NBER WORKING PAPER SERIES

AN EMPIRICAL ANALYSIS OF THE INTERCONNECTION QUEUE

Sarah Johnston
Yifei Liu
Chenyu Yang

Working Paper 31946
http://www.nber.org/papers/w31946

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2023

We thank Joseph Aldy, Jim Bushnell, Ken Gillingham, Will Gorman, Koichiro Ito, Ashley Langer, David Popp, Mar Reguant, Andrew Sweeting, Frank Wolak, and seminar and conference participants at BSE Summer Forum (Applied IO), Cal Poly, Cornell, Duke, Heartland Workshop, Maryland, MEA Meetings, NIU, NBER Decentralization Conference, NBER Economics of Innovation in the Energy Sector Conference, NBER Summer Institute (IO), Oberlin, Rice, UC Berkeley Power Conference, UC Davis, UPenn, UW-Madison, and Yale for helpful comments and discussions. We also thank Christian McDewell and four industry experts for insight into the electricity market. We also thank the numerous undergraduate research assistants at UW-Madison and the University of Maryland who helped with data collection. We gratefully acknowledge financial support from the National Science Foundation (SES-2215063), the NBER-Sloan Economics of Innovation in the Energy Sector Initiative, the University of Maryland Grand Challenges Grant Program, and the University of Wisconsin - Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

An Empirical Analysis of the Interconnection Queue
Sarah Johnston, Yifei Liu, and Chenyu Yang
NBER Working Paper No. 31946
December 2023
JEL No. D02,L0,Q00

ABSTRACT

Generators applying to connect to the U.S. power grid go through an interconnection queue. Most wind and solar generators that begin the process do not complete it. Using new data, we find that a long queue increases the average waiting time, and high interconnection costs are a key factor in a generator's decision to withdraw. We develop and estimate a dynamic model of the queue and quantify the effects of policy reforms. Our simulations indicate that reducing waiting times can significantly increase completions. An alternative queuing mechanism can therefore increase completed capacity by removing certain generators to reduce congestion. A flat entry fee has a similar effect. We also quantify the effects of reforming how interconnection costs are assessed. These policy reforms lead to a substantial reduction in carbon emissions.

Sarah Johnston
Department of Agricultural and
Applied Economics
University of Wisconsin-Madison
427 Lorch Street
Madison, WI 53706
sarah.johnston@wisc.edu

Yifei Liu
Department of Agricultural and
Applied Economics
University of Wisconsin-Madison
427 Lorch Street
Madison, WI 53706
yifei.violet.liu@wisc.edu

Chenyu Yang
University of Maryland
3115E Tydings Hall
7343 Preinkert Dr.
College Park, MD 20742
chny.yang@gmail.com

1 Introduction

Electricity production accounted for over a quarter of both U.S. and global carbon emissions in 2021 (CBO, 2022; IEA, 2022*a*,*b*). Many countries' climate goals center on transitioning to a low carbon electricity grid while simultaneously electrifying heating and transportation. Meeting these goals will require massive investment in wind and solar powered generators.

Yet, connecting an electricity generator to the U.S. power grid is increasingly difficult. The process, known as interconnection, takes several years. It can also be costly: connecting generators must often pay to upgrade the transmission infrastructure because the local grid is at capacity (Plumer, 2023). Renewable energy developers cite interconnection as the single biggest hurdle they face (Driscoll, 2022), and less than a quarter of the wind and solar generators that start the process complete it (Rand et al., 2021). We study the design of this interconnection process.

Interconnection works as follows. A generator wishing to connect to the transmission grid joins a waitlist known as the interconnection queue. The grid operator then conducts a series of engineering studies to determine whether the new generator will overload the grid, and, if so, the cost of the new equipment, such as transmission lines, that is necessary to resolve the overload ("interconnection cost"). The grid operator studies generators in the queue on a first-come, first-served basis. Generators in the queue must pay for the studies to remain in the queue and can drop out at any time. A generator usually needs three studies to arrive at a final cost. After the final study, the generator can connect by paying the interconnection cost, or it can leave the queue.

From an economic perspective, the current queuing process is far from optimal. Priority is by entry date, but the probability of completion is significantly different across generators with the same final interconnection cost. Thus, while grid operators struggle to keep up with an influx of interconnection requests from renewables, many studies are done for generators with low probabilities of completion. The cost of being in the queue is also low and may be insufficient to offset the congestion externality imposed on other generators.

We use novel data to study the design of this process. We hand collect these data for the PJM, the largest of the U.S. regional grid operators that serves 65 million people in parts of the Mid-Atlantic, Midwest, and Southern United States (PJM, 2021*b*). Our data cover all generator interconnection requests from 2008 through 2020: 4,085 requests and the 7,117 engineering studies for these requests.

Using this new dataset, we quantify a congestion externality: studies are returned more slowly when more generators are waiting in the queue. Starting in 2018, there was a dramatic rise in the number of generators queuing each year. This increase is explained by both a transition to renewables and renewable generators being much smaller on average than fossil fuel generators. U.S. grid operators struggled to keep up with demand for interconnection studies.[1] One of the most striking examples is PJM. It announced in 2022 that it would not review new interconnection requests again until 2026 while it works to clear its backlog (Howland, 2022). We use variation in a generator's queue position to estimate the effects of this backlog. We find a 10 percent increase in the number of higher queued generators reduces a generator's probability of receiving the third and final study by 5 percent, on average.

We next show that interconnection costs can be very high and are hard to predict. The distribution of interconnection costs has a long right tail. In the second study, the median interconnection cost is \$0.05 million per megawatt, and the 90th percentile is \$0.41 million per megawatt, roughly a quarter of the installation cost for wind and solar generators. We also find that observably similar generators can have very different interconnection costs, a finding consistent with project developers' complaints that interconnection costs are unpredictable (Caspary et al., 2021).

Finally, we show that interconnection costs are a key factor in generator decisions to withdraw from the queue. After every study, generators with high interconnection costs are more likely to withdraw from the queue. For example, a generator with a second study interconnection cost above 0.1 million per megawatt (33 percent of generators) is 49 percent more likely to withdraw from the queue prior to receiving the third study. We find similar results when we control for distance to the grid connection point, a proxy for permitting difficulties, and when we instrument for interconnection costs with the difference in costs across studies.

We then develop an empirical model of queuing to study the incentives of potential generators. We model withdrawal decisions as an optimal stopping problem. A generator waits in the queue for the necessary studies and forms beliefs about when the next study will arrive and what its interconnection cost estimate will be. The continuation value of a generator depends on its characteristics and the status of the queue. We also develop a tractable queuing equilibrium concept that accommodates the non-stationarity in the data.

We estimate the model in two steps. First, we use a rich set of generator and queue characteristics to jointly model the arrival of new studies and the update of interconnection costs. These estimates allow us to construct generator beliefs. Second, we embed these beliefs in the

[1] While we focus on the United States, this pattern of increased requests and grid connection challenges are widespread, including in Australia, France, Italy, Spain, and the U.K. (Mooney, 2023).

generator's optimal stopping problem, and use the observed withdrawal decisions and the dynamic model of queuing to recover the waiting cost and payoff for completing interconnection. The generator's decision to begin operation after receiving its final interconnection cost helps to identify the payoff function. Interim interconnection cost estimates in prior studies shift generator expectations of the final cost and help to identify the waiting costs.

Using the estimated model, we first find that reducing waiting time significantly increases completed capacity. A 10 percent increase in the study arrival probability increases total completed capacity by 4.0 GW, or 5.3 percent. Renewable capacity increases by 1.6 GW, or 4.6 percent. Reducing waiting not only shortens the deferment of payoffs, but also reduces a generator's waiting cost, such as the cost of leasing land, and exposure to other causes of withdrawals, such as the expiration of signed long term contracts.

This result motivates potential gains from modifying the queuing mechanism to reduce congestion. We solve for an alternative queuing mechanism that excludes a set of generators from the queue to maximize completed capacity. This mechanism removes a large share of generators below 100 megawatts in size. By facilitating the completion of larger generators, this mechanism increases completed capacity by 7.9 GW, of which 1.5 GW is renewable. We also consider other mechanisms that put a greater weight on renewable capacity.

We next consider a series of proposed reforms. We first consider a flat entry fee, which also screens out small generators. We find that an entry fee of 900,000 dollars per generator maximizes the completed capacity, adding 6.9 GW of capacity, but only 0.8 GW of renewable capacity. This policy also removes many more small generators than the queuing mechanisms we considered above. We find that increasing the fee to request later studies does not increase completed capacity. We also consider an alternative way of assessing interconnection costs. To reduce uncertainty, PJM plans to study large clusters of generators together and share the total interconnection costs equally on a per-megawatt basis. We show that, with a subsidy, assessing costs in this way can add more renewable capacity than other reforms.

The climate impact of these policy reforms is large. Using the U.S. Environmental Protection Agency's Avoided Emissions and Generation Tool (AVERT), we calculate the avoided carbon emissions from the added renewable capacity. At a social cost of carbon of $185 per ton, the implied annual benefit is 259 million dollars per GW of added renewable capacity.

Related Literature

In this paper, we study electricity transmission policy and formally account for the effects of interconnection queues. Although the interconnection process is a key step in entering the electricity market and has received some attention in research on energy policies (e.g., Gergen, Cannon Jr and Torgerson (2008); Alagappan, Orans and Woo (2011)), it has been rarely studied in the economics literature, likely due to a lack of data. In a related paper, Gonzales, Ito and Reguant (2022) study how the expansion of the transmission grid enables the entry of renewables. More broadly, in considering the economic implications of electricity transmission policy, this paper relates to papers studying the effects of transmission constraints on competition, emissions, and renewable energy investment (Wolak (2015); Ryan (2021); Davis and Hausman (2016); Fell, Kaffine and Novan (2021); LaRiviere and Lyu (2022); Doshi (2022)).

We also contribute to the literature on how public policy affects investment in renewable energy (see, e.g., Metcalf (2010); Hitaj (2013); Johnston (2019); Aldy, Gerarden and Sweeney (2023); Deschenes, Malloy and McDonald (2023)). More broadly, there are a number of papers in the energy and environmental literature that study investment and industry dynamics (Ryan (2012); Gowrisankaran, Reynolds and Samano (2016); Fowlie, Reguant and Ryan (2016); Blundell, Gowrisankaran and Langer (2020); Butters, Dorsey and Gowrisankaran (2021); Elliott (2021); Gowrisankaran, Langer and Zhang (2022); Abito et al. (2022); Covert and Sweeney (2022); Davis, Holladay and Sims (2022); Leisten and Vreugdenhil (2023)). The cited papers are unified in focusing on how environmental regulations interact with dynamic incentives in equilibrium.

Our study uses data from PJM, the largest U.S. transmission organization by the number of customers served. A number of papers (e.g., Mansur (2007, 2008); Bushnell, Mansur and Saravia (2008); Allcott (2012); Buchsbaum et al. (2022)) also use data from PJM but focus on market structure issues, while Linn and McCormack (2019) study exit by coal-fired power plants.

Finally, we contribute to the empirical literature on dynamic assignment mechanisms (e.g. Agarwal et al. (2021); Waldinger (2021); Verdier and Reeling (2021); Liu, Wan and Yang (2021)) by studying a queuing problem in a novel and important market. We develop a tractable queuing equilibrium concept for a non-stationary environment. Our equilibrium concept is similar to Weintraub et al. (2010), and we use a finite horizon assumption to capture the non-stationarity in the data (Igami (2017); Yang (2020)). To understand the equilibrium effects of congestion, we solve for alternative queuing mechanisms that maximize various planner objectives (Che and Tercieux (2021)).

2 PJM Interconnection Process

We first describe how PJM manages the interconnection process in more detail. The two main types of participants are developers and transmission owners. Developers (e.g., NextEra Energy) enter their generators (e.g., 2.5MW Front Royal Solar Field in Virginia) in the queue and pay the interconnection costs identified in the studies. These developers can be either independent producers (more common for renewable) or regulated utilities (more common for natural gas).

Transmission owners conduct the interconnection studies and construct the upgrades (Connell and McGill, 2020).[2] They are regulated utilities, and we assume they do not exercise market power, i.e., they charge the cost of upgrades in a competitive factor market.[3]

2.1 Entering the Queue

To enter a generator in the interconnection process, a developer must secure land sufficient to build the generator (PJM, 2021*f*). Developers must also pay a deposit to enter the queue or move on to the next stage. The deposit amount depends on the generator size and stage, with larger generators and later studies typically requiring higher deposits. For the median generator size of 20 MW, the three deposits would be 12,000, 10,000 and 50,000 dollars. Generators that withdraw have their deposit returned, less a 10 percent non-refundable portion and any study costs already incurred (PJM, 2021*c*).

[2]While new generators pay for transmission network upgrades through the interconnection process, other transmission investment is planned by the grid operator. In PJM, this transmission planning process is called RTEP, which stands for regional transmission expansion plan. The primary goal of the RTEP is to maintain reliability (PJM, 2021*d*). At a high-level, the two types of transmission investment are substitutes, but they are funded differently. Connecting generators pay for the network upgrades they trigger through the interconnection process. In contrast, electricity consumers pay for RTEP investment via higher transmission rates (PJM, 2021*d*). In our analysis, we treat RTEP investment as fixed. Twenty-seven billion dollars worth of RTEP investment was completed from 2008-2020 (PJM, 2023). Locations with recent RTEP investment are associated with increased entry and a moderate decrease in interconnection costs (Table F.2). While generator entry responds to RTEP, the RTEP process does not consider generators waiting in the queue or try to anticipate entry by generators (PJM, 2021*d*). More generally, larger-scale transmission infrastructure is challenging to build due to difficulties surrounding siting and permitting, overlapping jurisdictions, and cost allocation (Davis, Hausman and Rose, 2023).

[3]The transmission owner does not profit from this investment because upgrades paid for by connecting generators do not go into its rate base.

2.2 Queuing Rules

Generators can enter the queue at any time, but each year is divided into two windows. Generators that enter in the same 6-month window are put in the same cohort and will receive up to three studies (feasibility, system impact, and facility study) sequentially. Through these studies, generators learn increasingly accurate information about the costs of interconnection. To receive the next study, a generator must incur a cost, but it can freely leave the queue at any time. PJM may requires just one or two studies if it determines that a generator is not required to make significant network upgrades or share costs with other generators. After the last study is issued, the generator chooses to leave the queue or sign an interconnection service agreement in which it agrees to pay the final interconnection cost, thus completing the interconnection process.

The official timeline for the interconnection process is quite rigid. For generators that apply within the same time window, PJM starts conducting the first studies (feasibility studies) one month after the closing of the window. Within three months, generators are supposed to receive their first studies. At this point, generators have another month to decide whether to advance to the second study (system impact study). The second study then takes four months, at which point generators have one month to decide whether to request the third and final study (facility study). The third study takes 6 months. Finally, generators and PJM agree on final details and sign the interconnection service agreement, over a 6.5-month period (PJM, 2021*a*).[4]

Despite this timeline, significant delays in delivering studies can occur due to the number of backlogged interconnection requests and a lack of staff capacity (Shoemaker, 2021). Upon receiving a study, a generator has approximately one month to decide whether to pay the deposit to request the next study, but it has little control over when the transmission organization delivers the study. A solar developer in PJM recently lodged a complaint with Federal Energy Regulatory Commission (FERC) after waiting more than two years for the second study (Hale, 2021).

2.3 Timing of Interconnection within Project Development

Generators apply for interconnection early in project development. After a developer secures the land for the generator, the permitting and interconnection process occur simultaneously.

[4] This agreement also sets a commercial operation date. After signing the interconnection service agreement, the generator may suspend the process for up to 3 years (up to 1 year if the suspension has a negative impact on subsequent generators), though these suspensions are rare (PJM, 2021*e*).

Compared to permitting, interconnection delays and costs are perceived as the more important hurdle (RechargeNews (2021); Collier (2021)). Renewable energy generators sign long term contracts to sell the power while in the queue or once the interconnection process is completed. After the long term contract and interconnection service agreements are signed, the physical generator and interconnection facilities are constructed. This construction typically takes less than a year, far shorter than the time the generator spends in the queue.[5]

A slow interconnection process can adversely impact other steps in project development. Power buyers are willing to sign long term contracts a few years before a generator begins operation, but developers must apply for interconnection much earlier. Developers are thus left trying to forecast demand a few years in advance and enter generators into the queue accordingly. There have also been cases where developers with long term contracts are unable to begin operation as planned because they are still waiting for interconnection. A long wait may also cause a developer's option to lease to expire, prompting the developer to relinquish site control or renegotiate with landowners.

2.4 Proposed Reforms

There have been several recent efforts to reform the interconnection process, which a key regulator described as in "chaos" (Potter, 2021). In 2023, the Federal Energy Regulatory Commission (FERC), the federal regulator overseeing U.S. transmission policy, approved new rules governing the interconnection process. Broadly, the new rules try to transition interconnection queue priority from "first-come, first-served" to "first-ready, first-served". The rules increase the cost for generators to enter and remain in the queue, move from separate studies for each generator to studying several proposed generators together, and penalize transmission owners for delays in completing studies (Hale and Christian, 2023).

PJM plans to implement reforms consistent with FERC's proposed new rules. In addition to increasing the cost of entry,[6] PJM transmission owners will group generators into large clusters and study all generators in a cluster jointly, issuing each study at the cluster level rather than the generator level. Generators within each cluster will share the identified interconnection costs proportionally, thus reducing the variability in interconnection costs (PJM, 2022).

[5]This timeline is based on the timeline for a wind generator in AWEA (2019); wind and solar generators follow similar timelines.

[6]For example, PJM plans to require larger deposits and stricter site control requirements. Deposits will be partially a share of the generator's network upgrade costs that increases with each of the three study phases. Generators must also demonstrate site control throughout the process rather than only upon entering the queue.

Table 1: Summary Statistics

	Study 1		Study 2		Study 3	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Cost per MW	0.12	0.40	0.18	0.48	0.10	0.16
Wait time (mos.)	5.3	2.7	12.7	11.3	19.0	13.0
Size in MW	97	196	106	191	162	276
Uprate	0.21	0.41	0.23	0.42	0.16	0.37
Solar	0.60	0.49	0.60	0.49	0.59	0.49
Natural Gas	0.17	0.37	0.16	0.36	0.22	0.42
Wind	0.09	0.29	0.10	0.30	0.12	0.32
Battery	0.10	0.30	0.10	0.30	0.04	0.19
Coal, oil, diesel	0.02	0.13	0.01	0.12	0.01	0.09
Other	0.03	0.18	0.03	0.17	0.02	0.15
Cost sharing	0.04	0.19	0.60	0.49	0.60	0.49
Study 1 cost sharing	0.42	0.49	0.48	0.50	0.45	0.50
Receive engr. tests	0.82	0.39	0.88	0.32	0.04	0.21
Distance to substation (km)	3.81	5.99	3.71	6.18	3.40	5.17
Ordinance	0.28	0.45	0.31	0.46	0.30	0.46
N	4,083		2,433		672	

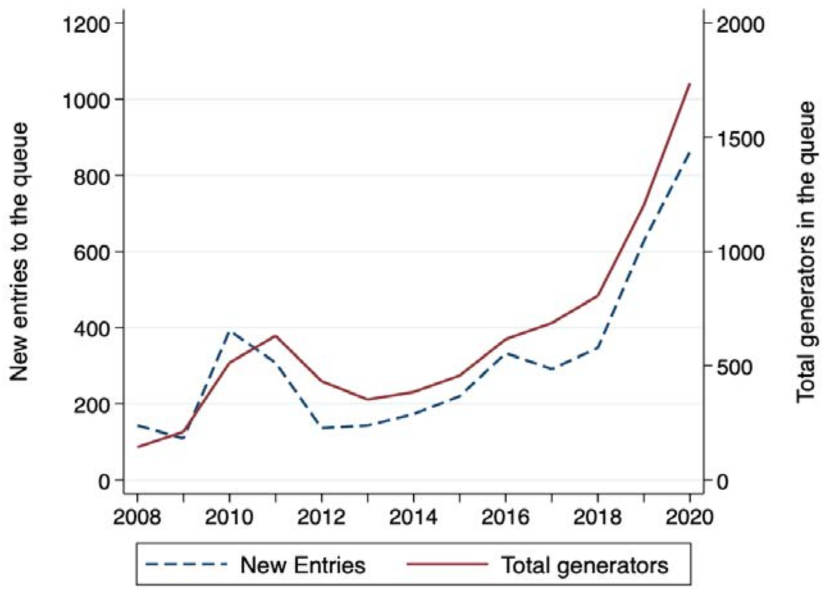
Generators entering the queue in 2008-2020. Costs in millions of 2020 dollars. Cost per MW is interconnection cost estimate divided by the generator's size in MW. Wait time for Study 1 is wait in months for the first study after joining the queue. Wait time for Study 2 is wait in months for second study after receiving the first study. Wait time for Study 3 is wait in months for third study after receiving the second study. Uprate is an indicator for a capacity increase to an existing generator. Cost sharing is an indicator for if a generator shares costs with other generators. Study 1 cost sharing is an indicator for if the first study mentions shared network upgrade costs. Receive engr. tests is an indicator for receiving any of three engineering tests: generator deliverability, multiple facility contingency, and short circuit analysis. Distance to substation is the distance to the nearest substation in km. Ordinance is an indicator for a local ordinance restricting renewable energy development.

3 Interconnection Request Data

Our main data are based on the 4,085 interconnection requests in PJM from 2008 to 2020. These data come from pdfs of 7,117 engineering studies done as part of the interconnection process. Because the formats are irregular, we hand collect these data. We start our sample in 2008 because data before 2008 have even more irregular formats, making it hard to identify the relevant costs. Summary statistics are presented in Table 1 and Appendix Table F.1.

The PJM queue is dominated by requests for solar, natural gas, and wind generators. These three fuels accounted for 82 percent of interconnection requests from 2008-2020. Natural gas generators tend to be much larger than wind and solar generators; they account for 14 percent of requests but 40 percent of requested capacity. Appendix Figure F.1 shows the proportion of new requests by fuel type in each of our sample years. In the rest of the paper,

Figure 1: Queue Size Over Time



Solid red line is annual avg. number of generators in the queue (measured from entry to final study receipt or withdrawal). Dotted blue line is total number of new entries to the queue by year.

we refer to wind and solar generators as "renewable", and the rest as "non-renewable".

The number of interconnection requests has increased over time. Figure 1 shows both the number of new interconnection requests each year (dotted line) and the average number of generators in the queue by year (solid line).[7] Requests increased dramatically starting in 2015. This increase was driven by renewable generators. Because renewables are smaller on average, the increase in requested capacity was less pronounced: the capacity of new requests in 2008 was 42 GW compared to 69 GW in 2020.[8]

3.1 Waiting times

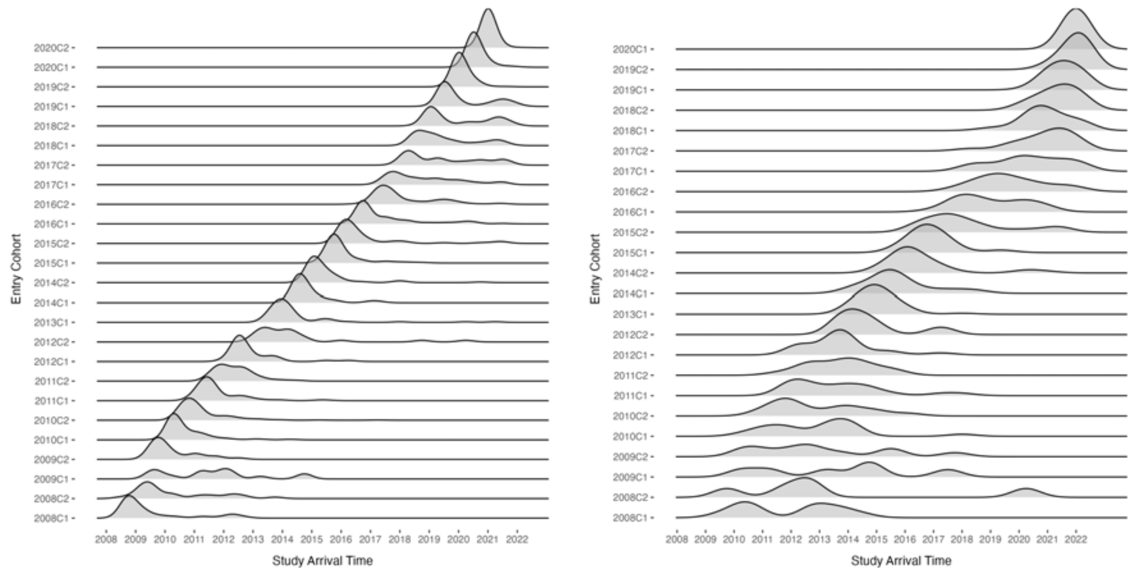
The official timeline is that the studies should take no more than 3, 4, and 6 months. Over three quarters of generators receive the first study within 6 months. The second study's arrival time is more variable; the mean wait time is 10 months with a standard deviation of 7 months. Finally, the mean wait for the third study is 15 months with a standard deviation of 9 months. The mean waiting time to receive the terminal study is 24 months.

[7] The queue size is based on the queue and withdrawal dates for all generators queued between 1997 and 2020.

[8] The spike in 2010 was due to an influx of solar generators. The likely cause was a temporary program that offered the federal investment tax credit for solar investment as a cash grant rather than a nonrefundable tax credit (Aldy, Gerarden and Sweeney, 2023).

Figure 2: Study Arrival Times by Date of Queue Entry

(a) Study 2 (b) Study 3



Panel (a) shows the distribution of arrival dates for the second study by entry cohort. Panel (b) shows the distribution of arrival dates for the third study by entry cohort.

While priority is by entry cohort, study arrival is stochastic. Figure 2 plots the distribution of arrival time of studies by entry cohort. Cohorts that entered the queue earlier are more likely to receive a study then generators in later cohorts that are waiting for the same study, but this is not guaranteed. Time spent in the queue increased over our sample, though not as dramatically as the number of generators. The mean wait time from entering the queue to receiving the terminal study was 20 months for generators queuing in 2008-2012 compared to 27 months for those queuing in 2013-2017. [9]

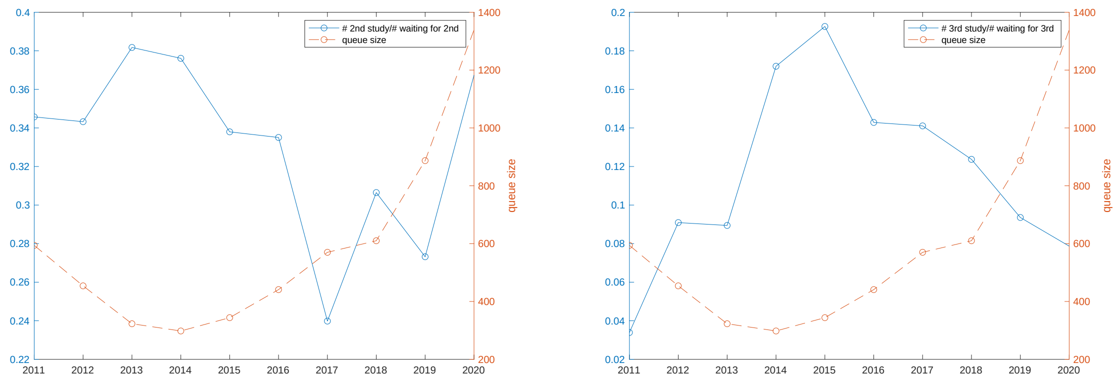
The growing queue mainly delays the third study. In Figure 3, we separately plot the proportions of generators eligible for the second study and the third study that receive one in each year. The arrival probability of the second study declines starting in 2017 but reverts to the 2011 level in 2020 despite a large queue. However, the third study arrival probability is highly negatively correlated with the queue size. In Appendix C.2, we formally estimate the arrival process after controlling for a large number of generator characteristics. The estimation leverages the cross-sectional variation in queue positions,[10] and we find evidence for a

[9] One reason wait times are variable is that these studies sometimes need to be revised. For example, 19 percent of generators in our sample had their second study revised. Revisions can be due to the connecting generator changing its request or nearby generators dropping out of the queue. We will model the arrival of the final version of each study, not the intervening versions which are not always posted by PJM.

[10] In period t, we compute a generator's queue position as the number of generators that entered the queue in the same or an earlier entry cohort and are still in the queue.

Figure 3: Queue Size and Study Arrival

(a) Second Study (b) Third Study



Panel (a) blue line shows annual proportion of generators waiting for the second study that receive it. Panel (b) shows annual proportion of generators waiting for the third study that receive it. Red lines show avg. number of generators in the queue by year (from first study receipt to final study receipt or withdrawal).

congestion externality: the probability a waiting generator receives its third study falls in the number of higher queued generators waiting for the third study.

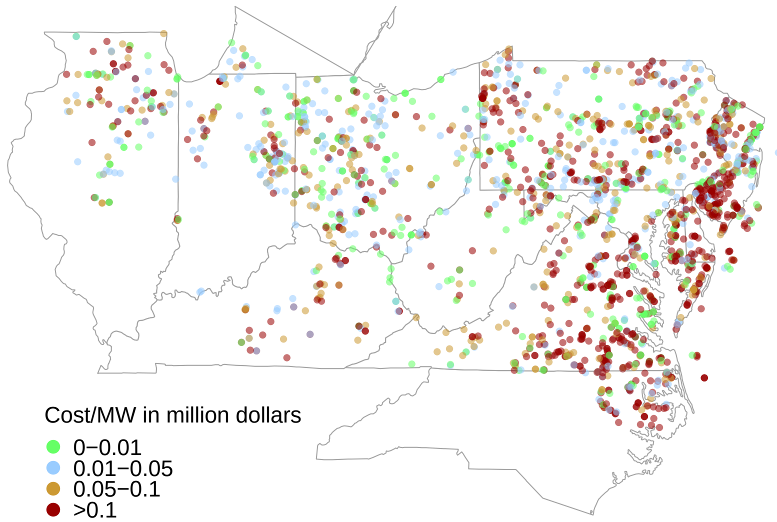
3.2 Interconnection Costs

Interconnection costs are often comprised of two components: the direct cost to connect the production facility to the grid, and the indirect cost of upgrading the network to avoid an overload. Both components may require building or upgrading lines, circuit breakers, and even a substation. PJM expressly states in studies that the interconnection costs do not consider permitting costs or rights of way. The first study (feasibility study) does only some of the required engineering tests. This study reports preliminary direct interconnection costs and whether the generator may share costs with other generators using common transmission infrastructure. The second study completes the remaining engineering tests, updates the direct interconnection costs, and reports the network upgrade costs. A generator may be solely responsible for network upgrades or share this cost with other generators. The third study details the engineering specification of all upgrades and provides a final update to the costs.

While the median interconnection cost estimate is close to zero, the cost distribution is right skewed. We use cost per megawatt (MW) as our measure of cost.[11] For the second study,

[11]There do not appear to be economies of scale for moderately sized generators. For generators from the 10th to 90th percentile in size, a 1 standard deviation increase in capacity is associated with 0.05 standard deviation decrease in the second study interconnection cost.

Figure 4: Costs by Location



Second study interconnection cost estimates by location. Costs in millions of 2020 dollars per MW.

thirty percent of generators have interconnection costs less than 0.01 million per MW. Yet, the 75th percentile of the cost distribution is 0.15 million per MW and the 90th percentile is 0.42 million per MW. For comparison, installation costs for wind and solar generators are roughly 1.5 million per MW. The median interconnection cost per MW is also higher for renewables (0.08 vs. 0.02).

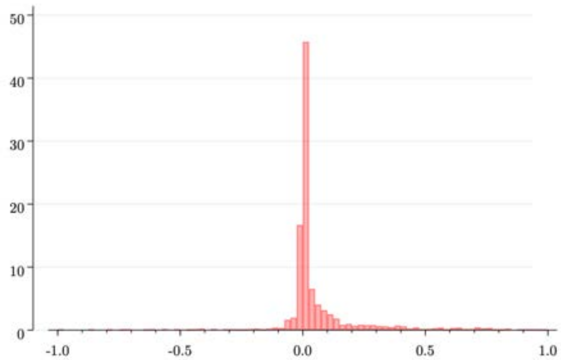
3.2.1 Interconnection Costs are Hard to Predict

There are several sources of uncertainty about interconnection costs. First, although a developer may be able to use engineering models to identify whether an interconnection will cause instability, it is hard to predict what remedies a transmission owner will require and the cost of these remedies, according to our interview with an industry expert. Second, the withdrawal of other generators sharing the cost of a network upgrade changes a generator's cost. After withdrawals, PJM may still require the upgrade, so the remaining generators' share of its cost increases. If PJM no longer requires the upgrade, the cost for the remaining generators will decrease. Third, due to the long waiting time, technologies like solar inverters evolve, and a developer may need to change the production equipment, causing the grid operator to reevaluate the generator and assess a different cost.

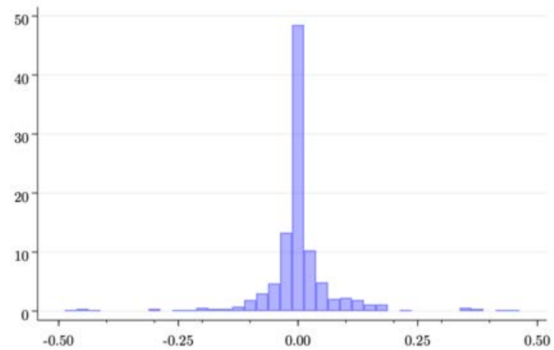
Observable generator characteristics explain some but far from all of the variation in interconnection costs. Appendix Table F.2 shows that a regression of having a low second study

Figure 5: Differences in Cost Estimates across Studies Within A Generator ($Million/MW)

(a) Study 2 - Study 1



(b) Study 3 - Study 2



Panel (a) shows a histogram of the difference in the first and second study interconnection cost estimates for generators queuing from 2011-2020 that received study 2. N =2,023; 50 generators with cost differences above 1 in magnitude excluded. Panel (b) shows a histogram of the difference in the second and third study interconnection cost estimates for generators queuing from 2011-2020 that received study 3. N =479; 3 generators with cost differences above 0.5 in magnitude excluded. Y-axis is the percent of observations in each bin. Costs are in millions of 2020 dollars per MW.

interconnection cost on characteristics such as size, state, fuel type, and year of entry has an R^2 of 0.41. Geography also does not explain much of this variation. Figure 4 plots the location of generators with study 2 cost estimates.[12] More generators are clustered along the more populated east coast. These generators have higher interconnection cost estimates, on average, but interconnection costs can vary substantially in the same geographic area, even within fuel type.[13]

We next show that changes in costs across studies are also hard to predict. Figure 5 shows that interconnection costs for the same generator do not systematically decrease across studies.[14] This pattern suggests that cost changes are not anticipated. If developers were able to predict how a generator's interconnection cost would evolve across studies, we would expect selection on this difference. Appendix Figure F.2 also shows that generators disproportionately withdraw from the queue in the two months after receiving a study, which suggests the studies provide new information.

[12]We focus on costs from the second study because the first study does not typically indicate a generator's contribution to shared network upgrade costs (see Appendix A for more detail and an example of the information available in the first study). This is a selected sample because generators with high interconnection costs are more likely to drop out after the first study.

[13]We observe a similar pattern when plotting the residuals after regressing costs on generator characteristics.

[14]Costs on average increase from the first to the second study because the second study includes the contribution to shared costs. For generators that do not share costs, the distribution of this cost difference is symmetric around zero.

3.2.2 High Interconnection Costs Lead to Withdrawals

We next test whether generators with high interconnection costs are more likely to withdraw from the queue. Specifically, we regress an indicator for withdrawing from the queue on an indicator for having a high (above 0.1 million per MW) interconnection cost. We define withdrawing as leaving the queue before the next study arrives, or before beginning operation for generators that have received their final study.

Table 2: Interconnection Costs on Probability of Withdrawing from the Queue

	Study 1		Study 2			Study 3	
	OLS	OLS	OLS	IV	IV	OLS	IV
Cost above 0.1m/MW	0.123***	0.123***	0.231***	0.293***	0.241***	0.113*	0.074
	(0.022)	(0.022)	(0.031)	(0.054)	(0.057)	(0.063)	(0.117)
Study 1 cost sharing		0.018			0.106***		
		(0.020)			(0.033)		
Log total for sharing		0.015***			0.020**		
		(0.005)			(0.009)		
Mean of dependent var.	0.28	0.28	0.43	0.43	0.43	0.55	0.55
F-statistic (instrument)			972		808		202
N	3,191	3,191	1,269	1,269	1,269	345	345

Generators queuing from 2011-2020; generators still active excluded. SEs in parentheses; clustered by substation. Dep. var. are indicators for projects withdrawing from the queue before receiving the next study or before beginning operation for generators with their final study. Cost above 0.1m/MW indicates if that study's interconnection cost estimate is above 0.1 million dollars per MW. Study 1 cost sharing indicates if the first study mentions shared network upgrade costs. Log total for sharing is the log of the total costs to be shared listed in Study 1 for generators with cost sharing. IV results instrument for having a high cost using the change in having a high cost across studies. All specifications control for size (5 bins), fuel type, state, uprate, and FE for the year of queue entry and the year the study is issued. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Across all three studies, we find that generators with a high interconnection cost estimate are more likely to withdraw from the queue (Table 2). For the first study, the OLS estimates in the first column imply that having a high interconnection cost is associated with an increase in the probability of withdrawal of 12 percentage points, or 44% at the mean withdrawal rate. The estimated effect is similarly large for other studies.[15]

We find similar estimates when we instrument for costs with the change in these costs across studies. Generators may have private information about costs. The resulting selection would attenuate the relationship between interconnection costs and withdrawals. To address this concern, we instrument for current interconnection costs with the innovation to these costs, i.e., the difference between the current cost and the cost in the previous study. This

[15] Our result is robust to using a continuous measure of costs or more cost bins (Appendix Table F.3).

instrument is valid because changes in interconnection costs across studies are hard to predict (Section 3.2.1). For the effect of the second study costs on withdrawals, the IV estimates are similar to the OLS estimate and statistically significant. For the effect of third study costs, the IV estimate is positive and similar in magnitude but not precisely estimated.[16]

3.3 Speculative Interconnection Requests

Requests for interconnection are inexpensive and have a high option value, so developers enter many more requests than the number of generators they expect to build. This phenomenon is commonly referred to as speculative interconnection requests. We take this behavior to mean that most potential generators were entered into the queue during the period we study.[17]

To understand whether strategic interactions are important, we collect data on the identities of developers. The name of the developer is listed in the studies starting with the second study. This name is often the limited liability corporation that owns the generator (e.g., "7 Bridges Solar, LLC"), and we use local news articles, regulatory documents, and developer websites to match each generator to its developer (e.g., "Longroad Energy"). Of the generators in our data, we are able to identify the developer for 39% overall, 52% that reached study 2, and 81% that reached study 3.

These data show that concentration is low. We match 1,574 generators to 383 unique developers, and the largest developer ("Invenergy") accounts for only 5.0% of these generators. This lack of concentration suggests that strategic interactions may be less important in this setting.

We next turn to the question of whether developers are submitting multiple interconnection requests with the intention of building only one generator. Our data suggest that

[16]In Appendix Table F.4, we also find these effects are not driven by permitting difficulties. Specifically, we show they are robust to including controls for whether the county has an ordinance restricting the siting of renewables and the distance to the point of interconnection.

[17]There is some debate about the quality of the marginal entries into the queue. The following quote sums up the perspective that these entries are potentially viable projects:

> *Maybe there are some people who carpet bomb the queues with speculative projects, but I think in general they appear to be speculative because people know it's going to take five years to get through the process, so you have to do that early on. It would be unwise to fully develop your site prior to entering a queue that you have no certainty on getting through, especially because so many things can change in five years.*- Boone Staples, director of transmission analysis at Tenaska (Penrod, 2022)

Another perspective is that these marginal requests are not viable; for example, FERC says its 2023 reforms will "discourage speculative, commercially non-viable interconnection requests" FERC Staff (2023). This concern may be especially relevant for requests put in by firms that have not previously developed projects, though we note that developers with projects in the queue can and do sell these projects to other developers.

modeling the generator's rather than the firm's problem is a good approximation. For cases where a developer had more than one generator in an entry cohort, either all or none of the generators were completed 71% of the time. Appendix Table F.5 reports this same fraction for the top 15 developers in our sample. Overall, we view these data as generally consistent with developers being willing to build all generators that are individually profitable.

4 A Model of the Interconnection Queue

To understand the effects of the queue backlog and quantify the impact of policy reforms, we develop a finite horizon, discrete-time, non-stationary dynamic model of the interconnection queue. We define a period to be a quarter. We assume that every generator enters the queue with a first study. The timing of the queue in each period is as follows:

1. At the beginning of a period t, the generator observes the cost estimate from the latest study and other time-varying information, such as the current calendar time, how many studies the generator has received, and whether certain engineering tests have been conducted.

2. The generator forms beliefs about whether the next study will arrive in the current period, the new cost estimate, and the other contents of the study. It decides whether to wait or withdraw.

3. For a generator that chooses to wait, a new study arrives with some probability. If the new study arrives,

   (a) For a generator with two studies, the new study is the final study. The generator decides whether to complete the interconnection or withdraw.

   (b) For a generator with one study,

      i. With some probability the new study is the final study. The generator observes the final cost estimate and decides whether to complete the interconnection or withdraw;

      ii. Otherwise, the generator observes the cost estimate and other contents of the study and decides whether to request the next study or withdraw.

   If no new study is issued, the generator continues to the next period.

4. Potential generators decide whether to enter the queue.

4.1 Generator Decisions in the Queue

4.1.1 Notation

We focus on a particular generator's decision and omit the generator subscript in our notation. A generator in calendar time period $t = 1, \ldots, T$ is associated with time-invariant generator characteristics x, such as the size of the generator. While in the queue, the generator incurs a waiting cost of $o_t(\tau, \tilde{\tau}, x)$, where $\tau = 1, \ldots, T_0$ is the number of periods a generator has been in the queue, and $\tilde{\tau}$ is the number of periods since receiving the previous study. Through its dependence on when the previous study was received, the waiting cost accounts for the study fee and deposit required by PJM to advance in the queue.[18] We assume that each generator waits a maximum of eight years ($T_0 = 32$). The last quarter is $T = 85$, corresponding with an end year of 2028.

We use $k \in \{1, 2, 3\}$ to indicate which study the generator has received. We use c to denote the interconnection cost estimate from the latest study, and z to denote other information from previous studies.[19] In our analysis, we focus on two sets of contents in the study and specify $z = \left(z^{\text{test}}, z^{\text{cost-sharing}}\right)$. The first component z^{test} is an indicator for whether PJM has conducted a set of engineering tests. These engineering tests quantify how much a new generator will overload the grid. The variable $z^{\text{cost-sharing}}$ is another indicator for whether PJM deems the generator part of a cost-sharing group collectively responsible for the same transmission upgrade.[20]

We use y_t to denote the equilibrium queue state that affects the cost and waiting time of a generator. The queue state has two main components that are both generator specific: (1) the generator's queue position, which depends on the number, composition, and actions of other generators with earlier entry times, and (2) the number and sizes of other generators in the same region. Given the "first-come-first-served" rule, the queue position is the main factor determining the arrival time of a study. A generator's outcome also depends on what other generators are present in the queue and geographically close. For example, PJM may jointly study generators that overload the same circuit breakers and ask them to share costs. We use a vector of variables in y_t to flexibly account for these effects. We discuss the detailed

[18]Deposits are paid in the first period after receiving the previous study.

[19]The information may be cumulative. For example, each study contains results from different engineering tests, and the generator needs to aggregate these results to form beliefs about the final interconnection cost.

[20]Our estimates indicate that cost-sharing has, at most, a moderate effect on both the costs and study arrival. We therefore do not explicitly model the interactions between generators sharing costs. We also note that this is likely a PJM-specific feature. For example, the way costs are shared in another grid operator SPP can lead to a waiting game since costs often substantially decline when generators in the cost-sharing cluster withdraw (Charles et al. (2023)).

specification of y_t in Appendix C.1.

We use $\pi_t(\tau, x)$ to denote the generator's expected discounted operating profit when it completes the interconnection process in period t after waiting for τ periods.[21] We assume that the support of $(k, c, z, \tau, \tilde{\tau}, t, x)$ is discrete.

4.1.2 Belief Assumptions

We start with the belief of a generator that has received two studies and is waiting for the third and last study. Given the latest cost estimate c and study information z, we assume the belief about the arrival of the third study with a cost estimate c' in period t is

$$r_t\left(c'; c, z, \tau, \tilde{\tau}, x\right) \equiv H_3\left(c'; c, z, \tau, \tilde{\tau}, y_t, x\right), \tag{1}$$

where the function H_3 is the probability of PJM issuing the new study given the current status of the queue and the generator state. This function represents PJM's "production process" of studies. We assume this process is invariant to the actions of the generator up to the arguments of the H_3 function, which flexibly include the generator's queue position, whether the generator is a renewable generator, and a large set of other time-varying queue characteristics as well as time fixed effects. We make similar assumptions below about the other arrival probability functions, also denoted by H but distinguished by subscripts.

To simplify a generator's belief, we assume that the generator belief $r_t(\cdot)$ depends on the generator characteristics and the information contained in previous studies, varies over time, and is consistent with the equilibrium queue state y_t. The underlying assumption is that a generator reacts to the state of the queue aggregated in the function $H_3(\cdot)$, but not the actions of individual generators. This "large-market" assumption helps to reduce a generator's state space and has been used to analyze other settings with many strategic players (e.g. Agarwal et al. (2021); Chen and Xu (2023)).

Similarly, the belief about the arrival of a second study that is the final study with a new cost estimate c' is

$$p_t\left(c'; c, z, \tau, \tilde{\tau}, x\right) \equiv H_2\left(c'; c, z, \tau, \tilde{\tau}, y_t, x\right), \tag{2}$$

where H_2 is the probability that a second study arrives and PJM deems a third study to be

[21]We can allow π to explicitly depend on the current and past queue states and account for how the equilibrium queue outcomes may affect the expected payoff. For example, existing entrants may decrease the expected profitability of the focal generator. In a robustness analysis, we estimate a profit function that depends on the completed generation capacity at the transmission owner territory level and find its effect to be negligible.

unnecessary.

Next, we assume that the belief about the arrival of a second study that is not the final study with a cost estimate c' and new information z' in the study is

$$q_t\left(c', z'; c, z, \tau, \tilde{\tau}, x\right) \equiv \tilde{H}_2\left(c', z'; c, z, \tau, \tilde{\tau}, y_t, x\right). \tag{3}$$

The function $\tilde{H}_2$ is the probability of the non-final second study arrival. In both $r_t(\cdot)$ and $p_t(\cdot)$, we assume that generators form beliefs about the costs which will directly enter the payoff function after receiving the final study. In $q_t(\cdot)$, the generator forms the beliefs about the cost c' and other contents z' of the next study. Both the current c and z affect the generator belief.

Finally, we assume that, when a generator enters the queue, the initial cost c and the study information z in the first study are assigned with probability $H_1(c, z; y_t, x)$. We assume the generator belief is consistent with this probability distribution:

$$v_t(c, z; x) = H_1(c, z; y_t, x). \tag{4}$$

4.1.3 Generator Decision

We start from the last period a generator can be in the queue, $\tau = T_0$. A generator that reaches this maximum waiting time receives the outside option and leaves the queue without completing interconnection. We assume the outside option (scrapping the project and relinquishing site control) is valued at $b_t(x) + \xi_t$ in period t, where $b_t(\cdot)$ varies over time and is a function of generator characteristics, and ξ_t is known to the generator but unobserved by the researcher.

For $\tau < T_0$, we first consider the case where the generator has received the final study cost estimate c in calendar time t. We assume the total cost to bring the generator online, including the costs of construction and equipment, is $g_t(x) + c + \varepsilon_t$, where $g_t(\cdot)$ represents how observed characteristics affect the cost, c is the interconnection cost from the final study, and ε_t is the generator-specific unobserved cost. Importantly, $g_t(\cdot)$ flexibly accounts for calendar time to capture exogenous trends such as the decrease in renewable installation costs and changes in subsidy policies. If the expected total profit exceeds the value of the outside option,

$$\pi_t(\tau, x) - g_t(x) - c - \varepsilon_t > b_t(x) + \xi_t,$$

the generator completes the interconnection. The expected value of reaching this stage is

$$\Pi_t(\tau, x, c) = E_{\varepsilon_t, \xi_t} \max\{\pi_t(\tau, x) - g(t, x) - c - \varepsilon_t, b_t(x) + \xi_t\}, \tag{5}$$

where we integrate over the random variables in the expectation sign's subscript.

We next consider the generator decision when it has two non-final studies. Suppose the generator's last study indicates a cost of c. The decision is whether to wait for the third study or withdraw. The option value of waiting depends on the probability $r_t(\cdot)$ of receiving a study next period, the waiting cost $o(\cdot)$, and the value of the outside option. The value of waiting is given by the following Bellman equation

$$\begin{aligned} W_t(c, z, \tau, \tilde{\tau}, x) = E_{\xi} \max \Bigg\{ & b_t(x) + \xi_t, \sum_{c'} r_t(c'; c, z, \tau, \tilde{\tau}, x) \cdot \Pi_t(\tau, x, c') \\ & + \left(1 - \sum_{c'} r_t(c'; c, z, \tau, \tilde{\tau}, x)\right) \cdot W_{t+1}(c, z, \tau + 1, \tilde{\tau} + 1, x) - o_t(\tau, \tilde{\tau}, x) \Bigg\}. \end{aligned} \tag{6}$$

where we take expectations over values of the outside options and the final study cost. We do not separately include a discount factor in addition to the waiting cost. The probability of staying in the queue is the probability that the first term in the maximand is lower than the second term.

Now we consider the decision when the generator decides whether to wait for the second study after entry. The generator may receive a second study that is the final study, a second study that is not the final study, or withdraw. The value of waiting takes into account the respective payoffs, the arrival probability $p_t(\cdot)$ of a final second study, and the arrival probability $q_t(\cdot)$ of a non-final second study:

$$\begin{aligned} V_t(c, z, \tau, \tilde{\tau}, x) = E_{\xi} \max \Bigg\{ & b_t(x) + \xi_t, \sum_{c'} p_t(c'; c, z, \tau, \tilde{\tau}, x) \cdot \Pi_t(\tau, x, c') \\ & + \sum_{c', z'} q_t(c', z'; c, z, \tau, \tilde{\tau}, x) \cdot W_{t+1}(c', z', \tau + 1, 1, x) \\ & + \left(1 - \sum_{c'} p_t(c'; c, z, \tau, \tilde{\tau}, x) - \sum_{c', z'} q_t(c', z'; c, z, \tau, \tilde{\tau}, x)\right) \\ & \cdot V_{t+1}(c, z, \tau + 1, \tilde{\tau} + 1, x) - o_t(\tau, \tilde{\tau}, x) \Bigg\} \end{aligned} \tag{7}$$

In the above, we can normalize the net mean profit π to be $\pi - g - b$. This normalization does not affect the choice probabilities of waiting or completion but simplifies the Bellman equations. With a slight abuse of notation, we still use π to denote the normalized mean profit and write our Bellman equations as

$$\Pi_t(\tau, x, c) = E_{\xi,\varepsilon} \max \left\{ \pi_t(\tau, x) - c - \varepsilon_t, \xi_t \right\}, \tag{8}$$

$$\begin{aligned} W_t(c, z, \tau, \tilde{\tau}, x) = E_{\xi} \max \Bigg\{ & \xi_t, \sum_{c'} r_t\left(c'; c, z, \tau, \tilde{\tau}, x\right) \cdot \Pi_t\left(\tau, x, c'\right) \\ & + \left(1 - \sum_{c'} r_t\left(c'; c, z, \tau, \tilde{\tau}, x\right)\right) \cdot W_{t+1}\left(c, z, \tau+1, \tilde{\tau}+1, x\right) - o_t\left(\tau, \tilde{\tau}, x\right) \Bigg\}. \end{aligned} \tag{9}$$

$$\begin{aligned} V_t(c, z, \tau, \tilde{\tau}, x) = E_{\xi} \max \Bigg\{ & \xi_t, \sum_{c'} p_t\left(c'; c, z, \tau, \tilde{\tau}, x\right) \cdot \Pi_t\left(\tau, x, c'\right) + \sum_{c', z'} q_t\left(c', z'; c, z, \tau, \tilde{\tau}, x\right) \\ & \cdot W_{t+1}\left(c', z', \tau+1, 1, x\right) + \Bigg(1 - \sum_{c'} p_t\left(c'; c, z, \tau, \tilde{\tau}, x\right) \\ & - \sum_{c', z'} q_t\left(c', z'; c, z, \tau, \tilde{\tau}, x\right)\Bigg) \cdot V_{t+1}\left(c, z, \tau+1, \tilde{\tau}+1, x\right) - o_t\left(\tau, \tilde{\tau}, x\right) \Bigg\} \end{aligned} \tag{10}$$

At the terminal period $\tau = T_0$ for each generator without its final study, the normalization implies that $W_t(c, z, \tau = T_0, \tilde{\tau}, x) = V_t(c, z, \tau = T_0, \tilde{\tau}, x) = 0$.

Finally, we model the entry decision as

$$E_{c,z}\left(V_t(c, z, \tau = 1, \tilde{\tau} = 1, x) \,|t, x\right) > \underline{c}_t(x), \tag{11}$$

where the left-hand side is the expected surplus from entering the queue, and $\underline{c}_t(x)$ is the entry cost. The expectation is taken over the belief of the first study cost c and information z conditional on the entry time t and x. The set of N potential entrants is denoted as $\mathcal{X} = \left\{(i, t_i, x_i)_{i=1}^{N}\right\}$, where t_i and x_i are the entry time and characteristics of generator i. A potential entrant i leaves after period t_i if it decides not to enter the queue.

To compute the value of entry, a potential generator may use information from past or current interconnection requests. To partially account for this information, we calculate the number of other generators that have withdrawn from the queue in the past two quarters within a 100km radius and the same transmission owner territory, and include it in the first study cost function. Our entry model does not directly account for the interconnection costs in recent studies of nearby generators. In Appendix B, we find these costs have at most a moderate effect on subsequent entry at the same point of interconnection, especially compared with other factors outside the queue, such as baseline transmission investment by the grid operator.

4.1.4 Queuing Equilibrium

We consider a finite-horizon queuing equilibrium, where the beliefs of the generators are consistent with the state of the queue in calendar time t. The finite horizon assumption allows us to capture the non-stationarity in the cost of wind turbines and solar panels and the increase in the number of entrants.

We use $\Psi_t(c, \tau, x)$ to represent the withdrawal probability of a generator after receiving the final study with a cost estimate c, $\Lambda_t(c, z, \tau, \tilde{\tau}, x)$ the withdrawal probability when it waits for the third study, $\Upsilon_t(c, z, \tau, \tilde{\tau}, x)$ the withdrawal probability when it waits for the second study, and $\Xi_t(x)$ the withdrawal probability of potential entrants. We use $m_t(c, z, \tau, \tilde{\tau}, k, x)$ to denote the fraction of generators waiting in period t with a cost estimate c, time-varying characteristics $\{z, \tau, \tilde{\tau}, k\}$, and time-invariant characteristics x. The equilibrium consists of (1) optimal withdrawal probabilities $\{\Psi_t, \Lambda_t, \Upsilon_t, \Xi_t\}_{t=1}^T$, (2) the beliefs about new studies $\{r_t, p_t, q_t, v_t\}_{t=1}^T$, (3) the composition of the queue in every period $\{m_t\}_{t=1}^T$, and (4) the number of waiting generators N_t in every period $\{N_t\}_{t=1}^T$. The equilibrium queue status that affects study arrival and costs is aggregated from the queue composition and queue size via a function S,

$$y_t = S\left(x, \{m_{t'}, N_{t'}, \Psi_{t'}, \Lambda_{t'}, \Upsilon_{t'}, \Xi_{t'}, r_{t'}, p_{t'}, q_{t'}, v_{t'}\}_{t'=1}^t\right).$$

A queuing equilibrium satisfies the following conditions:

1. Optimality conditions. The withdrawal probabilities $\{\Psi_t, \Lambda_t, \Upsilon_t, \Xi_t\}_{t=1}^T$ are consistent with the Bellman equations in (8), (9), (10) and (11).

2. Consistent beliefs. The generator beliefs $\{r_t, p_t, q_t, v_t\}_{t=1}^T$ about the arrival probabilities of new studies and their contents are consistent with (1), (2), (3) and (4).

3. Balance conditions. For every $t = 1, \ldots, T$,

   (a) The transition of the queue ($\tau > 1$).

i. For generators with two studies and waiting for the third study,

$$\begin{aligned}
&N_{t+1}m_{t+1}\left(c,z,\tau+1,\tilde{\tau}+1,k=2,x\right)\\
&=N_t m_t\left(c,z,\tau,\tilde{\tau},k=2,x\right)\cdot\left(1-\Lambda_t\left(c,z,\tau,\tilde{\tau},x\right)\right)\\
&\cdot\left(1-\sum_{c'}r_t\left(c';c,z,\tau,\tilde{\tau},x\right)\right).
\end{aligned}\tag{12}$$

ii. For generators just receiving the second study,

$$\begin{aligned}
&N_{t+1}m_{t+1}\left(c,z,\tau+1,\tilde{\tau}=1,k=2,x\right)\\
&=N_t\sum_{c',z',\tilde{\tau}'}m_t\left(c',z',\tau,\tilde{\tau}',k=1,x\right)\\
&\cdot\left(1-\Upsilon_t\left(c',z',\tau,\tilde{\tau}',x\right)\right)\cdot q_t\left(c,z;c',z',\tau,\tilde{\tau}',x\right).
\end{aligned}\tag{13}$$

iii. For generators with one study and waiting for the second study,

$$\begin{aligned}
&N_{t+1}m_{t+1}\left(c,z,\tau+1,\tilde{\tau}+1,k=1,x\right)\\
&=N_t m_t\left(c,z,\tau,\tilde{\tau},k=1,x\right)\cdot\left(1-\Upsilon_t\left(c,z,\tau,\tilde{\tau},x\right)\right)\\
&\cdot\left(1-\sum_{c',z'}q_t\left(c',z';c,z,\tau,\tilde{\tau},x\right)-\sum_{c'}p_t\left(c';c,z,\tau,\tilde{\tau},x\right)\right).
\end{aligned}\tag{14}$$

(b) The boundary condition (entry):

$$N_t m_t\left(c,z,\tau=1,\tilde{\tau}=1,k=1,x\right)=v_t\left(c,z;x\right)n_t\left(x\right)\left(1-\Xi_t\left(x\right)\right),\tag{15}$$

where $n_t(x)$ is the number of new generators with characteristics x and entry time t in the set of potential entrants $\mathcal{X}$.

In equilibrium, a generator has perfect foresight over the evolution of the queue but faces uncertainty over its own studies.

5 Identification and Estimation

There are two key sets of parameters in our model. First, we flexibly specify the components of queue state y_t (the S function) and directly estimate the functions that govern how PJM issues studies (H_1, H_2, $\tilde{H}_2$ and H_3) from data. These parameters govern the exogenous transition dynamics of the model between quarters. Appendix C provides estimation details. Our estimates imply that a 10% increase in the number of higher queued generators reduces a generator's probability of receiving the third study by 5%, on average.

Second, we use the estimated functions to construct beliefs and use the withdrawal and completion decisions to recover preferences, which are the parameters of the profit function $\pi_t(\tau, x)$, the waiting cost $o(\tau, \tilde{\tau}, t, x)$, and the distribution of ξ_t and ε_t. Importantly, we assume that the set of potential entrants $\mathcal{X}$ is the set of observed actual entrants. There are two main reasons for this assumption. First, as discussed in Section 2, developers face long waiting times and great cost uncertainty in the queue as well as an extremely low entry cost relative to the generator's installation cost. In response, developers enter a large number of generators into the queue even though many generators have a low probability of completion. Given this unique institutional fact, we think it is reasonable to assume that developers have exhausted all entry opportunities. Second, the utility-scale solar industry, which constitutes the majority of requests, significantly expanded during the latter half of our sample, but the expansion was due to a large decrease in the cost of solar panels that was unrelated to the queuing process. Given these considerations, we fix the entry probability to be 1 in estimation. In Section 5.2, we discuss the assumptions for estimating the entry costs.

5.1 Generator Preferences

5.1.1 Identification

We start with the identification of the payoff function. We take the functions H_1, H_2, $\tilde{H}_2$ and H_3 as given. In standard optimal stopping problems, it is often not possible to separately identify the waiting cost o and the payoff function π because a stopping decision (withdrawal) may be explained by a low payoff or a high waiting cost. Our case is different. We observe two types of decisions: the decision to wait while the generator is in the queue, and the decision to complete the interconnection when the final study is issued. The variation of the generator characteristics and final-study interconnection cost c identify the payoff function following the standard identification argument of binary choice problems (Manski (1988)).

Specifically, we specify the profit and waiting cost functions as

$$\pi_t(\tau, x) = \beta_\pi \cdot d_\pi(\tau, t, x), o_t(\tau, \tilde{\tau}, x) = \beta_o \cdot d_o(\tau, \tilde{\tau}, t, x),$$

where (β^π, β^o) are vectors of parameters, $(d_\pi, (\cdot), d_o(\cdot))$ are vectors of covariates based on current calendar time, waiting time and generator characteristics (size, fuel type, location, and entry time). Intuitively, to identify the effect of an indicator variable on profits, we exploit the variation of the withdrawal probability Ψ implied by the maximization problem in equation (8) across generators conditional on this variable being 0 vs 1, holding other variables fixed. For a continuous variable, we exploit the variation in Ψ across generators with different values of this variable. To see how the distribution of the unobservable is identified, consider, for example, a distribution of ξ_t known up to its variance. If the variance is large, the withdrawal probability will not significantly change with the interconnection cost. In the special case that ξ_t is mean zero and symmetric, the withdrawal probability will be close to 0.5 at a high variance. On the other hand, a small variance implies that withdrawal probabilities would be dramatically different for generators with small differences in interconnection costs. In our sample of generators that completed the interconnection or withdrew after receiving the last study, the median interconnection cost is $9,100/MW. The completion rate is 49.4% for generators below the median, and 28.8% for those above it.

The argument for identifying the waiting cost is constructive. Given the identified profit function π_t, we can use backward induction to construct the value function for any generator in its last period $\tau = T_0$. Then the only unknown parameter in the withdrawal probability at $\tau = T_0 - 1$ is the waiting cost. The withdrawal probability is identified from data, and inverting this probability identifies the waiting cost. We can apply the argument to $\tau, \tilde{\tau} = T_0 - 2, T_0 - 3, \ldots$ and any t. This identification argument relies on the normalization that the mean value of the outside option is 0, which means that we interpret the waiting costs as the difference between the actual waiting costs and the (potentially time-varying) value of the outside option. In the empirical analysis, we adopt a more parsimonious specification for the waiting cost to limit the number of parameters.

One may be concerned about unobserved heterogeneity across generators. In particular, generators at locations that have higher demand may be more likely to stay, and those at locations with higher unobserved costs of building infrastructure may be more likely to exit. To address this concern, we consider the following extension in a robustness analysis, where the profit function is

$$\pi_t(\tau, x) = \beta_\pi \cdot d_\pi(\tau, t, x) + \zeta_{\text{sub}}. \tag{16}$$

The unobservable ζ_{sub} is a substation-level random effect for the nearest substation with a normal distribution and unknown variance σ_{sub}. We use the panel structure of the data, where multiple generators enter near the same substation, to identify σ_{sub}. The intuition is that a larger variance implies stronger within substation correlations in the withdrawal decisions.

5.1.2 Estimation

We use maximum likelihood to recover generator preference parameters in the baseline model based on the 2011 to 2020 sample.[22] We use the estimated transition functions in Section C.2 as generators' beliefs in the dynamic programming problem. To estimate the extension with the random substation effects in (16), we use a simulated maximum likelihood approach.

Table 3: Profit Function Parameter Estimates ($1,000/MW)

	(1) Baseline	(2) Heterogeneity
Capacity (MW)		
10-20	52.30	45.42
	(12.10)	(11.16)
20-100	88.12	85.83
	(12.68)	(12.58)
>100	32.16	32.78
	(14.87)	(14.91)
Renewable	-100.31	-100.56
	(22.67)	(21.66)
$\sigma_{\text{substation}}$		69.85
		(6.97)

Estimated parameters for the profit function in $1,000 of 2020 dollars per MW. Renewable is an indicator for being a wind or solar generator. SE in parentheses. (2) Heterogeneity is a specification that includes random effects at the nearest substation level. $\sigma_{\text{substation}}$ is the standard deviation of this random effect. We also include entry year and year fixed effects at the 3-year level, state fixed effects, and fixed effects for the number of quarters in the queue. Appendix Table F.11 reports estimates for other profit function parameters.

Table 3 presents the estimates of the profit function, β^{π}. Column (1) is our baseline estimate. Column (2) includes substation-level unobserved heterogeneity. We also include state, entry year, and year fixed effects to account for policy and technology heterogeneity across space and time. The estimates of covariate parameters are largely similar across the two specifications. We find that generators with capacity between 20 and 100 MW have a higher net profit on a per megawatt basis than those below 10 MW, and the profit of larger generators is not necessarily higher. Renewable generators have a lower net profit per megawatt of capacity, consistent with their higher withdrawal rates and lower capacity factors.

Appendix Table F.7 presents the estimates for the waiting costs. We find that waiting costs are high in periods immediately after receiving the studies, and the additional cost of waiting for the third study is $3,500 per MW per quarter (e.g., $71,000 per quarter for a 20 MW

[22] We use the pre-2011 sample to compute the queue status (such as the number of higher-queued generators) and hold them fixed.

generator). These costs account for both the administrative costs of maintaining the queue position as well as economic costs such as extending site control. The estimates rationalize the data pattern that a high proportion of withdrawals occurred after the studies are issued (Appendix Figure F.2).

Appendix Table F.11 presents estimates for the standard deviation parameters for the unobservables. The estimated standard deviations of ξ_t, which reflect the unobserved profit and cost shocks after the generator learns the final study cost estimate, are large (greater than $\$100,000$ MW in most years), but have decreased since 2012. Renewables face considerably greater uncertainty than other generators. Unlike other generators, renewables that have been in the queue for more than two years face a higher standard deviation of $\tilde{\xi}_t$, which is consistent with our interviews that these generators often face additional uncertainties about securing a long term contract and sourcing the panels or turbines at the stage of completing the interconnection. These uncertainties are often much greater than ε_t, which is the quarter-to-quarter unobserved outside option as the generator continues to wait. The estimated standard deviation of ε_t is about $\$78,770$ per MW in the baseline estimates, comparable to that of the random effect in specification (2).

5.2 Entry Costs

To estimate the entry cost, we rely on a free-entry assumption:

Assumption. *The lowest expected profit from entering the queue in each year is* 0.

In effect, we assume that the marginal interconnection request just breaks even, which is reasonable given the large number of entrants. We define the expected profit as the expected surplus of entering the queue minus the entry cost in (11):

$$E_{c,z}\left(V_t\left(c, z, \tau = 1, \tilde{\tau} = 1, x\right)|t, x\right) - \underline{c}_t\left(x\right).$$

For simplicity, we assume that the entry cost (on a per megawatt basis) is the same across potential generators in a given year, but this assumption can be relaxed to allow the cost to vary based on observable characteristics, such as location and fuel type. For the entry cost in year j, we compute the lowest expected surplus from entering the queue in year j, defined as

$$\underline{c}_{j(t)} = \min_{x \in \mathcal{X}_j} E_{c,z}\left(V_t\left(c, z, \tau = 1, \tilde{\tau} = 1, x\right)|t, x\right),$$

where $\mathcal{X}_j$ is the set of potential entrants in year j, $j(t)$ denotes the year of the period t, and the value function is given by the Bellman equations (8) through (11).

This procedure gives us reasonable estimates for the entry costs. In Appendix Figure F.3, we present the distribution of the expected surplus by year. We estimate the entry costs are from $48,200 to $81,100 per megawatt. For perspective, annual lease payments for solar range from about $2,000 to $20,000 per megawatt (Parker et al., 2023), though an exclusive option to lease is cheaper. Our estimates of the economic costs of entry also account for other factors, such as deposits to enter the queue, negotiating leases with landowners, building community support, and starting the permitting process.

6 Equilibrium Simulation and Model Fit

We use the estimated model to simulate the queuing equilibrium defined in Section 4.1.4. The details of the simulation procedure are given in Appendix D. We provide a high level summary here.

Algorithm 1 Summary of Equilibrium Simulation

- Initialization
  - Use the observed outcomes in the data to compute the queue status y_t for period $t = 1, \ldots, T$.
  - Compute the distribution of study arrival, cost bins and information z using $H_1, H_2, \tilde{H}_2, H_3$ for each potential entrant i in $\mathcal{X}$ for each $\tau, \tilde{\tau}_i$ after the entry date t_i.
  - Solve the equations (8), (9), (10) and (11), and the associated withdrawal probabilities.
- Iteration
  1. Given the current study transition and withdrawal probabilities, compute the fractions of a generator i that withdraws, waits, and, if the final study arrives, completes the interconnection.
  2. Update the queue status y_t.
  3. Update the study transition and withdrawal probabilities.
  4. Iterate until convergence.

The procedure embeds three assumptions. First, by explicitly starting the iteration with the observed outcomes, we select a particular equilibrium that is naturally motivated by data. Second, we assume perfect foresight for the evolution of the queue status. This is not an overly restrictive assumption, as in our context, the long term increase in queue size is not surprising given the rapid decrease in the cost of renewable generators. At the same time, generators do not have perfect foresight over short term events such as the study arrival or future values of unobserved shocks. Third, we take the arrival time of potential entrants and their characteristics x as given.

To validate our model, we first compare the time series of aggregate investment it predicts to the data. Figure 6 shows the cumulative capacity that was completed, by queue year, for generators in our sample (queued between 2011 and 2020). The left panel shows the cumulative capacity for all generators, and the right panel shows the cumulative capacity for renewables. The solid black lines plot the total capacity of generators in the data that are completed, i.e., they have completed construction and started operation. For example, the solid black line in 2018 represents the total completed capacity, as of 2022, of generators that entered the queue by 2018. As we near the end of our sample, there are many generators that are waiting for their next studies or have received their final interconnection study but have not been completed or withdrawn from the queue. The dotted black lines represent the capacity of these generators (which include generators still under construction) plus the capacity of generators that are completed from data.

Our model is able to match these time series. The solid blue line plots the model's prediction for the total capacity *ever* completed, by queue year. For the early years, this should match the capacity actually completed as of 2022 (the solid black line); all generators queued in these early years have had time to finish construction and begin operation. The completed capacity for generators entering the queue in 2014 and 2018 is slightly under-predicted by the point estimates of the model but covered within the reasonably narrow confidence intervals. For later years, the blue line lies between the solid and dotted black lines as expected.[23]

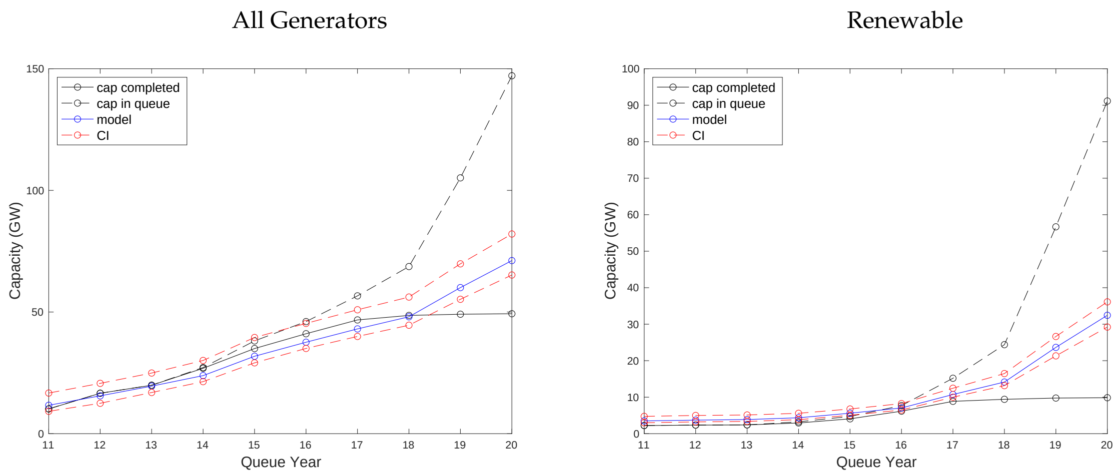
7 Counterfactual Simulations

We use our estimated model to conduct a series of simulations. In the first set of counterfactuals, we simulate the added capacity if the necessary studies are delivered faster. This exercise can be seen as quantifying the returns to increasing PJM's processing capacity.

Alternatively, we may hold PJM's processing capacity fixed and ask the following questions: are there too many generators in the queue? If so, which ones should be removed

[23]In Appendix Figure F.4, we show the model fit for the growing queue sizes.

Figure 6: Cumulative Completed Capacity as of 2022, Model and Data



The solid black lines (cap completed) represent the total cumulative capacity of generators in our sample that have started operation by 2022. For example, the capacity in 2018 represents the total completed capacity of generators that entered the queue in 2018 or earlier. The dotted black lines (cap in queue) represent this capacity plus the capacity of generators that have not withdrawn as of 2022. The blue line is the simulated cumulative capacity that will ever be begin operation, by queue year. The red line represents the confidence intervals for the model prediction.

to maximize the completed capacity? This simulation considers alternative queuing mechanisms. Specifically, we can reduce congestion by removing generators that are less likely to complete or add less capacity. The reduced congestion in turn expedites the studies for the remaining generators, reducing their chance of withdrawal and potentially increasing total completed capacity.

We next consider several proposed policy reforms. Given we find a substantial congestion externality, we focus on reforms that increase the cost of entering or staying in the queue. We also simulate the effects of grouping large clusters of generators together and charging each the average per-megawatt interconnection cost for the cluster.

7.1 Reducing Study Delays

As discussed in Section 2.2, reducing study delays may have a first-order impact on increasing completed capacity. A structural reason for the long waiting time is grid operators' limited capacity to process generator requests. Grid operators told us they are limited by engineering staff capacity and that they had difficulty hiring electrical engineers, whose supply is inelastic in the short run due to the training involved. In principle, grid operators could offer higher wages, hire more engineers, and speed up the process. We simulate five scenarios in which the study arrival probabilities of the second and third studies are increased by 5% to 25%.

We find large increases in completed capacity from reducing waiting times. Figure 7 reports the simulation results. The $x-$axis shows the average waiting time from the first to the last study conditional on not withdrawing from the queue before the last study, and the $y-$axis is the added capacity. A 10% increase in the study arrival probabilities reduces waiting time by about 2%,[24] and adds 4.0 GW of total capacity, of which 1.6 GW is renewable.[25] These additions correspond to increases of 5.6% and 4.6%, respectively. The added capacity increases quickly in the study arrival probability for all generators; for example, a 25% increase adds nearly 9 GW of new capacity.

We interpret these large gains as being due to reductions in both the cost and uncertainty from waiting in the queue. Generators incur costs like maintaining the option on the land, but also face many risks, such as the expiration of signed purchase agreements (Penrod, 2022). Our model captures the first with the waiting costs and the second with the idiosyncratic unobservable ε_t. We estimate the standard deviation of ε_t to be larger than the difference in waiting costs due to generator size, fuel type, or time.

7.2 Alternative Queuing Mechanisms

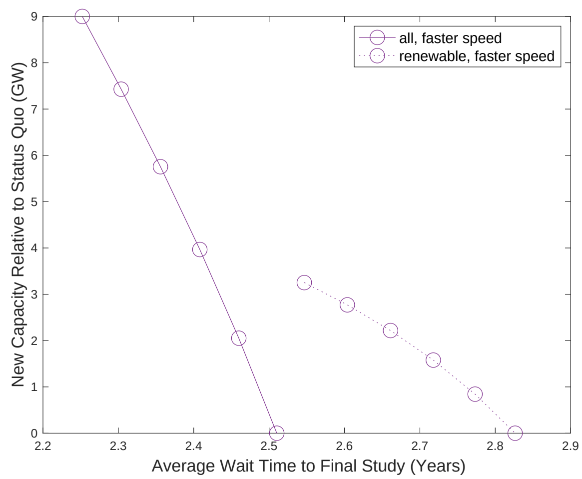
We next take the processing ability of PJM as given and consider alternative queuing mechanisms. We search for, for each potential entrant, a weight between 0 and 1 to maximize the total completed capacity. We treat each generator as a continuous mass, and the weight represents the share that the grid operator includes in the queue.[26] These weights modify the queue status. For example, the queue position of a generator is given by the sum of the weights of higher-queued generators, and we can move up the queue position of a lower-queued gener-

[24] A 10% increase in the study arrival does not result in a 10% decrease in the waiting time for two reasons: (1) we plot the conditional average waiting time given that the generators do not withdraw, and generators that reach their last studies receive studies faster, which limits the scope for faster study delivery to reduce their waiting time, and (2) faster study delivery increases the value of waiting, so more generators wait in equilibrium, increasing congestion.

[25] A developer could save the time of going through the queue again by entering at the same location with multiple identical requests across time. If a prior generator fails to complete the interconnection even when the interconnection cost is low, the generator can immediately use the queue position of a subsequent generator at the same location and expect similar interconnection study results. We could over-predict the completed capacity if this practice is widespread. We do not find this practice to be of first-order importance: locations that saw consecutive entry of generators that have the same fuel sources and similar capacities contribute to 0.12 GW of added total capacity and 0.07 GW of added renewable capacity.

[26] An alternative is to solve the integer programming problem of which generators to keep in the queue, but this optimization problem would be much harder.

Figure 7: Faster Study Delivery



Y-axis is new capacity added relative to the status quo. X-axis is the average wait time from first study issue date to the receipt of the final study. Each point on the line corresponds to either the status quo or an increase in the arrival probability of the second and third studies. From right to left, the points on each line are the status quo, 5%, 10%, . . ., 25%. The solid line shows the change in total generation capacity, while the dotted line shows the change in renewable generation capacity.

ator by reducing these weights. We consider the following maximization problem:

$$\max_{\boldsymbol{w}} \sum_{t,c,\tau,x} \rho_t(c,\tau,x) \cdot (1 - \Psi_t(c,\tau,x)) \cdot \text{cap}_x, \tag{17}$$

where $\boldsymbol{w}$ is the vector of weights, $\rho_t(c,\tau,x)$ is the number of generators with characteristics x and final study cost estimate c in period t after having waited τ periods, and cap_x is the capacity of these generators. This quantity ρ is based on the equilibrium queue status and generator withdrawal probabilities, both of which depend on the weights $\boldsymbol{w}$.[27]

A planner may value some generators more than others. For example, a planner may favor non-renewable generators to improve grid reliability, or renewables for their environmental benefits.[28] We also consider algorithms that maximize the following, more flexible objective

[27] In this exercise, we view the H functions as structural parameters reflecting the processing constraints when PJM conducts studies. For example, a grid operator that faces multiple waiting generators may have to study them sequentially, finishing one generator before moving on to the next. Alternatively, it may be able to parallelize and simultaneously work on multiple generators at the same time. We estimate the study process through the H functions. If the study process is strictly sequential, and a lower-queued generator has high capacity but might leave the queue after prolonged waiting, then moving up this generator's rank and studying it first may increase the total completed capacity.

[28] It is not obvious the planner would put a higher weight on renewables given these generators are already subsidized at the state and federal level. A grid operator we spoke with was concerned that too little dispatchable generation capacity was being completed to maintain reliability in the face of policy-induced coal generator retirements.

function:

$$\max_{w} \sum_{t,c,\tau,x} \rho_t \left(c,\tau,x\right) \cdot \left(1 - \Psi_t \left(c,\tau,x\right)\right) \cdot \text{cap}_x \cdot \theta_x, \tag{18}$$

where θ_x reflects the planner's preference for type x generators. We consider three sets of planner preference parameters, where θs on renewables and other generators are $(0.5, 0.5)$, $(0.75, 0.25)$ and $(0.9, 0.1)$. Maximizing the objective in (18) under the preference parameters of $(0.5, 0.5)$ amounts to maximizing (17). The maximization is also subject to optimality conditions 8 through 11 and a set of modified balance conditions in Appendix D.3.

In our implementation, a generator has the same priority as other generators with the same fuel type, in the same entry year range $\{< 2015, [2015, 2018), \geq 2018\}$ and the same size range $\{< 20\text{MW}, [20\text{MW}, 50\text{MW}),\ [50\text{MW}, 100\text{MW}), \geq 100\text{MW}\}$. We use this discretization to simplify the optimization problem, but we could also allow the weights to differ across finer categories of generator characteristics or compute a dynamic queuing rule with time-varying weights.

We find that these alternative queuing mechanisms meaningfully increase completed capacity. In Table 4, when a planner equally values renewable and non-renewable generation capacity (Column (1)), we find an increase of 7.94 GW in total capacity, 1.48 GW of which is renewable. When the planner prefers renewables (Columns (2) and (3)), the completed renewable capacity increases but the total added capacity is considerably smaller.

Table 4: Added Capacity in Alternative Queuing Mechanisms (GW)

	Planner Preference Parameter $(\theta_{\text{renewable}}, \theta_{\text{non-renewable}})$		
	$(0.5, 0.5)$	$(0.75, 0.5)$	$(0.9, 0.1)$
Added Capacity	7.94	4.09	0.08
Renewable	1.48	2.66	3.02
Non-Renewable	6.46	1.43	-2.94

Table reports new capacity added relative to the status quo for generators in our sample. The model-predicted total completed capacity in the status quo is 71.15 GW, 32.44GW of which is renewable.

The queuing mechanisms remove many small generators from the queue. Table 5 shows the percentage of removed generators within each fuel-size group. With equal preferences for the renewable and non-renewable generators, many generators below 100 MW are screened out. As the planner prefers renewable capacity more, the mechanism removes fewer renewable generators with a capacity between 20 and 100 MW but still removes more than 45% of renewable generators under 20MW. More medium sized non-renewable generators are also removed. Under all preference parameters, fewer but larger generators complete interconnection.

Table 5: Percentage of Generators Removed by Alternative Queuing Mechanisms

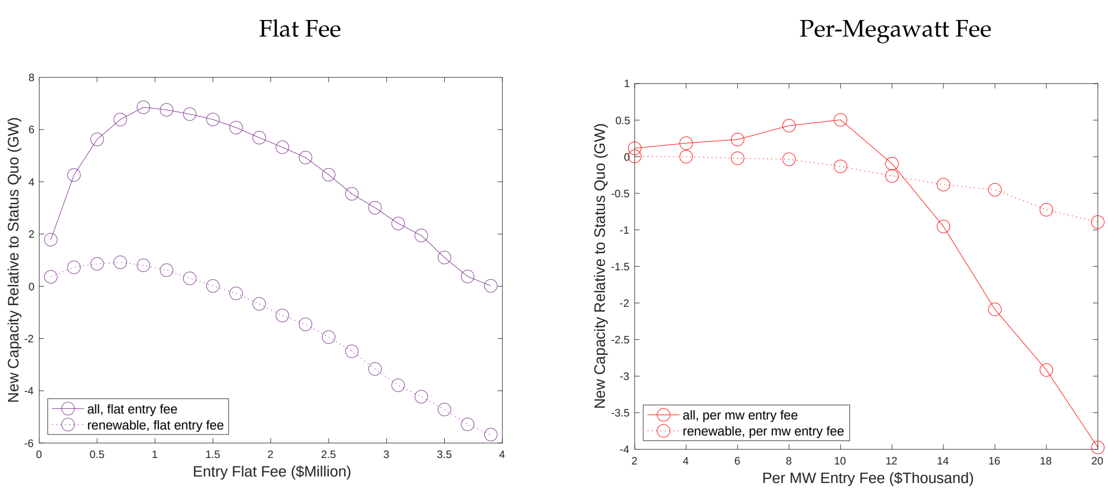
	Planner Preference Parameter $(\theta_{\text{renewable}}, \theta_{\text{non-renewable}})$		
	$(0.5, 0.5)$	$(0.75, 0.5)$	$(0.9, 0.1)$
Renewable Capacity (MW)			
<20	72.31	45.02	67.06
20-100	51.76	25.73	23.81
>100	1.19	0.00	0.00
Non-Renewable Capacity (MW)			
<20	55.97	56.24	56.24
20-100	57.12	77.17	73.08
>100	0.00	8.52	28.05

Table reports the share of generators in a group that are removed by each queuing mechanism. For example, 72.31% of generators whose sizes are below 20MW are removed when the planner preference parameters for the renewable and non-renewable generators are 0.5 and 0.5.

7.3 Entry Fees

We consider two types of entry fees, defined as the cost to enter the queue (obtain the first study). The first type is a flat per generator fee. The second type is assessed per megawatt.

Figure 8: Added Capacity with Entry Fee



Y-axis is new capacity added relative to the status quo. X-axis is the fee to enter the queue and receive the first study. The solid line shows the change in total generation capacity, the dotted line shows the change in renewable generation capacity.

For both types of entry fee, we find an inverted-U relationship between the added capacity and fee levels (Figure 8). The flat fee that maximizes completed capacity is $900,000. It would add over 6.85 GW of capacity, but only 0.8 GW of renewable capacity. It also decreases the number of completed generators by 40%.[29] The per-megawatt fee achieves a much smaller effect, adding at most 0.5 GW of capacity. None of the generators added by the per-megawatt fee is a renewable generator.

Table 6 breaks down which potential generators decide not to enter the queue at the flat fee level that maximizes completed capacity. Compared with the alternative queuing mechanisms above, the fee screens out many more small generators and more renewable generators compared to non-renewable generators resulting in a lower completed capacity.

Table 6: Generators that No Longer Enter the Queue with a Flat Fee=$900,000

% of Group	Entry Year			
	<2013	2013-2015	2016-2018	>2018
Renewable Capacity (MW)				
<20	92.11	98.18	94.77	95.98
20-100	69.27	70.59	30.63	40.22
>100	1.92	0.00	0.00	0.00
Non-Renewable Capacity (MW)				
<20	93.70	85.06	91.73	87.97
20-100	29.07	22.92	44.68	12.82
>100	0.90	0.00	0.00	0.00

Table reports the share of generators in a group that no longer enter the queue when there is a $900,000 per generator entry fee. For example, 92.11% of generators that entered the queue from 2008-2012 and whose sizes are below 20MW no longer enter the queue.

We also simulate an increase in the fees for later studies. We again consider both flat and per megawatt fees, and assume the same increase in fee is levied for both the second and third study. Screening at this later stage is less effective: completed capacity falls for all fee levels. Appendix E.1 provides more details.

7.4 Sharing Interconnection Costs Equally Across Large Clusters of Generators

As discussed in Section 2.4, PJM plans to reform the process of assessing interconnection costs to reduce cost uncertainty. Specifically, PJM will designate several large clusters and study all generators in each cluster at once and assign the total network upgrade costs of the cluster to each generator based on capacity.[30]

Our next counterfactual approximates this cluster-level cost allocation. We first assume

[29] In PJM's 2022 reforms, the collected entry fees will be refunded to generators that complete interconnection. We do not account for these refunds in our simulations.

[30] Cluster study may also reduce the time PJM needs to process each study; we abstract from this effect.

each transmission owner territory and entry year combination is a cluster.[31] Rather than endogenously determine the total costs to be shared after each study, we consider a policy of pre-committing to a set per-megawatt interconnection cost. We do so for computational tractability. Therefore, relative to the proposed reform, the policy we simulate leads to a greater reduction in uncertainty.

We consider three assumptions on the total cluster-level cost in each cluster: (1) the total interconnection costs of completed generators; (2) the total final study interconnection costs of generators; (3) the total study 2 interconnection costs of generators. For each, we arrive at the per-megawatt cost by dividing the total costs observed in the data by the total size of the corresponding group of generators, e.g., total study 2 costs by the total size of generators in that cluster that reached study 2. These per-megawatt costs are fixed in the simulation. The "true" interconnection costs, as predicted by the H functions, will likely differ from these pre-committed costs. If the revenue is less than "true" interconnection costs, a subsidy is required to sustain the cost commitment.

Table 7 shows that this policy can increase completed capacity, but requires a subsidy to do so. For example, pre-committing to the average cost for generators that reached study 2 (row 3) adds 2.3 GW of total capacity, at a subsidy cost of 0.71 billion dollars. We also find that committing to lower costs has decreasing returns: the added GW per billion dollars of subsidy falls as we decrease the pre-committed cost. While reducing uncertainty may reduce the incentive to enter the queue, we find that entry falls by less than one percent under each of the three cost assumptions. In Appendix E.2, we study the effects of directly subsidizing interconnection costs. The results suggest that this type of cluster-level cost allocation without subsidy would reduce completed capacity.

The increases in capacity are mainly from the renewables. In fact, for the costs in rows (2) and (3), renewable capacity increases while other capacity decreases. Because renewable generators have higher interconnection costs, non-renewable generators subsidize renewable generators under a policy of sharing costs equally.

We note that entry may switch from locations where expected costs rise to locations where expected costs fall. We partially account for this behavior because an entrant whose expected value from entering the queue is lower than the entry cost will not enter. Yet, we do not account for increased entry in locations where expected costs fall, a behavior that would increase the amount of subsidy required.

[31] As of November 2023, PJM had not released the geographic regions it will use to form clusters.

Table 7: Cluster Study

	Added Capacity (GW)		Subsidy ($Billion)
	Total	Renewable	
Cluster-level Cost Assumption			
(1) Completed Generators	7.74	5.23	4.48
(2) Final Study Costs	3.13	3.65	1.41
(3) Study 2 Costs	2.28	2.72	0.71

Clusters are at the transmission operator territory by year level. Row (1) Completed Generators assumes the pre-committed cost per MW for the cluster is the average per MW cost for completed generators in that transmission owner territory and year in the data. Row (2) Final Study Costs the same but for the final study costs in the data. Row (3) Study 2 Costs the same but for the second study costs in the data.

7.5 Climate Impacts

We found that several policy reforms would significantly increase renewable generation capacity, an increase that would offset electricity production from fossil fuel generators. To quantify the approximate impact of this investment on carbon emissions, we use the U.S. Environmental Protection Agency's Avoided Emissions and Generation Tool (AVERT). This tool translates changes in generation capacity into changes in CO_2 emissions using data on regional demand, production costs,and market dispatch (US EPA, 2022). We value these emissions using the social cost of carbon of $185 per ton of CO_2 in Rennert et al. (2022).

Additional renewable generation capacity would produce large social benefits from avoided carbon emissions. For example, increasing the study arrival probability by 10% results in 1.58 GW more renewable capacity. The AVERT tool calculates that adding 1.58 GW of utility-scale renewable generation capacity to the Mid-Atlantic region in 2022 would decrease annual CO_2 emissions by 2.14 million metric tons.[32] The implied annual social value of this reduction is 394 million dollars. The more conservative social cost of carbon used by the U.S. government ($51 per ton) implies a social value is 109 million dollars. Our other simulations find gains in renewable capacity in the range of 1-3.5 GW. AVERT shows that CO_2 emissions are offset roughly at the rate of 1.4 million metric tons per gigawatt in this range.

8 Conclusion

We use novel data from the largest transmission operator in the United States to study the interconnection queue. We find there is a congestion externality. If a generator has more

[32]We assume the composition of this 1.58 GW increase is one third wind and two thirds solar, matching the roughly 2 to 1 ratio of solar to wind capacity in our sample.

generators ahead of it in the queue, the probability it receives the third and final study in a given quarter falls. We also find that interconnection costs are hard to predict and are a key factor in generators' decisions to withdraw from the queue.

We next study policy reforms using a dynamic model that accounts for equilibrium effects. We find that reducing study delays can substantially increase investment. Given the congestion effects, we next ask whether removing some generators can increase the amount of generation capacity completed. Alternative queuing policies prioritize larger generators and increase completed capacity. A flat fee to enter the queue has a similar effect, though this increase disproportionately comes from non-renewable generators. If paired with a subsidy, a cluster-study approach that allocates the total interconnection cost for a large group of generators equally on a per-megawatt basis better targets renewable generators. These increases in renewable capacity have substantial environmental benefits.

References

Abito, Jose Miguel, Christopher R Knittel, Konstantinos Metaxoglou, and André Trindade. 2022. "The Role of Output Reallocation and Investment in Coordinating Externality Markets." *International Journal of Industrial Organization*, 83(102843).

Agarwal, Nikhil, Itai Ashlagi, Michael A Rees, Paulo Somaini, and Daniel Waldinger. 2021. "Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys." *Econometrica*, 89(1): 37–76.

Aguirregabiria, Victor, and Pedro Mira. 2007. "Sequential estimation of dynamic discrete games." *Econometrica*, 75(1): 1–53.

Alagappan, Lakshmi, Ren Orans, and Chi-Keung Woo. 2011. "What drives renewable energy development?" *Energy policy*, 39(9): 5099–5104.

Aldy, Joe, Todd Gerarden, and Richard Sweeney. 2023. "Investment versus Output Subsidies: Implications of Alternative Incentives for Wind Energy." 10(4).

Allcott, Hunt. 2012. "The smart grid, entry, and imperfect competition in electricity markets." National Bureau of Economic Research.

AWEA. 2019. "Wind Powers America Annual Report 2019." `https://www.awea.org/resources/publications-and-reports/market-reports`.

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin. 2007. "Estimating dynamic models of imperfect competition." *Econometrica*, 75(5): 1331–1370.

Blundell, Wesley, Gautam Gowrisankaran, and Ashley Langer. 2020. "Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations." *American Economic Review*, 110(8): 2558–2585.

Buchsbaum, Jesse F., Catherine Hausman, Johanna L. Mathieu, and Jing Peng. 2022. "Spillovers from Ancillary Services to Wholesale Energy Markets." NBER Working Paper No. 28027.

Bushnell, James B, Erin T Mansur, and Celeste Saravia. 2008. "Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets." *American Economic Review*, 98(1): 237–266.

Butters, R. Andrew, Jackson Dorsey, and Gautam Gowrisankaran. 2021. "Soaking up the sun: Battery investment, renewable energy, and market equilibrium." NBER Working Paper 29133.

Caspary, Jay, Michael Goggin, Rob Gramlich, and Jesse Schneider. 2021. "Disconnected: The Need for a New Generator Interconnection Policy." Americans for a Clean Energy Grid Report.

CBO. 2022. "Emissions of Carbon Dioxide in the Electric Power Sector." urlhttps://www.cbo.gov/system/files/2022-12/58419-co2-emissions-elec-power.pdf.

Charles, Dan, Nick Fountain, Willa Rubin, and Sally Helm. 2023. "Green energy gridlock." *Planet Money*, [Online; accessed on 2023-10-16].

Chen, Yanyou, and Daniel Yi Xu. 2023. "A structural empirical model of r&d, firm heterogeneity, and industry evolution." *The Journal of Industrial Economics*, 71(2): 323–353.

Che, Yeon-Koo, and Olivier Tercieux. 2021. "Optimal queue design." 312–313.

Collier, Rob. 2021. "Interconnection Slowdown: What's Causing It and What You Can Do." textitLevelTen Energy Blog. url=https://www.leveltenenergy.com/post/interconnection-slowdown.

Connell, Jason, and Susan McGill. 2020. "Interconnection Process Overview." `https://www.pjm.com/-/media/committees-groups/task-forces/iprtf/postings/interconnection-process-overview.ashx`.

Covert, Thomas R., and Richard L. Sweeney. 2022. "Winds of Change: Estimating Learning by Doing without Cost or Input Data." Working Paper.

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

Davis, Lucas W., Catherine Hausman, and Nancy L. Rose. 2023. "Transmission Impossible? Prospects for Decarbonizing the US Grid." NBER Working Paper 31377.

Davis, Rebecca J., J. Scott Holladay, and Charles Sims. 2022. "Coal-Fired Power Plant Retirements in the United States." In *Environmental and Energy Policy and the Economy, Vol. 3.* , ed. Matthew J. Kotchen, Tatyana Deryugina and James H. Stock. Chicago:University of Chicago Press.

Deschenes, Oliver, Christopher Malloy, and Gavin McDonald. 2023. "Causal effects of Renewable Portfolio Standards on renewable investments and generation: The role of heterogeneity and dynamics." *Resource and Energy Economics*, 75(101393).

Doshi, Gaurav. 2022. "Wiring America: The Short- and Long-Run Effects of Electricity Grid Expansion." Working Paper.

Driscoll, William. 2022. "Interconnection delays and costs are the biggest barrier for utility-scale renewables, say developers." *PV Magazine*. `https://pv-magazine-usa.com/2022/02/14/interconnection-delays-and-costs-are-the-biggest-barrier-for-utility-scale-renewables-say-develop`

Elliott, Jonathan T. 2021. "Investment, Emissions, and Reliability in Electricity Markets."

Fell, Harrison, Daniel Kaffine, and Kevin Novan. 2021. *American Economic Journal: Economic Policy*, 13(2): 241–272.

FERC Staff. 2023. "Staff Presentation: Improvements to Generator Interconnection Procedures and Agreements."

Fowlie, Meredith, Mar Reguant, and Stephen P Ryan. 2016. "Market-based emissions regulation and industry dynamics." *Journal of Political Economy*, 124(1): 249–302.

Gergen, Michael J, George D Cannon Jr, and Shannon D Torgerson. 2008. "A modest proposal: A market-based approach to generation interconnection process reform." *The Electricity Journal*, 21(9): 8–18.

Gonzales, Luis E., Koichiro Ito, and Mar Reguant. 2022. "The Dynamic Impact of Market Integration: Evidence from Renewable Energy Expansion in Chile." Working Paper.

Gowrisankaran, Gautam, Ashley Langer, and Wendan Zhang. 2022. "Policy Uncertainty in the Market for Coal Electricity: The Case of Air Toxics Standards." NBER Working Paper 30297.

Gowrisankaran, Gautam, Stanley S Reynolds, and Mario Samano. 2016. "Intermittency and the value of renewable energy." *Journal of Political Economy*, 124(4): 1187–1234.

Hale, Zach. 2021. "Solar developer adds to pile of complaints targeting PJM study delays."

Hale, Zach, and Molly Christian. 2023. "FERC approves 'historic' rule to clear backlog of US generation projects."

Hitaj, Claudia. 2013. "Wind power development in the United States." *Journal of Environmental Economics and Management*, 65: 394–410.

Howland, Ethan. 2022. "FERC approves PJM's 'first- ready, first-served' interconnection reform plan, steps to clear backlog."

IEA. 2022*a*. "Electricity Market Report - January 2022." `https://www.iea.org/reports/electricity-market-report-january-2022`.

IEA. 2022*b*. "Global Energy Review: CO2 Emissions in 2021." `https://www.iea.org/reports/global-energy-review-co2-emissions-in-2021-2`.

Igami, Mitsuru. 2017. "Estimating the innovators dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998." *Journal of Political Economy*, 125(3): 798–847.

Johnston, Sarah. 2019. "Nonrefundable Tax Credits versus Grants: The Impact of Subsidy Form on the Effectiveness of Subsidies for Renewable Energy." *Journal of the Association of Environmental and Resource Economists*, 6(3): 433–460.

LaRiviere, Jacob, and Xueying Lyu. 2022. "Transmission constraints, intermittent renewables and welfare." *Journal of Environmental Economics and Management*, 112(102618).

Leisten, Matthew, and Nicholas Vreugdenhil. 2023. "Dynamic Regulation with Firm Linkages: Evidence from Texas." Working Paper.

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

Liu, Tracy, Zhixi Wan, and Chenyu Yang. 2021. "The efficiency of a dynamic decentralized two-sided matching market."

Manski, Charles F. 1988. "Identification of binary response models." *Journal of the American statistical Association*, 83(403): 729–738.

Mansur, Erin T. 2007. "Do oligopolists pollute less? Evidence from a restructured electricity market." *The Journal of Industrial Economics*, 55(4): 661–689.

Mansur, Erin T. 2008. "Measuring welfare in restructured electricity markets." *The Review of Economics and Statistics*, 90(2): 369–386.

Metcalf, Gilbert. 2010. "Investment in energy infrastructure and the tax code." In *Tax policy and the economy, Vol. 24.* , ed. Jeffrey Brown. Chicago:University of Chicago Press.

Mooney, Attracta. 2023. "Gridlock: how a lack of power lines will delay the age of renewables." *Financial Times*.

Parker, Nick, Sarah Johnston, Bryan Leonard, and Justin Winikoff. 2023. "Renewable Energy on American Indian Land." Working Paper.

Penrod, Emma. 2022. "Why the energy transition broke the U.S. interconnection system."

PJM. 2010. "OPEN ACCESS TRANSMISSION TARIFF."

PJM. 2021*a*. "Connecting to the Grid FAQs: How long does the interconnection process take?" https://learn.pjm.com/three-priorities/planning-for-the-future/connecting-grid/how-long-does-the-interconnection-process-take.

PJM. 2021*b*. "Forward Together: 2020 PJM Annual Report." https://services.pjm.com/annualreport2020/.

PJM. 2021*c*. "PJM Manual 14A: New Services Request Process." 29 ed., https://www.pjm.com/-/media/documents/manuals/m14a.ashx.

PJM. 2021*d*. "PJM Manual 14B: PJM Region Transmission Planning Process." 51 ed.

PJM. 2021*e*. "PJM Manual 14C: Generation and Transmission Interconnection Facility Construction." 14 ed.

PJM. 2021*f*. "PJM Manual 14G:Generation Interconnection Requests." 7 ed., https://www.pjm.com/-/media/documents/manuals/m14g.ashx.

PJM. 2022. "PJM Interconnection, L.L.C., Docket No. ER22- -000." https://pjm.com/directory/etariff/FercDockets/6726/20220614-er22-2110-000.pdf.

PJM. 2023. "Project Status & Cost Allocation." (Accessed Jan 31, 2023).

Plumer, Brad. 2023. "The U.S. Has Billions for Wind and Solar Projects. Good Luck Plugging Them In." *The New York Times*. https://www.nytimes.com/2023/02/23/climate/renewable-energy-us-electrical-grid.html.

Potter, Ellie. 2021. "FERC's Christie says agency should examine interconnection queue 'chaos'."

Rand, Joseph, Mark Bolinger, Ryan H. Wiser, Seongeun Jeong, and Bentham Paulos. 2021. "Queued Up: Characteristics of Power Plants Seeking Transmission Interconnection As of the End of 2020." https://emp.lbl.gov/publications/queued-characteristics-power-plants.

RechargeNews. 2021. "Texas solar outlook clouded by grid, storm regulation." *Reuters Events*. https://www.reutersevents.com/renewables/solar-pv/texas-solar-outlook-clouded-grid-storm-regulation.

Rennert, Kevin, Frank Errickson, Brian C. Prest, Lisa Rennels, Richard G. Newell, William Pizer, Cora Kingdon, Jordan Wingenroth, Roger Cooke, Bryan Parthum, David Smith, Kevin Cromar, Delavane Diaz, Frances C. Moore, Ulrich K. MÃŒller, Richard J. Plevin, Adrian E. Raftery, Hana SevcÃkova, Hannah Sheets, James H. Stock, Tammy Tan, Mark Watson, Tony E. Wong, and David Anthoff. 2022. "Comprehensive evidence implies a higher social cost of CO2." *Nature*, 610: 687–692.

Ryan, Nicholas. 2021. "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market." *American Economic Journal: Microeconomics*, 13(2): 202–42.

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

Shoemaker, Jason. 2021. "Interconnection Queue Status Update." `https://www.pjm.com/-/media/committees-groups/committees/pc/2021/20210831/20210831-item-07-queue-status-update.ashx`.

US EPA. 2022. "Avoided Emissions and Generation Tool (AVERT)." `https://www.epa.gov/avert/download-avert`.

Verdier, Valentin, and Carson Reeling. 2021. "Welfare effects of dynamic matching: An empirical analysis." *The Review of Economic Studies, forthcoming*.

Waldinger, Daniel. 2021. "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation." *American Economic Review*, 111(8): 2660–96.

Weintraub, Gabriel Y, C Lanier Benkard, Przemysław Jeziorski, and Benjamin Van Roy. 2010. "Nonstationary oblivious equilibrium." Working Paper.

Wolak, Frank A. 2015. "Measuring the competitiveness benefits of a transmission investment policy: The case of the Alberta electricity market." *Energy Policy*, 85: 426–444.

Yang, Chenyu. 2020. "Vertical structure and innovation: A study of the SoC and smartphone industries." *The RAND Journal of Economics*, 51(3): 739–785.

A Allocation of Shared Network Upgrades

There are two main types of interconnection costs: direct connection costs and network upgrades. While we collect costs separately for each of the major categories listed in the study, it is not always easy to distinguish between direct costs and network upgrades, and we combine both types into one cost measure for our analysis.

Network upgrades are often shared among multiple generators. A generator's contribution to these shared costs is typically first reported in the second (system impact) study. The first (feasibility) study will note that the generator may be responsible for a contribution toward these costs and will often report the total. Figure A.1 shows a typical example. The generator will likely be responsible for a portion of the 17.75 million, and an estimate of its exact responsibility will be provided in the next study. For this example, the first study cost we use in estimation is 7.85 million. This generator will also have the variable *s1 cluster* set to one because the first study indicates the generator may be responsible for additional costs and lists a non-zero total for these costs.

Figure A.1: Example cost breakdown for the first study

The AD1-034 project will be responsible for the following costs:

Description	Total Cost
Attachment Facilities	$ 1,550,000
Direct Connection Network Upgrades	$ 5,500,000
Non Direct Connection Network Upgrades	$ 800,000
Total Costs	**$ 7,850,000**

In addition, the AD1-034 project may be responsible for a contribution to the following costs:

Description	Total Cost
New System Upgrades	$ 0
Previously Identified Upgrades	$ 17,500,000
Total Costs	**$ 17,750,000**

B Prior Generator Cost and Entry

In our analysis, we do not consider how interconnection studies provide information that affects the entry decisions of future generators. Specifically, high costs in a location may cause entrants to substitute to other locations. Alternatively, suppose a recent upgrade to the infrastructure made by one generator lowers the cost for the next generator. Then, locations where a generator with a high interconnection cost may be about to complete interconnection would be desirable. PJM's cost-sharing makes this second channel less likely.[33]

To assess the empirical relevance of this concern, we first look at the region-level time-series of interconnection costs and entry. Figure B.1 shows the time series of mean study 2 costs for the nineteen transmission owner territories in PJM. The time period is the half-year. We focus on the study 2 cost because it is an important source of information: this is the first study to include a generator's contribution to shared network upgrade costs, and it is more widely available than study 3. High costs do not appear to result in either higher or lower entry in subsequent periods.

We next estimate the effect of prior study costs on entry at the same point of interconnection. The estimation sample consists of all substation-quarters where the second study cost of a higher-queued generator is available, i.e., locations where a prior entry has made it to at least the second study. The outcome variable is an indicator for whether entry occurs in quarter t at this location.[34] We regress this variable on whether the most recent study 2 cost at that substation exceeds \$0.1 million/MW. We also include three measures of grid-operator transmission investment (RTEP) affecting the substation,[35] the voltage of the transmission line at the substation, and the locational marginal price (LMP). To construct our measure of the LMP, we use a quarterly average of hourly prices for peak hours (7am-11pm) for three days each month (10$^{\text{th}}$, 20$^{\text{th}}$, and 30$^{\text{th}}$, if available, else the 28$^{\text{th}}$). We condition on state-year-quarter fixed effects to control for changes in state-level incentives and time trends.

We estimate that prior costs have a small effect on entry relative to other factors. These estimates are reported in Table B.1. The effect of the prior generator's study 2 cost is not precisely estimated, with a 95% confidence interval from -0.006 to 0.0179, which are -8% to 25% of the mean entry rate. Baseline RTEP investment has a larger effect: cumulative investment greater than \$100,000 is associated with 30% increase in the entry rate. Supplemental RTEP investment and cumulative investment since the most recently completed generator have small effects. Generators also tend to enter at higher voltage substations. Finally, a one standard deviation (\$12/MWh) increase in the LMP is associated with a 160% increase in the mean

[33]Empirically, we find little evidence a costly completed interconnection reduces the cost for the subsequent generator once we control for whether the subsequent generator is an uprate.

[34]If a generator requests to interconnect by tapping into a line instead of at a substation, we assign the request to the nearest substation.

[35]We use data from PJM on the planned and completed transmission investment. These data list which substations will be affected by each investment.

entry rate.

Table B.1: Effect of Interconnection Costs on Entry at the Same Substation

Most Recent Study 2 Cost>$0.1 million/MW	0.0059
	(0.0061)
Cumulative RTEP Baseline Investment>$0.1 million	0.0228
	(0.0081)
Cumulative RTEP Supplemental Investment>$0.1 million	0.0011
	(0.0080)
Cumulative RTEP Since Most Recently Completed Generator>$0.1 million	-0.0014
	(0.0087)
Voltage>230kV	0.0234
	(0.0065)
LMP ($/MWh)	0.0097
	(0.0004)
Mean of dependent var.	0.0729
N	17,428

Observations are at the substation by quarter level for 2008-2020. Sample is all substation-quarters where a prior second study interconnection cost is available (N = 17,428; 1981 unique substations). Dependent variable is an indicator for entry into the queue at that substation in that quarter. Cumulative RTEP Baseline Investment>$0.1 million is an indicator for if the cumulative amount of baseline transmission investment completed since 2008 and affecting that substation is greater than $100,000. Cumulative RTEP Supplemental Investment>$0.1 million is an indicator for if the cumulative amount of supplemental transmission investment completed since 2008 and affecting that substation is greater than $100,000. Cumulative RTEP Investment>$0.1 million is an indicator for if the cumulative amount of either type of RTEP investment since the last connecting generator at the substation was completed is greater than $100,000. Voltage>230kV is an indicator for if the voltage of the substation is greater than 230 kV. LMP is a quarterly average of the nearest locational marginal price at peak hours in 2020 $/MWh. Regression includes state-year-quarter fixed effects. SE in parentheses, clustered by substation.

C Study Arrival Functions

C.1 Queue State y_t

For the S function that aggregates the current and past evolution of the queue to the queue state y_t, we include two sets of vectors. The first set describes the queue position. In period t, we compute each generator's queue position as the number of generators that entered the queue in the same or an earlier entry cohort and are still in the queue. We include the queue position and its interaction with whether generator i is waiting for the third study. We also allow for nonlinear and time-varying effects of the queue position by including its higher order terms and interactions with time fixed effects.

The second set of variables accounts for the effects of other local generators in the queue. As discussed in Section 2, the local transmission owner usually conducts the tests and issues the studies on behalf of PJM. A large local backlog may also affect the the local transmission owner's ability to conduct studies. Furthermore, multiple generators in the same location may share the costs of transmission upgrades. We therefore also include the number and the total capacity of local generators. We consider two definitions of being local to a generator. One is to be in the same transmission owner service territory, and the other is to be within a 100km radius of the generator.[36] For similar reasons, withdrawals of local generators may lead to re-study, which changes costs and delays the studies for remaining generators. We therefore also include the local (as defined above) withdrawals in the past two quarters. Thus, the queue status y_t is specific to a generator's location and entry time.

C.2 Transition Dynamics

The functions H_1, H_2, $\tilde{H}_2$ and H_3 describe the timing and information of a new study. These functions are directly identified and estimable from data (e.g., Aguirregabiria and Mira (2007); Bajari, Benkard and Levin (2007)). We focus on the specification of these functions in this section. The input of the functions includes the cost assessment c from the most recent study, the study information z from previous studies, the time since the generator entered the queue τ, the time since the most recent study was issued $\tilde{\tau}$, the current calendar time t, generator-specific characteristics x, and the current queue state y_t.

The outputs of the H functions differ. The functions H_2 and H_3 generate the probability that a new study that is the final study (whether it is the second or third study) will be issued in the current period, and that the new cost is c. The function $\tilde{H}_2$ generates the probability that a new study that is not the final study will be issued and that the updated cost and information are c and z. The function H_1 generates the cost and information for the first study.

Given the complex set of decisions in the queuing process, we simplify our model by considering two processes: (1) the arrival of a new study and the cost distribution, and (2)

[36] In engineering models, PJM measures distances in "electric distances" based on impedance, which requires detailed knowledge about the physical distance and capability of the connecting transmission lines. We find that a physical distance of 100km is a reasonable cutoff to account for the effect of nearby generators. We randomly sampled 100 generators that PJM identifies as contributing to the same violation with at least one other generator. Then, for each of 100 sampled generators, we compute the average distance to the corresponding cost-sharing generators. We find that 86% of the average distances are within 100km.

conditional on receiving the second study, PJM's decision to expedite the interconnection requests, i.e., require only two studies, and to update the information z. We then combine these processes to produce the H functions.

C.2.1 New Study Arrival and Cost

We start by specifying a flexible probability function for receiving a new study with a cost c. We discretize the cost and consider a joint probit (for study arrival) and ordered probit (for the cost level) model. We divide the interconnection costs (in million \$/MW) into $L = 4$ bins, with bin $\ell \in C_\ell = \{[0, 0.01], (0.01, 0.05], (0.05, 0.20], (0.20, \infty]\}$. We specify the latent variables governing study arrivals and costs as

$$u^{\star\text{arrive}} = \beta^{\text{arrive}} \cdot d_1(\ell, z, k, x, y_t),$$
$$u^{\star\text{cost}} = \beta^{\text{cost}} \cdot d_2(\ell, z, k, x, y_t),$$

where the $d_{(\cdot)}(\cdot)$ functions are vectors of flexible polynomials of the characteristics, and the $\beta^{(\cdot)}$s are vectors of parameters. We use cost bin ℓ in place of c (from the most recent study) to make it clear that we model the cost as a discrete variable. The probability that the new study arrives and the new cost estimate is in bin ℓ' is defined as

$$h_t\left(\ell'; \ell, z, \tau, \tilde{\tau}, k, x, y_t(x)\right) = \Pr\left(0 < u^{\star\text{arrive}} + \epsilon_t^{\text{arrive}}, \mu_{\ell'} \leq u^{\star\text{cost}} + \epsilon_t^{\text{cost}} \leq \mu_{\ell'+1}\right), \quad \text{(C.1)}$$

where $\mu_1 = -\infty, \mu_2 = 0$, and $\mu_2 \leq ... \leq \mu_L < \mu_{L+1} = \infty$ are a series of parameters. We allow the normally distributed errors $\epsilon_t^{\text{arrive}}$ and ϵ_t^{cost} to be correlated.

C.2.2 PJM Decision to Expedite Interconnection Requests and Update Information z

PJM determines whether to expedite a generator based on the size and generation type and whether certain types of violations are identified in the first study (PJM (2010)). Instead of fully modeling PJM's rules and engineering simulations for these decisions, we simplify the analysis by using a flexible probit model based on generator characteristics. We use $p_t^{\text{final}}(\ell, z, x, y_t)$ to denote the predicted probability the second study is the final study.

We apply the same approach to the other two PJM decisions, which we track as the cumulative information $z = \left(z^{\text{test}}, z^{\text{cost-sharing}}\right)$. For $z^{\text{test}} \in \{0, 1\}$, we focus on a set of three tests: generator deliverability, multiple facility contingency, and short circuit analysis. These tests are jointly conducted to determine how much the generator will overload the grid. PJM may choose to conduct these tests in the first or the second study depending on the generator's characteristics and the current local transmission network conditions. We also consider whether the generator shares costs with other generators, $z^{\text{cost-sharing}} \in \{0,1\}$. To determine this variable, PJM conducts additional tests (short circuit dynamic analysis and system protection analysis) to identify generators responsible for the same upgrade. Both variables can inform the arrival time of the next study and the cost. We denote the corresponding predicted probabilities as $p_t^{\text{test}}(\ell, z, x, y_t)$ and $p_t^{\text{cost-sharing}}(\ell, z, x, y_t)$.

C.2.3 The H Functions

We combine these processes to yield the following transition functions. We set

$$H_3 = h_t\left(\ell'; \ell, z, \tau, \tilde{\tau}, k = 2, x, y_t\right)$$

and

$$H_2 = h_t\left(\ell'; \ell, z, \tau, \tilde{\tau}, k = 1, x, y_t\right) \cdot p_t^{\text{final}}\left(\ell, z, x, y_t\right).$$

For the probability of receiving a non-final second study, we assume (omitting arguments of the p_t^{final}, p_t^{cluster} and p_t^{test} functions) that:

$$\begin{aligned}\tilde{H}_2\left(\ell', z^{\text{test}\prime}, z^{\text{cluster}\prime}; \ell, z, \tau, \tilde{\tau}, y_t, x\right) = &\left(z^{\text{cluster}\prime} p_t^{\text{cluster}\prime} + \left(1 - z^{\text{cluster}\prime}\right)\left(1 - p_t^{\text{cluster}\prime}\right)\right) \\ &\cdot \left(z^{\text{test}\prime} p_t^{\text{test}\prime} + \left(1 - z^{\text{test}\prime}\right)\left(1 - p_t^{\text{test}\prime}\right)\right) \cdot \left(1 - p_t^{\text{final}}\right) \\ &\cdot h_t\left(\ell'; \ell, z, \tau, \tilde{\tau}, k = 1, x, y_t\right).\end{aligned}$$

For H_1, we assume that the cost in the first study is determined by a flexible ordered probit function of x and y_t. We set z to be what we observe in the first studies in the data.[37] Although our specification restricts the unobservables to be independent, we do allow PJM decisions on issuing studies, costs, expediting generators, and conducting tests to be correlated through a large set of flexibility specified observables.

C.2.4 Identification and Estimation

The main identifying assumption for PJM decisions is that any generator characteristics that drive the withdrawal decision but are not in $(c, z, \tau, \tilde{\tau}, k, t, y_t, x)$ are independent of the unob-

[37] If either z is 1, then the indicator will continue to be 1 in subsequent studies. Our results are robust to setting $z = (0, 0)$ and assuming only study 2 updates this information.

servables in the probit and the ordered-probit models above underlying the process of PJM issuing studies. Interviews with industry experts suggest that the additional motives for withdrawing a generator that we do not control for, such as a local authority failing to approve a proposed loan program for solar generators, are largely unrelated to PJM's studies. It is hard for both PJM and generators to predict which generators will withdraw first.[38]

We next focus our discussion on identifying how congestion affects new study arrival. The key variation is the change in a generator's queue position. A generators initial queue position is determined by the number of generators in the queue at the time of entry. The position advances as a result of the completion and withdrawals of higher-queued generators. Given that these movements are driven by other generators' decisions that are generally hard to predict, we use both the cross-sectional variation in the queue position across generators as well as the changes in the queue position as a generator advances to identify how the queue position affects the study arrival.

We use maximum likelihood to estimate these transitions. Table C.1 presents parameter estimates of the covariates related to queue positions. The first parsimonious specification is our baseline estimate. The second specification explores whether the number of larger and higher queued generators has a different effect on the study arrival. In the third specification, we consider the entire queue size and use a richer set of other generator covariates.

Conditional on how long a generator has been waiting, we find that the queue position has a generally small effect on the arrival probability of the second study but a negative effect on the arrival probability of the third study. The number of large, higher-queued generators ($>$ 20MW) has little additional effect conditional on the total number of higher-queued generators. These findings are based on the implied average marginal effect of increasing the number of higher queued generators, calculated as the marginal change in the study arrival probability for every generator-quarter observation waiting for that study. A 10% increase in higher-queued generators reduces the probability of receiving the third study by 5% on average, and this effect is similar across generator sizes and fuel types (Appendix F.8).[39]

[38]One may be concerned that these unobserved local shocks may cause withdrawals that slow down studies of nearby generators. In the third column of Appendix Table F.9, we include the size, completion and withdrawal of nearby generators in the queue, and our results do not change.

[39]Our findings are robust to including a rich set of other covariates (Appendix Table F.9). The estimates of the ordered probit function for the interconnection costs are in Table F.10, and in Table F.6, we present the estimates of p_t^{final}, p_t^{test}, and p_t^{cluster}.

Table C.1: New Study Arrival Probit Model: Queue Position Parameters

	(1) Baseline	(2) Congestion Effects of Higher-Queued Large Generators	(3) Additional Queue and Generator Characteristics
Queue Position			
ln (# Higher Queued Generators)	-0.10	-0.10	-0.10
	(0.04)	(0.04)	(0.05)
ln (Total # Generators in the Queue)			-0.19
			(0.09)
Generator i Waiting for Third Study $\times$			
ln (# Higher Queued Generators)	0.74	0.49	0.48
	(0.15)	(0.43)	(0.24)
ln (# Higher Queued Generators)^2	-0.11	-0.11	-0.10
	(0.02)	(0.05)	(0.02)
ln (# Higher Queued Generators > 20MW)		0.23	
		(0.40)	
ln (# Higher Queued Generators > 20MW)^2		0.02	
		(0.06)	
ln (Total # Generators in the Queue)			-0.10
			(0.11)
ln (Higher Queued Generators Capacity MW)			0.19
			(0.14)
Generator i Waiting for Third Study $\times$			
ln (# Higher Queued Generators) $\times$Entry Year			
2013-2015	0.16	0.15	0.12
	(0.02)	(0.02)	(0.03)
2016-2018	0.16	0.15	0.15
	(0.02)	(0.02)	(0.03)
>2018	0.04	0.02	0.04
	(0.03)	(0.03)	(0.04)

Parameter estimates related to the queue position for the new study arrival probit. # Higher Queued Generators is the number of generators that entered the queue in the same or an earlier entry cohort and are still in the queue for that quarter. SE in parentheses. Appendix Table F.9 presents estimates of the parameters for other covariates.

D Equilibrium Simulation

D.1 Preliminaries

Each potential generator i is associated with a queue date t_i. If the generator decides to enter the queue, it receives the first study. In the calendar time periods $t_i+1, t_i+2, \ldots, t_i+T_0$, if the generator has not received the final study, the generator decides whether to wait or continue. At the end of t_i+T_0, if the generator has not received the final study, the generator leaves the queue.

In the simulation, we solve for a vector of values $n_i(\ell, z, \tau, \tilde{\tau}, k)$, which is the fraction of generator i with a cost estimate in bin ℓ (Appendix C.2.1), study information z, having waited

τ periods, $\tilde{\tau}$ periods from the last study, and waiting for kth study. We do not need the x in the argument because the simulation is done for each potential entrant i in the set $\mathcal{X}$. We also do not need the calendar time index because the calendar time at τ is $t_i + \tau$. In our solution, $n_i(\ell, z, \tau, \tilde{\tau}, k) \in [0, 1]$ is a fraction. Ignoring the integer constraint greatly simplifies the simulation process. In particular, we treat, for example, uncertain arrivals of studies as fractions of the waiting generator receiving studies over multiple periods. Removing stochasticity allows us to tractably model the equilibrium. For the study information z^{test}, we set the variable to 1 in all subsequent studies if PJM performs the corresponding tests in a study.

D.2 Simulation Steps

- Initialization

  - We start with an initial guess of $n_i(\ell, z, \tau, \tilde{\tau}, k)$, where n_i is the fraction of a potential generator i in a particular state and ℓ indicates the ℓth cost bin. We use the observed data as the starting point. For example, if a generator i with 1 study, cost $c \in C_\ell$ and information z waited 2 periods and exited the queue, then $n_i(\ell, z, \tau = 1, \tilde{\tau} = 1, k = 1) = 1$, $n_i(\ell, z, \tau = 2, \tilde{\tau} = 2, k = 1) = 1$, and 0 for any other input values.
  - Use the initial n_i, compute the probabilities implied by $H_1, H_2, \tilde{H}_2$ and H_3. The only additional input we need is the queue state for each i. We use y_{it} to denote i's queue state. All of the queue states we include in the H functions are computable from the observed outcomes in data. For example, the queue size at the calendar time t, is calculated as

$$\sum_k \sum_{\tilde{\tau}} \sum_\tau \sum_\ell \sum_z \sum_i n_i(\ell, z, \tau, \tilde{\tau}, k) \mathbb{1}(t_i + \tau = t).$$

  - Given the predicted probabilities (v_i, p_i, q_i, r_i) based on the H functions, we solve the Bellman equations and the corresponding withdrawal probabilities for each generator i, $\Upsilon_i, \Lambda_i, \Psi_i$ and Ξ_i for each combination of $(\ell, z, \tau, \tilde{\tau})$. To simplify notation, we omit queue state y_{it} and generator characteristics x_i from the arguments of these functions.

- Iteration

  1. Use the study transition probabilities predicted by H and withdrawal probabilities probabilities $(\Upsilon_i, \Lambda_i, \Psi_i, \Xi_i)$ to update n_i. Specifically, for those with two studies

and waiting for the third study and for any ℓ and z,

$$\begin{aligned}
&n_i\left(\ell, z, \tau+1, \tilde{\tau}+1, k=2\right) \\
&= n_i\left(\ell, z, \tau, \tilde{\tau}, k=2\right) \cdot \left(1-\Lambda_i\left(\ell, z, \tau, \tilde{\tau}\right)\right) \\
&\cdot \left(1-\sum_{\ell'} r_i\left(\ell'; c, z, \tau, \tilde{\tau}\right)\right).
\end{aligned}$$

In the first period after receiving the second study,

$$\begin{aligned}
&n_i\left(\ell, z, \tau+1, \tilde{\tau}=1, k=2, x\right) \\
&= \sum_{\ell', z', \tilde{\tau}'} n_i\left(\ell', z', \tau, \tilde{\tau}', k=1\right) \\
&\cdot \left(1-\Upsilon_i\left(\ell', z', \tau, \tilde{\tau}'\right)\right) \cdot q_i\left(\ell, z; \ell', z', \tau, \tilde{\tau}'\right).
\end{aligned}$$

For those with one study and waiting for the second study and for any ℓ and z,

$$\begin{aligned}
&n_i\left(\ell, z, \tau+1, \tilde{\tau}+1, k=1\right) \\
&= n_i\left(\ell, z, \tau, \tilde{\tau}, k=1\right) \cdot \left(1-\Upsilon_i\left(\ell, z, \tau, \tilde{\tau}\right)\right) \\
&\cdot \left(1-\sum_{\ell', z'} q_i\left(\ell', z'; \ell, z, \tau, \tilde{\tau}\right)-\sum_{\ell'} p_i\left(\ell'; \ell, z, \tau, \tilde{\tau}\right)\right).
\end{aligned}$$

The entry process implies that

$$n_i\left(c, z, \tau=1, \tilde{\tau}=1, k=1\right)=v_i(c, z) \cdot \left(1-\Xi_i\right).$$

2. Update the queue status y_t, using n_i as the number of generators in a particular state.
3. Update the study transition and withdrawal probabilities.
4. Iterate until convergence.

D.3 Balance Conditions in Alternative Mechanisms

We use $w_{t,x}$ to denote the weight for a generator of characteristics x and entry time t. The balance conditions are:

1. For generators with two studies and waiting for the third study,

$$
\begin{aligned}
& N_{t+1} m_{t+1}(c, z, \tau+1, \tilde{\tau}+1, k=2, x) \\
& = N_t m_t(c, z, \tau, \tilde{\tau}, k=2, x) \cdot (1-\Lambda_t(c, z, \tau, \tilde{\tau}, x)) \\
& \cdot \left(1-\sum_{c'} r_t(c'; c, z, \tau, \tilde{\tau}, x) w_{t-\tau, x}\right).
\end{aligned} \tag{D.2}
$$

2. For generators just receiving the second study,

$$
\begin{aligned}
& N_{t+1} m_{t+1}(c, z, \tau+1, \tilde{\tau}=1, k=2, x) \\
& = N_t \sum_{c', z', \tilde{\tau}'} m_t(c', z', \tau, \tilde{\tau}', k=1, x) \\
& \cdot (1-\Upsilon_t(c', z', \tau, \tilde{\tau}', x)) \cdot q_t(c, z; c', z', \tau, \tilde{\tau}', x) w_{t-\tau, x}.
\end{aligned} \tag{D.3}
$$

3. For generators with one study and waiting for the second study ($\tau > 0$),

$$
\begin{aligned}
& N_{t+1} m_{t+1}(c, z, \tau+1, \tilde{\tau}+1, k=1, x) \\
& = N_t m_t(c, z, \tau, \tilde{\tau}, k=1, x) \cdot (1-\Upsilon_t(c, z, \tau, \tilde{\tau}, x)) \\
& \cdot \left(1-\sum_{c', z'} q_t(c', z'; c, z, \tau, \tilde{\tau}, x) w_{t-\tau, x}-\sum_{c'} p_t(c'; c, z, \tau, \tilde{\tau}, x) w_{t-\tau, x}\right).
\end{aligned} \tag{D.4}
$$

4. The boundary condition (entry):

$$
N_t m_t(c, z, \tau=1, \tilde{\tau}=1, k=1, x) = v_t(c, z; x) n_t(x) (1-\Xi_t(x)) w_{t, x}. \tag{D.5}
$$

E Additional Counterfactual Results

E.1 Additional Detail on Study Fees

Figure E.1 shows that the added capacity falls at every level for both types of study fees. There are two key distinctions from the entry fee, which we find can increase completed capacity. First, the additional study fees may cause large generators with a high interconnection cost identified in study 1 to drop out, but small generators with a low study 1 cost to stay in the queue. Second, generators that wait long enough to request the third study are those that are more likely to complete interconnection, and screening at this stage becomes less effective. As a result, these fees do not sufficiently protect the incentives of large generators to remain in the queue and lead to an overall reduction in the completed capacity.

E.2 Subsidizing Interconnection Costs

The cluster study counterfactuals reduce uncertainty, re-allocate costs across generators, and subsidize these costs. To understand the role of subsidizing costs in increasing completed capacity, we next consider a subsidy that is a percentage of a generator's interconnection costs, holding entry fixed. For different subsidy levels, we compute the effect on completed capacity as well as the total subsidy cost.

We find that, as in the cluster cost-allocation simulations, subsidies for interconnection costs disproportionately benefit renewable generators and have diminishing returns. Figure E.2 reports counterfactual investment and the corresponding subsidy cost for subsidy levels of 10%, 20%, ..., 100%. The solid line shows the increase in total capacity, and the dashed line shows the increase in renewable capacity. For all subsidy levels, more than half of the resulting increase in capacity is renewable capacity.

Compared to the cluster cost-allocation simulations (Table 7), directly subsidizing interconnection costs results in more total and renewable capacity. For example, subsidizing interconnection costs at 50% requires approximately 4 billion dollars but increases total capacity by 13 GW, of which 7 GW is renewable capacity. In contrast, a similar level of subsidy for cluster-level cost allocation (4.5 billion dollars) adds just 8 GW of total capacity and 5 GW of renewable capacity. This result that direct subsidies are more effective is likely because, compared to sharing costs equally, subsidizing a percentage of costs gives generators with high "true" costs a greater incentive to withdraw. Overall, these results suggest that equalizing costs across generators without subsidy would reduce completed capacity.

F Additional Figures and Tables

Table F.9: New Study Arrival Probit Model: Additional Estimates

	(1) Baseline	(2) Large Generators	(3) More Covariates
Last Study Cost ($Million/MW)			
0.01-0.05	0.11	0.11	0.12
	(0.04)	(0.04)	(0.04)
0.05-0.20	0.05	0.06	0.07
	(0.04)	(0.04)	(0.04)
>0.20	-0.01	-0.00	-0.02
	(0.05)	(0.05)	(0.06)
Has Tests (z^{test})	-0.12	-0.12	-0.13
	(0.04)	(0.04)	(0.04)
Cost Sharing ($z^{\text{cost-sharing}}$)	0.04	0.03	0.01
	(0.05)	(0.05)	(0.05)
ln # Projects in the Queue			
<10km			0.00
			(0.02)
<100km			0.11
			(0.04)
Same TO	-0.03	-0.03	-0.03
	(0.02)	(0.02)	(0.02)
Withdrawn in the Past 2 Quarters, <100km			-0.04
			(0.04)
Withdrawn in the Past 2 Quarters, Same TO			-0.02
			(0.01)
Generator i Waiting for Third Study $\times$			
ln # Projects in the Queue$\times$			
Same TO	-0.14	-0.13	-0.13
	(0.03)	(0.03)	(0.04)
Withdrawn in the Past 2 Quarters, <100km			-0.08
			(0.07)
Withdrawn in the Past 2 Quarters, Same TO			-0.00
			(0.01)
Wind/Solar	-0.04	-0.04	0.35
	(0.04)	(0.04)	(0.11)
Wind/Solar$\times$Entry Year			
2013-2015			0.52
			(0.17)
2016-2018			0.31
			(0.20)
>2018			0.17
			(0.23)
Wind/Solar$\times$Year			
>2013-2015			-0.60
			(0.16)
>2016-2018			-0.51
			(0.20)
>2018	0.06	0.07	-0.42
	(0.05)	(0.05)	(0.23)
Wind/Solar$\times$Capacity (MW)			
10-20			-0.23
			(0.10)
20-100			-0.19
			(0.10)
100-500			-0.12
			(0.11)
>500			-0.35
			(0.16)
Capacity (MW)			
10-20			0.21
			(0.08)
20-100			0.13
			(0.08)
100-500	-0.73	-0.73	-0.53
	(0.08)	(0.08)	(0.12)
>500	-0.63	-0.64	-0.34
	(0.10)	(0.10)	(0.13)
Capacity>100$\times$Entry Year			
>2013-2015			-0.02
			(0.16)
>2016-2018			-0.16
			(0.19)
>2018			-0.34
			(0.21)
Capacity>100$\times$Year			
>2013-2015	0.65	0.64	0.60
	(0.10)	(0.10)	(0.15)
>2016-2018	0.60	0.60	0.66
	(0.10)	(0.10)	(0.20)
>2018	0.52	0.52	0.70
	(0.09)	(0.09)	(0.22)

We additionally include entry year fixed effects, year fixed effects, number of quarters in the queue fixed effects and their interactions with whether the generator is waiting for its third study, state fixed effects, indicators for whether the generator is an uprate, waiting for its third study, and study 1 indicates the generator might later share costs with others.

Table F.10: Ordered Probit Model: Cost Estimates

	(1) Baseline	(2) Large Generators	(3) More Covariates
Last Study Cost ($Million/MW)			
0.01-0.05	0.49	0.49	0.51
	(0.08)	(0.08)	(0.09)
0.05-0.20	1.24	1.24	1.21
	(0.08)	(0.08)	(0.08)
>0.20	2.34	2.34	2.31
	(0.09)	(0.09)	(0.09)
Has Tests (z^{test})			0.07
			(0.07)
Cost Sharing ($z^{\text{cost-sharing}}$)	-0.17	-0.17	-0.23
	(0.10)	(0.10)	(0.12)
ln Total MW in the Queue			
<10km			0.03
			(0.03)
<100km			-0.01
			(0.07)
Same TO	-0.03	-0.03	-0.02
	(0.02)	(0.02)	(0.03)
PJM			0.08
			(0.13)
Withdrawn in the Past 2 Quarters, <100km	0.02	0.02	0.01
	(0.01)	(0.01)	(0.01)
Withdrawn in the Past 2 Quarters, Same TO			0.00
			(0.02)
Generator i Waiting for Third Study ×			
ln Total MW in the Queue			
PJM			0.01
			(0.01)
Withdrawn in the Past 2 Quarters, <100km	-0.02	-0.02	-0.02
	(0.01)	(0.01)	(0.02)
Withdrawn in the Past 2 Quarters, Same TO			-0.00
			(0.02)
Wind/Solar	0.45	0.45	0.33
	(0.09)	(0.09)	(0.18)
Wind/Solar×Entry Year			
>2013-2015			0.42
			(0.32)
>2016-2018			0.59
			(0.41)
>2018	-0.48	-0.48	0.08
	(0.20)	(0.20)	(0.46)
Wind/Solar×Year			
>2013-2015			-0.03
			(0.30)
>2016-2018			-0.43
			(0.42)
>2018	0.09	0.09	-0.28
	(0.19)	(0.19)	(0.49)
Capacity (MW)			
10-20			-0.08
			(0.09)
20-100			-0.01
			(0.10)
100-500			-0.51
			(0.24)
>500			-0.46
			(0.27)
Capacity>100×Entry Year			
>2013-2015	0.02	0.02	0.89
	(0.15)	(0.15)	(0.32)
>2016-2018	-0.13	-0.13	0.93
	(0.12)	(0.12)	(0.43)
>2018	0.14	0.14	1.22
	(0.09)	(0.09)	(0.47)
Capacity>100×Year			
>2013-2015			-0.26
			(0.32)
>2016-2018			-0.44
			(0.44)
>2018			-0.55
			(0.50)

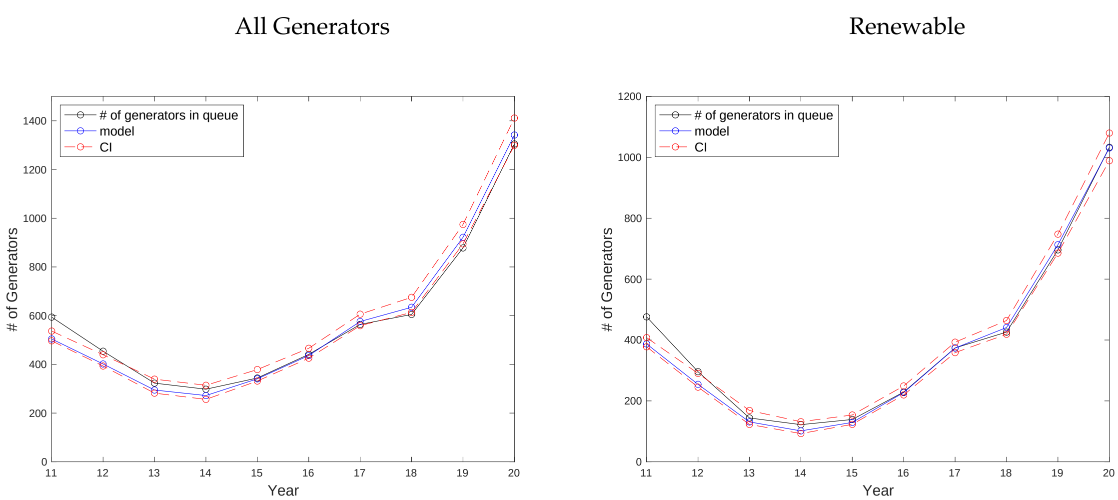
We additionally include entry year and year fixed effects at a 3-year level, state fixed effects, number of quarters in the queue fixed effects up to 2 years and interactions with whether the generator is waiting for the third study, indicators for whether the generator is an uprate, waiting for its third study, study 1 indicates the generator might later share costs with others, any upgrade was completed during the same substation, the most recent interconnection at the same substation, and any RTEP was completed at the same substation.The ordered probit cutoffs are $[0, \hat{\mu}_2 = 0.93, \hat{\mu}_3 = 1.24]$, with (μ_2, μ_3)'s standard errors at $(0.04, 0.08)$ in the baseline.

Table F.11: Standard Deviation Estimates for Unobservables ($1,000/MW)

ξ_t	Year			
	<2013	2013-2015	2016-2018	>2018
Renewable				
Waiting Time≤2 Years	877.03	240.11	381.58	471.48
	(120.45)	(86.62)	(56.52)	(40.67)
Waiting Time>2 Years	1294.16	342.51	613.45	546.47
	(211.91)	(73.18)	(79.83)	(40.36)
Non-Renewable				
Waiting Time≤2 Years	666.73	320.08	342.46	247.62
	(152.71)	(52.92)	(47.10)	(36.46)
Waiting Time>2 Years	354.98	324.09	227.92	95.21
	(340.11)	(79.61)	(74.43)	(28.36)
ε_t		78.77		
		(4.80)		

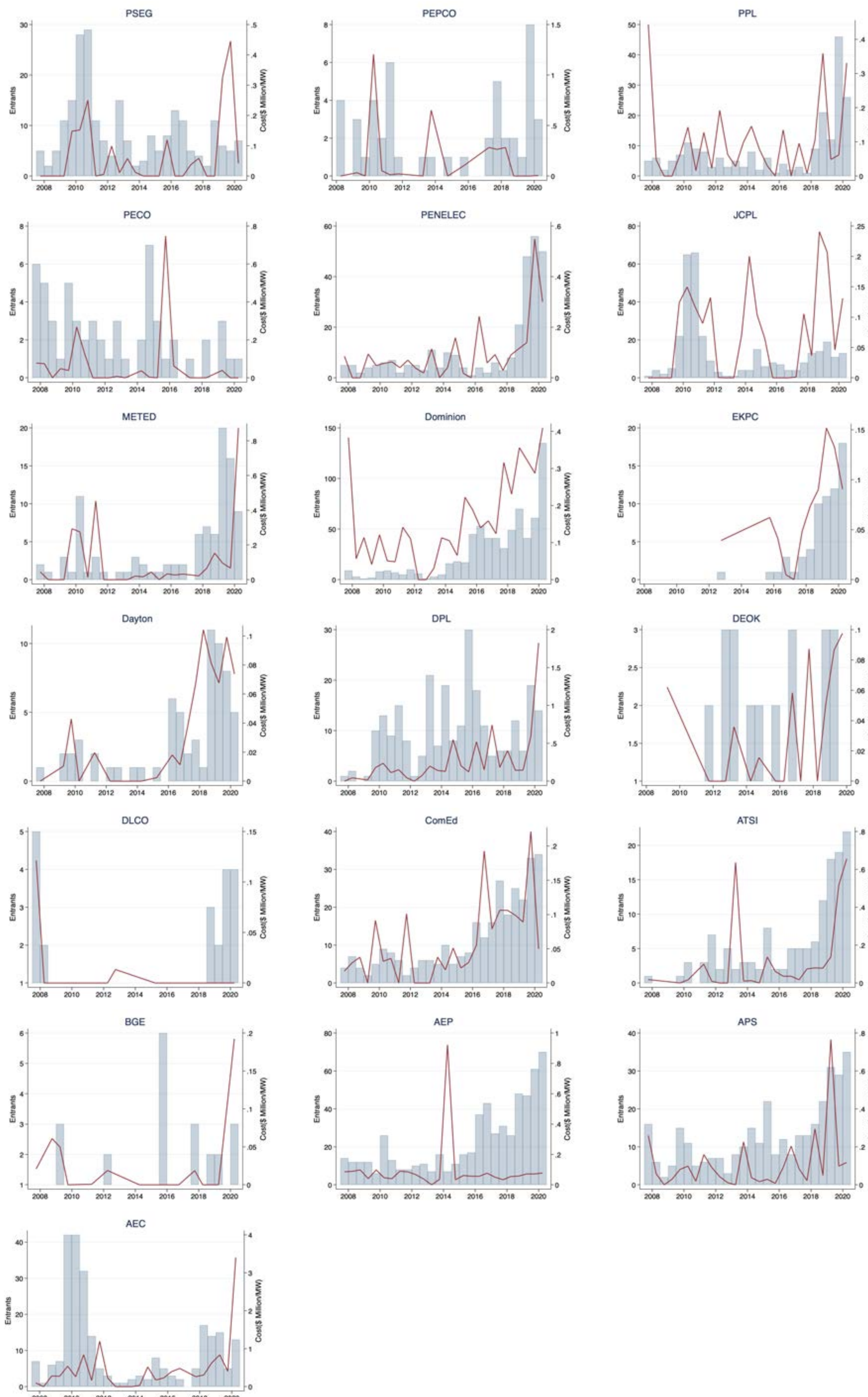
Estimates of the standard deviations of the unobservables. SEs in parentheses.

Figure F.4: Number of Generators Waiting in Queue 2011-2020, Model and Data



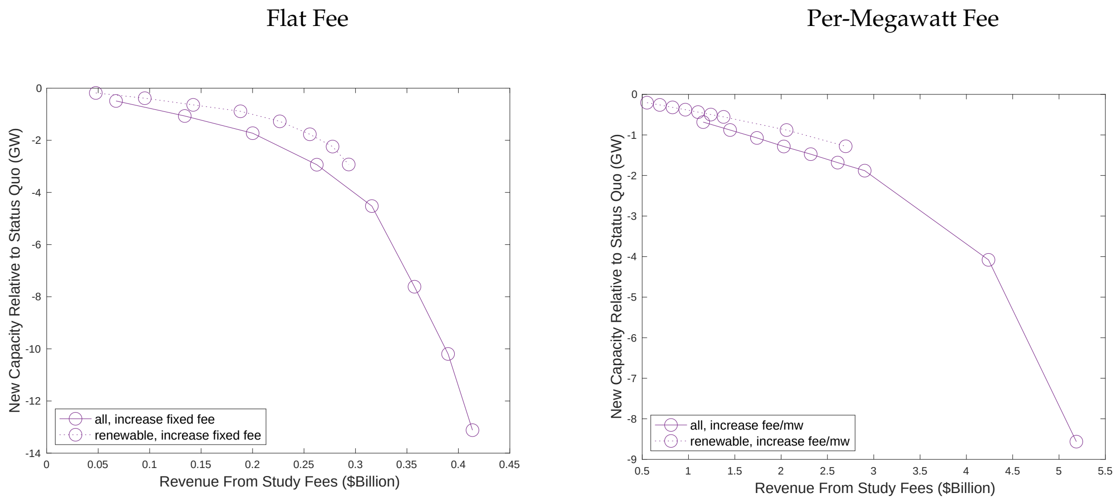
For year $t = 2011, \ldots, 2020$, the solid black lines (number of generators waiting in queue) at t represents the number of generators that entered the queue in or after 2008 and had not withdrawn by t. The blue line is the simulated prediction. The red line represents the confidence intervals for the model prediction.

Figure B.1: Mean Study 2 Costs and Entry in Each Transmission Owner Territory



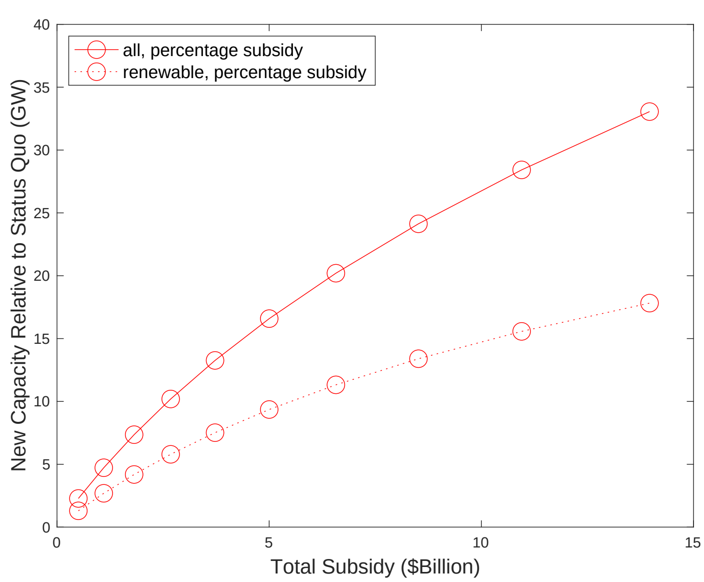
The bar graph (left $y-$axis) shows the number of entrants in a transmission owner territory by year, and the red line (right $y-$axis) shows the average study 2 costs for studies issued in that territory by year.

Figure E.1: Added Capacity with Increased Study Fee



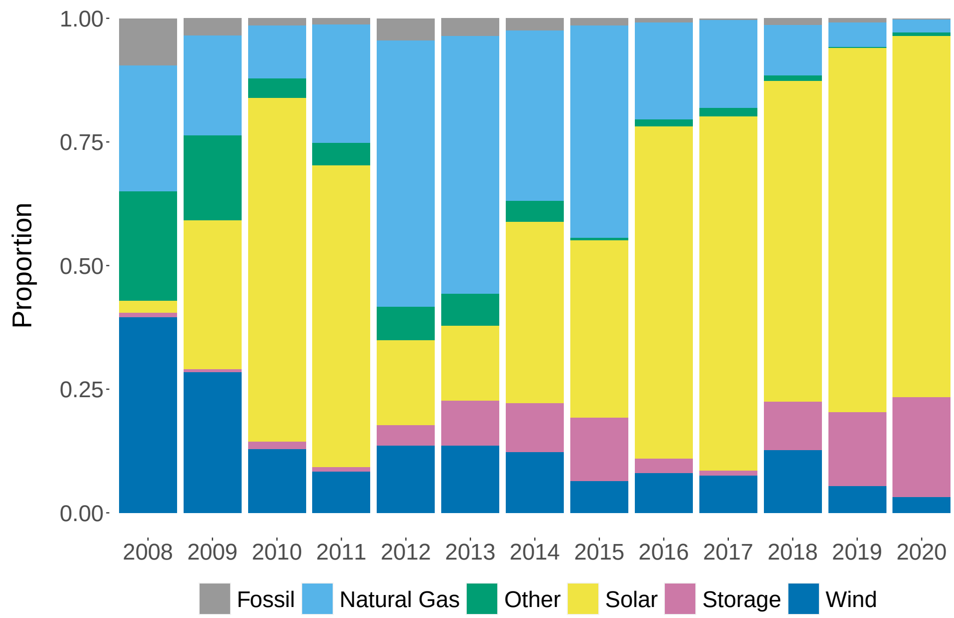
Y-axis is new capacity added relative to the status quo. X-axis is the increase in fees for the second and third studies (we assume both are increased by the same amount). The solid line shows the change in total generation capacity, the dotted line shows the change in renewable generation capacity.

Figure E.2: Subsidizing Interconnection Costs



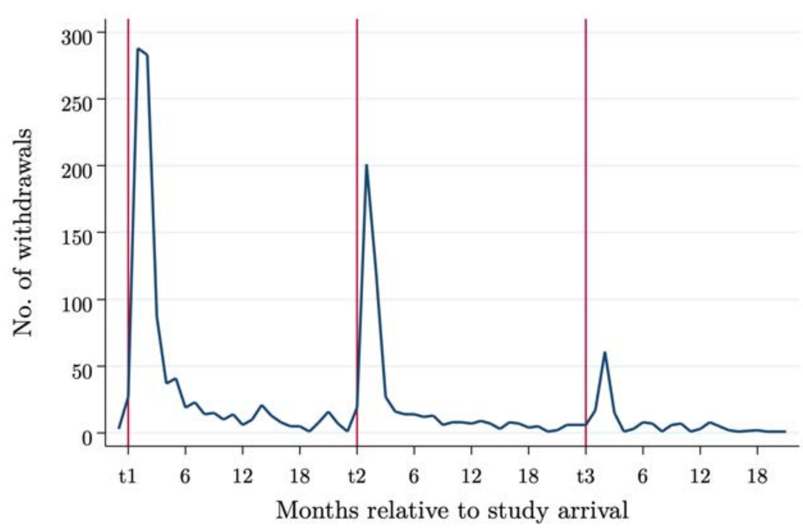
Y-axis is new capacity added relative to the status quo. X-axis is the total cost of the subsidy. The solid line shows the change in total generation capacity, while the dotted line shows the change in renewable generation capacity. From left to right, the points on each line are subsidies for interconnection costs of 10%, 20%, . . ., 100%.

Figure F.1: Fuel type of generators entering the queue



Proportion of new generator interconnection requests by fuel type. Fossil is coal, oil, and diesel; other is biomass, nuclear, hydro, and wood.

Figure F.2: Withdrawals Relative to Study Arrival



Number of generators withdrawing from the queue in each month relative to the arrival of their most recent study. t1 is the arrival of the first study. t2 is the arrival of the second study. t3 is the arrival of the third study. Excludes generators that withdrew two or more years after the arrival of their most recent study.

Table F.1: Full Summary Statistics

	Study 1		Study 2		Study 3	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Cost per MW	0.12	0.40	0.18	0.48	0.10	0.16
. . . ≤0.01m	0.41	0.49	0.27	0.45	0.16	0.37
. . . (0.01m, 0.05m]	0.20	0.40	0.21	0.41	0.32	0.47
. . . (0.05m, 0.1m]	0.14	0.35	0.17	0.38	0.23	0.42
. . . >0.1m	0.25	0.43	0.34	0.48	0.29	0.46
Wait time (mos.)	5.3	2.7	12.7	11.3	19.0	13.0
Size in MW	97	196	106	191	162	276
. . . ≤10	0.27	0.45	0.17	0.37	0.09	0.28
. . . 10-20	0.26	0.44	0.27	0.44	0.23	0.42
. . . 20-100	0.25	0.43	0.33	0.47	0.36	0.48
. . . 100-500	0.17	0.38	0.20	0.40	0.22	0.42
. . . >500	0.05	0.21	0.05	0.21	0.10	0.30
Uprate	0.21	0.41	0.23	0.42	0.16	0.37
Revision	0.04	0.20	0.19	0.39	0.18	0.39
Requested energy in MW	97	197	105	191	162	276
Requested capacity in MW	66	174	68	164	121	259
Solar	0.60	0.49	0.60	0.49	0.59	0.49
Natural Gas	0.17	0.37	0.16	0.36	0.22	0.42
Wind	0.09	0.29	0.10	0.30	0.12	0.32
Battery	0.10	0.30	0.10	0.30	0.04	0.19
Coal, oil, diesel	0.02	0.13	0.01	0.12	0.01	0.09
Other	0.03	0.18	0.03	0.17	0.02	0.15
Cost sharing	0.04	0.19	0.60	0.49	0.60	0.49
Study 1 cost sharing	0.42	0.49	0.48	0.50	0.45	0.50
Receive engr. tests	0.82	0.39	0.88	0.32	0.04	0.21
Distance to substation (km)	3.81	5.99	3.71	6.18	3.40	5.17
Build new substation	0.22	0.41	0.27	0.44	0.47	0.50
Ordinance	0.28	0.45	0.31	0.46	0.30	0.46
Prior RTEP Investment (0m, 60m]	0.47	0.50	0.48	0.50	0.44	0.50
Prior RTEP Investment >60m	0.18	0.38	0.18	0.39	0.20	0.40
N	4,083		2,433		672	

Generators entering the queue in 2008-2020. Costs in millions of 2020 dollars. Cost per MW is interconnection cost estimate divided by the generator's size in MW. Wait time for Study 1 is wait in months for the first study after joining the queue. Wait time for Study 2 is wait in months for second study after receiving the first study. Wait time for Study 3 is similarly defined. Size in MW is our measure of size which is the maximum of the requested energy and requested capacity. Uprate is an indicator for a capacity increase to an existing generator. Revision is an indicator for if the study was revised. Requested energy in MW is the associated energy resource for the interconnection request. Requested capacity in MW is the associated capacity resource for the interconnection request; all capacity resources must participate in PJM's capacity auction. Cost sharing is an indicator for if a generator shares costs with other generators. Study 1 cost sharing is an indicator for if the first study mentions shared network upgrade costs. Receive engr. tests is an indicator for receiving any of three engineering tests: generator deliverability, multiple facility contingency, and short circuit analysis. Distance to substation is the distance to the nearest substation in km. Ordinance is an indicator for a local ordinance restricting renewable energy development. Prior RTEP Investment is RTEP transmission investment made within 18 months of the issue date of a focal project and located within 10km of it.

Table F.2: Predictors of Interconnection Costs

	IHS Cost	Low Study Cost
	(1)	(2)
Project Size Bin 1 (10MW, 20MW]	-0.049**	0.009
	(0.025)	(0.027)
Project Size Bin 2 (20MW, 100MW]	-0.020	-0.134***
	(0.026)	(0.027)
Project Size Bin 3 (100MW, 500MW]	-0.028	-0.168***
	(0.027)	(0.030)
Project Size Bin 4 >500MW	-0.056*	-0.210***
	(0.030)	(0.051)
Distance To Substation:	0.003	-0.007
	(0.004)	(0.005)
Ordinance	0.029	-0.043*
	(0.019)	(0.022)
Uprate	-0.055***	0.494***
	(0.021)	(0.029)
Fuel-Natural Gas	-0.044	0.104
	(0.053)	(0.078)
Fuel-Solar	-0.026	-0.122
	(0.054)	(0.080)
Fuel-Wind	-0.031	-0.034
	(0.060)	(0.084)
Fuel-Storage	-0.028	-0.051
	(0.055)	(0.081)
Fuel-Other	-0.032	0.129
	(0.063)	(0.098)
Prior RTEP Investment (0m, 60m]	-0.012	0.035**
	(0.015)	(0.017)
Prior RTEP Investment >60m	-0.015	0.056**
	(0.022)	(0.026)
Observations	2,432	2,432
R-Squared	0.122	0.421

Projects queuing from 2008-2020. SEs in parentheses; clustered by substation. In column (1), Dep. var. is the inverse hyperbolic sine transformation of Study 2 interconnection cost estimate (mean 0.14). In column (2), Dep. var. is indicator for Study 2 cost estimate less than 0.01m/MW (mean 0.27). Prior RTEP Investments is measured as investments made within 18 months of the issue date of a focal project and located within 10km of it. Controls for state, the year of queue entry and the year the study is issued. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table F.3: Withdrawal Regressions with Alternative Cost Measures

	Study 1		Study 2			Study 3	
	OLS	OLS	OLS	IV	IV	OLS	IV
A. Continuous Measure							
IHS cost	0.290***	0.290***	0.347***	0.350***	0.290***	0.457**	0.092
	(0.054)	(0.054)	(0.057)	(0.074)	(0.072)	(0.179)	(0.427)
B. Cost bin specification							
Low cost	-0.042	-0.043*	-0.046			0.023	
	(0.025)	(0.025)	(0.050)			(0.076)	
Mid cost	-0.011	-0.010	-0.046			0.110	
	(0.027)	(0.027)	(0.052)			(0.086)	
High cost	0.107***	0.106***	0.198***			0.174**	
	(0.027)	(0.027)	(0.046)			(0.083)	
Study 1 shared		X			X		
Mean dep. var.	0.28	0.28	0.43	0.43	0.43	0.43	0.43
First stage F				215	195		12
N	3,191	3,191	1,269	1,269	1,269	345	309

Generators queuing from 2011-2020; generators still active excluded. SEs in parentheses; clustered by substation. Dep. var. are indicators for projects withdrawing from the queue before receiving the next study or before beginning operation for generators with their final study. IHS cost is the inverse hyperbolic sine transformation of the study's interconnection cost estimate. IV results instrument for the IHS of cost using the change in the IHS of cost across studies. Low cost bin indicates an interconnection cost between 0.01 and 0.05 million dollars per MW. Mid cost bin indicates an interconnection cost between 0.05 and 0.1 million dollars per MW. High cost bin is a cost greater than 0.1 million dollars per MW. Study 1 shared indicates if the regression includes two variables from Study 1: an indicator for if the study mentions shared network upgrade costs and the log of the total costs to be shared. All specifications control for size (5 bins), fuel type, state, uprate, and FE for the year of queue entry and the year the study is issued. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table F.4: Withdrawal Regressions with Alternative Controls

	Study 1		Study 2			Study 3	
	OLS	OLS	OLS	IV	IV	OLS	IV
A. Original specification							
High cost	0.123***	0.123***	0.231***	0.293***	0.241***	0.113*	0.074
	(0.022)	(0.022)	(0.031)	(0.054)	(0.057)	(0.063)	(0.117)
Study 1 shared		X			X		
Mean dep. var.	0.28	0.28	0.43	0.43	0.43	0.55	0.55
First stage F				972	808		202
N	3,191	3,191	1,269	1,269	1,269	345	345
B. Including permitting controls							
High cost	0.123***	0.123***	0.230***	0.293***	0.241***	0.115*	0.074
	(0.022)	(0.022)	(0.031)	(0.054)	(0.057)	(0.064)	(0.118)
Study 1 shared		X			X		
Mean dep. var.	0.28	0.28	0.43	0.43	0.43	0.55	0.55
First stage F				948	788		199
N	3,191	3,191	1,269	1,269	1,269	345	345
A. Including substation FE							
High cost	0.108***	0.107***	0.275***	0.267***	0.246***	0.255*	0.347*
	(0.031)	(0.031)	(0.049)	(0.072)	(0.075)	(0.134)	(0.176)
Study 1 shared		X			X		
Mean dep. var.	0.26	0.26	0.41	0.41	0.41	0.43	0.43
First stage F				365	326		18
N	2,413	2,413	759	759	759	138	138

Generators queuing from 2011-2020; generators still active excluded. Panel C. excludes generators that were the only generator connecting at that substation. SEs in parentheses; clustered by substation. Dep. var. are indicators for projects withdrawing from the queue before receiving the next study or before beginning operation for generators with their final study. Permitting controls are two variables: the log of the distance from the generator location to the nearest substation in kilometers and whether the county had an ordinance restricting renewable energy development. Study 1 shared indicates if the regression includes two variables from Study 1: an indicator for if the study mentions shared network upgrade costs and the log of the total costs to be shared. All specifications control for size (5 bins), fuel type, uprate, and FE for the year of queue entry and the year the study is issued. Panels A. and B. include state FE; Panel C. include substation FE. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table F.5: Top 15 Developers and Selective Completion

Developer	No. of generators	Cohorts with a generator	Avg. gen. per cohort	Cohorts, excl. active	Frac. all or 0 completed
Invenergy	79	18	4.4	8	0.75
PSEG	53	17	3.1	17	0.53
Community Energy	52	17	3.1	11	0.64
Dominion	49	19	2.6	16	0.81
LS	39	18	2.2	12	0.83
EDF	30	12	2.5	6	0.83
Apex	28	13	2.2	6	0.67
NextEra	27	11	2.5	5	1
SunEnergy1	27	9	3	5	0.4
Urban Grid	27	6	4.5	0	-
EffiSolar	26	5	5.2	5	0.6
EDP	25	13	1.9	9	0.78
IMG	24	8	3	8	0.63
AEP	23	11	2.1	10	0.9
Exelon	20	10	2	9	0.89

Summary statistics for the top 15 developers in our sample by number of generators. Cohorts with a generator is the number of queue entry cohorts (out of 33) that the developer has a generator in. Cohorts, excl. active is the same but excludes cohorts where, as of August 2023, at least one of the developer's generators is still active, i.e., it has neither begun operation or withdrawn. Frac. all or 0 completed is the mean of a developer-cohort level indicator for if either all or none of the generators in that cohort were completed; this mean is taken across cohorts without active generators.

Table F.6: PJM Decisions $\left(p_t^{\text{final}}, p_t^{\text{test}}, p_t^{\text{cluster}}\right)$ and Study 1 Cost Distribution H_1

	Probit			Ordered Probit
	Study 2 is Final	Test	Cost Sharing	Study 1 Cost
Last Study Cost ($Million/MW)				
0.01-0.05	-0.47	-0.09	0.32	
	(0.11)	(0.27)	(0.09)	
0.05-0.20	-0.84	-0.12	0.31	
	(0.11)	(0.30)	(0.09)	
>0.20	-1.43	-0.13	-0.28	
	(0.15)	(0.30)	(0.14)	
Capacity (MW)				
10-20	-0.64	0.63	0.35	-0.34
	(0.10)	(0.21)	(0.11)	(0.04)
20-100	-1.26	0.35	0.70	-0.12
	(0.13)	(0.24)	(0.12)	(0.05)
>100	-2.03	0.90	0.85	-0.17
	(0.15)	(0.31)	(0.13)	(0.05)
Renewable	-0.16	0.94	-0.23	0.81
	(0.09)	(0.26)	(0.08)	(0.04)
ln Total MW in the Queue				
Withdrawn in the Past 2 Quarters, <100km	-0.09	0.04	0.01	0.03
	(0.01)	(0.02)	(0.01)	(0.00)
Withdrawn in the Past 2 Quarters, Same TO	-0.15	-0.00	-0.00	0.04
	(0.01)	(0.03)	(0.01)	(0.01)

We additionally control for state, year and entry year fixed effects. For the ordered probit, we estimate μ_3 and μ_4 to be 0.58, 1.57, with standard errors of 0.017 and 0.027.

Table F.7: Waiting Cost Parameter Estimates ($1,000/MW/Quarter)

	(1) Baseline	(2) Heterogeneity
Renewable	2.83	2.08
	(1.73)	(1.33)
Renewable×		
Capacity>100MW	-1.75	-1.58
	(0.40)	(0.33)
Entry Year>2013	-2.93	-1.86
	(1.69)	(1.37)
Year		
2013-2015	-1.29	-1.12
	(3.34)	(2.47)
2016-2018	-3.90	-2.85
	(2.71)	(2.13)
>2018	4.92	3.12
	(2.08)	(1.73)
$\tilde{\tau}$ Quarters after Receiving		
$\tilde{\tau} = 1$, Study 1	30.41	22.81
	(8.40)	(6.24)
$\tilde{\tau} = 2$, Study 1	32.24	25.96
	(9.57)	(7.23)
$\tilde{\tau} = 1$, Study 2	-19.42	-16.31
	(9.32)	(7.37)
$\tilde{\tau} = 2$, Study 2	93.66	74.32
	(10.37)	(8.14)
ln Capacity×		
$\tilde{\tau}$ Quarters after Receiving		
$\tilde{\tau} = 1$, Study 1	-1.07	-0.76
	(2.32)	(1.71)
$\tilde{\tau} = 2$, Study 1	-0.40	-0.62
	(2.56)	(1.95)
$\tilde{\tau} = 1$, Study 2	-1.00	-0.36
	(2.40)	(1.91)
$\tilde{\tau} = 2$, Study 2	-7.52	-6.62
	(2.44)	(1.89)
Waiting for Study 3	3.56	2.99
	(0.65)	(0.52)

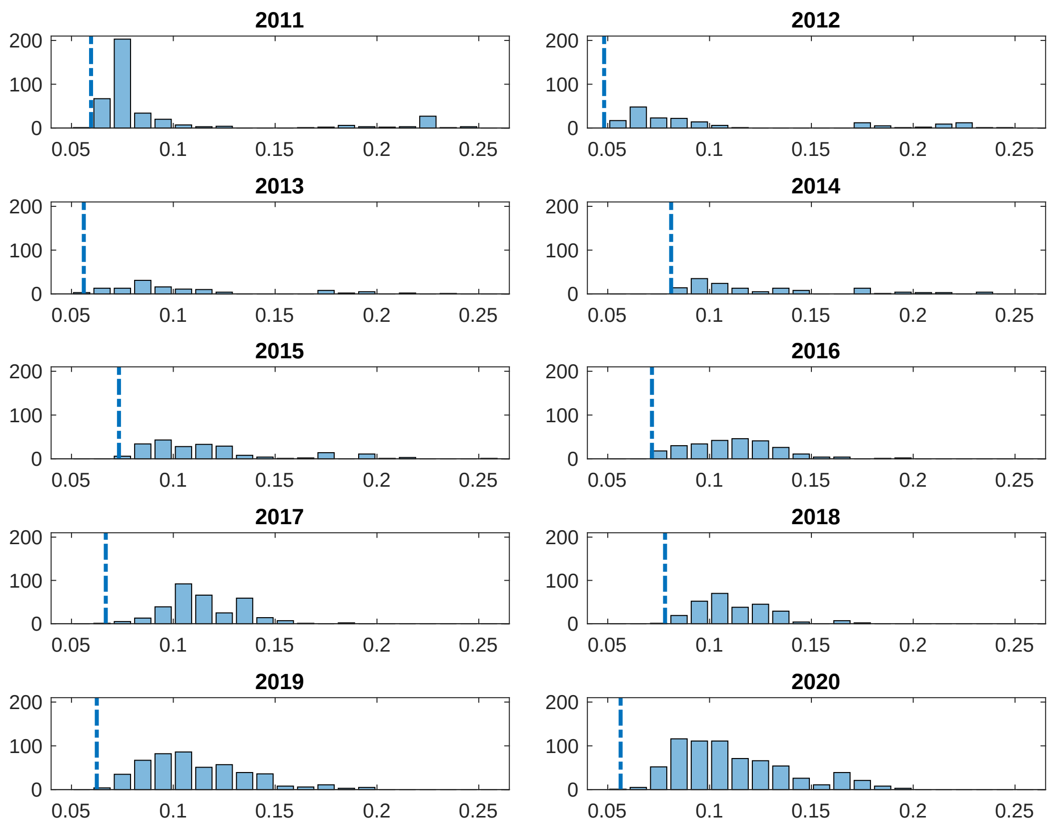
We additionally include entry year and year fixed effects at the 3-year level, state fixed effects and fixed effects for the number of quarters in the queue.

Table F.8: Marginal Effects of Increasing the Queue Sizes on Study 3 Arrival Probability

	Avg Arrival Prob	ΔQueue Position $\downarrow$=10%,	
		$\Delta \Pr$ new study	$\Delta \Pr$ new study SE
Renewable, $\leq$ 20MW	0.068	-0.004	<0.001
Renewable, $>$ 20MW	0.053	-0.003	<0.001
Non-Renewable, $\leq$ 20MW	0.093	–0.004	<0.001
Non-Renewable, $>$ 20MW	0.071	-0.003	<0.001

To calculate the $\Delta \Pr$ new study we increase the number of higher queued generators by 10%.

Figure F.3: Expected Surplus of Entering the Queue and Cost of Entry



Plots of the distribution of the expected surplus from entering the queue by year. The vertical line indicates the lowest surplus of generators that entered in each year.