up to $8.5 billion, or about 6%, which we see for a BAT capacity of 11,000 MW.

5.3. Dynamic analysis with optimal investment

The dynamic analysis allows us to compare outcomes allowing for optimal investment. In the model we use to estimate investment costs, the top ten firms in Table 3 make strategic investment decisions. For computational reasons, we instead analyze the simpler modeling scenarios summarized in Table 8 in the counterfactual simulations that follow. In the first scenario (SOCPLAN), a social planner chooses the level of investment that maximizes social welfare. The second scenario (NST) is one with non-strategic investment; investment takes place if the marginal benefit is higher than the marginal cost of capital. In the third scenario (FCI), the ten strategic firms fully coordinate investment acting as one firm. In the final scenario (LFG), which adds an element of competition to the third scenario, the ten strategic firms are assigned into two coalitions and decide on optimal investment maximizing the coalitions' profits in a leader-follower game discussed in detail below.

In all modeling scenarios, we maintain the set of plants that we considered during estimation preserving plant-level heterogeneity. Similar to our discussion in the static analysis, all dollar amounts are in 2010$. The key message that the inefficiency inherent in separate CO_2 markets is mitigated, and can even be eliminated, is robust across all models of optimal investment behavior.

5.3.1. Social planner and nonstrategic investment

The social planner chooses investment to maximize the present discounted value (PDV) of social welfare. The social welfare is the sum of consumer surplus and profits of strategic firms minus the damages from CO_2 emissions calculated using the SCC. As a reminder, the consumer surplus pertains to the buyers in the wholesale electricity market. In what follows, we compare the social-planner (SOCPLAN) scenario to the scenario with non-strategic investment (NST).

The steady-state (2030) BAT capacity is 60,000 MW implying a social welfare of $1129 billion in the case of the social planner (Table 7). Its counterpart in the case of nonstrategic investment is lower than the socially optimal level whether we model a single CO_2 market or separate CO_2 markets. However, the BAT capacity with separate CO_2 markets is closer to the socially optimal level. In particular, there are 43,000 MW of BAT capacity with a single CO_2 market and 48,000 MW with separate CO_2 markets, which is an increase of 11.6%.

The electricity generation cost with separate CO_2 markets (116 billion) is highly comparable to that with a single CO_2 market (117 billion), but at the expense of a higher investment cost, namely, $37 billion as opposed to $33 billion, an increase of about 12%. The stronger incentives to invest in BAT capacity with separate CO_2 markets eliminate the difference in social welfare between the two regulatory scenarios implying a total surplus of $\$1,102$ billion that is qualitatively similar to its counterpart with a single CO_2 market ($1100 billion).

5.3.2. Strategic investment

Fully-coordinated investment behavior In this modeling scenario (FCI), the ten strategic firms fully coordinate investment to maximize the (sum of the) PDV of their profits. With a single CO_2 market, the steady-state BAT capacity is suppressed to 5000 MW raising the average electricity price to $97 per MWh. In contrast, its counterpart with separate CO_2 markets is 11,000 MW, and the average electricity price is $92 per MWh. The much lower BAT capacity when firms coordinate investment to maximize industry profits compared to the nonstrategic case is driven by the desire to raise electricity prices by keeping capacity low.

The electricity generation cost is lower with separate CO_2 markets ($220 billion) than with a single CO_2 market ($243 billion). Although the investment cost is higher with separate CO_2 markets ($6 billion) than with a single CO_2 market ($2 billion), the difference in the investment costs between the two regulatory scenarios is smaller than the difference in the generation costs. As a result, the total (generation plus investment) cost is 7.8% lower with separate CO_2 markets.

The emissions are lower with separate CO_2 markets (284 million tons) than with a single CO_2 market (296 million tons), contrary to what one would expect with leakage. This reduction in emissions stems from investment in power plants that emit less per MWh of electricity than the existing plants.

Finally, the social welfare with separate CO_2 markets is $1048 billion, which is higher than its counterpart with a single CO_2 market ($1036 billion). The finding that social welfare with separate CO_2 markets is qualitatively similar to its counterpart with a single CO_2 market echoes our finding in the case of nonstrategic investment.

The leader-follower game In the last modeling scenario (LFG), we relax the assumption of fully coordinated investment by introducing competition. For computational reasons, we study a two-firm (leader-follower) perfect-information investment game. We create two coalitions of strategic firms by allocating their existing plants into two groups with plant characteristics balanced between the two. We treat one coalition as the leader (invests first) and the other as the follower (invests second).

Each coalition decides strategically on investment taking into account profits earned from the plants it owns and how investment changes endogenous state variables, including the BAT capacity of all firms in both coalitions. We maintain the assumption of competitive behavior in both the electricity and CO_2 markets and solve the stage game by finding the market clearing prices. With a competitive wholesale electricity market, the equilibrium quantity and price are not affected by our assumption on the number of investing firms, conditional on the set of plants in the market.

Our results confirm that introducing competition weakens firms' incentives to strategically withhold investment in order to raise electricity prices. The steady-state BAT capacity is 16,000 MW with a single CO_2 market, and it is 21,000 MW with separate CO_2 markets. Although these capacity levels are lower than their socially optimal counterpart (60,000 MW), they exceed their coordinated-investment levels of 5000 MW and 11,000 MW, respectively. Moreover, in the leader-follower game, the players are able to raise electricity prices above the socially-optimal levels but not as much as when they fully coordinate their investment decisions. The electricity prices are $96 per MWh with a single CO_2 market and $88 per MWh with separate CO_2 markets.

The electricity generation cost with separate CO_2 markets ($189 billion) is lower than with a single CO_2 market ($221 billion) by 14.4%. Although the sign of the difference is the same as the one in the previous models of optimal investment behavior, the magnitude of the difference is much larger. Accounting for the higher investment cost with separate CO_2 markets, the total (generation plus investment) cost ($211 billion) is also lower by about 8.3% than its counterpart with a single CO_2 markets ($230 billion).

Finally, the social welfare with separate CO_2 markets ($1077 billion) is qualitatively similar to its counterpart with a single CO_2 market ($\$1,068$ billion) echoing our findings for the other models of optimal investment behavior. The inefficiency is eliminated because the level of investment with separate CO_2 markets is closer to the social optimum compared to the level of investment with a single CO_2 market.

6. Conclusion

In this paper, we show that correcting an environmental externality in the absence of coordinated regulation across jurisdictions entails social welfare that is qualitatively similar to the one when coordinated regulation is possible. The main driving force behind our finding is investment by firms participating in an integrated product market, which mitigates, and even eliminates, the inefficiencies that emerge in the absence of coordinated regulation.

Our workhorse is a dynamic structural model of production and investment for the largest wholesale electricity market in the world, the Pennsylvania–New Jersey–Maryland (PJM) Interconnection. The environmental regulation of interest entails targets for carbon dioxide (CO_2) emissions from electricity generation achieved via a market-based mechanism with and without coordination across states participating in PJM. In the case of coordinated regulation, there is a single PJM-wide CO_2 market. With uncoordinated regulation, there are separate CO_2 markets, one for each of the participating states.

Our model preserves the rich plant-level cost heterogeneity in the data while being tractable enough to evaluate market outcomes across the two regulatory scenarios. We achieve tractability by assuming that market participants invest in the best available technology (BAT) at the time of the investment, which is consistent with the current interpretation of the Clean Air Act. In our setup, CO_2 emissions from BAT capacity are assumed to be exempt from the targets and, hence, the location of firms' investment is irrelevant—only the total amount of investment matters. This is an important assumption for computational feasibility but comes at the cost of generalizability; it is conceivable that emissions from BAT capacity are not exempt from the targets.[29] To the extent that emissions from BAT are not exempt from the target, its cost advantage will be smaller, weakening the incentives to invest. While investment levels are likely to change, investment incentives will still be stronger with separate CO_2 markets than with a single CO_2 market.

Given the recent developments in U.S. environmental policy, the future of federal regulations aiming to curb CO_2 emissions in the electric power sector seems to be unclear. An important question that can be answered using our framework is whether states have unilateral incentives to adopt emission restrictions in the absence of any federal mandate. For example, Abito (2019) uses our model to analyze the implications of Pennsylvania's decision to unilaterally join the Regional Greenhouse Gas Initiative for PJM. The potential benefit of unilateral adoption would be to provide incentives for investment in more efficient capacity, as in the case of uncoordinated regulation. It is also important to emphasize the potential benefits for consumers in states that do not adopt any emissions regulations because more efficient capacity may decrease electricity prices for the whole region. Any careful analysis should take into account the interaction between the product and environmental markets and the adjustments that occur beyond the short-run, such as investment in new capacity.

CRediT authorship contribution statement

Jose Miguel Abito: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Christopher R. Knittel:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Konstantinos Metaxoglou:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **André Trindade:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijindorg.2022.102843

References

Abbring, J., Campbell, J., 2010. Last-in first-out oligopoly dynamics. Econometrica 78 (5), 1491–1527.
Abito, J. M., 2019. Self-imposed emission limits: is there a case for pennsylvania and RGGI?
Abito, J.M., 2020. Measuring the welfare gains from optimal incentive regulation. Rev. Econ. Stud. 87 (5), 2019–2048.
Abito, J.M., Chen, C., 2021. How much can we identify from repeated games? Econ. Bull. 41 (3), 1212–1222.
Adilov, N., 2012. Strategic use of forward contracts and capacity constraints. Int. J. Ind Organ. 30 (2), 164–173.
Allaz, B., Villa, J., 1993. Cournot competition, futures markets and efficiency. J. Econ. Theory 59, 1–16.
Bajari, P., Benkard, C., Levin, J., 2007. Estimating dynamic models of imperfect competition. Econometrica 75 (5), 1331–1370.
Bajari, P., Chan, P., Krueger, D., Miller, D., 2013. A dynamic model of housing demand: estimation and policy implications. Int. Econ. Rev. 54 (2), 409–442.
Berry, S., 1992. Estimation of a model of entry in the airline industry. Econometrica 60 (4), 889–917.
Borenstein, B., Bushnell, J., Wolak, F., 2002. Measuring market inefficiencies in California's restructured wholesale electricity market. Am. Econ. Rev. 92 (5), 1376–1405.
Borenstein, S., Bushnell, J., Wolak, F.A., Zaragoza-Watkins, M., 2019. Expecting the unexpected: emissions uncertainty and environmental market design. Am. Econ. Rev. 109 (11), 3953–3977.
Bresnahan, T., Reiss, P., 1990. Entry in monopoly markets. Rev. Econ. Stud. 57 (4), 531–553.
Bushnell, J., Flagg, M., Mansur, E., 2017a. Capacity markets at a crossroads.
Bushnell, J., Holland, S., Hughes, J., Knittel, C., 2017. Strategic policy choice in state-level regulation: the EPA's clean power plan. Am. Econ. J. 9 (2), 57–90.
Bushnell, J., Ishii, J., 2007. An equilibrium model of investment in restructured electricity markets.
Bushnell, J., Mansur, E., Saravia, C., 2008. Vertical arrangements, market structure, and competition: an analysis of restructured US electricity markets. Am. Econ. Rev. 98 (1), 237–266.
Chen, J., Doraszelski, U., Harrington Jr., J.E., 2009. Avoiding market dominance: product compatibility in markets with network effects. RAND J. Econ. 40 (3), 455–485.
Cicala, S., 2015. When does regulation distort costs? Lessons from fuel procurement in US electricity generation. Am. Econ. Rev. 105 (1), 411–444.
Davis, L., Muehlegger, E., 2010. Do americans consume too little natural gas? An empirical test of marginal cost pricing. RAND J. Econ. 41, 791–810.
Dixon, H., 1985. Strategic investment in a competitive industry. J. Ind. Econ. 33, 205–212.
Doraszelski, U., Escobar, J.F., 2010. A theory of regular Markov perfect equilibria in dynamic stochastic games: genericity, stability, and purification. Theor. Econ. 5 (3), 369–402.
EPA, 2015. By the numbers: cutting carbon pollution from power plants.

[29] Section A.5 discusses an interesting commitment problem with the New Source Complements (NSC) in the Clean Power Plan, which adjust state-level emissions targets to account for projected future investment. With the NSCs, emissions from BAT capacity are also subject to CO_2 prices in our model. For the scenario with fully coordinated investment and a single CO_2 market, we find that there is zero investment in BAT capacity because of one-sided commitment. The EPA commits to an adjusted target based on expected investment, but does not condition these targets on whether investment actually materializes. In effect, existing capacity now faces a more lax emissions target that dampens the incentive to invest.

Ericson, R., Pakes, A., 1995. Markov-perfect industry dynamics: a framework for empirical work. Rev. Econ. Stud. 62, 53–82.
Fowlie, M., 2009. Incomplete environmental regulation, imperfect competition, and emissions leakage. Am. Econ. J. 1:2, 72–112.
Fowlie, M., 2010. Emissions trading, electricity restructuring, and investment in pollution abatement. Am. Econ. Rev. 10 (3), 837–869.
Fowlie, M., Muller, N., 2019. Market-based emissions regulation when damages vary across sources: what are the gains from differentiation? J. Assoc. Environ. Resour. Econom. 6 (3), 593–632.
Fowlie, M., Reguant, M., Ryan, S., 2016. Market-based emissions regulation and industry dynamics. J. Polit. Econ. 124 (1), 249–302.
Goulder, L., Jacobsen, M., van Benthem, A., 2012. Unintended consequences from nested state and federal regulations: the case of the Pavley greenhouse–gas-per-mile limits. J. Environ. Econ. Manag. 63, 187–207.
Goulder, L. H., Stavins, R., 2010. Interactions of state and federal climate change policies. NBER Working Paper 16123.
Gowrisankaran, G., Reynolds, S.S., Samano, M., 2016. Intermittency and the value of renewable energy. J. Polit. Econ. 124 (4), 1187–1234.
Green, R., Newbery, D., 1992. Competition in the british electricity spot market. J. Polit. Econ. 100 (5), 929–953.
Harsanyi, J.C., 1973. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. Int. J. Game Theory 2 (1), 1–23.
Hausman, C., Muehlenbachs, L., 2019. Price regulation and environmental externalities: evidence from methane leaks. J. Assoc. Environ. Resour. Econom. 6 (1), 73–109.
Holland, S.P., Yates, A.J., 2015. Optimal trading ratios for pollution permit markets. J. Public Econ. 125, 16–27.
Hortacsu, A., Puller, S., 2008. Understanding strategic bidding in multi-unit auctions: a case study of the texas electricity spot market. RAND J. Econ. 39 (1), 86–114.
IWG, 2013. Technical support document: technical update of the social cost of carbon for regulatory impact analysis under executive order 12866. Interagency Working Group on Social Cost of Carbon, United States Government, May 2013, Revised November 2013.
Judd, K.L., 1998. Numerical Methods in Economics. MIT press.
Mansur, E.T., 2007. Upstream competition and vertical integration in electricity markets. J. Law Econ. 50 (1), pp.125–156.
Muller, N.Z., Mendelsohn, R., 2009. Efficient pollution regulation: getting the prices right. Am. Econ. Rev. 99 (5), 1714–1739.
Mundell, R.A., 1957. International trade and factor mobility. Am. Econ. Rev. 47 (3), 321–335.
Myatt, J., 2017. Market power and long-run technology choice in the us electricity industry.
PJM, September 2016. EPA's final clean power plan compliance pathways economic and reliability analysis. PJM Interconnect..
Ryan, S., 2012. The costs of environmental regulation in a concentrated inudstry. Econometrica 80 (3), 1019–1061.
Samuelson, P.A., 1948. International trade and the equalisation of factor prices. Econ. J. 58 (230), 163–184.
Spees, K., Newell, S., Carlton, R., Zhou, B., Pfeifenberger, J., 2011. Cost of new entry estimates for combustion-turbine and combined-cycle plants in PJM.
Wolak, F., 2000. Empirical analysis of the impact of hedge contracts on bidding behavior in a competitive electricity market. Int. Econ. J. 14 (2), 1–39.
Wolfram, C., 1998. Strategic bidding in a multiunit auction: an empirical analysis of bids to supply electricity in england and wales. RAND J. Econ. 29 (4), 703–725.