WORKING PAPER

December 2017 ■ RFF WP 17-27

Vertical Market Power in Interconnected Natural Gas and Electricity Markets

Levi Marks, Charles F. Mason, Kristina Mohlin, and Matthew Zaragoza-Watkins

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

RESOURCES FOR THE FUTURE

Vertical Market Power in Interconnected Natural Gas and Electricity Markets

Levi Marks, Charles F. Mason, Kristina Mohlin

and Matthew Zaragoza-Watkins*

Abstract

New England is at the leading edge of an energy transition in which natural gas is playing an increasingly important role in the US electricity generation mix. In recent years, the region's wholesale natural gas and electricity markets have experienced severe, simultaneous price spikes. While frequently attributed to limited pipeline capacity serving the region, we demonstrate that such price spikes have been exacerbated by some gas distribution firms scheduling deliveries without actually flowing gas. This behavior blocks other firms from utilizing pipeline capacity, which artificially limits gas supply to the region and drives up gas and electricity prices. We estimate that capacity withholding increased average gas and electricity prices by 38% and 20%, respectively, over the three-year period we study. As a result, customers paid $3.6 billion more for electricity. While the studied behavior may have been within the firms' contractual rights, the significant impacts in both the gas and electricity markets underscore the need to improve regulation and coordination as these two energy markets become increasingly interlinked.

*(Marks) University of California, Santa Barbara. Email: levimarks@umail.ucsb.edu. (Mason) University of Wyoming. Email: bambuzlr@uwho.edu. (Mohlin) Environmental Defense Fund. Email: kmohlin@edf.org. (Zaragoza-Watkins, corresponding author) Vanderbilt University. Email: matthew.zaragoza-watkins@vanderbilt.edu. We are grateful for thoughtful comments from Severin Borenstein, Jon Coifman, Scott Holladay, Erin Mansur and Michael Mehling. We thank seminar participants at EDF, Harvard University, MIT, UC Santa Barbara and UT Knoxville, and participants at the 2017 ASSA Meetings, CESifo, IAEE and the EAERE Summer Conference. Finally, we want to especially acknowledge Greg Lander, Naim Jonathan Peress, and Natalie Karas for invaluable guidance as we worked through the institutional details of these markets. All remaining errors are our own.

1 Introduction

Natural gas has replaced coal as the predominant electricity generation resource in the United States and provided more than a third of the country's utility-scale electricity generation in 2016.[1] With half the burn-point emissions of other fossil fuels and physical properties that make it ideal for balancing out the intermittency of renewables, its share of the generation mix is only expected to increase in the immediate future.[2] With over 50% of its total generation already coming from gas, New England is at the leading edge of this transition and an ideal environment in which to explore issues that may arise from the growing interdependencies between gas and electricity markets.

In recent years, New England's wholesale natural gas and electricity markets have experienced severe, concurrent price spikes. During the months of extreme cold that marked the winter of 2013-14 (i.e., the "Polar Vortex"), for example, New England gas prices averaged $17.86 per MMBtu (million British Thermal Units) —almost four times the Henry Hub price—and reached a record high of $78/MMBtu on January 22, 2014.[3] These extreme price spikes have been commonly attributed to limited pipeline capacity serving New England, and this "scarce capacity" narrative has been used in recent proposals to expand natural gas pipeline capacity serving the region.[4]

Limited pipeline capacity is indeed partly responsible for these extreme prices. But we also find strong evidence that two firms that held significant shares of the contracts to flow gas on the Algonquin Gas Transmission Pipeline—one of the two major pipelines serving New England—regularly restricted capacity to the region by

[1] See EIA 2017b.
[2] See EIA 2017a.
[3] The natural gas price at Henry Hub in Lousiana is the benchmark price for natural gas traded in the United States and is considered reflective of the commodity value without transportation costs.
[4] See, for example, Rose *et al.* 2014 and ICF 2015.

scheduling deliveries without actually flowing gas. These unusual scheduling practices tied up capacity that, in a well-functioning market, should have been released, or would have otherwise made available, to other shippers. Instead, significant quantities of pipeline capacity went unutilized on many of the coldest days of the year, pushing up the price of gas.

While most shippers had little incentive to sacrifice revenue from gas sales by withholding capacity, the two firms observed to withhold capacity also own large portfolios of electric generation units located in the region, giving them an incentive to increase gas prices in order to *raise rivals' costs* (Salop & Scheffman, 1983). That is, by restricting sales of a necessary input to production for their downstream competitors in the wholesale electricity market, the capacity-withholding firms increased the quantity of electricity their largely non-gas units were called upon to generate and the price those units earned.

In this paper, we analyze three recent years of scheduling data on Algonquin for evidence of firms withholding pipeline capacity in this manner. We find clear patterns of withholding at a subset of delivery nodes operated by Avangrid and Eversource (henceforth referred to as Firm A and Firm B), the only two firms operating on the pipeline with substantial assets and operations in both the gas distribution market and the electricity generation market. Using a panel data regression model, we empirically demonstrate these nodes were disproportionately served by specialized types of contracts that allow firms to call for gas on demand and to make large adjustments, without notice, in the last few hours of the day, two necessary conditions for executing a penalty-free withholding strategy. Over our three year study period, aggregate withholding at these few nodes reduced the pipeline's effective capacity by approximately 50,000 MMBtu per day, on average. On 37 days, over 100,000 MMBtu of capacity—about 7% of the pipeline's total daily capacity and about 28% of the daily capacity that is typically used to supply gas-fired generators—was withheld at these nodes.

This behavior significantly impacted both natural gas and electricity prices in New England. We employ an instrumental variables model to estimate a counterfactual gas price series, finding that gas prices were $1.68/MMBtu (38%) higher on average during our entire study period and $3.82/MMBtu (68%) higher during the winters. We proceed to construct a simulation model of New England's wholesale electricity market and use our estimated counterfactual gas price series as an input to estimate the effect on the region's electricity market. We find that electricity prices were about $10/MWh (20%) higher on average over our study period due to capacity withholding. Our simulation predicts that underutilized pipeline capacity ultimately resulted in a transfer from New England electricity ratepayers to generators (and their fuel suppliers) of about $3.6 billion over the course of our study period, about half of which occurred during the particularly cold winter of 2013-14. While the studied behavior may have been within the two firms' contractual rights, the significant impacts in both the gas and electricity markets show the need to consider improvements to market design and regulation as these two energy markets become increasingly interlinked.

The remainder of this paper is structured as follows: Section 2 reviews the relevant literature on raising rivals' costs and similar market power scenarios in electricity markets. In section 3 we provide background on the three markets that comprise our institutional setting. Section 4 presents the theoretical framework that forms the basis for our empirical strategy. Our empirical analysis is broken into two parts: Section 5 presents our investigation into firms' patterns of capacity withholding and section 6 presents our estimations of gas and electricity market impacts. Section 7 concludes.

2 Literature Review

Market power refers to the ability of a firm (or group of firms) to raise and maintain price above the level that would prevail under competition. Like all network utilities, energy transportation infrastructure is characterized by large initial capital investments and spatial differences in supply and demand that create an environment susceptible to the exercise of market power. When transmission constraints bind, they effectively segment the network into a set of smaller markets wherein firms that don't own or control a significant share of total assets across the network may have significant local abilty to set prices (Borenstein *et al.* , 1995). So far, this situation has mostly been studied in the context of electricity markets and much less so for gas markets.

Network congestion fluctuates with demand, meaning markets may be highly concentrated at some times and highly competitive at others. Consequently, Borenstein *et al.* (1999) discourage applying traditional measures of market concentration such as the Herfindahl-Hirsch Index (HHI) to electricity markets. Instead, they suggest modeling energy markets to investigate whether firms employ strategic behavior in their production decisions. Borenstein & Bushnell (1999) employ this method to predict significant potential for market power at the outset of the deregulation of the California electricity market and again to empirically verify ex-post the majority contribution of market power to California's extremely costly 2000 energy crisis (Borenstein *et al.* , 2002).

The instance of market power discussed in this paper stems from the contracts that serve as property rights to natural gas transportation capacity, which are in many ways analogous to transmission rights in electricity markets. Joskow & Tirole (2000) analyze the interaction between transmission rights and market power in electricity markets. They present a model of a two-node grid, where an upstream node with many competitive, low-cost generators is separated by a single transmission line from a downstream node where a single firm controls more expensive generation resources. Different marginal costs lead the independent system operator (ISO) to pay different prices at each node, which enhances efficiency in ideal conditions but also introduces the possibility of gaming the system. In this setting, if the downstream generator obtains physical transmission rights (which allocate capacity for generators to use to transmit electricity at no additional cost), inefficiency may arise. Under some realistic conditions, the downstream generator finds it more profitable to use physical rights to withhold transmission capacity to increase the downstream node's price, leading to welfare losses due to productive inefficiency. Further, the downstream generator is incentivized to acquire all the physical rights so they can simultaneously decide transmission capacity to and production at the downstream node.

Joskow and Tirole's analysis provides an interesting parallel to our setting, where firms that own downstream electric generation and pipeline capacity rights are, under some conditions, incentivized to use those rights to tie up capacity. Interestingly, Joskow & Tirole (2000) advocate for adapting the capacity release regulations of the gas pipeline industry to the electricity market to mitigate the potential abuse of physical rights in this manner. However, our study clearly shows that capacity release rules as they stand are insufficient to overcome the incentives toward inefficiency that are created by physical transportation constraints.

Cremer & Laffont (2002) adapt Joskow and Tirole's two-node, two-producer electricity model to natural gas to show similar results, although their model is limited in the depth to which it incorporates the institutional differences of the gas market. A much more heavily studied area of market power in natural gas is the supply-side market concentration in the European gas market, which imports a majority of its gas from only three countries – Russia, Norway, and Algeria (see e.g., Lise & Hobbs

2009, Boots *et al.* 2003, Holz *et al.* 2008).

While the ability to influence prices emerges from the physical capacity constraint in our setting, the firm's primary incentive to withhold capacity comes from vertical integration across the gas and electricity markets. One commonly-studied concern in the literature on vertical market power is *foreclosure* (sometimes also termed *raising rivals costs*), wherein a vertically-integrated firm instructs its upstream entity to restrict sales of a necessary production input to its downstream entity's competitors to increase the prices and market share enjoyed by that arm of the firm (e.g., Hart *et al.* 1990, Ordover *et al.* 1990).

Adapting the concept of raising rivals costs specifically to energy markets, Hunger (2003) raises the concern that a merger between a gas company and an electricity generation firm may incentivize it to withhold gas from the generation market to raise the wholesale electricity price received by its generators. Withholding is profitable if its impact on the firm's revenues in the electricity market, determined by the level of generation capacity and the elasticity of the generation supply curve, exceeds the opportunity cost of not selling the gas to other generators. Vazquez *et al.* (2006) expands on this opportunity to exert market power in the context of examining a real-world merger in Spain between a dominant natural gas firm and an electricity firm with a large quantity of gas-fired generation resources. In their model, a monopolistic gas producer restricts output beyond the level required to capture monopolistic rents in the power market in order to increase the wholesale electricity price and the revenues of their generators in that market. In this paper, we expand the theory developed by Vazquez *et al.* (2006) and Hunger (2003) by integrating a careful consideration of the role of transmission constraints and rights to capacity, adapted from the literature on market power in electricity markets, and empirically identify a real-world example of this scheme at play in New England.

3 Background: Three Interconnected Markets

3.1 The market for natural gas transportation

The modern US market for natural gas transportation was established through a series of reforms implemented by the Federal Energy Regulatory Commission (FERC) in the 1980s and 90s. These reforms effectively separated the gas transportation market from the physical commodity market by requiring interstate pipeline companies to sell their transportation services through long-term contracts for pipeline capacity.[5] Under this regime, capacity purchasers enter into multi-year contracts with a pipeline at FERC-regulated rates, which are designed to allow pipeline companies to earn a "just and reasonable" rate of return on their investment (FERC, 2017). Local gas distribution companies (gas utilities or LDCs which in turn provide gas to retail residential, commercial and industrial customers) have tended to be the largest subscribers of pipeline capacity, procuring sufficient contracts to meet retail customer demand.[6] As regulated utilities that are able to pass procurement costs through to their ratepayers, LDCs assume little commercial risk associated with entering into these long-term contracts. Other purchasers of long term capacity contracts include industrial facilities and gas marketers (taking speculative positions). Gas-fired electric generators, which represent an increasing fraction of wholesale gas demand, have tended not to purchase long term contracts , because the cyclicality of their demand—both daily and seasonal

[5] Previously, interstate pipeline companies would buy gas from producers at the wellhead, transport it to centers of demand, and sell it for a single price incorporating both their cost of the gas itself and their cost plus allowed profit from transportation. Following the reforms, the pipeline companies do not take ownership of the gas at any point. For a detailed history on the restructuring of the natural gas transportation market, see Oliver & Mason (2018).

[6] In addition, when the pipelines were converted from merchant to transportation-only entities, their firm sales contracts, which were almost exclusively with LDCs, were converted to firm transportation contracts. Outside of rare situations like those in Arizona, Florida, and Louisiana, in particular, there were almost no end-users or electricity generators with preexisiting firm sales contracts with interstate pipelines to be converted to firm transportation as part of industry restructuring.

depending on conditions in the wholesale electricity market—has made procuring long-term pipeline capacity contracts cost-prohibitive.

In 2008, FERC amended the rules governing long-term contracts to allow contract holders to sell temporary use of their pipeline capacity at unregulated prices on a secondary "capacity release market."[7] Additionally, pipeline operators may use the capacity release market to allocate unreserved capacity or sell unreserved capacity on an interruptible basis at an associated volumetric rate up to a maximum FERC-set rate. These policies are designed to promote a more liquid market for gas transport and to allow pipeline operators to efficiently allocate scarce capacity. Secondary capacity release sales can last anywhere from several hours to a year. Most on Algonquin fall in the range of a few days to a few weeks.

Firms that hold capacity rights on a pipeline are known as "shippers."[8] They exercise their capacity rights by electronically submitting "nominations" to the pipeline company on a daily basis. Nominations consist of an intake "receipt" point, an outflow "delivery" point, and a scheduled daily quantity of gas to flow.[9] This quantity must be flowed at a roughly even rate over the course of the 24-hour gas day, which runs from 9 a.m. til 9 a.m. Central Time the following day.[10] To induce shippers to judiciously manage nominations and flows, differences between scheduled nominations and actual flows incur imbalance penalties upon the shipper, which can be more or less severe

[7] While the capacity release market had been in existence since 1997, it was not until 2008 that capacity with a duration of a year or less could be sold into the market at an uncapped price.

[8] In many cases, the "shipper" that manages gas and capacity purchases is a separate subsidiary arm of their parent company. For simplicity, we refer to an LDC itself, its parent energy firm, and its shipping arm interchangeably throughout the paper.

[9] Net a small percentage that is skimmed by the pipeline operator to power compression stations.

[10] The precise rule for most contracts is that the gas must be flowed over a period lasting from 16 to 24 hours. For example, a capacity owner on Algonquin holding 24,000 MMBtu would generally be entitled to flow between 1,000 MMBtus and 1,500 MMBtus per hour as specifically set forth in their service agreement. Note that currently, the only service agreements providing for an associated 6% hour (like the 1,500 MMBtus per hour in the example) were those that were converted from sales agreements during restructuring.

depending on the size and nature of the infraction. Imbalance penalties are generally more severe when there is less slack available in the system to compensate, as in the winter.[11]

Shippers are able to make adjustments to their nominations during the gas day. FERC requires pipelines to offer a minimum of three "intraday" scheduling cycles, though some pipelines (including Algonquin) offer more frequent scheduling opportunities. On Algonquin and a few other lines, (like Transcontinental Gas Pipeline serving the east coast and New York), the last intraday cycle generally occurs a few hours before the end of the gas day and is commonly know as the "clean up" cycle. During this cycle, shippers match their scheduled nominations to their actual flows.

In addition to firm contracts, some pipelines (including Algonquin) offer "no-notice" contracts, which are a form of legacy contract generally only available to LDCs. On Algonquin, a no-notice contract is tied to storage capacity and service on Texas Eastern Transmission Company (which is owned and operated by same parent company as Algonquin). Together the storage service and no-notice contract allow an LDC to adjust its scheduled flows without prior notice, and to flow their total scheduled quantity of gas on an uneven hourly basis and over a period of less than a full 24 hours. FERC allows these contracts on the basis that they are necessary in order for LDCs to reliably serve their retail customers.

3.2 The wholesale natural gas market

In New England, gas is typically traded without the benefit of an exchange where bids and asks can be matched and settled for market-wide price determination. Instead, buyers and sellers must search for willing and able counter-parties and negotiate prices. This is partly because sales on New England's wholesale spot market frequently involve

[11] See Appendix A.2 for further detail on imbalance penalties.

delivery to a specific pipeline node, effectively re-bundling the physical commodity with the transportation service.[12]

Gas-fired generators purchase the vast majority of their gas on the wholesale spot market, on a “delivered to their location” basis, because their energy needs are typically much more variable and less predictable than those of LDCs and therefore not well served by long-term contracts. Most LDCs are both consumers and marketers of gas. Because they must hold sufficient long-term contracts to reliably supply their gas heating rate-paying customers, they find themselves, on all but the coldest winter days, with excess capacity rights. That excess capacity can be used either to ship gas to the region and sell it on the spot market or can be sold directly on the capacity release market.[13] Independent marketers do not themselves use gas, but instead hold long-term contracts in anticipation of profiting from short-run sales to the other firm types (primarily generators).[14]

Gas utilities (LDCs) are typically regulated monopolies. By regulatory design, they are allowed to make a fixed rate of return on their shareholder’s capital investments. Ratepayers finance the LDCs’ purchases of capacity rights which protect against gas price shocks. Hence, LDCs are subject to rules that limit their ability to profit

[12] Spot market prices therefore incorporate the wellhead price of the gas, the cost of the pipeline capacity needed to transport it, and the shadow price of the pipeline capacity constraint, which captures the difference in prices between the receipt and delivery regions due to differences in available supply when the pipeline is at full capacity (Cremer *et al.* , 2003). Because the prices of primary contracts for capacity are regulated but capacity-release and spot-market prices are not, the owners of capacity are able to extract congestion rents when capacity is scarce (Oliver *et al.* , 2014).

[13] Indeed, FERC regulations intended to ensure pipelines are fully utilized require them to sell any capacity (that is not scheduled by firm shippers) to interruptible contract shippers requesting access to that capacity by means of nominations to use it. In Section 5, we show how LDCs circumvent the intent of this rule in New England by initially nominating more capacity than they intend to use and then adjusting their scheduled quantity downward at the end of the gas day.

[14] Another set of participants in the spot market for gas are asset managers who act as third party agents and/or principles (depending on the Asset Management Agreement’s terms) for for contract holders. These independent marketers hold long-term contracts either as principal directly with the pipeline; or as replacement shipper under a long-term capacity release transaction.

from their excess contracts. These "revenue-sharing" rules are set by public utility commissions and vary across states.[15] In general, they require LDCs to return a certain percentage of revenues from capacity release and spot market sales (sometimes referred to as "non-firm margin" sales) to their ratepayers.

For the regulator, choosing an appropriate revenue-sharing rule is a balancing act between protecting ratepayers and allowing the LDC to keep enough profit such that they are incentivized to transact their excess capacity efficiently. It is generally held that LDCs require little incentive to efficiently market their excess capacity. If the incremental cost of marketing an additional unit of excess capacity is small, then the LDC's share of the profit from marketing that additional unit of excess capacity can be similarly small without significantly distorting the LDCs behavior. However, this reasoning fails to consider the incentives of firms that earn profit in other interconnected markets. In particular, for firms that operate in natural gas supply and delivery markets as well as electricity markets, shrinking the incentive to efficiently market excess capacity could have the unintended consequence of increasing the relative weight those firms assign to profits earned in the market where their incentives are not diminished. Indeed, it is likely that revenue-sharing rules in New England have contributed to the extent, and location, of capacity withholding. In Connecticut, for example, the state where most of the capacity-withholding behavior we observe takes place, LDCs must return 99% of non-firm margin sales to ratepayers, while in Massachusetts the revenue-sharing rule distributes 90% to ratepayers, and in Rhode Island the rule is slightly more complicated but works out to about 83%.[16]

[15] Revenue-sharing rules sometimes also vary across firms within states, though this is not the case for the three New England states that are the focus of this analysis.

[16] Specifically, Rhode Island's Public Utilities Commission requires the single LDC that operates in the state to return 100% of the first $1 million of all non-firm margin sales to ratepayers and 80% of all additional revenues.

3.3 The wholesale electricity market

In 1999, New England became one of the first regions to implement a competitive wholesale electricity market. Under the current structure, energy (i.e., the electricity commodity) is traded in the *day-ahead* market, which is operated by the Independent System Operator of New England (ISO-NE).[17] The market takes the form of a first-price auction, in which generators bid in quantities of energy they will supply at given prices for each hour of the day. In an idealized setting, competition incentivizes generators to bid in their true marginal costs of production, allowing ISO-NE to utilize the lowest-cost combination of generation resources to use to meet demand.[18] As illustrated in Figure 1, the most expensive unit of generation required to meet demand (which is almost perfectly inelastic in the short run) sets the wholesale electricity price paid to all generators called upon operate.[19]

While virtually all generators competitively bid into ISO-NE's day-ahead market,[20] the way in which they earn profits depends on whether they are considered to be

[17] ISO-NE additionally operates a *real-time* electricity market that balances short-term fluctuations in supply and demand, a forward capacity market, markets for transmission rights, and markets for system services such as regulation and reserves. Although many of these markets are affected by gas pipeline underutilization, we focus our analysis on the day-ahead energy market because it is the most significant in terms of trading volumes and generally sets price expectations for other markets.

[18] In practice, the first-price auction structure and local market power due to transmission constraints give generators some ability to mark up their bids above marginal cost (Kim, 2016). We do not address markup in this paper as it is not the focus of our analysis.

[19] This market-clearing price corresponds to the cost of the energy itself, which ISO-NE terms the "energy price." Line losses, transmission constraints, and spatially heterogeneous demand imply different values of energy in different areas of the grid, which ISO-NE accounts for by adjusting the energy price up or down at various nodes in the network to construct "locational marginal prices" (LMPs). LMPs are typically within a few cents to a few dollars of the energy price. Reserve requirements, heterogeneous ramping rates for various generation technologies, and other physical properties of the system further complicate the true dispatch procedure employed by ISO-NE. Throughout this paper, we model the day-ahead market in a simplified setting without these considerations, which successfully captures the relevant dynamics between the wholesale gas and electricity markets without getting weighed down in detail.

[20] A few hundred small-scale solar, wind, run-of-river hydro, and landfill gas facilities representing less than 1% of New England's total capacity do not participate in the day-ahead market and instead sell energy directly to utilities or commercial customers through bilateral contracts.

regulated or *merchant unregulated* assets. Merchant unregulated generators are typically owned by independent power generation firms that do not operate any transmission lines or distribution services.[21] These generators, which represent about 85% of New England's total capacity, pay all of their own fuel and capital costs and retain all of their own revenues.[22]

The regulated generation assets that make up the other 15% are typically owned by electricity distribution utilities. In some cases, regulated plants exist because they provide reliability services that are not economical under the current market structure but which are necessary to provide reliable service (i.e. peaking plants in urban load pockets). In other cases, they are holdovers from the regulated environment two decades prior. Instead of profiting based on the outcomes of the wholesale electricity market, the electric utilities that own these generation resources are entitled to make a fixed rate-of-return on the capital investment that goes into them, much in the same way they profit from their distribution assets.

New England is heavily reliant on gas-fired electric generation, with gas supplying about half of all electricity generated in the region. As shown in Figure 1, natural gas occupies the middle portion of the bid supply curve, and consequently a gas-fired plant is the marginal generator about three-fourths of the time. Accordingly, higher wholesale gas prices usually imply higher electricity prices, and this effect is amplified at higher prices due to the convexity of the bid-supply curve.

As will be discussed in detail in Section 5, the gas LDCs engaged in capacity withholding are owned by parent energy firms that also own hundreds of megawatts

[21] One notable exception is 76.5 MW of unregulated wind generation owned by one of the two firms observed to withhold gas pipeline capacity on Algonquin.

[22] This accounting is admittedly more complex for generators that enter into bilateral contracts with utilities and commercial customers; however, because the prices in these contracts are set by expectations of wholesale market prices, we are able to focus on day-ahead market outcomes without loss of generality.

of regulated generation capacity in New England, and one of the two additionally owns about 75 megawatts of unregulated capacity. These parent firms are primarily gas and electricity distribution utilities, but each have meaningful margins in the wholesale electricity market. As the markets for gas and electricity become increasingly interdependent, new opportunities emerge for firms with operating arms in both markets to exert market power. In the next section, we formalize the relationships between the three markets described here in order to demonstrate how a vertically-integrated firm may benefit by using its contracts for capacity to reduce availability on the pipeline rather than to actually transport gas.

4 Theoretical framework

In this section we present a graphical discussion of the incentives to withhold pipeline capacity.[23] The central point developed in this section is that firms that operate in both the pipeline transportation market and the electricity market have a clear incentive to restrict gas deliveries during periods of scarcity (raising the wholesale gas price), in order to capture rents (and raise rivals costs) in the electricity market. The incentive for these firms to withhold gas, rather than sell it in the wholesale market, is amplified for gas LDCs by revenue-sharing rules that require gas utilities to dividend most of the profit from non-firm-margin sales back to their ratepayers. We consider how state-level variation in these rules impacts which nodes firms use to withhold supply. Finally, we consider how the spatial nature of the pipeline network contributes to where in the system firms are likely to withhold capacity, noting that firms will have a stronger incentive to withhold capacity at points where it is likely to have the greatest impact on the wholesale electricity price.

[23] An algebraic presentation is relegated to the Appendix.

Based on the institutional features discussed in Section 3, suppose firms may operate in three vertically related markets: Furthest upstream is the market for natural gas pipeline transportation, which is used to deliver gas to a wholesale gas market, in which LDCs sell gas to electric generators. Second is the wholesale gas market (in which LDCs sell gas after serving retail gas demand). Third is the wholesale electricity market in which some generating units are gas-fired and others are not.

Our firm of interest operates as a seller in both the wholesale electricity market as well as in the wholesale and retail gas markets. It owns electric generating units (for simplicity we can assume non-gas fired). Through its LDC operating arm, the firm also holds pipeline capacity – which, as with the other LDCs – is a source of market power in the wholesale gas market during periods of scarcity. Similar to the other LDCs, this firm is required to first serve demand in the retail gas market, and may after that sell any excess pipeline capacity in the wholesale gas market under the same regulated rates of return and profit sharing rules, respectively. Unlike the other LDCs on the system, this firm's incentives derive from the interaction of its positions in the electricity and gas markets.

We start by considering the impact of a reduction in total deliveries in the gas transportation market upon the wholesale gas market. We interpret this reduction in flows as arising because one firm chooses to overschedule, *i.e.* it schedules for larger deliveries in the day ahead timely cycle than are ultimately executed the next day. From the firm's perspective, the important consideration for assessing the impact on the wholesale gas price is the degree of total excess pipeline capacity prior to overscheduling.

Figure 2 illustrates the interaction between the pipeline transport and wholesale gas markets. We model the supply curve of gas transport as Leontief in nature: marginal costs are constant (reflecting a constant per-unit commodity price and a constant cost of transporting gas along the pipeline) so long as scheduled deliveries fall below total pipeline capacity. The wholesale price of natural gas will then equal marginal cost. However, when scheduled deliveries rise to the level of maximum available capacity there is no possibility to increase transportation capacity at any cost in the short-run, as illustrated by the vertical turn in the supply curve. In this situation, the wholesale price of gas will be determined by the level of demand (rather than the marginal cost), i.e., by how much potential buyers are willing to pay for those last units of capacity, and thus make it possible for sellers of pipeline capacity to extract scarcity rents.

When the quantity demanded at the (constant) marginal cost is less than maximum available capacity so that there is a lot of excess pipeline capacity, as in the left panel of Figure 2, overscheduling will have little impact upon the wholesale price of natural gas. But when the demand curve shifts so that the quantity demanded at marginal cost is close to the maximum available capacity and the pipeline comes closer to its capacity constraint, as in the right panel of Figure 2, overscheduling can induce an increase in the wholesale price of natural gas; the magnitude of this increase in price depends on the level of demand compared to available capacity (i.e., excess capacity) and the magnitude of overscheduling. The smaller is the initial level of excess capacity (*e.g.,* as a result of large retail demand for gas – perhaps as a result of colder than expected temperatures) the easier it is to force wholesale gas prices up.

Any increases in natural gas wholesale prices will have a derivative effect upon electricity markets, as illustrated in Figure 3. Two factors are in play here: the initial interaction of supply and demand in the electricity market, and the degree to which increased natural gas wholesale prices shift the supply curve for electricity. We model the demand for electricity as perfectly inelastic, with wholesale electricity prices reflecting the marginal cost of supplying electricity at the market-clearing quantity (as depicted in the left panel). The industry aggregate marginal cost, in turn, reflects the incremental cost of the marginal producer. This marginal cost curve rises slowly over a significant range of quantities. As electricity sales rise, ever-less efficient (and therefore more expensive) sources of electricity supply are called on, and so the marginal cost curve becomes steeper and steeper.

Importantly, many of these more costly sources of electric generation are gas-fired. Increases in gas wholesale prices are likely to cause an upward shift in the electricity supply curve, particularly at higher quantities. If demand intersects supply in the elastic region, or if the overscheduling in the gas transport market does not induce much of an increase in wholesale natural gas prices, then the impact on electricity prices is inconsequential (left panel). But if demand intersects supply in the less elastic portion and wholesale gas prices rise significantly as a result of overscheduling in the gas transport market, then there will be a noticeable increase in electricity prices (right panel).

Of course, any overscheduling that leads to unused pipeline capacity means that there are foregone profits from not supplying the wholesale gas market. This effect is illustrated as the light rectangle in the left panel of Figure 4. These profits are only partly foregone by the overscheduling LDC, who is subject to a revenue sharing rule: most of these foregone earnings would have had to be returned to the LDC gas customers. As such, the ultimate amount of foregone profits for the firm's shareholders is substantially smaller, as illustrated by the darker (small) rectangle in the left panel. By contrast, any increases in profits in a separate market need not be subject to this regulatory effect. In particular, if the firm has holdings in the wholesale electricity market then any extra profits that arise therein as a result of the higher wholesale gas prices are retained; this effect is illustrated in the right panel of Figure 4.[24]

[24] The scenario illustrated in this Figure implicitly assumes the firm is not subject to cost of service regulation in the electricity market. In the markets we analyze empirically in the next section, one firm is subject to rate-of-return regulation also on its electricity generation holdings. But

The model sketched above articulates the differentially large incentive a firm that holds positions in both the electricity and gas markets can have to influence the price of wholesale natural gas, especially in the presence of a revenue sharing rule. We also wish to explore the spatial nature of incentives when these integrated firms are located at different points along the pipeline. To this end, we imagine two delivery points (nodes), one upstream and one downstream. There are retail gas customers located at both nodes; we presume the market downstream is larger than the market upstream (*e.g.,* on the Algonquin pipeline, which flows from South to North, the upstream node might be Hartford and the downstream node Boston). As above, the LDCs are obliged to meet all retail gas demand at that node. One important distinction to the model above is that a firm located at the downstream node has the right to sell gas in either the upstream or the downstream wholesale gas market, whereas a firm located at the upstream node can only sell gas in the upstream market.[25]

With minor adaptation, the raising rivals' costs arguments described above can be applied here. The key point is that the incentive to raise costs by influencing the price of delivered gas is larger for a firm located upstream than downstream, for two reasons: first, because there is necessarily less spare pipeline capacity upstream it takes less withholding to engender any particular level of increase in delivered price.

then there is a separate, and subtle, motive for pushing up electricity prices. If this rate-of-return regulation requires a minimum level of operation of that generation capacity and the firm in question holds high-cost electricity generation capacity there is a risk that it will not be able to meet this stipulation. By forcing up electricity prices, the firm may be able to convert its high-cost units into economically viable units, thereby accessing a revenue stream that might otherwise be unavailable. This problem will be particularly acute if the firm has established costly capital, such as expensive scrubbers, only to find these units priced out of the market at most points in time.

[25] One can think of these firms holding contracts for delivery, with one firm holding a contract guaranteeing delivery to the upstream node, and another firm holding a contract allowing delivery downstream or any other points further upstream. With this interpretation, the latter could choose to withdraw gas at either point, while the former firm would be obliged to remove gas only upstream. This case is most operative when the system approaches scheduling near to full contractual entitlements. At these times, restrictions to receipts and deliveries along primary paths cause the delivery rights of the contracted capacity to the upstream location to be the terminus of their firm rights.

Second, because of the additional pipeline tariff a firm must pay to utilize the segment between the upstream and downstream points, the upstream firm has a natural cost advantage over the downstream firm. As we noted above, any such cost advantage is the root source of motives to influence markets by raising input prices (and thereby raising rivals' costs).

The final point we wish to make relates to demand shocks, as might arise in inclement weather conditions. In weather conditions that raise demand, we expect to see a larger impact in the downstream market than in the (comparatively smaller) upstream market. If these shocks are anticipated when flows are scheduled, as seems likely, then the disproportionate increase in downstream demand will have spillover effects in the upstream market, because an increase in scheduled deliveries downstream also raises the amount of gas shipped to or through the upstream point they must reduce spare capacity upstream. In essence, the increased downstream demand creates conditions where it is easier to hold upstream markets hostage – and where it takes less intervention to force wholesale gas prices up through overscheduling.

5 Detecting Capacity Withholding

In this section, we show empirically that capacity withholding occurred on the Algonquin pipeline, and that it resulted in significant levels of pipeline capacity going unused in aggregate. Moreover, withholding occurred at only a small subset of LDC-designated delivery nodes operated by just two parent energy companies that also own significant generation capacity in the region. These nodes are primarily located in Connecticut, where revenue-sharing rules are least generous to LDCs, and they are disproportionately served by no-notice contracts that allow shippers to make large schedule adjustments without incurring imbalance penalties.

We detect capacity withholding by analyzing hourly-level scheduling data for all 117 delivery nodes on the Algonquin pipeline for every day in our three-year study period.[26] This reveals a unique pattern exhibited by a handful of nodes wherein shippers consistently reserve more capacity than they use to actually flow gas. These shippers avoid incurring imbalance penalties designed to prevent this type of overscheduling by reducing their scheduled quantity at the last moment, such that their final scheduled quantity matches what they actually flowed.

It is striking to observe these patterns at all, but we note that withholding increases in both magnitude and variance in the winter months, when greater demand for gas for both heating and generation increases the value of the fixed stock of pipeline capacity. As illustrated in the previous section, revenue-sharing regulations limit LDCs' ability to extract these congestion rents through the gas transportation market or the wholesale gas market, thereby making the less-efficient extraction pathway—raising the wholesale electricity price—comparatively more attractive. While the available data are insufficient to conclusively determine whether this is indeed the motivation of the withholding firms, we demonstrate that the observed patterns of withholding are consistent with this explanation and inconsistent with several other possibilities.

5.1 Analysis of Scheduling Patterns

Each contract for capacity gives the shipper holding it the right to use a certain amount of space along the pipeline between one or more specifically listed *receipt* (intake) nodes and one or more specifically listed *delivery* (outflow) nodes. To actually

[26] Hourly scheduled quantities for all nodes going back three years are publicly available through the Algonquin pipeline's FERC-mandated electronic bulletin board, generating about seven million node-hour observations. Note that we only observe scheduled quantities; actual flows are known only to the pipeline company and individual nodal operators.

exercise this right, the shipper must electronically submit a *nomination* to the pipeline company, which states the quantity of gas they intend to move, where it will enter the pipeline, and where it will exit. This capacity scheduling process is carried out on a daily basis for each gas day, which runs from 9 a.m. til 9 a.m. the following day.[27] Importantly, capacity is nominated not as a rate of flow, but rather as a total quantity to be transported over the course of the gas day at a roughly constant rate.

Shippers must submit their initial nominations by the close of the *timely cycle*, which occurs at 1 p.m. the day before the gas day, in order to be guaranteed the capacity provided by their contracts.[28, 29] In contrast to the majority of interstate pipelines, Algonquin allows shippers to adjust their nominations on an hourly basis over the 44-hour scheduling period, which begins with the timely cycle at 1 p.m. the day before and concludes at 9 a.m. at the end of the gas day.[30] While schedule changes are observed at all hours, the vast majority of adjustments are made within three specific windows: In the eight hours following the initial nomination at the timely cycle; between 6 a.m. and 6 p.m. encompassing the start the gas day; and between 6 a.m. and 8 a.m., just before the end of the gas day.

Figure 5 illustrates the scheduling pattern at a typical LDC-designated node, which

[27] The gas day and all associated scheduling times are in Central Clock Time for all interstate pipelines to facilitate harmonization of the gas transportation industry across the US.

[28] If shippers neglect to nominate by the timely cycle, the pipeline may use their contracted capacity to allow other shippers to move gas to other points on the pipeline that are not their specifically listed nodes. These nominations, which are called *secondary in-path*, *secondary out-of-path*, and *interruptible* nominations, are used by LDCs and independent marketers to sell gas on the wholesale market to generators located at other parts of the pipeline. On a day when the pipeline is fully scheduled, the shipper that did not nominate on time will only be able to utilize pipeline capacity if another shipper will adjust its nomination downward later in the scheduling period to free up some capacity.

[29] *No notice* contracts are exempt from this requirement: For the most part, provided a non-zero nomination is submitted in the Timely cycle, these contracts enable scheduled quantities to be adjusted at any time during the scheduling period with guaranteed approval.

[30] Most other pipelines allow schedule adjustments at just five specific times following the initial timely cycle nomination: A *late cycle* in the evening the day before the gas day, three *intraday cycles* during the gas day, and a final *cleanup cycle* near the end of the gas day.

is characterized by frequent, relatively large adjustments in either direction in either of the earlier two scheduling windows and less frequent, relatively small adjustments in the final window. Each delivery node's pattern is unique, but the vast majority of LDC-designated nodes can be broadly characterized by this description.

The substantial, bidirectional schedule changes made in the first two common time windows are consistent with LDCs getting better information about their expected retail demand, generators getting better information about electricity market conditions, and gas being traded on the wholesale market, either directly or through independent marketers, to efficiently allocate capacity between these two firms types.[31] On a cold day when pipeline capacity is fully scheduled, a downward adjustment at one node in either of these first two windows is therefore accompanied by an increase at another node, and the pipeline remains fully scheduled. In contrast, the schedule adjustments made in the window between 6 a.m. and 8 a.m. at the end of the gas day represent shippers matching the node's final scheduled daily nomination to the quantity of gas that was actually delivered to it, in what the industry refers to as the *clean up* or *true up* cycle. Beyond accurate bookkeeping, this adjustment is necessary for shippers to avoid the accounting imbalance penalties assessed for monthly deviations between scheduled and actual flows in excess of 5%.[32] The relatively small, less frequent, bi-directional end-of-day adjustments observed in Figure 5 are consistent with an LDC shipper that nominates capacity with the intent to use their entire nomination to transport gas to customers. In this case, the adjustments reflect only differences between their prediction of retail customer demand and realized demand, which will be minimal given previous opportunities for adjustment if predictions are accurate in expectation.

[31] For simplicity, we disregard the two industrial end users directly connected to Algonquin here as they account for less than 1% of the market.

[32] See Appendix A.2 for further detail on imbalance penalties.

We focus our attention on ten LDC-designated delivery nodes that do not exhibit the same pattern, and instead are observed to make large, consistently negative schedule adjustments in the final three hours of the gas day. Figure 6 illustrates an example of this distinct scheduling pattern, which is exhibited very prominently at six nodes that are clear outliers, and to a lesser extent at four additional nodes.

These large, consistently negative adjustments just before the end of the gas day are consistent with an LDC shipper that intentionally nominates capacity in excess of its predictions of its customers' total daily demand. In this case, the adjustments incorporate the differences between predicted and realized demand *plus* the offloading of overscheduled capacity to avoid imbalance penalties. Capacity cannot be double booked, and the pipeline company manages nominations such that the scheduled flows through any point on the pipeline do not exceed safe operating limits. Thus, a negative adjustment in the clean up cycle indicates capacity that was scheduled but not utilized to ship gas to that node. When aggregate nominations reach the pipeline's capacity constraint and the negative schedule adjustment is not accompanied by a positive adjustment at another node, the negative adjustment in the cleanup cycle corresponds to capacity that went unused across the entire system for that gas day.

The six nodes we clearly observe making consistent, substantial negative adjustments in the final hours of the gas day (hereafter referred to as *downscheduling*) are all operated by shippers owned by just 2 out of the 27 parent energy firms that operate delivery locations, which we will refer to as Firm A and Firm B. To establish that the scheduling patterns at these nodes systematically differ from the rest of the distribution, we plot each node's average schedule changes over each hour of the 44-hour scheduling period in Figure 8. We note also that following these six nodes, the four nodes with the next largest negative schedule adjustments in the final hours of the gas day are also operated by Firms A and B, suggesting they may be involved as well (see Table 1). Across the entire system, aggregate schedule adjustments in the final three hours of the gas day averaged -48,493 MMBtu over our three-year study period and -51,152 MMBtu in the winters. On 37 days in the study period, the aggregate adjustment exceeded -100,000 MMBtu, which is roughly 7% of the pipeline's total capacity[33] and roughly 28% of the total supply to electricity generators connected to the Algonquin pipeline.

Examining scheduling patterns separately for the winter season and the rest of the year reveals a behavioral difference between these two firms: Firm A consistently downschedules in both the summer and winter seasons, while Firm B engages in downscheduling primarily during the winter season.[34] In general, the suspect Firm A nodes appear to be exhibiting a blanket policy of always withholding excess capacity to the extent their contract holdings enable them to do so, whereas the suspect Firm B nodes appear to "turn on" this policy at some point toward the beginning of the winter season and then turn it off again sometime in the spring. This difference could be interpreted as suggestive evidence that the incentives for Firm A to increase electricity prices are stronger than those for Firm B.[35] However, we refrain from placing much weight on this interpretation, because the decision to withhold or release excess capacity is relatively inconsequential in the summer. Blocking capacity will not affect gas and electricity prices when the pipeline is not fully scheduled, and capacity rights cannot be used to extract scarcity rents through the capacity release or wholesale gas spot market on warm days when the pipeline is uncongested. In contrast, as explained in the theory section, this behavior will raise gas and electricity prices when the pipeline is constrained, and when scarcity rents are available, revenue-sharing

[33] Measured at the Stony Point compression station, which is the most frequent bottleneck for deliveries to New England.

[34] We define winter as between December 1 and March 31 following the delineation used by Algonquin Gas Transmission (Spectra, 2016).

[35] We explore this possibility further in Section 5.4.

mechanisms will diminish capacity rights holders' ability to extract them through capacity release or wholesale gas market sales. We proceed to explore how these mechanisms relate to observed downscheduling in the following sections.

5.2 Geography of Suspect Nodes

Spatially, we observe that eight of the ten most frequently downscheduling nodes, including five of the six clear outliers, are located in close proximity to one another in Connecticut (see Figure 9). Importantly, this section of the pipeline is downstream of its primary bottleneck at the Stony Point compression station, meaning nominations to these locations can exclude others from delivering gas to New England when the pipeline is congested.[36] However, all nodes downstream of Stony Point, including those in Rhode Island and Massachusetts, have the same capability to influence the pipeline's effective capacity constraint. While there are some potential alternative explanations,[37] we believe downscheduling behavior occurs primarily at nodes in Connecticut because of the strength of the state's extra-marginal revenue sharing rules for LDCs.[38]

While it is relatively straightforward to see in the data that downscheduling primarily occurs at LDC-designated nodes in Connecticut, we formally test this

[36] In order to keep gas flowing at a high rate across long distances, interstate pipelines have compression stations every 50 to 100 miles that effectively break the pipeline into a series of segments. On the Algonquin pipeline, Stony Point is the compression station that is most frequently scheduled up to its operating capacity first, and it is located downstream of all of Algonquin's major Western receipt points.

[37] One potential alternative explanation is that there is significantly less generation capacity located in Connecticut, meaning there is less demand for natural gas in that state. This may combine with frictions for using contracts to deliver gas further downstream than its listed delivery nodes to make excess contracts delivering gas to Connecticut less valuable for capacity release or spot market sales and comparatively more valuable for capacity overscheduling (see Appendix A.3 for a more detailed explanation).

[38] Recall from Section 3.2 that LDCs in Connecticut are required to return 99% of non-firm margin (i.e. capacity release and gas spot market) sales to ratepayers, versus 90% in Massachusetts and a slightly more complicated rule that works out to about 83% in Rhode Island.

hypothesis using the following regression model:

$$D_{it} = \alpha_0 + \beta_1 LDC_{it} + \beta_2 CT_{it} + \beta_3 (LDC_{it} * CT_{it}) \\ + \alpha_1 HDD_{it} + \alpha_2 HDD_{it}^2 + \alpha_3 W_t + \lambda_t + \varepsilon_{it}$$

Downscheduling D_{it} is defined as node i's scheduled daily quantity at 6 a.m. on gas day t (three hours before the end of gas day) minus its scheduled daily quantity at 9 a.m. at end of that gas day. LDC_{it} and CT_{it} are binary indicators for whether node i is an LDC or located in Connecticut, respectively.[39] While neither of these variables change over the study period, we use node-days as the unit of analysis to allow for the inclusion of time-varying controls, and we adjust standard errors for clustering at the node level.[40] These controls include temperature in the form of heating degree days (HDD_{it} and HDD_{it}^2),[41] an indicator for whether the day is a weekend (W_t), and quarter fixed effects (λ_t) to capture seasonal variation common to all nodes.

Results are presented in Columns (1) and (2) of Table 3. We find that downscheduling is concentrated at LDC-designated nodes with significance at the 5% level. Column (2) demonstrates that downscheduling occurs primarily at such nodes located in Connecticut, as the coefficient on the interaction is an order of magnitude greater

[39] We elect to use a binary indicator for Connecticut for ease of interpretation, especially given our frequent use of this variable for interactions in subsequent regressions. Our results are robust to using a continuous variable for the revenue-sharing mechanism (see Table A1 in the Appendix).

[40] The variable we use for contracts in subsequent regressions does vary across nodes over time; however, given our sample size, the inclusion of node fixed effects diminishes our statistical power to the extent that we are unable to draw any meaningful conclusions. In essence, some nodes engage in downscheduling behavior and some do not, and the objective of this section is to empirically explore the factors that contribute to this behavior to the extent that the available data enable us to do so. We acknowledge that this is primarily a correlational analysis, and that while our findings provide supporting evidence for our hypotheses about the institutional conditions and firm incentives that drive capacity withholding, we do not causally identify these mechanisms.

[41] *Heating degree days* refers to the number of degrees below 65 Fahrenheit and is commonly used as a better proxy for heating demand than temperature. We average this variable across Connecticut, Rhode Island, and Massachusetts, using data from the National Climatic Data Center.

than the coefficient on either indicator variable. This is consistent with the hypothesis that revenue sharing rules being less generous to LDCs in Connecticut distort their incentives to use their excess pipeline capacity efficiently in the upstream market.[42]

5.3 Ability to Avoid Imbalance Penalties

As described in Section 3.1, shippers must pay accounting imbalance penalties to the pipeline company if their scheduled flows deviate from their actual flows in excess of 5% on a monthly basis.[43] Additionally, shippers must pay *Operational Flow Order* (OFO) imbalance penalties if they cause a physical imbalance in the system on days where the pipeline issues an OFO warning.[44] Overscheduling could potentially be executed under regular firm-service contracts. But to avoid imbalance penalties, the shipper would need to source gas from a supplier that is complicit in injecting an amount of gas into the pipeline that differs from what has been scheduled. Any storage provider or producer could potentially play this role,[45] but the on-demand storage service offered by the Texas Eastern pipeline provides a convenient tool for shippers using Algonquin to accomplish this without involving a third party.[46] This on-demand

[42] We acknowledge, of course, that any other characteristic of Connecticut could be driving this result. In particular, we believe one other significant contributing factor may be lower demand for natural gas for generation in Connecticut (discussed in detail in Appendix A.3). We have focused on differential revenue sharing rules for the moment because it is more straightforward, but the results of this subsection also pick up the effect of alternative mechanisms such as reduced generation capacity.

[43] This threshold gives shippers some room to block capacity on occasion without later making adjustments in the clean up cycle, but frequently withholding large shares of excess capacity would require the downscheduling behavior we observe.

[44] OFO warnings are extremely frequent on the Algonquin pipeline during the winter.

[45] It would be circuitous for a marketer to be involved here, as they would need their supplier to be complicit in injecting less gas than was scheduled.

[46] The second party is the pipeline operator: Both Texas Eastern and its storage service are operated by the same parent energy firm that operates the Algonquin pipeline. It is a near certainty the pipeline operator is aware of the scheduling practices on its pipelines that result in underutilized capacity. However, as pipeline companies make 99% of their revenues from fixed-charge payments associated with the quantity of reserved firm-service and only 1% on the use of these capacity contracts, their incentives lead them to favor constructing new pipeline capacity – to sell more

storage service does not automatically inject gas into the pipeline when nominations are made, but instead is used by the pipeline company to balance pressure across the entire system.

Furthermore, if the shipper is transporting gas under a no notice contract, the pipeline guarantees the shipper's ability to make changes to their schedule at any point during the scheduling period. Such an arrangement facilitates overscheduling by ensuring that any downscheduling adjustments in the final few hours will be automatically approved by the pipeline company. [47] As shown in Figures 10 and 11, both Firm A and Firm B hold sufficient *no notice* contracts and sufficient contracts originally sourcing gas from on-demand storage locations on Texas Eastern to engage in the levels of downscheduling we observe without incurring imbalance penalties.

Contract rights vary on a roughly quarterly basis on the Algonquin pipeline, allowing us to exploit temporal as well as spatial variation. Graphically, Figures 12-14 show a relationship between these two firms' holdings of *no notice* contracts sourcing gas from the Texas Eastern pipeline (NN from TE contracts) and downscheduling over time at the segment level. For most segments of the pipeline, there appears to be no relationship between downscheduling behavior and NN from TE contracts. However, for the segment between Cromwell and Chaplin, where two suspect nodes are located, aggregate downscheduling is of roughly the same order of magnitude as Firm A's and Firm B's holdings of NN from TE contracts delivering to that segment. Between Oxford and Cromwell, where six suspect nodes are located, their NN from TE contract

contracts – rather than ensuring existing capacity is fully utilized. In New England, this incentive took the concrete form of the proposed Access Northeast pipeline expansion project for which the pipeline operator, Firm B, and another New England gas utility are co-developers.

[47] A requirement that interstate pipeline companies offer no notice contracts was included in FERC Order 636, the policy that mandated the unbundling of gas transportation service from the physical commodity, at the request of LDCs, who argued that continuation of their no notice service would also be needed in the new market structure to ensure they could reliably serve unexpected fluctuations in demand.

holdings appear to be an approximate upper bound on the level of downscheduling that occurs. We examine these apparent trends in greater detail using a set of specifications that introduce NN from TE contracts as an independent variable, the most inclusive of which is the following:

$$
\begin{aligned}
D_{it} = \alpha_0 + \beta_1 LDC_{it} + \beta_2 CT_{it} + \beta_3 C_{it} \\
&+ \beta_4(LDC_{it} * CT_{it}) + \beta_5(LDC_{it} * C_{it}) + \beta_6(CT_{it} * C_{it}) + \beta_7(LDC_{it} \times CT_{it} \times C_{it}) \\
&+ \alpha_1 HDD_{it} + \alpha_2 HDD_{it}^2 + \alpha_3 W_t + \lambda_t + \varepsilon_{it}
\end{aligned}
$$

Here, C_{it} is the total quantity of NN from TE contracts delivering gas to node i that are in effect on day t.[48] Column (3) of Table 3 indicates that NN from TE contract holdings are not necessarily a predictor of downscheduling across all LDC nodes, but Columns (4) and (5) demonstrate that they are an extremely strong predictor of downscheduling behavior at nodes located in Connecticut. In particular, for nodes in Connecticut, every additional MMBtu of NN from TE contracts is associated with an average increase in downscheduling of 0.42 MMBtu.[49]

[48] These data are available through "Index of Customers" reports, which interstate pipelines are required to make publicly accessible on their reporting web sites for the previous three years. As both contracts are reported separately for each the two pipelines, it impossible to confirm that a particular nomination or contract on Algonquin is used in combination with a particular contract on Texas Eastern. We therefore consider the "from TE" component of "NN from TE" to be proxy for a delivery right sourcing from storage from Texas Eastern, noting that both Firm A and Firm B hold sufficient *no notice* contracts on Texas Eastern sourcing gas from on-demand storage sites to engage in the levels of downscheduling we observe, as shown in Figure 11.

[49] Both *no notice* contracts and contracts sourcing gas from Texas Eastern are independently strong predictors of downscheduling behavior at nodes in Connecticut. Because these two contract characteristics overlap so heavily for Firms A and B, we are unable to empirically separate whether one or the other is critical to downscheduling.

5.4 Electricity Market Incentives

In Section 4, we outlined the incentives by which a firm owning generation capacity may benefit from higher electricity prices, and the connection between electricity prices and downscheduling in the gas transportation market. Here, we explore how generation capacity ownership relates to downscheduling behavior seen in the data.

Table 2 demonstrates that the only two firms that consistently engage in withholding behavior are also the two firms that hold the most and third-most generation capacity in New England among the LDCs served by the Algonquin pipeline. We note also that the firm holding the second most generation capacity holds no *no notice* contracts on Algonquin and an order of magnitude fewer contracts of any kind (including regular firm service contracts) than either Firm A or Firm B, meaning that while this firm benefits from higher electricity prices, its ability to affect the Algonquin pipeline's effective capacity constraint is very limited.

To bring electricity generation ownership into our empirical model, it is necessary to restrict our sample to LDCs only, as power plant nodes connected to Algonquin are of course operated by firms holding large quantities of generation capacity in the region. We therefore remove the LDC indicator from our previous specifications and replace it with a measure of generation capacity ownership. We run five separate specifications that mirror those of Table 3, the most inclusive of which is the following:

$$\begin{aligned}D_{it} = \alpha_0 &+ \beta_1 MW_{it} + \beta_2 CT_{it} + \beta_3 C_{it} \\ &+ \beta_4(MW_{it} \times CT_{it}) + \beta_5(MW_{it} \times C_{it}) + \beta_6(CT_{it} \times C_{it}) + \beta 7(MW_{it} \times CT_{it} \times C_{it}) \\ &+ \alpha_1 HDD_{it} + \alpha_2 HDD_{it}^2 + \alpha_3 W_t + \lambda_t + \varepsilon_{it}\end{aligned}$$

Here, MW_{it} is an indicator for whether the parent firm operating node i owns at least

100 megawatt (MW) of any type of generation capacity in the region.[50] Columns (1) and (2) of Table 4 demonstrate a significant correlation between owned generation capacity and downscheduling behavior. In particular, nodes operated by firms that own at least 100 MW of generation capacity in New England downschedule over 1000 MMBtu more than nodes that do not. We find that this relationship is again driven primarily by nodes in Connecticut.

Interacting contract holdings with generation capacity ownership and state allows us to investigate nodes that have both the ability and the incentive to withhold capacity. Column (5) of Table 4 demonstrates that such nodes are clear outliers from the rest of the distribution in terms of their schedule adjustments in the clean-up cycle. With a high degree of statistical confidence, we find that nodes operated by firms owning at least 100 MW of generation capacity downschedule about 40% on average of their total NN from TE contract holdings if they are located in Connecticut, but only

[50] We elect to use an indicator here to capture incentives created by ownership of both regulated and unregulated generation capacity. Our choice of 100 MW as the threshold is admittedly somewhat arbitrary, but we believe the reader will agree that it effectively separates firms with significant generation capacity in the region from those without (see Table 2). Beyond making the interpretation of interactions more difficult, using a continuous variable for the parent firm's generation capacity ownership is problematic for this set of regressions because the incentive pathways are quite different for different types of capacity. These pathways depend especially on whether the capacity is merchant unregulated or regulated, and additionally on many other characteristics, such as the fuel type, whether it is baseload or peaking, and the age of the facility. We demonstrate this with a robustness check wherein we perform the same set of regressions, but using the continuous variable for MW of generation owned rather than the binary one, producing generally erratic results (see Table A2 in the Appendix). We understand these counterintuitive results to be driven by the fact that Firm A downschedules more than Firm B, yet Firm B owns more generation capacity in New England, and we take this as suggestive evidence that, by merit of its directness, the incentive pathway for merchant unregulated generation is much stronger than the derivative pathways for regulated generation. Indeed, when we use a continuous measure of merchant unregulated MW owned as our independent variable of interest, our results are once again intuitively signed and highly significant (see Table A3 in the Appendix).

about 2% of their NN contract holdings if they are located outside of Connecticut.[51,52]

The available data enable us to detect downscheduling behavior and explore correlations with various mechanisms and incentives that may determine it. Our results are consistent with the hypothesis that these scheduling practices are representative of intentional capacity withholding intended to raise gas and electricity prices. However, we are unable to prove this conclusively without a source of plausibly exogenous variation to electricity market incentives.

One potential alternative explanation, for example, is that these two firms are exercising risk aversion by reserving an upper bound on the capacity they think they might need to ensure they will have access to it if demand turns out to be higher than expected. Although we cannot concisely reject this hypothesis given our data, we find it unlikely for two reasons: First, The two firms engaging in downscheduling are the only two firms that have a significant LDC presence on Algonquin and also have strong incentives to have higher electricity prices. Nodes operated by other firms appear to be able to consistently do a very good job of predicting the next day's demand. Second, by their very nature, the *no notice* contracts held by Firms A and

[51] To challenge our understanding that *no notice* contracts sourcing gas from Texas Eastern in particular are requisite for systematic withholding, we re-run this set of regressions using contracts of any type and point of origin in their place. Table A4 in the Appendix presents these results, which are characteristically similar but generally smaller in magnitude. For example, the coefficient on the triple interaction of contracts, capacity, and Connecticut is still significant at the 1% level, but about 2.5 times smaller than when using *no notice* contracts from Texas Eastern, suggesting that these contracts are indeed particularly useful for downscheduling. However, the available data are insufficient to eliminate the alternative possibility that by happenstance some other unobserved factor causes these two LDCs that operate in Connecticut and also own significant generation capacity to hold NN from TE contracts in large quantities.

[52] While the ten most-downscheduling nodes are all operated by Firm A or Firm B, several independent marketers also manage contracts delivering gas to these locations. (The nodal operator manages actual flows and is typically responsible for the majority of the node's deliveries, but it is not necessary to operate a node to make deliveries to it. Independent marketers in particular use contracts to deliver gas to the region but do not operate nodes on the Algonquin pipeline.) To confirm that Firms A and B are indeed responsible for the withholding we observe, we run this set of regressions first considering only contracts held by them and then considering only contracts held by other firms. The results, presented in Table A5, clearly demonstrate that contracts held by Firms A and B drive the downscheduling we observe.

B guarantee their ability to ramp up capacity usage at any point during the gas day, making it unnecessary for them to overbook capacity in this manner.

We therefore proceed to explore these firms' incentives further by modeling the effect of downscheduling behavior on gas and electricity prices, enabling us to compare their lost revenues in the gas market to their increased revenues in the electricity market, and estimate costs and distributional impacts in the process. Our methodology and results are presented in the next section.

6 Distributional and Welfare Effects

Next, we investigate the impact of downscheduling on wholesale gas prices and simulate the welfare effects on the wholesale electricity market, which market we believe bore the majority of the incidence of the downscheduling behavior. First, we use an instrumental variables approach to estimate the elasticity of demand for natural gas in the New England wholesale spot market, and use our estimated demand elasticity to construct a counterfactual Algonquin City Gate (ACG) price series.[53] Then, we use the observed and counterfactual gas prices to simulate the effects on economic dispatch in the New England wholesale electricity market, which we use to calculate the welfare and distributional impacts on generators and consumers.

6.1 The impact of downscheduling on natural gas prices

6.1.1 Elasticity of demand for natural gas

Ideally, we would directly measure the effect of downscheduling on the the Algonquin City Gate (ACG) price, the main price index for wholesale gas transactions in

[53] Here, we suppose that all other marketers would not change their quantity supplied in response to decreases in downscheduling by other firms. That is, residual supply is inelastic.

New England.[54] Unfortunately, downscheduling is correlated with demand, through temperature, on a seasonal and daily basis. That is, these firms primarily engage in downscheduling on colder days and during the winter months when capacity is more likely to be constrained. Moreover, we believe day-to-day variation in downscheduling during the winter is somewhat driven by the quantity of excess contracts each firm has available (i.e., after supplying demand from residential and commercial heating customers), which is largely driven by temperature. We therefore use an instrumental variables approach to address this endogeneity.

Typically, one would use the price of the same good in another market as an instrument for the price of a good in the market of interest (see for example, Hausman (1996) or Nevo (2001)). Our instrument for the ACG price is the Henry Hub (HH) price. The HH, which is located in Erath, LA, is a major distribution node and the primary pricing point for natural gas futures traded on the New York Mercantile Exchange. The HH price is determined by macroeconomic price shifters and unlikely to be affected by local supply and demand shifters affecting the ACG price, satisfying the exclusion restriction. Our instrumental variables specification is as follows:

$$D_t = \alpha_0 + \beta_1 P_t^{ACG} + \alpha_1 HDD_t + \alpha_2 HDD_t^2 + \alpha_3 W_t + \lambda_t + \epsilon_t$$

D_t is the natural log of quantity demanded (by electric generators) in the ACG wholesale gas market on day t and P_t^{ACG} is the natural log of the ACG price instrumented by the natural log of the HH price. We include controls for temperature using heating degree days (HDD_t and HDD_t^2), an indicator for whether the day is a weekend (W_t), and month-of-year fixed effects λ_t. The parameter β_1, captures the instrumented price elasticity of demand for ACG gas. To account for heteroskedasticity

[54] During our study period, the ACG was constructed by Platts, which collected price data by surveying market participants about their recent transcations.

and autocorrelation, we use Newey and West's optimal lag selection criteria to specify the covariance matrix. Because we are primarily interested in this relationship on cold days, when downscheduling is most likely to have impacted the ACG price, we restrict the estimation to days with positive heating degrees (*i.e.* days where the maximum temperature is below sixty-five degrees farenheit).

Table 5 summarizes the results of our instrumental variables approach. As a benchmark, Column (1) reports estimates from an OLS regression of the natural log of the ACG spot market quantity on natural log of the ACG price and covariates. The ACG price coefficient is -0.24 with a standard error of 0.02. Column (2) reports the results for a first-stage regression of the natural log of the ACG price on the natural log of the HH price and the same covariates described above. The t-statistic on the natural log of the HH price is 27.19 and the joint F-statistic is 209.27, suggesting the HH price is a strong instrument. Column (3) reports results for the main IV regression. The coefficient on the ACG price is -0.27 with a standard error of 0.02. Column (4) tests for endogeneity in the relationship between the HH price and the New England HDD terms. Only the coefficient on HDD is significant at the 10% level, though the magnitude (0.008) is quite small relative to the HH price coefficient in Column 3 (1.13), which is significant at the 1% level.

Our OLS and instrumented estimates of the elasticty of demand fall squarely within the range of estimates found in (Davis & Kilian, 2011) (*i.e.* , -0.1 to -0.34). For the calculations and simulations that follow, we use the instrumented coefficient and standard error.

6.1.2 Parameterizing the demand function

Next, we plug in our estimated price elasticity of demand into a constant-elasticity demand function: $D(p) = kp^{-0.27}$. Here, quantity demanded D is a function of a multiplier k, the ACG price p, and the elasticity of demand, which we estimated in the previous section to be -0.27. Using the daily quantity of gas demanded by electric generators (the primary source of spot-market demand) and the ACG price, we can solve for the vector of daily ks. Substituting this vector of ks back into the constant-elasticity demand function fixes the relationship between quantity and price, which allows us to calculate a counterfactual vector of prices from a counterfactual vector of quantities.

6.1.3 Constructing counterfactual quantities and prices

To construct a vector of counterfactual daily quantities, we begin by summing the downscheduled quantities at the ten nodes operated by Firms A and B where downscheduling occured most frequently and intensively. As above, downscheduled quantities are measured as the change in scheduled quantities during the last three hours of the gas day, less the average fraction, across all other nodes, of capacity downscheduled in the last three hours of the same gas day. To account for downstream capacity constraints, the daily downscheduled quantity is bounded from above by unused capacity at the compression station in Burillville, RI.[55] Next, we add this to daily deliveries to nodes serving electric generators, our measure of observed quantity demanded, to arrive at a counterfactual vector of quantities. Substituting the counterfactual vector of quantities into the parameterized demand function yields a counterefactual vector of prices. We bound our counterfactual price vector from below using the vector of Texas Eastern Zone M3 (TEM3) prices. The price of gas at TEM3, which sits at the junction of the Texas Eastern and Algonquin pipelines, captures the cost of delivering

[55] The Burrillville compression station has a very high average rate of capacity utilization and is the last potential bottleneck before the Algonquin pipeline branches into a nodal network. Unused capacity at Burillville is measured as the difference between end-of-day scheduled quantity and daily operational capacity.

gas to – or the price paid by buyers procuring gas for takeaway at – New England's doorstep.[56]

6.1.4 Comparing observed and counterfactual prices

Differences between the series of observed and counterfacual prices measures the impact of withholding. The average price difference in our three-year sample is \$1.68 (39%). The price difference during only the winter months of each year is \$3.82 (68%). While these price differences are significant, their effects are amplified in the wholesale electricity market. Next, we simulate the impact of withholding on wholesale electricity prices, substituting our counterfactual series of gas prices for the observed price series.

6.2 Electricity Market Impacts

Measuring the effect of capacity withholding on the wholesale electricity price is crucial both to understanding firms' incentives to withhold and to determining the resultant costs borne by ratepayers. We accomplish this by simulating New England's wholesale electricity market under both observed and counterfactual gas price series and comparing the two. Broadly, our estimation procedure consists of a.) re-constructing ISO-NE's wholesale electricity auction market, b.) non-parametrically estimating a distribution of counterfactual generator supply bids using our counterfactual gas price series, and c.) clearing the market a large number of times using draws from the counterfactual bid distribution.

[56] That is, using the price at TEM3 as a lower bound allows us to identify and correct for instances when our counterfactual quantity implies a slack capacity constraint. Failing to account for this would overstate the effect of capacity withholding.

6.2.1 New England's Wholesale Electricity Market

In New England's deregulated wholesale electricity market, prices and allocations are determined by a day-ahead auction administered by ISO-NE. At 10 a.m. the day before each operating day,[57] generators submit supply bids stating how much generation they will produce for what price for each hour of the day, and distribution utilities submit demand bids stating how much energy they will require at each hour of the day. The market-clearing price (expressed in $/MWh) is the price bid in by the generator providing the most expensive unit of generation required to meet demand (the "marginal generator"). All generators that bid in generation at prices lower than the clearing price are paid the clearing price for each MWh of energy they produce.[58]

Our re-construction and clearing of the day-ahead market is a departure from ISO-NE's actual dispatch procedure in two relatively minor ways. First, we exclude exports to or imports from the neighboring New York and Canadian energy markets. Second, our model does not incorporate no-load fees[59] or startup costs.[60] We believe

[57] Unlike the gas day, ISO-NE's electricity market operating day simply runs from 12 a.m. to 11:59 p.m. Eastern Time.

[58] In addition to the day-ahead market, ISO-NE also operates a *real-time* auction market that adjusts for unforeseen fluctuations in supply and demand throughout the operating day. We focus our analysis on the day-ahead market as 95% of ultimately consumed energy is traded there (Kim, 2017) and real-time prices generally closely track those of the day-ahead market. Additionally, we note that a substantial amount of energy is actually traded outside of the wholesale market through bilateral contracts between generators and utilities. We do not consider these transactions in our analysis as prices for these contracts are set by expectations of wholesale market prices.

[59] In addition to the principle variable cost component of their bids, generators can also optionally submit a no-load fee. No-load fees are a fixed sum to be paid to the generator for each hour it is called upon to operate that are designed to capture its fixed operating costs (such as labor). Incorporating no-load fees into the simulation model would incorporate a preference toward more fully using a smaller number of resources rather than spreading generation out across a larger number of resources, which would be more costly. This implies that by omitting no-load fees, our model produces downward-biased market clearing prices in both the real-world scenario and our counterfactual scenario without capacity withholding. In general, no-load fees are highest for peaking oil units. Because capacity withholding results in substitution from gas to oil, and oil units have higher no-load fees, the downward bias our model introduces by omitting no-load fees will likely be higher for the real-world scenario, which implies that our final estimates of the cost incurred by electricity ratepayers may be biased downward as well.

[60] Similarly, generators can choose to include cold-, intermediate-, and/or hot-start fees. These

assuming zero no-load fees and startup costs is reasonable for our application because a.) although they compose a substantial portion of total power costs paid by consumers, they do not enter the wholesale market price, as it is determined by the variable cost component of generators' bids, and b.) the key input we are adjusting—the gas price—is fundamentally a component of generators' variable costs.

6.2.2 Estimating Generator Bid Functions

A single generator is allowed to offer different quantities of generation for different prices, and accordingly their supply bid takes the form of a step function of up to ten price-quantity pairs. For example, a typical fuel-switching oil and gas generator might offer an initial 100 MW for \$30/MWh, another 40 MW for \$32/MWh, and a final 30 MW for \$80/MWh, with the price jump reflecting the switch from gas to oil. Because the size and number of steps can vary across days, even within an individual generation unit, it is infeasible to parametrically estimate a relationship between the gas price and generators' bids. We therefore use a non-parametric approach that employs a variant of nearest neighbor matching combined with a resampling procedure. Our approach is a novel extension of Reguant (2014)'s technique for estimating generators' expectations of competitors' bid functions.

Using notation loosely following Reguant (2014), we define the bid function generator i actually submitted on day t as b_{it}^{0}.[61] Each bid function consists of 24 sets of up to

startup fees are paid to generators each time they switch from being offline or in standby mode to operating. Assuming zero startup costs greatly simplifies the model by making it independently solvable for each hour rather than a dynamic problem. Additionally, while we observe the various startup costs as submitted in generators' bids, the associated times for each type of start vary by generator and are not made publicly available.

[61] Complete generator supply bid data covering our study period is publicly available in an anonymized format from ISO-NE's website. We de-anonymize these bids for robustness checks and to estimate revenues of Firms A and B; however, our main simulation model does not require knowledge of generator fuel types, technology types, or other characteristics to generate a counterfactual electricity price series.

10 quantity-price pairs $\langle p_{iths}, q_{iths} \rangle$, where h indexes hour and $s = 1, ..., 10$ indexes step. Instead of imposing a structure on the price-quantity pairs, we estimate the entire bid function b_{it} conditionally on electricity market demand (D_t^E) and temperature (HDD_t) using a two-step matching procedure (denoted below as f_m).[62] We employ this procedure to estimate two bid functions for each generator for each day—one corresponding to their expected bid given the actual observed gas price (b_{it}^1) and the other corresponding to their expected bid given the counterfactual gas price we estimate in the preceding section (b_{it}^2):[63]

$$b_{it}^1 = \mathbb{E}_i[b_{it}(p_t^{G,obs})|D_t^E, HDD_t]$$
$$b_{it}^2 = \mathbb{E}_i[b_{it}(p_t^{G,cf})|D_t^E, HDD_t]$$

The purpose of the first stage of the two-step procedure is to construct a sub-sample of potential match days that are highly similar to the target day in relevant electricity market conditions that determine generator bids. In particular, we isolate the 5% of days in our three-year study period that are most similar to day t in temperature and electricity market demand using Mahalanobis distance. For generators that were online during the entire three-year period, this generates a subsample of 56 days that are highly similar to the target day in these two key determinants of generators' bids.

The second stage serves to select days that experienced a gas price that was either very close to the gas price of the target day (for estimating b_{it}^1) or to our estimated counterfactual gas price for the target day (for estimating b_{it}^2). Specifically,

[62] For D_t^E, we use the predicted peak demand for the day-ahead market, which is published at the same time bids must be submitted for the day-ahead market and therefore represents market expectations of demand at that time.

[63] While we actually observe generator i's real bid function on day t, i.e. their actual bid function given the actual gas price, we estimate (b_{it}^1) in order to employ the same simulation methodology in both our actual and counterfactual scenarios in order to employ the same simulation methodology to both and remove any bias introduced by the methodology that is common to both.

in the second stage we select from the first stage subsample the three days that experienced wholesale gas prices most similar to that of the observed or estimated counterfactual gas price.[64,65] Each predicted bid function ($\widehat{b}^1_{it}$ and $\widehat{b}^2_{it}$) then consists of three of generator i's actual bids $\{b_{i\tau_1}, b_{i\tau_2}, b_{i\tau_3}\}$ from three days (τ_1, τ_2, and τ_3) that are highly similar to t in temperature, electricity demand, and either gas price (for $\widehat{b}^1_{it}$) or our estimated counterfactual gas price (for $\widehat{b}^2_{it}$):

$$\widehat{b}^1_{it} = f^1_m(b_{it}(p^{G,obs}_t)|D^E_t, HDD_t)$$
$$\widehat{b}^2_{it} = f^2_m(b_{it}(\widehat{p^{G,cf}_t})|D^E_t, HDD_t)$$

Our identifying assumption for f^2_m is that days with lower gas prices are not systematically different than days with higher gas prices in other ways that affect generator bids. This assumption could be violated if, for example, the price of another fuel such as coal happened to also be much lower on days with a lower gas price. If this were the case, our estimate for the costs to consumers of withholding behavior would be picking up both the effect of a lower gas price and a lower coal price.[66]

We perform the second stage for b^2_{it} multiple times in order to incorporate the statistical uncertainty of our estimated counterfactual gas price series into our final results. Each time we draw $\widehat{p^{G,cf}_t}$ from a distribution of estimated counterfactual prices $\widehat{P^{G,cf}_t}$ constructed by incorporating the uncertainty associated with our IV estimates in section 6.2.[67] We draw 100 complete counterfactual price series and perform the

64 We exclude the same day for estimation of b^1_{it}

65 We use three days here because it is again roughly 5% of the subsample of days for generators that participated in the day-ahead auction for every day of our study period. For generators that participated in the day-ahead market for less than 60 days of our study period (ten low-capacity generators representing less than .01% of generator-days in our sample), the first step does not generate at least 3 potential match days. To keep our methodology straightforward, we simply input their actual bid function when clearing the electricity market.

66 In the future, we intend to include energy prices in the first stage to control for this particular potential violation.

67 Specifically, we adjust our point estimate for the elasticity of demand by a random shock drawn

second stage matching procedure for each, generating 100 sets of 3 match days for each generator for each day.

6.2.3 Simulating the Electricity Market

We begin by simply re-constructing and clearing the wholesale auction market using actual generator bids b_{it}^0. This entails combining all price-quantity pairs submitted by all generators into an aggregate bid supply function and determining the most expensive step required to meet that hour's demand, recovering the price from that step as the output.

In comparing the generated price series to the true day-ahead price series, we assess that our model introduces a relatively small degree of noise and bias by assuming away trade, no-load costs, and startup fees. As the first panel of Figure 15 demonstrates, the distribution of prices produced by our simulation model using b_{it}^0 is highly similar to the distribution of actual day-ahead prices. Hour-by-hour, our price series tracks the actual price series fairly well, with a coefficient of correlation of .958. Our simulation-generated prices have a slight upward bias, with the average clearing price of \$48.05 being slightly higher than the actual average clearing price of \$47.53. However, when we calculate total annual energy cost to consumers ($\widehat{e}_0$),[68] we find that our simulation slightly underestimates the annual energy cost at \$6.248 billion versus the true value of \$6.267 billion.

Next, we re-construct and clear the wholesale auction market using our estimated distribution for actual generator bids. We do this by randomly drawing one of the three

from a Gaussian distribution of our estimated error, calculate the counterfactual gas price series using that adjusted elasticity, and lower-bound it ex-post with the Texas Eastern M3 price.

[68] By simply multiplying demand by the clearing price for each hour, aggregating over our study period and dividing by 3. Note that while energy costs account for the lions share of ratepayer electricity bills, they additionally include fixed costs such startup costs, no-load fees, distribution costs, and administrative fees that we do not attempt to estimate.

bid functions in $\widehat{b}_{it}^1$ for each generator for each hour, using those draws to construct aggregate bidsupply curves, and clearing the market for each hour. We perform this step 100 times to create a distribution of electricity prices for each hour, the means of which compose our point estimate for the actual electricity price.

While this is a most certainly a roundabout approach to estimating the actual electricity price for each day, it enables us to assess the level of uncertainty and/or bias that is introduced by our matching-resampling procedure, and to correct for that bias to some extent. As the second panel of Figure 15 shows, the distribution of simulated prices using our matching-resampling procedure is somewhat more disparate from the actual price distribution than the distribution using actual bids, which is notably flatter, but still tracks it fairly well. The coefficient of correlation between the two is .925, and the average clearing price is slightly lower than the actual at \$46.91. To calculate energy costs to ratepayers, rather than using the mean of our estimated price series, we independently multiply demand by clearing price for each hour, sum these hourly costs, and divide by three to create a distribution of a total annual energy costs $\widehat{E}_1$. The mean of this distribution, \$6.109 billion, is our estimate for the annual energy cost paid by ratepayers using $\widehat{b}_{it}^1$, $\widehat{e}_1$. This figure is lower than the actual energy cost by \$158 million, which we note is about an order of magnitude smaller than our final predicted estimates for the cost to ratepayers of withholding activity.[69]

Finally, we re-construct and clear the electricity market using our estimated distribution of counterfactual generator bids in our no-withholding scenario. Recall that each bid function $\widehat{b}_{it}^2$ consists of 100 sets of 3 potential match days, each set of 3

[69] In other words, while it's clear from the discrepancy that our matching-resampling procedure introduces some bias, we can be reasonably confident that our final estimates are accurately signed and correct in magnitude because this bias is small relative to our final estimated damages from pipeline capacity withholding. Moreover, we correct for bias that's common to both $\widehat{E}_2$ and $\widehat{E}_1$ by using the difference between the two, rather than the difference between $\widehat{E}_2$ and the actual actual cost, for our final damages estimate.

using a different draw from our generated distribution of counterfactual gas prices. For each of the 100 sets, we clear the market 10 times,[70] randomly selecting generator bid functions from one of the three potential match days each time as before, constructing a distribution consisting of 1000 counterfactual electricity price series.

The first panel of Figure 16 shows that our estimated distribution of counterfactual electricity prices not only has a taller peak in the 40-60$/MWh range, but is also noticeably shifted to the left. The average clearing price estimated by our model is $37.57, about $10 lower than the actual average price as well as our two estimates of it. Figure 17 demonstrates that the price differential was largely driven by the "polar vortex" of 2013-14, when electricity prices were $40.67 (51%) higher on average due to capacity withholding. Our model estimates that electricity prices were $17.50 (28%) higher due to capacity withholding during the winter of 2014-15, $4.81 (17%) higher during the winter of 2015-16, and $3.57 (13%) higher during the summers in our study period.[71,72]

We again calculate total energy costs independently for each of the 1000 iterations to construct the distribution of annual energy costs $\widehat{E}_2$. We then subtract $\widehat{E}_2$ from $\widehat{E}_1$ to produce a distribution of estimated costs to electricity ratepayers due to pipeline capacity withholding $\widehat{D}$, which incorporates the uncertainty associated with our wholesale gas price estimates and, to an extent, the uncertainty introduced by our sampling procedure. The full distribution of simulation results is shown in the second panel of Figure 16. The mean, $\widehat{d}$ = \$1.214 billion, is our estimated annual cost to consumers (\$3.642 billion over three years), and the 95% confidence interval is

[70] We are currently only using 10 iterations here due to computation times; we intend to ultimately do 100 iterations at this step as well

[71] We again define winter in this context as December 1 - March 31 following the Algonquin pipeline's tariff.

[72] We attribute the limited impact of withholding on prices during the winter of 2016-17 to generally warmer temperatures (37°F on average versus 28°F and 27°F during the other two winters).

$1.076-$1.381 billion ($3.228-$4.143 billion over three years).

7 Conclusion

To date, most growth in gas deliveries for electric generation has utilized legacy pipeline infrastructure. Much of this infrastructure was built to supply heating customers, with institutions designed to manage heating demand. As demand for gas for electric generation grows, gas-fired electric generators will increasingly find themselves in competition with legacy pipeline customers for scarce pipeline capacity. This new competitive environment is likely to affect the incentives of firms, especially those that operate in both markets, in unforeseen ways. Therefore, it will be critical that the institutions governing the trade and transport of gas, both as a heating source and for electric generation, are harmonized and structured efficiently.

In this paper, we identify a major inefficiency spanning the natural gas transportation and wholesale electricity markets. We quantify the extent to which two firms withheld pipeline capacity and detail the institutional arrangements that allowed these firms to execute their withholding strategy. Using an instrumental variables approach, we identify the effect of withholding on the wholesale natural gas price, and we employ an electricity dispatch model to trace that impact through the wholesale electricity market. During our study period, this withholding of pipeline capacity resulted in a transfer from electricity ratepayers to electricity generators (and their fuel suppliers) of $3.6 billion.

While only the firms in question can explain their actions and intentions, we have shown that it is consistent with the exercise of market power and inconsistent with the most plausible alternative explanations, namely risk aversion on the part of these firms. Furthermore, we have demonstrated that the decreased supply of gas to New England generators caused by this behavior increased revenues in the wholesale electricity market for these (and other) firms by margins that likely exceeded the opportunity cost of using the capacity to supply the gas spot market, especially when accounting for state revenue-sharing rules.

Just as transmission constraints create opportunities for market power in the electricity sector, so too can capacity constraints create opportunities for market power in the natural gas transportation industry. As natural gas prices have fallen, demand for gas to power the electric grid has steadily increased, creating new rivals and tying closer the gas transportation and electricity markets. In this context, it has become increasingly important that already existing pipeline capacity is optimally utilized not only to protect the interests of gas and electricity ratepayers, but also to ensure that unbiased price signals lead to an efficient level of new pipeline development. Pipeline market reforms that facilitate more flexible contracting mechanisms, more frequent scheduling cycles, and act to prevent capacity withholding, or impose a cost for capacity withholding and create a publicly-available record of capacity withholding; all of which will serve to better align the gas transport and electricity markets, could help to create more liquid markets in which firms find it more difficult to exert market power.

While the analysis here has focused on identifying the exercise of market power on one particularly congested pipeline serving New England, severe bottlenecks and vertically integrated firms coexist in many other natural gas transportation markets. To what extent capacity withholding has led to pipeline underutilization in other regions is an important area for future study.

References

Boots, Maroeska G, Rijkers, Fieke AM, & Hobbs, Benjamin F. 2003. GASTALE: An oligopolistic model of production and trade in the European gas market. *Energy Research Centre of the Netherlands (ECN), ECN*, **35**(1).

Borenstein, Severin, & Bushnell, James. 1999. An empirical analysis of the potential for market power in California?s electricity industry. *The Journal of Industrial Economics*, **47**(3), 285–323.

Borenstein, Severin, Bushnell, James, Kahn, Edward, & Stoft, Steven. 1995. Market power in California electricity markets. *Utilities Policy*, **5**(3), 219–236.

Borenstein, Severin, Bushnell, James, & Knittel, Christopher R. 1999. Market power in electricity markets: beyond concentration measures. *The Energy Journal*, 65–88.

Borenstein, Severin, Bushnell, James B, & Wolak, Frank A. 2002. Measuring market inefficiencies in California's restructured wholesale electricity market. *The American Economic Review*, **92**(5), 1376–1405.

Cremer, Helmuth, & Laffont, Jean-Jacques. 2002. Competition in gas markets. *European Economic Review*, **46**(4), 928–935.

Cremer, Helmuth, Gasmi, Farid, & Laffont, Jean-Jacques. 2003. Access to pipelines in competitive gas markets. *Journal of Regulatory Economics*, **24**(1), 5–33.

Davis, Lucas W, & Kilian, Lutz. 2011. The allocative cost of price ceilings in the US residential market for natural gas. *Journal of Political Economy*, **119**(2), 212–241.

EIA. 2017a (January). *Annual Energy Outlook with Projections to 2050*. US Energy Information Administration.

EIA. 2017b (September). *Short-Term Energy Outlook*. US Energy Information Administration.

FERC. 2017. *Cost-of-Service Rate Filings*. https://www.ferc.gov/industries/gas/gen-info/rate-filings.asp, note = Accessed: 2017-07-31.

Hart, Oliver, Tirole, Jean, Carlton, Dennis W, & Williamson, Oliver E. 1990. Vertical integration and market foreclosure. *Brookings papers on economic activity. Microeconomics*, **1990**, 205–286.

Hausman, Jerry A. 1996. Valuation of new goods under perfect and imperfect competition. *Pages 207–248 of: The economics of new goods*. University of Chicago Press.

Holz, Franziska, Von Hirschhausen, Christian, & Kemfert, Claudia. 2008. A strategic model of European gas supply (GASMOD). *Energy Economics*, **30**(3), 766–788.

Hunger, David. 2003. Analyzing gas and electric convergence mergers: A supply curve is worth a thousand words. *Journal of Regulatory Economics*, **24**(2), 161–173.

ICF. 2015. New England Energy Market Outlook: Demand for Natural Gas Capacity and Impact of the Northeast Energy Direct Project. *ICF International Inc.*

Joskow, Paul L, & Tirole, Jean. 2000. Transmission rights and market power on electric power networks. *The Rand Journal of Economics*, 450–487.

Kim, Harim. 2016. Heterogeneous Impacts of Cost Shocks, Strategic Bidding and Pass-Through: Evidence from the New England Electricity Market.

Lise, Wietze, & Hobbs, Benjamin F. 2009. A dynamic simulation of market power in the liberalised European natural gas market. *The Energy Journal*, 119–135.

Nevo, Aviv. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, **69**(2), 307–342.

Oliver, Matthew, & Mason, Charles. 2018. An overview of interstate gas pipeline regulation in the United States. *Progress in Economics Research*, **39**(01). (forthcoming, in press).

Oliver, Matthew E, Mason, Charles F, & Finnoff, David. 2014. Pipeline congestion and basis differentials. *Journal of Regulatory Economics*, **46**(3), 261–291.

Ordover, Janusz A, Saloner, Garth, & Salop, Steven C. 1990. Equilibrium vertical foreclosure. *The American Economic Review*, 127–142.

Reguant, Mar. 2014. Complementary bidding mechanisms and startup costs in electricity markets. *The Review of Economic Studies*, **81**(4).

Rose, Judah, Gerhardt, David, Karp, John, Brock, Frank, Sakya, Trisagni, & Shrestha, Mohit. 2014. Of Polar Vortexes and PJM Price Spikes. *ICF International Inc.*

Salop, Steven C., & Scheffman, David T. 1983. Raising Rivals' Costs. *The American Economic Review*, **73**(2).

Spectra. 2016. *FERC Gas Tariff of Algonquin Gas Transmission, LLC.* Sixth revised volume no. 1 edn.

Vazquez, Miguel, Berzosa, Ana, & Vazquez, Carlos. 2006. An analysis of the oligopolistic effects of the integration between electricity and gas suppliers in the Spanish market. *Available at SSRN 1262743.*

Tables

Table 1: Average schedule change in the last three hours of the gas day by node in MMBtu. Six nodes are clear outliers from the rest of the distribution, and the next four nodes that downschedule the most are also operated by either Firm A or Firm B. (return)

Rank	Schedule Change in Last 3 Hours	Schedule Change (Winter Only)	Node Operator	Node Type
1	-18,444	-17,865	Firm A	LDC
2	-10,576	-9,281	Firm A	LDC
3	-7,116	-7,766	Firm A	LDC
4	-3,889	-2,529	Firm A	LDC
5	-3,808	-8,426	Firm B	LDC
6	-2,401	-4,963	Firm B	LDC
7	-861	-286	Firm A	LDC
8	-711	-1,645	Firm B	LDC
9	-563	-59	Firm A	LDC
10	-479	-975	Firm B	LDC
11	-348	18	Firm L	Generator
12	-250	-594	Firm M	Generator
13	-229	-261	Firm C	LDC
...	...	...	...	...
116	395	449	Firm N	Generator
117	653	1,017	Firm K	LDC

Table 2: The two firms observed to consistently engaged in downscheduling are also the two that hold the most and third-most generation capacity in New England out of the 11 parent energy companies that operate LDC-designated nodes on Algonquin. We include firms' contract holdings on Algonquin as a proxy for the relative size of their natural gas operations on the pipeline, which reflects their ability to affect prices by withholding capacity. Firm I holds a large quantity of unregulated generation capacity, but its LDC operations on Algonquin are extremely limited, which limits ability to constrict pipeline capacity. Firm G appears to have both some ability and some incentive to withhold pipeline, though significantly less than the two firms we observe downscheduling. We note also that Firm A appears to withhold slightly more on average that it holds NN from TE contracts (with the "no notice" aspect the binding constraint). This suggests that "no notice" contracts may facilitate downscheduling, but are not absolutely required for it.(return)

LDC	Schedule Change in Last 3 Hours (MMBtu)	Schedule Change (Winter Only) (MMBtu)	Generation Capacity (MW)	Unregulated Capacity (MW)	All Contracts (MMBtu)	NN from TE Contracts (MMBtu)
Firm A	-41,506	-37,903	232	77	185,300	39,200
Firm B	-8,832	-18,320	1,177	0	632,400	105,100
Firm C	-225	-262	33	33	166,500	0
Firm D	-4	-10	0	0	32,000	0
Firm E	5	8	0	0	1,300	800
Firm F	7	36	0	0	24,500	0
Firm G	17	38	116	116	42,200	19,700
Firm H	21	56	23	0	7,600	5,000
Firm I	108	146	1,008	1,008	12,500	0
Firm J	177	227	10	0	768,900	167,100
Firm K	298	3	0	0	156,700	44,300

Not shown: 14 electricity generation firms and 2 industrial end users that operate nodes on Algonquin

Table 3: Relationships between downscheduling in the final 3 hours of the gas day and LDC status, whether the node is located in Connecticut (as a proxy for strength of LDC extra-marginal revenue sharing mechanisms), "no notice" contracts delivering gas from the Texas Eastern pipeline, and interactions. Coefficients on Contracts and CT×Contracts are omitted in columns (3) and (5) because only LDCs hold "no notice" contracts on Algonquin. (return)

	(1)	(2)	(3)	(4)	(5)
LDC	585.0**	114.2**	315.1		99.53**
	(240.6)	(48.20)	(207.9)		(45.61)
CT		61.95		236.9	61.95
		(89.38)		(375.3)	(89.38)
Contracts			0	0.00319	0
			(.)	(0.00344)	(.)
LDC×CT		1325.8**			212.1
		(647.5)			(473.8)
LDC×Contracts			0.0502		0.00214
			(0.0521)		(0.00344)
CT×Contracts				0.419***	0
				(0.119)	(.)
LDC×CT×Contracts					0.417***
					(0.122)
HDD	7.630	7.630	7.630	7.630	7.630
	(4.660)	(4.660)	(4.660)	(4.660)	(4.660)
HDD^2	-0.266**	-0.266**	-0.266**	-0.266**	-0.266**
	(0.129)	(0.129)	(0.129)	(0.129)	(0.129)
Weekend	83.48***	83.48***	83.48***	83.48***	83.48***
	(28.87)	(28.87)	(28.87)	(28.87)	(28.87)
Quarter FE	Yes	Yes	Yes	Yes	Yes
N	133,029	133,029	133,029	133,029	133,029

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Relationship between downscheduling and generation capacity ownership for LDC-designated nodes on Algonquin. "MW" here is a binary indicator for whether the node operator's parent firm owns at least 100 MW of generation capacity in New England. (return)

	(1)	(2)	(3)	(4)	(5)
MW	1012.1**	101.5**	322.9		-18.28
	(445.2)	(43.46)	(386.9)		(21.29)
CT		2.959		274.4	-18.29
		(20.81)		(465.6)	(12.45)
Contracts			-0.00259	0.00225	-0.00262
			(0.00271)	(0.00346)	(0.00269)
MW×CT		1623.8**			501.9
		(768.1)			(551.2)
MW×Contracts			0.156		0.0190***
			(0.116)		(0.00574)
CT×Contracts				0.417***	0.00435
				(0.122)	(0.00447)
MW×CT×Contracts					0.395***
					(0.120)
HDD	9.814*	9.814*	9.814*	9.814*	9.814*
	(5.537)	(5.537)	(5.537)	(5.537)	(5.537)
HDD^2	-0.328**	-0.328**	-0.328**	-0.328**	-0.328**
	(0.155)	(0.155)	(0.155)	(0.155)	(0.155)
Weekend	83.48**	83.48**	83.48**	83.48**	83.48**
	(33.89)	(33.89)	(33.89)	(33.89)	(33.89)
N	109,152	109,152	109,152	109,152	109,152

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

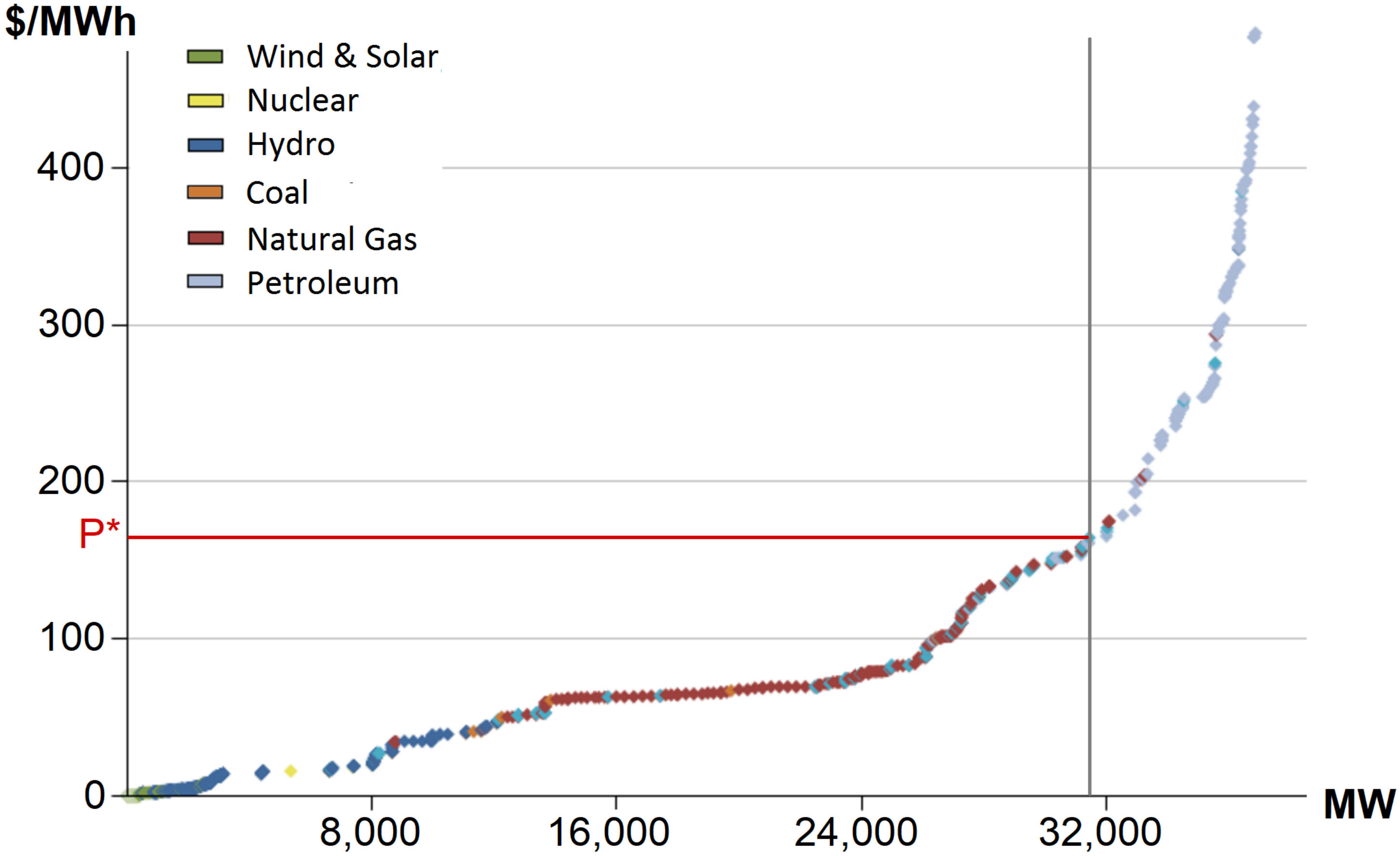
Table 5: Estimating the elasticity of demand for pipeline natural gas using OLS and Instrumental Variables, where the Algonquin City Gate (ACG) price is instrumented with the Henry Hub (HH) price. (return)

	(1)	(2)	(3)	(4)
VARIABLES	log_quantity	log_price_agt	log_quantity	log_price_henry
log_price_agt	-0.239***		-0.266***	
	(0.0197)		(0.0210)	
hdd	-0.00178	0.0172***	-0.00139	0.00768*
	(0.00314)	(0.00325)	(0.00313)	(0.00426)
hdd_2	-2.42e-05	0.000403***	-8.00e-06	5.94e-05
	(6.93e-05)	(6.92e-05)	(6.77e-05)	(7.68e-05)
weekend	-0.0932***	-0.0477	-0.0945***	-0.00230
	(0.0182)	(0.0309)	(0.0181)	(0.0266)
log_price_henry		1.129***		
		(0.0403)		
Constant	13.17***	-0.388***	13.19***	0.771***
	(0.0550)	(0.0612)	(0.0558)	(0.0752)
Observations	795	795	795	795
R-squared	0.660	0.788	0.659	0.089

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figures

Figure 1: An illustration of New England's day-ahead electricity market, which is constructed and cleared for each hour of each operating day. Generators bid in their marginal costs of generation (plus markup when they have market power) which can be used to construct a market "bid supply" curve ranking generation resources from lowest to highest cost. Electricity distribution utilities bid in their predicted levels of demand to construct a market demand curve (shown here as perfectly inelastic as there is very little operational demand response in New England). The most expensive generation resource required to meet demand for a given hour sets the wholesale price received by all generators called upon to operate. (return)



Note: The underlying data corresponds to marginal cost of generation and capacity rather than bid supply offers. We are not able to use the latter here in a straightforward manner because bid data is anonymized; however, the curve here roughly matches the distribution of generators' price and quantity supply offers to the day-ahead market by fuel type and serves the purpose of illustrating the bid supply curve.

Figure 2: Impact of overscheduling capacity on the wholesale gas market on warmer days when the pipeline is uncongested (left panel) and on colder days when it is fully scheduled (right panel) (return)

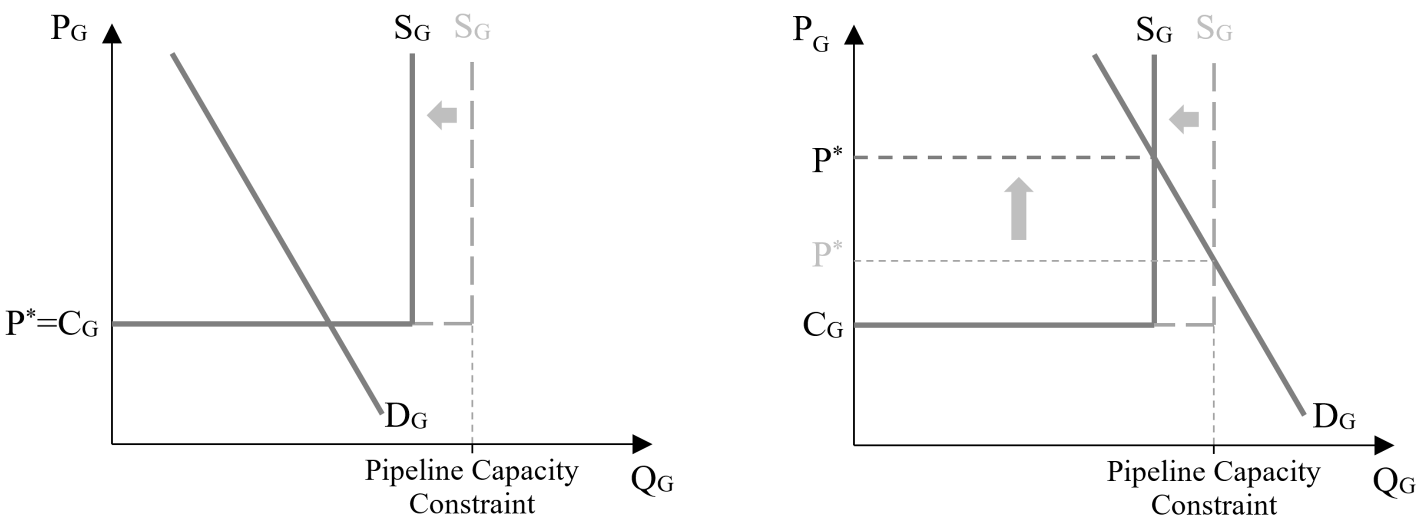


Figure 3: Impact of overscheduling capacity on the wholesale electricity market on warmer days when the pipeline is uncongested (left panel) and on colder days when it is fully scheduled (right panel) (return)

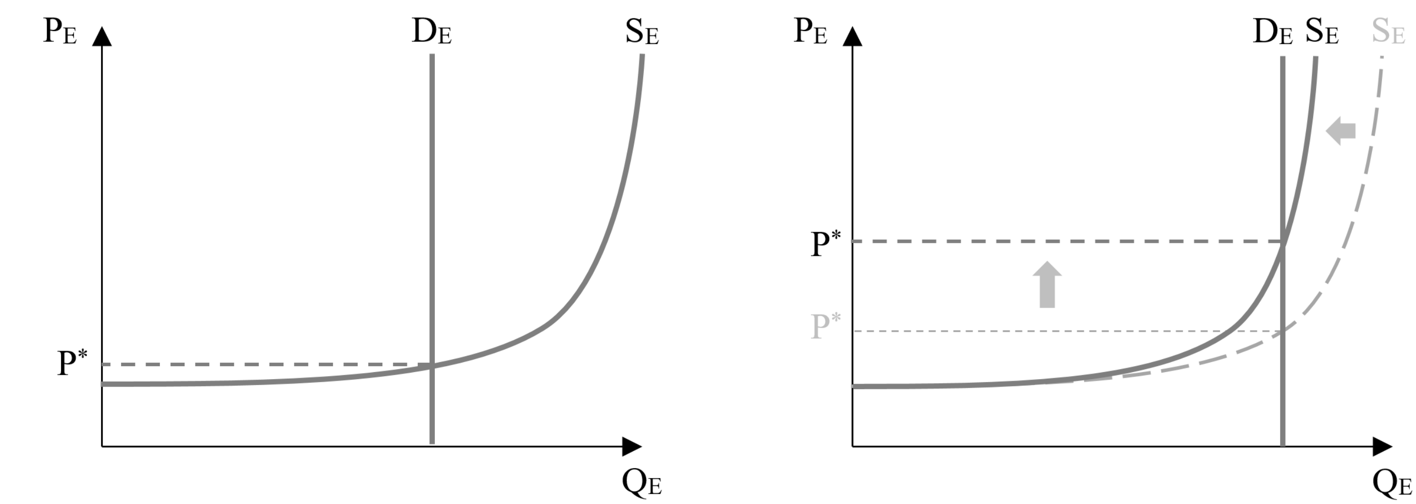


Figure 4: Impact of overscheduling capacity on the wholesale gas market (left panel) and on the electricity market (right panel). The dark shaded region in the left panel represents gas market revenues the withholding LDC sacrifices by letting their capacity go unused, which are restricted by revenue-sharing rules. The dark shaded area in the right panel corresponds to the additional revenues earned by the LDC's generation capacity from a higher wholesale electricity price. (return)

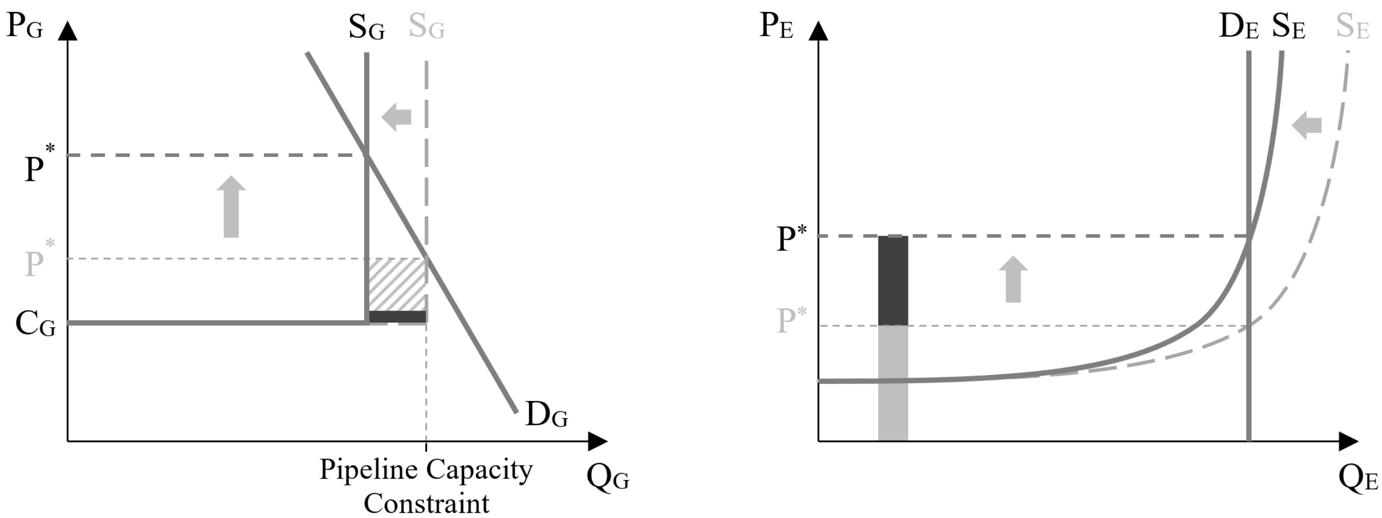


Figure 5: The scheduling pattern of a typical LDC delivery node. Each line represents one gas day in our three year study period. Winter is defined as December 1 through March 31 following the delineation used by Algonquin Gas Transmission. The X-axis covers the 44-hour scheduling period and the Y-axis is the total daily quantity of gas scheduled for delivery to that node at a given time. Line color represents the Algonquin Citygate price, wherein redder lines indicate higher-priced days. The scheduling pattern at this node and at most other LDC delivery nodes on Algonquin is characterized by most adjustments being made shortly after the timely cycle or around the start of the gas day with some slight balancing either direction in the final hours on some days. We constructed equivalent graphs for all 117 delivery nodes on the Algonquin pipeline and have made these publicly available our the online data appendix. (return)

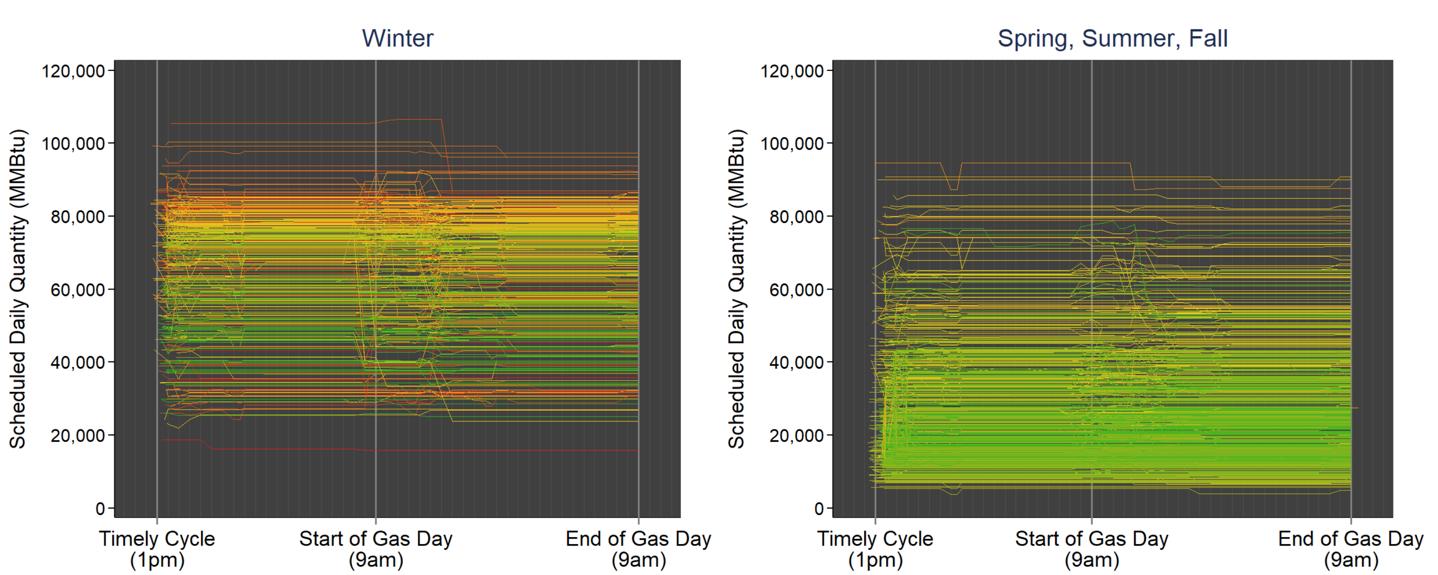


Figure 6: The scheduling pattern of an LDC delivery node that consistently downschedules its nomination in the final three hours of the gas day. This pattern is exhibited clearly at four nodes and to a lesser degree at two additional nodes, all of which are operated by Firm A. The y-axis corresponds to a total daily quantity rather than a flow rate, meaning these large negative schedule adjustments correspond to unused pipeline capacity. For example, if this node schedules 72,000 MMBtu at the beginning of the scheduling period, it is indicating to the pipeline company that it will be flowing gas at a rate of 72,000/24=3,000 MMBtu per hour over the course of the gas day and the pipeline company then reserves that capacity for them. When it reduces its scheduled quantity to 48,000 MMBtu three hours before the end of the gas day, it is not reducing its rate of flow at that time, but rather indicating to the pipeline company that it had been flowing gas at a rate of 2,000 MMbtu per hour over the gas day. (return)

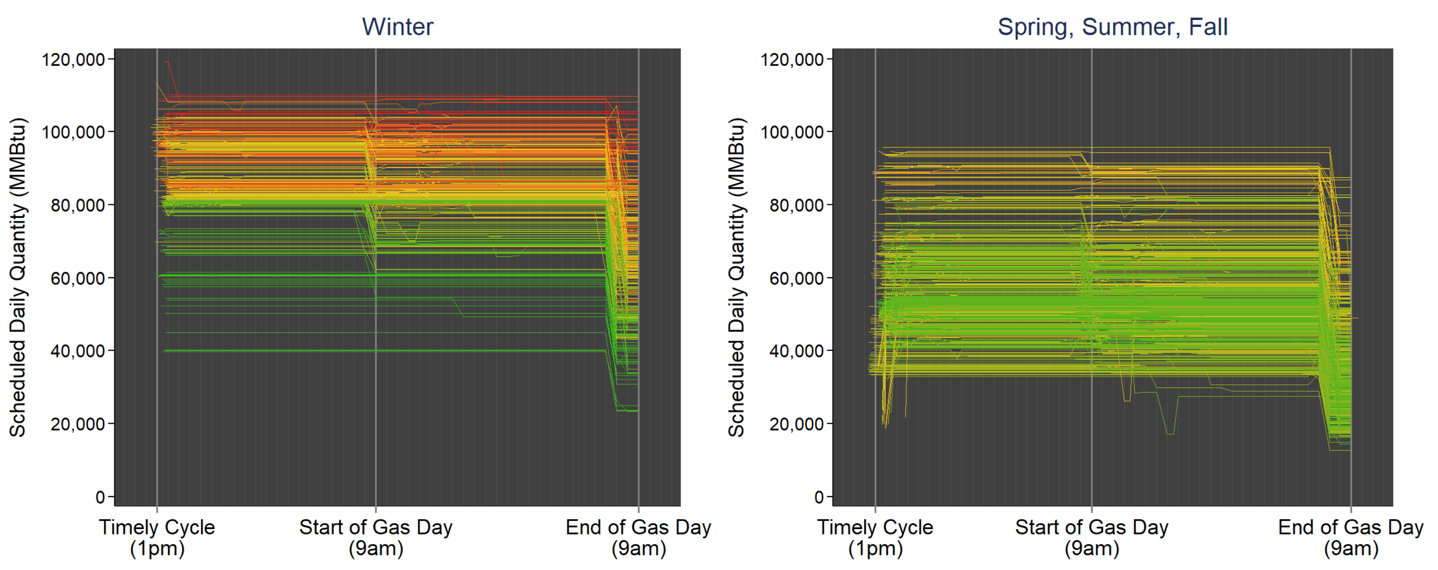


Figure 7: The scheduling pattern at a node operated by Firm B, which is characteristically similar to that of the suspect nodes operated by Firm A except that downscheduling behavior is concentrated in the winter when prices are high. We observe this pattern very clearly at two nodes and to a lesser extent at two additional nodes, all of which are operated by Firm B. (return)

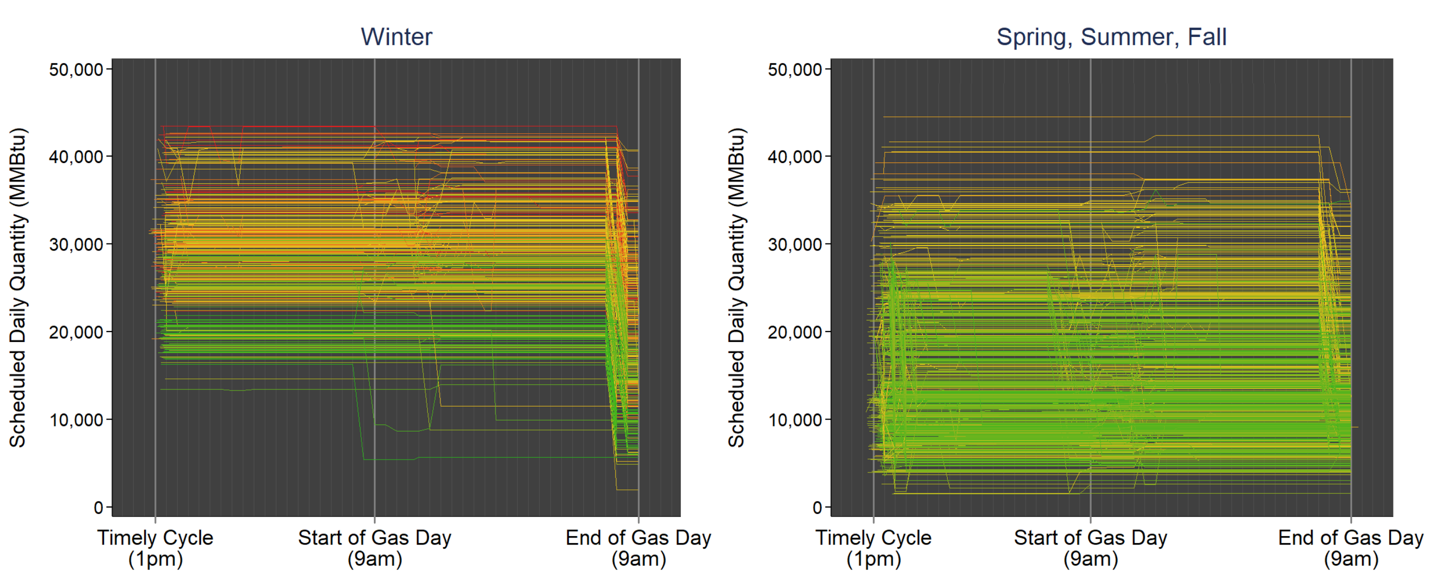


Figure 8: Schedule change from the previous hour over the scheduling period. Each line represents the average behavior of one the 117 delivery nodes on Algonquin. Six nodes operated by either Firm A or Firm B are clear outliers from the rest of the distribution in consistently making large negative schedule adjustments in the final hours of the gas day. Nodes operated by Firm A engage in this practice year round, while nodes operated by Firm B primarily perform these schedule adjustments primarily in the winter. (return)

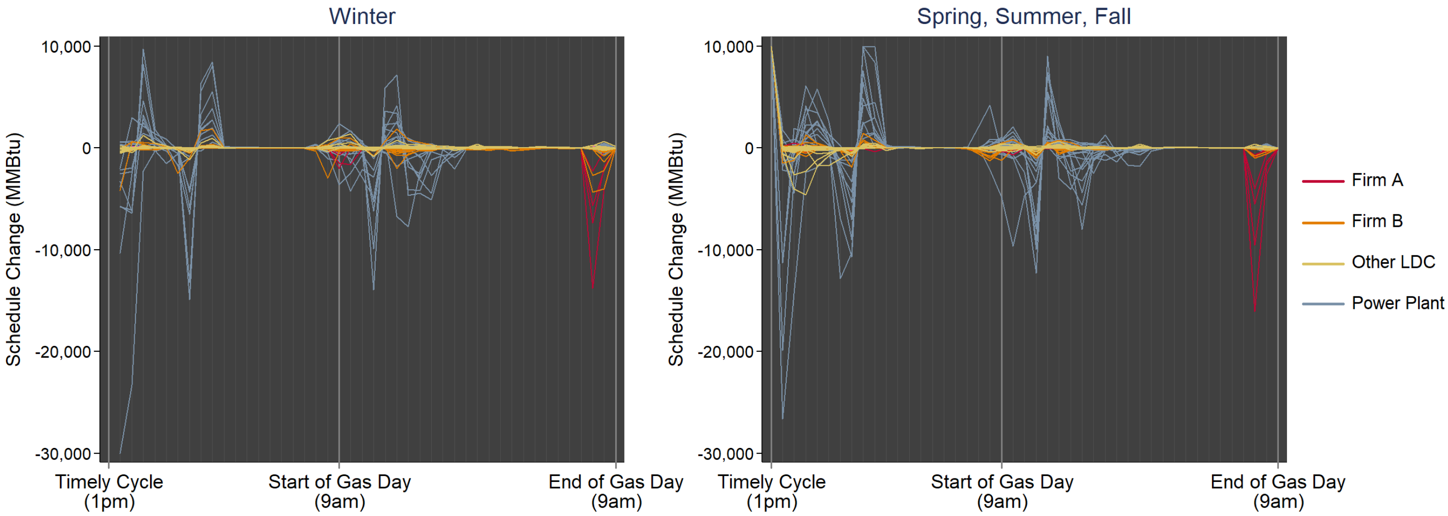


Figure 9: The locations of the 10 nodes that downschedule the most on average. Eight are located in Connecticut, where revenue sharing rules are strongest. (return)

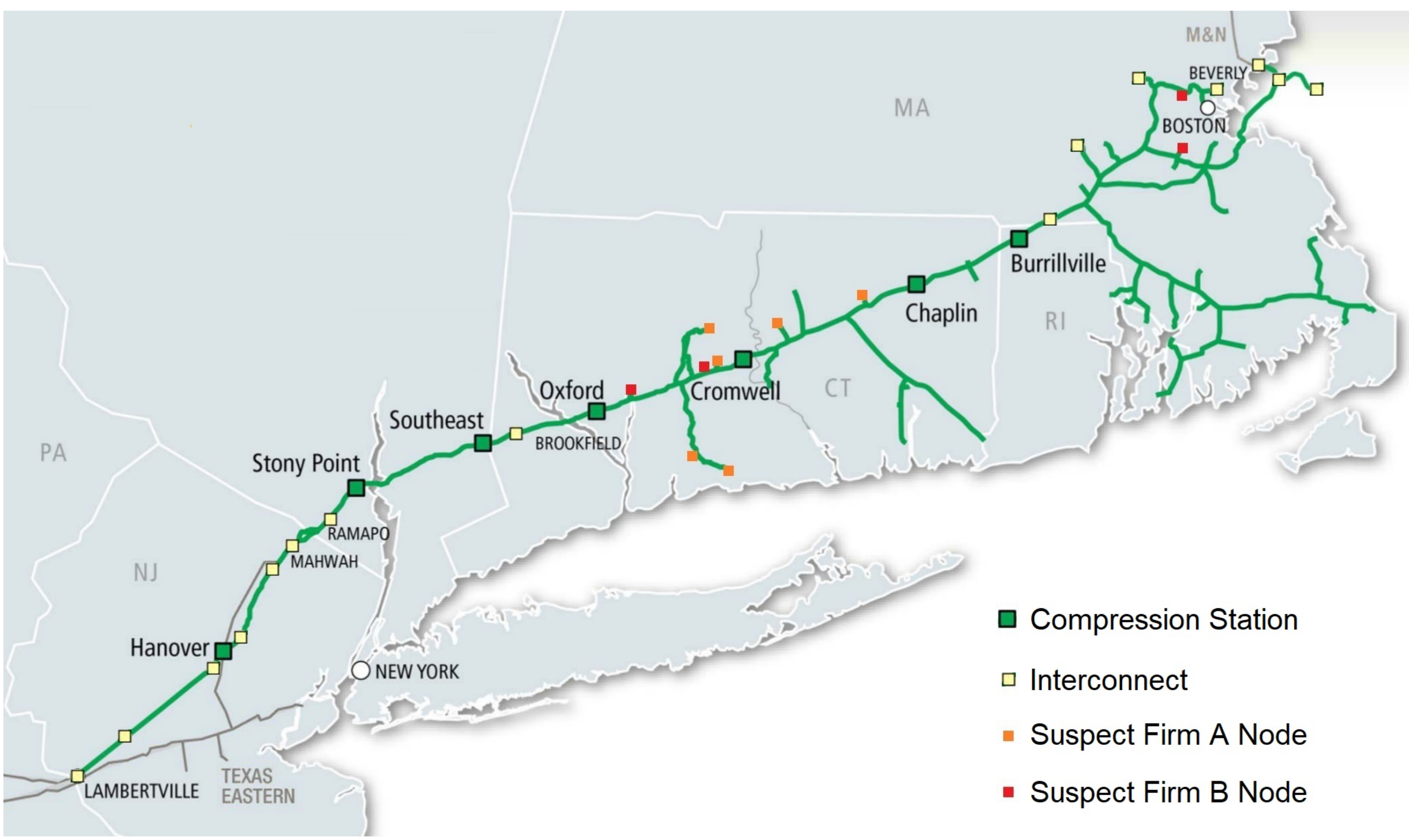


Figure 10: Average holdings of all contracts for capacity on the Algonquin pipeline from 2012 through 2016. Each bar represents a shipper; shippers are grouped by parent company for Firms A, B, and C, all of whom transport a majority of their gas through Algonquin's interconnects with the Texas Eastern Pipeline. Many of these contracts are no-notice. (return)

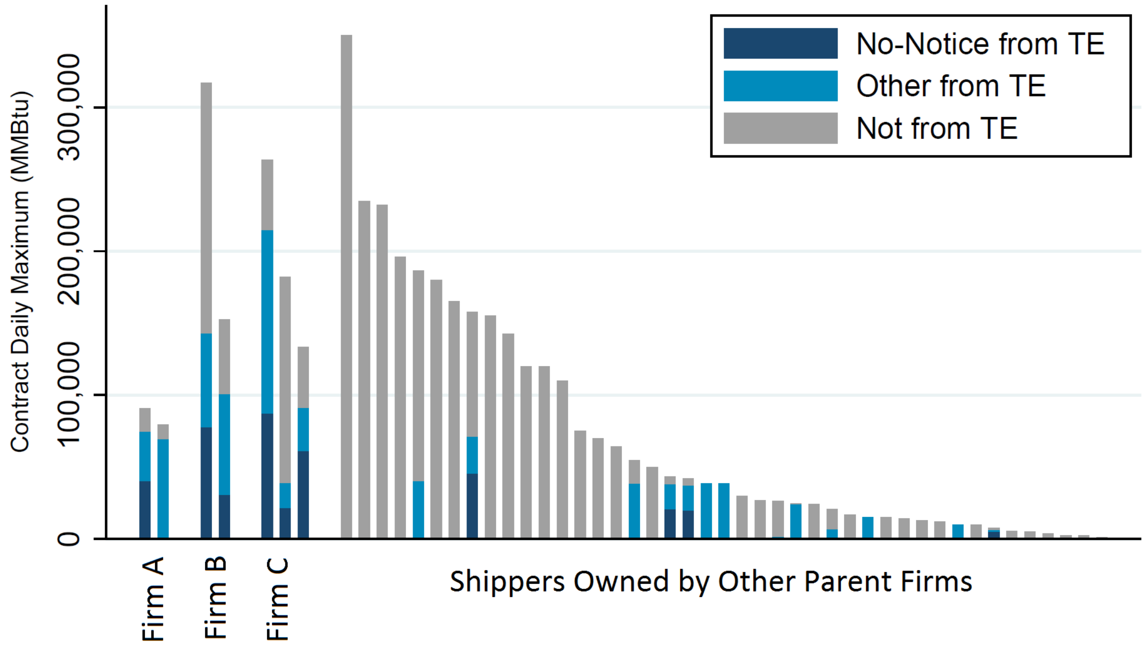


Figure 11: Average holdings of all contracts for capacity on the Texas Eastern pipeline for gas delivered to its two interconnects with the Algonquin pipeline. Of the gas sourced from Texas Eastern by Firms A, B, and C, much of it comes from storage. The upper limits of the aggregate downscheduling behavior observed (around 100,00- MMBtu) roughly match the sum of Firm A's and Firm B's no-notice contracts sourcing gas from storage on Texas Eastern. (return)

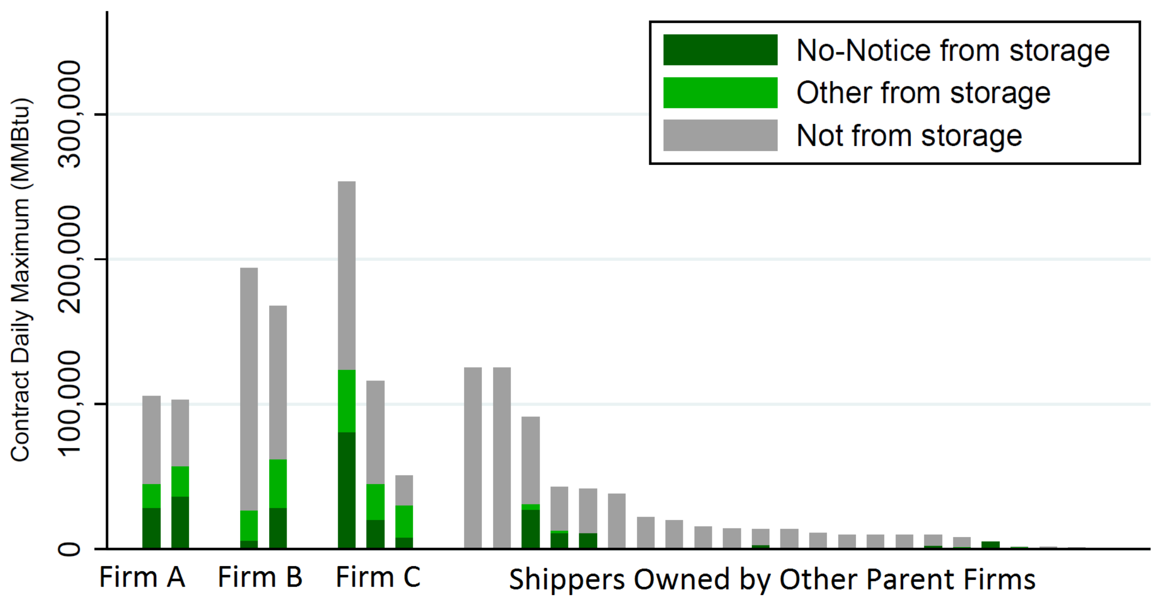


Figure 12: Aggregate downscheduling and contract positions of Firms A and B over time for the segment between the Oxford and Cromwell compression stations, in which six of the ten nodes that downschedule the most are located. These two firms' holdings of NN from TE contracts roughly correspond to an upper bound on the amount of downscheduling that occurs in this segment. Note: "All Contracts" is truncated at 350,000 in this series of charts for ease of presentation. (return)

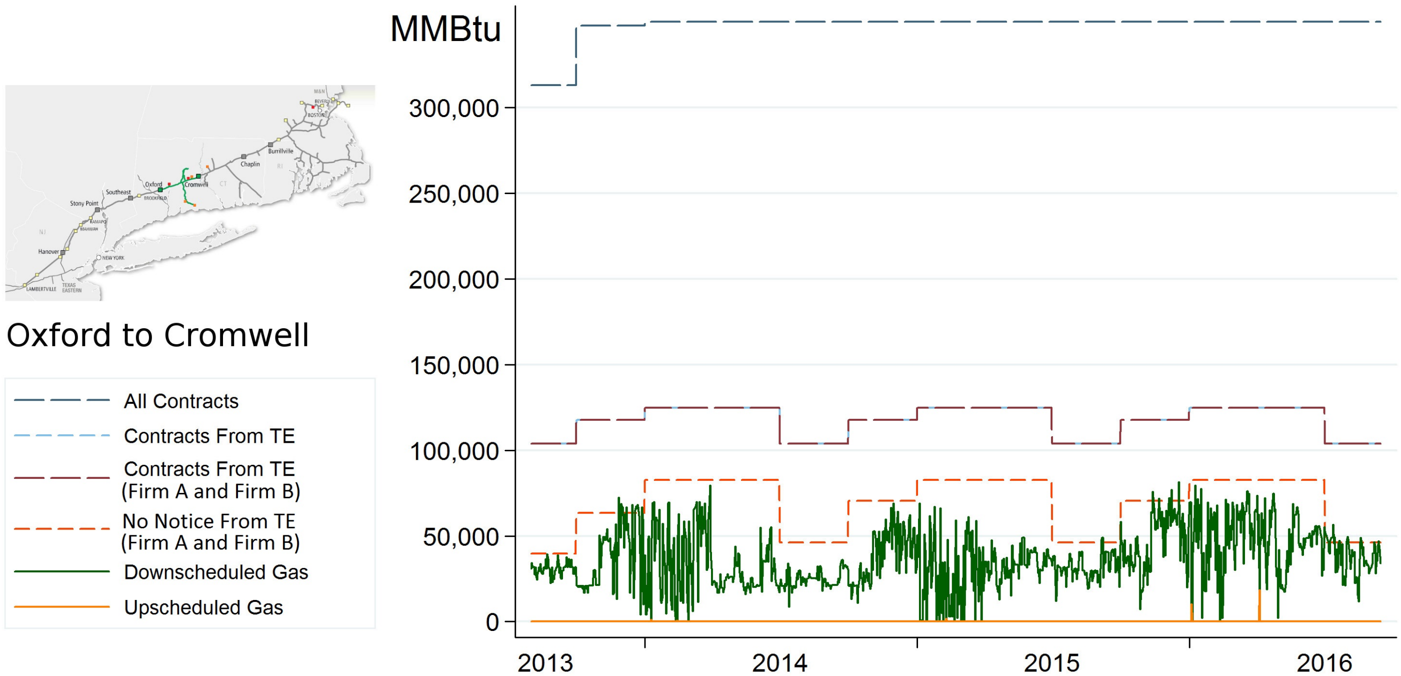


Figure 13: Aggregate downscheduling and contract positions of Firms A and B over time for the segment between the Cromwell and Chaplin compression stations, in which two of the ten nodes that downschedule the most are located. The level of downscheduling behavior is of roughly the same order of magnitude as these two firms' holdings of NN from TE contracts delivering gas to this segment. (return)

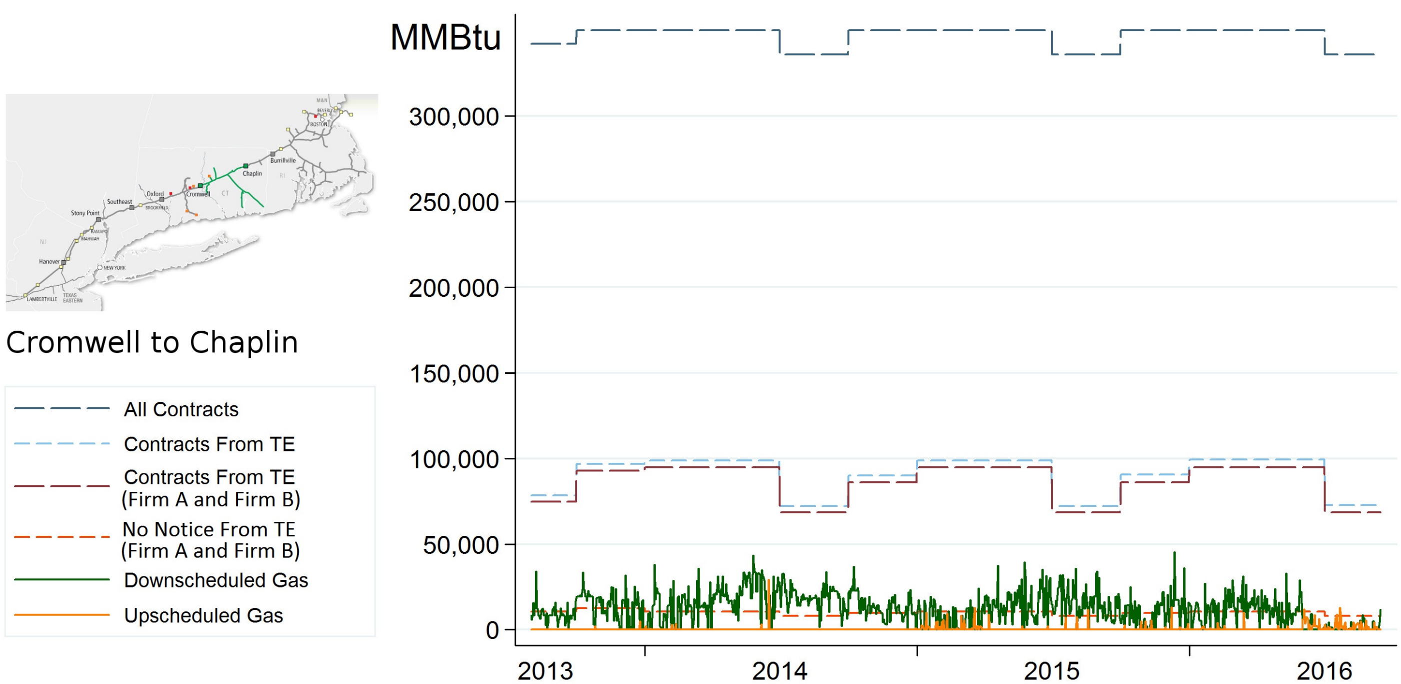


Figure 14: Aggregate downscheduling and contract positions of Firms A and B over time for the pipeline's "J System," which includes the Boston metropolitan area and serves many large electricity generators in addition to heating demand, and in which two of the ten nodes that downschedule the most are located. There is no obvious relationship here between downscheduling and NN from TE contracts held by Firms A and B. While they are not particularly noteworthy, for completeness we have made equivalent graphs for the other seven segments of the Algonquin pipeline available in the Online Data Appendix. (return)

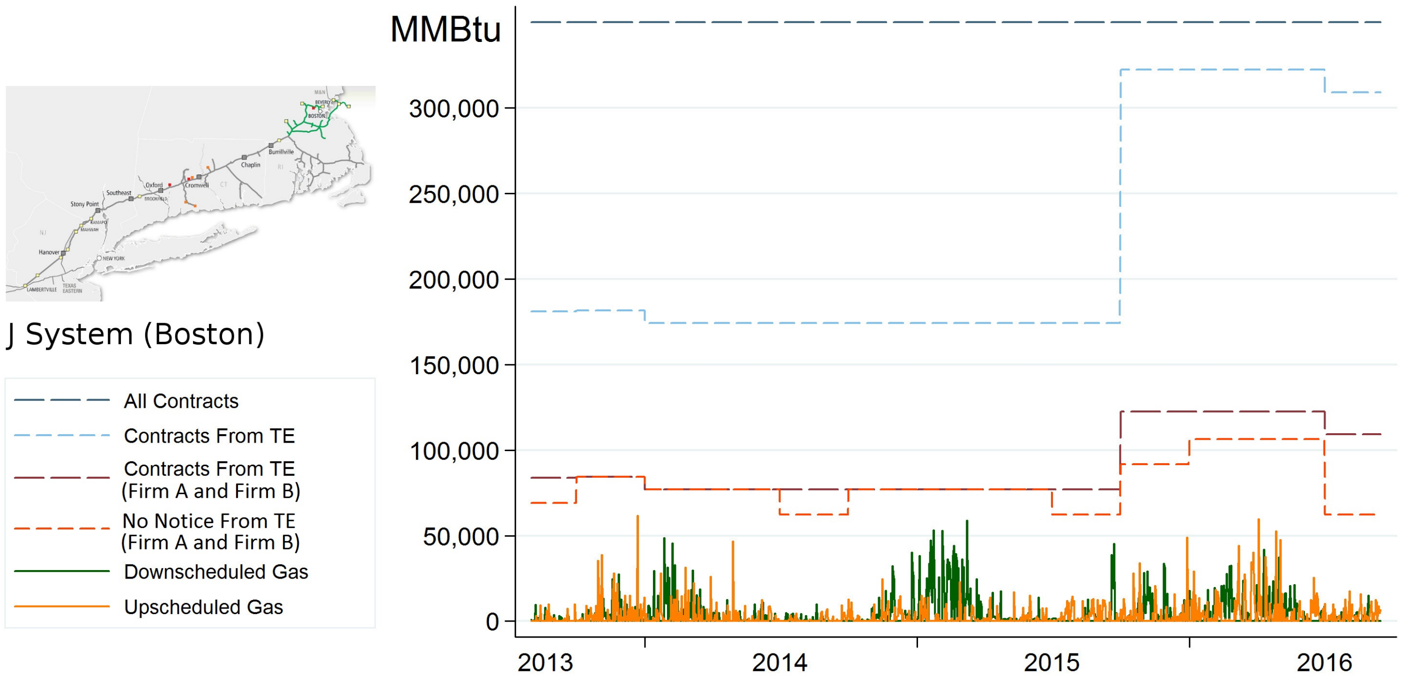


Figure 15: Distribution of hourly wholesale electricity prices from our re-construction of the market using generators' actual bids (left) and using our matching-resampling procedure with the actual gas price (right). Note that distributions are truncated at $400/MWh for readability. (return)

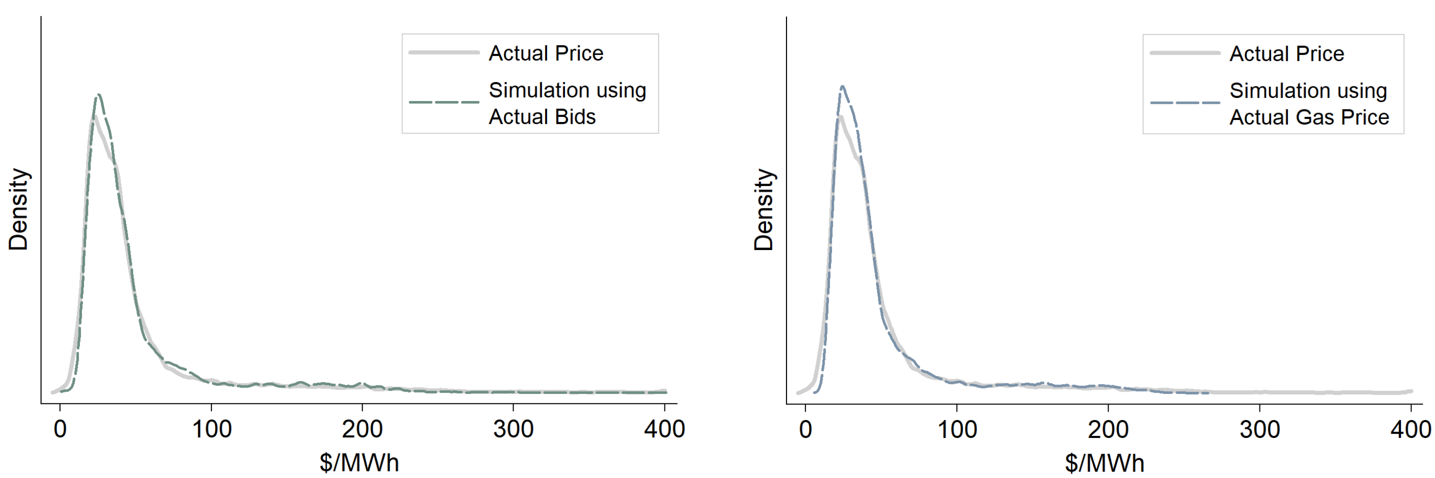


Figure 16: Left: The distribution of prices simulated using the counterfactual gas price is skewed left from that of actual prices and of the simulation using the actual gas price. Note that the distribution of actual prices is truncated at \$400/MWh for readability. Right: The distribution of estimated average annual electricity costs borne by ratepayers due to withholding activity. (return)

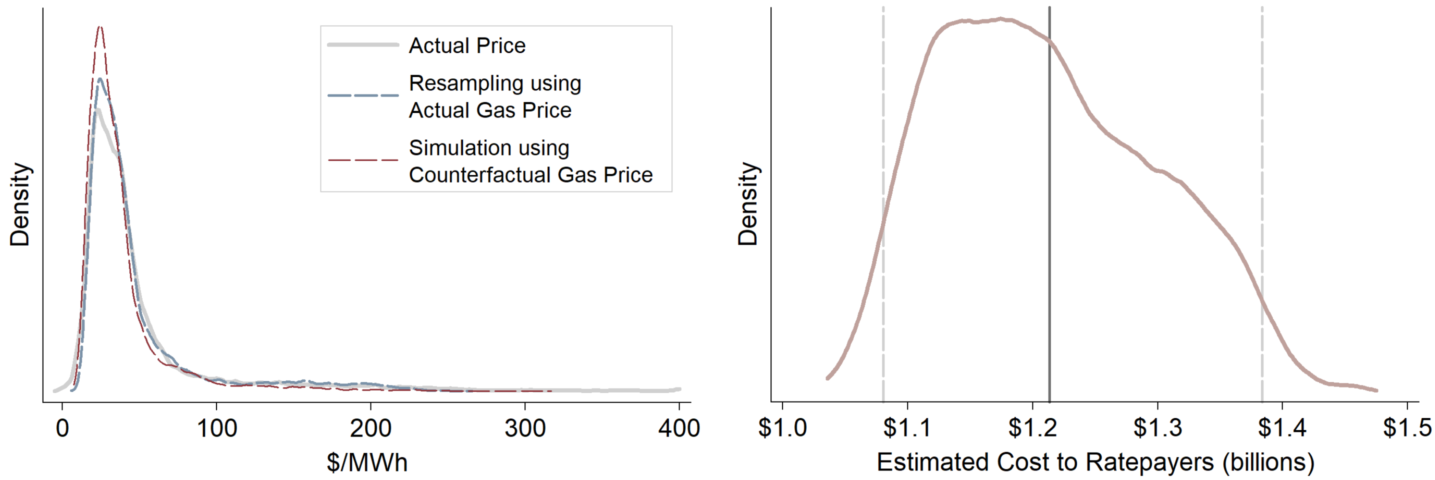
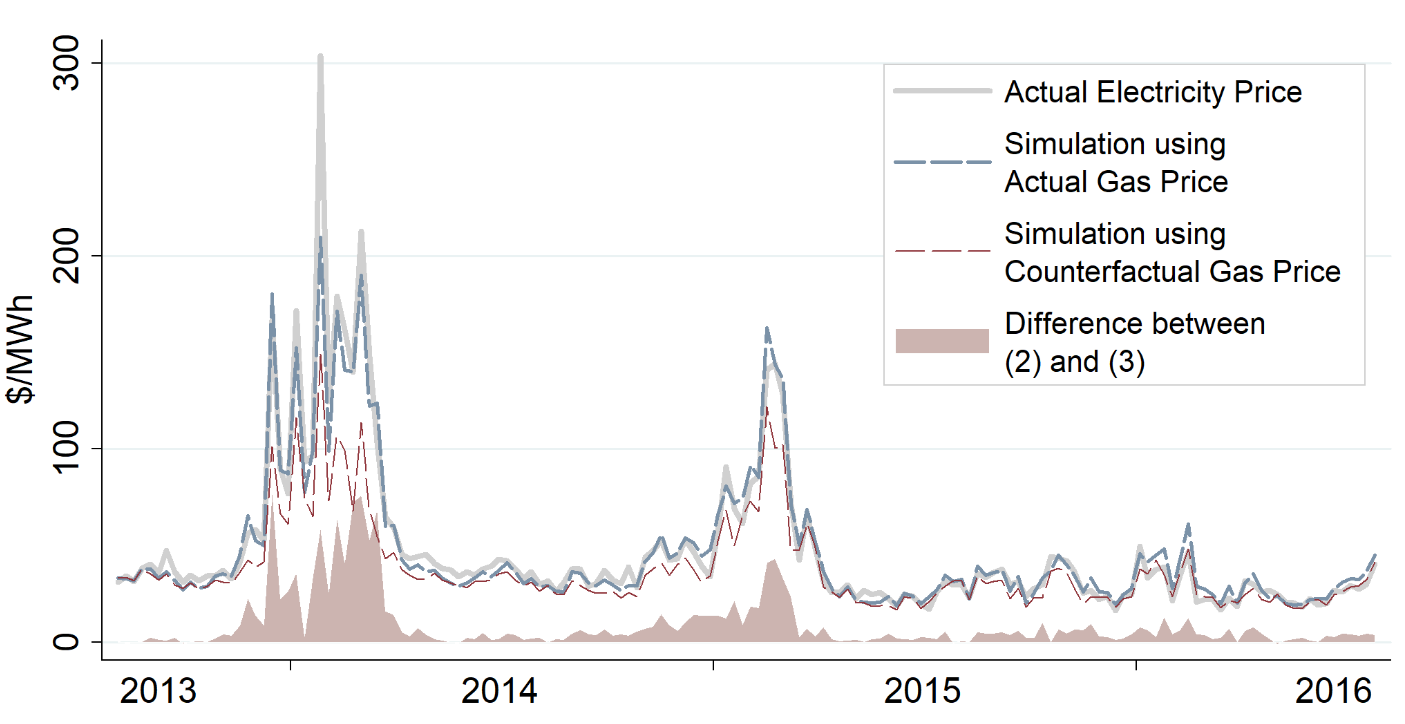


Figure 17: The strongest impacts of pipeline capacity withholding were realized during the winters of 2013-14 and 2014-15. (return)



Appendix

A.1 Robustness Checks

Table A1: Robustness check using the actual revenue sharing rule for each state. We consider the revenues retained by the LDC to focus on firm incentives (1% for CT, 10% for MA, 17% for RI) and take the inverse (1 for CT, 0.1 for MA, 0.0588 for RI) to facilitate interpretation of the interactions. The resultant variable ("Sharing") captures the feature that the incentive to withhold may be comparatively 10 times larger than the incentive to use capacity efficiently in the upstream market in Connecticut versus Massachusetts. Results are highly similar in sign and magnitude to our main set of regressions, where we elected to use a binary indicator for Connecticut to further simplify presentation. Specifications using other transformations of the pass-through parameter produce results that are consistent with our hypotheses, but generally much more difficult to interpret. (return)

	(1)	(2)	(3)	(4)	(5)
LDC	585.0**	-23.22	315.1		57.00
	(240.6)	(85.03)	(207.9)		(69.45)
Sharing		68.66		278.0	68.66
		(97.85)		(412.6)	(97.85)
Contracts			0	-0.0360***	0
			(.)	(0.0119)	(.)
LDC×Sharing		1462.6**			255.4
		(714.1)			(520.9)
LDC×Contracts			0.0502		-0.0366***
			(0.0521)		(0.0122)
Sharing×Contracts				0.457***	0
				(0.131)	(.)
LDC×Sharing×Contracts					0.454***
					(0.134)
N	133,029	133,029	133,029	133,029	133,029

	(1)	(2)	(3)	(4)	(5)
MW	1012.1**	-79.05	322.9		-72.08
	(445.2)	(97.80)	(386.9)		(65.01)
Sharing		2.759		324.5	5.041
		(21.07)		(512.1)	(34.56)
Contracts			-0.00259	-0.0365***	-0.00136
			(0.00271)	(0.0121)	(0.00182)
MW×Sharing		1804.7**			532.3
		(853.4)			(613.0)
MW×Contracts			0.156		-0.0266*
			(0.116)		(0.0143)
Sharing×Contracts				0.454***	-0.0150
				(0.134)	(0.0258)
MW×Sharing×Contracts					0.459***
					(0.135)
N	109,152	109,152	109,152	109,152	109,152

All regressions include quarter fixed effects and controls for temperature and day of week

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Robustness check using a continuous variable for MW owned by the node operator's parent firm ("MW_C"). We understand these counterintuitive results to be driven by the fact that Firm A downschedules more than Firm B, yet Firm B owns more generation capacity than Firm A, and we take this as suggestive evidence that the direct incentive pathways for merchant unregulated capacity (which Firm A owns and Firm B does not) are stronger than the indirect ones for regulated capacity. (return)

	(1)	(2)	(3)	(4)	(5)
MW_C	-0.0107	0.126***	-0.138		-0.00800
	(0.194)	(0.0464)	(0.157)		(0.0212)
CT		2503.0*		274.4	867.0
		(1329.7)		(465.6)	(928.3)
Contracts			0.0387	0.00225	-0.00280
			(0.0545)	(0.00346)	(0.00262)
MW_C×CT		-1.695			-0.634
		(1.082)			(0.737)
MW_C×Contracts			0.0000365		0.0000182***
			(0.0000345)		(0.00000482)
CT×Contracts				0.417***	0.612***
				(0.122)	(0.0751)
MW_C×CT×Contracts					-0.000384***
					(0.0000830)
N	109,152	109,152	109,152	109,152	109,152

All regressions include quarter fixed effects and controls for temperature and day of week

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Our results are robust to using MW of merchant unregulated generation owned by the node operator's parent firm ("Merchant") as the dependent variable of interest. Coefficient estimates are characteristically similar in sign and significance but smaller in magnitude as the continuous variable for MW owned has a larger scale than a binary one. For example, the interpretation of the triple interaction in column (5) would be that for each MW of generation owned by the parent firm of a node in Connecticut, that node will downschedule an additional .004 MMBtu on average for each 1 MMBtu of NN from TE contracts they own (additional to the effects of the other six variables and interactions). (return)

	(1)	(2)	(3)	(4)	(5)
Merchant	0.563	-0.0950***	0.0129		-0.0748***
	(0.871)	(0.0290)	(0.301)		(0.0158)
CT		253.4		274.4	-120.2**
		(188.8)		(465.6)	(56.18)
Contracts			0.00882	0.00225	0.00201
			(0.0133)	(0.00346)	(0.00354)
Merchant×CT		45.75**			26.99*
		(22.85)			(15.88)
Merchant×Contracts			0.00385		-0.0000315
			(0.00248)		(0.0000283)
CT×Contracts				0.417***	0.176***
				(0.122)	(0.0462)
Merchant×CT×Contracts					0.00398***
					(0.000952)
N	109,152	109,152	109,152	109,152	109,152

All regressions include quarter fixed effects and controls for temperature and day of week

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Robustness check using contracts of any type delivering gas from any point of origin ("Contracts$_A$") instead of just "no notice" contracts delivering gas from Texas Eastern. Coefficients on variables involving contracts are characteristically similar in sign and significance, but generally smaller in magnitude, suggesting that "no notice" contracts from Texas Eastern are indeed particularly useful for downscheduling. (return)

	(1)	(2)	(3)	(4)	(5)
MW	1012.1**	101.5**	-155.6		20.78
	(445.2)	(43.46)	(480.7)		(37.57)
CT		2.959		-612.3***	-53.37***
		(20.81)		(203.4)	(19.74)
Contracts$_A$			-0.00219*	-0.00124	-0.00233*
			(0.00122)	(0.00113)	(0.00124)
MW×CT		1623.8**			-656.3**
		(768.1)			(260.6)
MW×Contracts$_A$			0.0695		0.00366
			(0.0459)		(0.00230)
CT×Contracts$_A$				0.153***	0.00229
				(0.0197)	(0.00401)
MW×CT×Contracts$_A$					0.149***
					(0.0168)
N	109,152	109,152	109,152	109,152	109,152

All regressions include quarter fixed effects and controls for temperature and day of week

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Robustness checks using NN from TE contracts held by Firms A and B only ("Contracts$_{EA}$") and held by other firms only ("Contracts$_{-EA}$"). The coefficient on Contracts$_{EA}$ is omitted in column (5) of the first table because all of Firm A's and Firm B's NN from TE contracts either deliver to a node they operate (thus activating the MW indicator) or to a node in Connecticut. The coefficient on the triple interaction in the second table is omitted because only Firms A and B hold NN from TE contracts that deliver to nodes in Connecticut operated by a firm that owns significant generation capacity. (return)

	(1)	(2)	(3)	(4)	(5)
MW	1012.1**	101.5**	345.0		5.054
	(445.2)	(43.46)	(368.0)		(24.98)
CT		2.959		381.9	1.872
		(20.81)		(442.9)	(20.83)
Contracts$_{EA}$			0.0376	0.0191***	0
			(0.0908)	(0.00464)	(.)
MW×CT		1623.8**			477.4
		(768.1)			(552.1)
MW×Contracts$_{EA}$			0.133		0.0189***
			(0.139)		(0.00480)
CT×Contracts$_{EA}$				0.403***	0.0315
				(0.119)	(0.0194)
MW×CT×Contracts$_{EA}$					0.366***
					(0.111)
N	109,152	109,152	109,152	109,152	109,152

	(1)	(2)	(3)	(4)	(5)
MW	1012.1**	101.5**	1021.6**		86.49**
	(445.2)	(43.46)	(457.0)		(41.72)
CT		2.959		1484.8**	-15.74
		(20.81)		(690.3)	(12.08)
Contracts$_{-EA}$			-0.00313	-0.00471*	-0.00321
			(0.00269)	(0.00273)	(0.00269)
MW×CT		1623.8**			1634.2**
		(768.1)			(768.2)
MW×Contracts$_{-EA}$			-0.0623**		-0.00476
			(0.0306)		(0.00380)
CT×Contracts$_{-EA}$				-0.526*	-0.000313
				(0.288)	(0.00324)
MW×CT×Contracts$_{-EA}$					0
					(.)
N	109,152	109,152	109,152	109,152	109,152

All regressions include quarter fixed effects and controls for temperature and day of week

Standard errors in parentheses (clustered at the node level)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: There is limited generation capacity and thus limited demand from generators in the spot market for natural gas within and upstream of segments 5 and 6, where 8 of the 10 most-withholding nodes are located. (return)

Segment	Segment Name	Generation Capacity	Upstream Capacity
1	Lambertville to Hanover	0	0
2	Hanover to Stony Point	0	0
3	Stony Point to Southeast	0	0
4	Southeast to Oxford	6.3	6.3
5	Oxford to Cromwell	441.6	447.9
6	Cromwell to Chaplin	761.2	1209.1
7	Chaplin to Burrillville	861.8	2070.9
8	Burrillville to J and G Systems	447.8	2548.7
9	J System (Boston)	3430.6	5979.3
10	G System (Rhode Island)	2583.4	5132.1

Source: SNL and EIA Energy Mapping System

A.2 Imbalance penalties

In order to promote the efficient utilization of pipelines, FERC regulations require interstate pipeline companies to charge two types of imbalance penalties. The first is called an OFO imbalance penalty, which is assessed when shippers cause a physical imbalance on the system on a day when there is an "Operational Flow Order" (OFO) in effect. The pipeline company issues an OFO by electronically notifying all of its customers that the pipeline has reached or is about to reach its capacity constraint at at least one bottleneck on the system (usually a compression station). On a congested day, this usually happens well before the gas day starts, but it may be issued or updated during the gas day. Once this notice has been issued, any shipper that causes a physical imbalance in the system by withdrawing in excess of 2% more or less gas at their delivery node than they had injected at their receipt node is assessed a penalty equal to three times the Algonquin Citygate price times the quantity of the infraction.[73] This penalty is so severe because causing a physical imbalance on the system can be extremely harmful on days when the pipeline is near capacity constraint. If a customer withdraws more than they inject, other customers will lose service and be unable to withdraw the quantity they had scheduling using their rights. If a customer injects more than they withdraw, pressure levels could build up at a bottleneck to unsafe levels.

The second is a accounting imbalance penalty, which is assessed symmetrically for deviations in either direction between the quantity of gas the shipper had scheduled to flow through the pipeline and what they actually flowed on a monthly basis. Specifically, shippers are assessed a penalty that ranges from 1.1 to 1.5 times the Algonquin Citygate price times the quantity of the infraction on scale that increases

[73] See Spectra (2016) Section 26

with the size of the infraction.[74] This penalty is much less severe because it does not affect the safety and reliability of the the system, but still substantial—especially for LDCs, who are regulated such that they can pass the cost of gas through to their customers but cannot pass through any imbalance penalties they incur.

A.3 Alternative explanation for spatial heterogeneity

Alongside stronger revenue-sharing rules in Connecticut, lower demand for natural gas for electricity generation in the state may interact with capacity scheduling frictions to encourage greater withholding there. Contracts for capacity guarantee the holder's ability to transport gas to the delivery node listed in the contract. However, they may also be used to deliver gas to other nodes if capacity is available. This flexible service, termed "secondary" nominations, can be reliably used to transport gas to nodes in close proximity to the contracted delivery node (*i.e.* in the same segment) or upstream of that node. However, secondary nominations attempting to transport gas further downstream than the contracted delivery node are considerably less reliable. The capacity may be physically unavailable due to a bottleneck further downstream when the pipeline is highly congested during the winter, the pipeline company is not obligated to deliver to secondary locations and has incentive to err toward caution to reliably supply primary nominations, and uncertainty regarding approval of downstream secondary nominations persists until at least the timely cycle and potentially into the scheduling period. Therefore, the further downstream a contract's primary delivery point is, the more valuable it will be for spot market sales of bundled gas and transportation to generators.

The persistent uncertainty is highly relevant to gas-fired generators because the

[74] To be specific, the penalty is $1.1 \times P_{ACG} \times Q$ for deviations between 5% and 10%, $1.2 \times P_{ACG} \times Q$ for 10-15%, $1.3 \times P_{ACG} \times Q$ for 15-20%, $1.4 \times P_{ACG} \times Q$ for 20-25%, and $1.5 \times P_{ACG} \times Q$ for deviations in excess of 25% (see Spectra (2016) Section 25)

deadline to submit bids to the wholesale electricity market is at 10am the day before, which is three hours before the pipeline's timely cycle deadline.[75] A generator bidding based on an uncertain downstream secondary nomination therefore accepts some risk of being unable to acquire the gas needed to meet its load obligation. Dual-fired generation units can hedge this risk by burning petroleum if necessary, potentially at a net economic loss in the short term, and will weigh the likelihood of this outcome in their bidding strategies. Gas-only units have no such flexibility and face severe penalties from the ISO if they fail to meet their bidden load obligations. Contracts delivering gas upstream of the majority of gas-fired generation are thus less valuable for selling gas to generators, either directly or through the capacity release market, because of both a reduced quantity of gas that is actually delivered and uncertainty regarding when deliveries will be successful.

As shown in Table A6, there is significantly less gas-fired generation capacity connected to fifth and sixth segments of the Algonquin pipeline (where eight of the 10 suspect nodes are located) in comparison to the segments further downstream that serve Massachusetts and Rhode Island. When an LDC-affiliated shipper has excess capacity delivering gas to these segments after supplying their heating demand, if they want to sell gas to generators they will need to rely primarily on secondary nominations and accept some uncertainty of making the sale and a reduced price that incorporates generator uncertainty. However, by scheduling phantom transportation to their own primary delivery locations, they can artificially further constrict an existing bottleneck on the pipeline to reduce overall supply of gas to the region with certainty.

[75] Bids must be submitted to both the day-ahead and real-time electricity markets by 10am the day before the operating day. Bids to the real-time market may subsequently be adjusted; however, 95% of the total energy is traded in the day-ahead market and prices in the real-time market generally closely track those of the day-ahead market.