Raising Rivals' Costs: Vertical Market Power in Natural Gas Pipelines and Wholesale Electricity Markets

Levi Marks, Charles Mason, Kristina Mohlin

and Matthew Zaragoza-Watkins

January 1, 2017

Abstract

In recent years, New England has experienced severe, contemporaneous price spikes in its natural gas and wholesale electricity markets. Although these spikes are commonly attributed to limited pipeline capacity serving the region, we demonstrate that they have been exacerbated by firms with long-term contracts for pipeline capacity scheduling for deliveries without actually flowing gas. We analyze firms' scheduling patterns on the Algonquin pipeline and identify the institutional conditions that enable and incentivize this capacity-withholding behavior. We find that some firms are able to offset the opportunity cost of unused capacity by increasing the price of the gas they do sell in the spot market and by increasing the interconnected wholesale electricity price, which increases the revenues of infra-marginal generation resources owned by their parent energy companies. Finally, we employ an economic dispatch model to estimate the welfare losses, emissions consequences, and distributional impacts of capacity withholding over the period from 2014-2016.

1 Introduction

In recent years, New England has experienced severe, contemporaneous price spikes in the wholesale electricity and natural gas markets. During the polar vortex of 2013-14, natural gas prices reached record levels, as the basis differential between the Algonquin Citygate trading hub in Massachusetts and Henry Hub in Louisiana (the principal upstream natural gas trading point in the US) regularly exceeded $30/MMBtu and reached a record high of $74/MMBtu on January 22, 2014. These extreme price spikes have been commonly attributed to limited pipeline capacity serving New England (see e.g., EIA 2013, Rose *et al.* 2014). This "scarce capacity" narrative has also been used to support recent proposals for expanding natural gas pipeline capacity in New England (see ICF 2015).

Limited pipeline capacity is indeed largely responsible for these extreme prices. However, aside from the direct effect, scarce capacity also enables a form of market power that exacerbates the high prices. On cold days when the pipeline is at or near its physical capacity constraint, firms holding contracts for pipeline capacity can further restrict supply to the region by scheduling to use capacity without actually flowing gas. When the pipeline is congested, other shippers are unable to respond by increasing their supply to the wholesale gas spot market, so total supply is decreased and spot market prices increase. While most firms operating on the Algonquin pipeline have little incentive to sacrifice spot market sales by withholding capacity, those that also own generation assets in the region have an incentive to increase gas prices to raise electricity prices.[1] Furthermore, spatial variation in the concentration of electricity generators makes some contracts considerably less valuable for selling gas to generators through the spot market, yet just as valuable for tying up capacity to increase gas and electricity prices.

The purpose of this paper is to test the hypothesis that some companies systematically withheld pipeline capacity from the New England secondary pipeline capacity markets during peak demand periods in the years 2013-2016. To do this, we investigate the scheduling patterns of the companies that owned capacity on the Algonquin pipeline. We compare differences in scheduled and delivered natural gas quantities across firms with stronger and weaker incentives to exercise market power. Additionally, using an instrumental variables approach, we estimate the impact of these withholdings on the natural gas spot price. Finally, using a dispatch model, we estimate the impact in the downstream wholesale electricity market and calculate the associate welfare loss and distributional impacts.

Our results show that two firms with large holdings of both contracts for pipeline capacity as well as inframarginal electric generating capacity systematically withheld capacity from the secondary pipeline capacity markets. This behavior reduced the amount of gas available to generators on the Algonquin pipeline by roughly 14% on average during winter months. The estimated impact on downstream electricity prices were on average XX. This implies ZZ of transfers from ratepayers to electric generators and a deadweight loss

[1] In New England's wholesale electricity market, a natural gas-fired plant is usually the marginal generation resource, so gas and electricity prices are closely tied. All generators receive the same wholesale price determined by the marginal generator's bid, net a small level of spatial variation due to transmission costs.

estimated to YY.

The remainder of this paper is structured as follows. Section 2 reviews the related literature and section 3 describes the market features and rules for natural gas transportation relevant to our analysis. Section 4 presents the theoretical framework which forms the basis for our empirical strategy. Section 5 presents the empirical analysis of companies' withholding patterns. Section 6 presents the simulations where we estimate the price elasticity of gas demand and a resulting counterfactual daily price series for the Algonquin citygate gas price, and lastly the dispatch model and results for the impact of the higher gas prices in the downstream electricity market. Section 7 discusses the results and concludes the paper.

2 Literature Review

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 a significant share of total assets across the network may have significant local market power (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 is enmeshed in 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) lay the groundwork for 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 find 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. Furthermore, the downstream generator is incentivized to acquire as many of the physical rights as possible so they can simultaneously decide transmission capacity and production in the downstream node.

Joskow and Tirole's analysis provides an interesting parallel to our setting, where firms operating downstream of the pipeline's bottleneck own transportation rights and are, under some conditions, incentivized to use those rights to tie up capacity rather than to transport the resource. Interestingly, Joskow & Tirole (2000) advocate for adapting the capacity release regulations of the gas transportation 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 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 manipulate prices emerges from the physical capacity constraint in our setting, the firm's primary incentive to withhold capacity comes from vertical integration in the gas and electricity markets. One commonly-studied concern in the litertaure on vertical power is *foreclosure* (sometimes also termed "raising rivals costs"), wherein a vertically integrated firm instructs its upstream component to restrict sales of a necessary input to production to it's downstream component'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 recieved 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 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 The Market for Natural Gas Transportation

3.1 The pipeline capacity markets

In the 1980s and early 90s a series of reforms initiated by the Federal Energy Regulatory Commission (FERC) effectively decoupled the gas transportation service provided by the pipeline from the buying and selling of the physical commodity. The reforms required pipelines to offer transportation-only services to all types of customers. This enabled local distribution companies (LDCs) and end users to purchase gas directly from wellheads and market hubs and ship it to themselves using contracts for pipeline capacity. Under this structure, the pipeline companies themselves do not own gas at any point in the transportation process. Their revenues come only from selling contracts for pipeline capacity at FERC-regulated prices intended to guarantee a fair rate of return. The entities that purchase contracts, which can be either LDCs, electricity generation or industrial end users, or independent marketers, are known as shippers. While the prices shippers pay to initially purchase contracts from the pipeline company are regulated, they are able to sell short-term usage of their contracts to other shippers on a secondary "capacity release market" at unregulated prices.

Shippers use capacity acquired either through the primary or secondary markets to transport gas they have purchased at a wellhead or market hub receipt point to a different segment of the pipeline, and then either use the gas themselves or market it to others. Electricity generators and industrial end users of gas typically burn all the gas they ship. In contrast, LDCs typically use their contracts to ship gas beyond what's needed to supply their residential and business heating customers and compete with independent marketers to sell it to other demanders in the delivery region. In New England and many other parts of the US, the end users purchasing gas from independent and LDC-affiliated marketers in these spot transactions are usually gas-fired electricity generators, who less commonly hold long-term contracts for capacity because high variability in the electricity market makes their day-to-day operating schedule much less consistent.

The prices for these spot transactions within the delivery region incorporate the wellhead price of the gas, the cost of the contracts for capacity used to transport it, and the shadow price of 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 contracts for capacity are regulated but the prices for secondary capacity trades and spot natural gas transactions are not, owners of long-term contracts are able to extract rents from scarce capacity (Oliver *et al.* , 2014).

When FERC first decoupled transportation from sales in the early 1990s and created a capacity release program, the Commission was concerned that the extent of competition in the secondary market would not be sufficient and put in a rate ceiling. However, in 2008 FERC removed the price ceiling for short-term capacity release transactions of one year or less with the motivation that the rate ceiling worked against the interests of short-term shippers, because with the rate ceilings in place, a shipper looking for short-term capacity on a peak day who was willing to offer a higher price in order to obtain it, could not legally do so (FERC, 2008). The challenge facing FERC, illustrated by its removal of the price ceiling for the secondary market, is striking the right balance between allowing for prices to signal scarcity while protecting against the risk of market power and price manipulation. This analysis is a demonstration of the difficulties in striking that balance.

3.2 Scheduling for pipeline capacity

Pipeline capacity is scheduled on a daily basis, meaning that for each gas day (which runs from 9am to 9am the following day), shippers must choose a total quantity of gas to flow over the entire gas day at an approximately equal rate. The shipper "nominates" capacity by electronically submitting a proposed daily quantity of gas to the pipeline operator. Each nomination requires an explicit receipt point where the gas will enter the pipeline, a delivery point where the gas will be withdrawn, and ownership of capacity (acquired through either the primary or secondary markets) covering the entire path between the two points. The scheduling period for each gas day is broken into periods called "cycles". In order to have priority when they schedule capacity, the shipper must submit their initial nomination by the end of the "timely" cycle at 1pm the day before. For most pipelines, adjustments can be made during one "late" cycle (which closes at 6pm the day before) and two or three "intraday" cycles that take place during the actual gas day.

The Algonquin pipeline is fairly unique in that it allows shippers to make adjustments to their nomination schedules on an hourly basis at any time during the scheduling period, subject to their (the pipeline operator's) ability to make the change given the nominations of other customers and the operating conditions of the system. If the shipper is transporting gas under a special kind of contract for capacity called "no notice," the pipeline guarantees their ability to make changes to their schedule at any point during the scheduling period, which they accomplish either by bumping secondary capacity or holding reserve capacity at the outset.[2] If at any point in the gas day there is unused space on the pipeline after all holders of primary and secondary capacity have made their nominations, the pipeline company can sell the extra capacity as "interruptible" service, meaning the pipeline will stop the flow of gas if a primary or secondary capacity holder increases their nomination.[3]

3.3 Imbalance penalties

If some shippers draw gas from the pipeline in excess of their nomination, other customers will not be able to draw the gas they had scheduled. Conversely, if shippers inject more gas into the pipeline than they had scheduled (or inject the scheduled amount but withdraw out less), pressure may build to unsafe levels. [4] In order to ensure the pipeline operates smoothly, if the pipeline approaches

[2]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 no notice contracts would be needed in the new market structure to ensure they could reliably serve unexpected fluctuations in demand.

[3]Leftover capacity for interruptible service is almost certainly very rare on the Algonquin pipeline–still need to confirm this, though.

[4]The pipeline appears somewhat concerned but generally much less concerned about shippers taking less gas than they had scheduled. The shippers do not have incentives to take less than they have scheduled under normal conditions, the pipeline company doesn't care how much of the pipeline actually gets used, and it doesn't have immediate consequences for other customers like taking too much does. The language of OFO warnings changed to emphasize penalties for deviations in either direction rather than just taking too much in

its capacity constraint the pipeline company will issue an Operational Flow Order (OFO) warning that authorizes severe imbalance penalties.[5] While an OFO is in effect, any shipper that withdraws more or less than a 2% deviation from what they are having injected at the receipt point over the course of the gas day is charged a penalty equal to three times the cost of the gas of the deviation. Importantly, this penalty is assessed based on creating a physical imbalance in the system: If the shipper draws less gas from the pipeline than they had scheduled but also has less gas injected at the receipt point, they are not assessed the OFO penalty. The pipeline company also assesses a separate monthly imbalance penalty to firms whose aggregate actual flows of gas for the month differ from their aggregate scheduled flows. These penalties come into effect for deviations in excess of 5% and are much less severe than OFO penalties, ranging from 1.1 to 1.5 times the cost of the gas depending on the severity of the infraction.

In the context of our analysis it is useful to consider what a shipper who wishes to withhold pipeline capacity must do to avoid both types of imbalance penalties – both to avoid the charges themselves and to avoid arousing the suspicion of regulators. Because monthly accounting imbalance penalties are based on the end-of-day scheduled quantities, shippers can avoid them by reducing their scheduled daily quantity of gas to what they actually flowed in the final hours of the gas day. As the process of re-allocating capacity to another shipper takes 3 hours, a well-timed reduction will not leave enough time for

2013.

[5]Need to include a statistic about how common these are on winter days (pretty much every day)

another shipper owning secondary capacity to increase their nomination. The shipper can avoid physical OFO penalties by sourcing gas from a receipt point that is complicit in not actually injecting gas into the pipeline. An independent gas producer or storage site could fill this role, but such an accomplice is actually not needed in the case of the Algonquin pipeline as its operational structure already contains a built-in system for shippers to avoid injecting unused gas. Algonquin receives about a third of its gas from the Texas Eastern pipeline, a much longer pipeline that bring gas from Texas and other productive regions in the South to the Northeast. Unlike the Millenium and Tennessee pipelines, which are Algonquin's other two major sources of gas at its Western end (See Figure 3), the Texas Eastern pipeline possesses a great deal of storage capacity in the form of depleted reservoirs in Appalachia. The other notable feature of Texas Eastern is that is owned by the same parent energy company as the Algonquin pipeline, Spectra. This enables a high degree of coordination wherein Texas Eastern does not draw gas from its house-managed storage sources automatically but instead uses them to balance pressure between itself and the Algonquin pipeline. The consequence of this is that if a shipper on Algonquin sources their gas from storage on Texas Eastern, they are able to withdraw less than they scheduled without causing a physical imbalance and incurring corresponding OFO penalties.

4 Theoretical framework

In this section we model the incentives of firms with vertical arrangements in the pipeline gas transport and wholesale electricity markets, which are connected via the wholesale gas market. Firms that operate in both markets have a clear incentive to restrict supply in the pipeline transport market during periods of scarcity, in order to increase the wholesale gas price, which raises rivals costs (and captures rents) in the electricity market. The incentive for these firms to withhold gas, rather than sell it in the wholesale market, is likely amplified by a common form of regulation which requires LDCs to dividend, back to their ratepayers, most of the profit from selling wholesale gas. Finally, we consider how the spatial nature of the pipeline network determines where in the system firms are likely to withhold capacity, noting that firms will withhold capacity at points where it is likely to have the greatest impact of the wholesale electricity price.

Based on the institutional features discussed above, we envision three types of firms, operating in four relevant markets. The markets are: a market for pipeline transport (T), which is used to deliver gas to a wholesale gas market (W), in which LDCs sell gas to electric generators, which gas those generators combust to supply electricity to a wholesale electricity market (E), in which some generating units are gas-fired and others are not. There is also a retail gas market (R), in which LDCs serve retail gas demand. Type 1 firms own gas-fired electric generating units; they do not own pipeline transport capacity and operate solely in market E. Type 2 firms are gas LDCs; they are capacity holders in the pipeline transport market, in which they have some market power during periods of scarcity. They are required to use their pipeline capacity to serve gas demand in market R, and may use excess pipeline capacity to transport gas for sale in market W. When selling gas in market R they receive a certain regulated rate of return, which we model as a percentage α of their variable operating costs.[6] When selling gas in market W they receive a certain share β of the profit, with the remaining profit returning to their ratepayers. Finally, type 3 firms operate in all four markets; they are gas LDCs that also own electric generating units, which are not gas-fired. To simplify the presentation, we assume this firm owns electric generating capacity with zero marginal cost and fixed output, y_E^3. Like other LDCs they own pipeline transport capacity which they are required to use to serve demand in market R, and may use to sell gas in market W. Unlike other LDCs, their incentives derive from the interaction of their positions in markets T and E. To simplify, we suppose there is only one such firm.

Next, we solve for the equilibrium conditions in the downstream markets, beginning with the wholesale electricity market. To simplify, we assume market E has linear inverse residual demand for gas-fired generation (*i.e.*, net of deliveries from *inframarginal* generation, y_E^3, owned by the type 3 firm), with $p_E = a_E - Y_E$, where Y_E is the amount of gas-fired generation demanded. Gas-fired generation is supplied by an atomistic fringe of type 1 firms, which produce electricity using a technology with constant returns to scale, with respect to gas, which is the sole short-run variable input. For convenience, we choose units so that producing one unit of electricity requires one unit of

[6] The rate of return could be calculated on the basis of capital costs or operating costs. To simplify the presentation we abstract from the former and concentrate on the latter.

gas, which implies the competitive fringe has marginal cost, p_W, the cost of procuring gas in market W. Thus, the equilibrium condition in market E is $Y_E = a_E - p_W$.

As demand in market W derives entirely from gas-fired electric generation, market W also has linear inverse demand, with $p_W = a_E - Y_W$, where $Y_W = Y_E$ is the total quantity of wholesale gas demanded. The LDCs (*i.e.*, the type 2 firms and the type 3 firm) are oligopolists in the pipeline transport market, so they supply gas to market W at marginal cost, $p_u + p_T$, where p_u is the cost of procuring gas upstream of the pipeline and p_T is the price of transport service on the pipeline (*i.e.*, in market T). Thus, the equilibrium condition in market W is $Y_W^* = a_E - (p_u + p_T)$.

Demand for pipeline transport derives from markets R and W. Recall, LDCs are required to serve retail demand before using any pipeline capacity to serve demand in market W. Suppose demand for gas in market R is exogenous, inelastic and firm-specific, with $Y_T^r = \Sigma a_R^i = a_R$. Demand from the wholesale gas market is $Y_T^w = a_E - (p_u + p_T)$, which implies $Y_T = a_R + a_E - (p_u + p_T)$.

Next, suppose the marginal cost of providing pipeline transport can be expressed in terms of the ratio of used and unused capacity:

$$c(Y_T) = \frac{Y_T}{K - Y_T}$$

where K is exogenously determined pipeline capacity; this captures the idea that when there is sufficient slack in the system the market behaves "as if" it is competitive. With less slack, marginal cost is steeply convex.

Supply in market T is characterized by the first-order conditions of the oligopolistically competitive LDCs. Suppose, N of these firms, one of which is the type 3 firm, have symmetric residual pipeline capacity holdings, after serving market R. The profit function for each of the $N-1$ type 2 firms is:

$$\pi_i = (a_E + a_R - p_u - \Sigma y_T^{-i} - y_T^i)y_T^i - \frac{\Sigma y_T^{-i} + y_T^i}{K - \Sigma y_T^{-i} - y_T^i} y_T^i$$

and the first-order conditions are:

$$\frac{\partial \pi_i}{\partial y_T^i} = a_E + a_R - p_u - \Sigma y_T^{-i} - 2y_T^i - \frac{K(y_T^{-i} + 2y_T^i) - (y_T^i + y_T^{-i})^2}{(y_T^{-i} + y_T^i - K)^2} = 0;$$

manipulation yields:

$$\frac{(a_E + a_R - p_u)(Y_T - K)^2 + Y_T(Y_T - K)}{K} = y_T^i$$

This expression implicitly defines the reation fuctions for all type 2 firms in market T. However, the type 3 firm's profit maximizing strategy will be different, because each unit of gas delivered to the wholesale gas market decreases the downstream electricity price, which reduces that firm's revenues in the downstream electricity market, where it owns zero-marginal-cost generation, y_E^3. The type 3 firm therefore has an incentive to supply less capacity in the pipeline transport market. The firm's profit function is:

$$\pi_i = (a_E + a_R - p_u - \Sigma y_T^{-i} - y_T^i)y_T^i + (a_E + a_R - \Sigma y_T^{-i} - y_T^i)y_E^3 - \frac{\Sigma y_T^{-i} + y_T^i}{K - \Sigma y_T^{-i} - y_T^i} y_T^i$$

and the first-order condition for the type 3 firm is:

$$\frac{\partial \pi_i}{\partial y_T^i} = a_E + a_R - p_u - \Sigma y_T^{-i} - 2y_T^i - y_E^3 - \frac{K(y_T^{-i} + 2y_T^i) - (y_T^i + y_T^{-i})^2}{y^- i_T + y_T^i - K} = 0;$$

manipulation yields:

$$\frac{(a_E + a_R - p_u - y_E^3)(Y_T - K)^2 + Y_T(Y_T - K)}{K} = y_T^i$$

which implicitly defines the reaction function for the type 3 firm.

We can see by comparing the reaction function of the type 3 firm with those of the type 2 firms that $y_T^2 > y_T^3$ because of the higher implicit marginal cost of supplying pipeline capacity for firm 3 (*i.e.*, the lost revenue in market E). In response to the smaller deliveries from the type 3 firm, firms of type 2 supply greater capacity than they would in a symmetric equilibrium. Nevertheless, total supply Y_T is still lower by a factor of $y_E^3/2(n+1)$ compared to the case if no firms owned generation capacity in the downstream electricity market.

Incorporating a profit-sharing rule drives a wider wedge between the reaction functions of the different types of firms. Suppose (as is frequently the case in New England) that LDCs are allowed to keep only a fraction, β, of the profit from selling gas in the wholesale market, with the residual returned to their ratepayers. It's clear that applying β to the profit function of the type 2 firms will not change their first-order condition. For the type 3 firm, however, applying the profit-sharing rule increases the relative weight the type 3 firm places on profits in market E. In the extreme (*i.e.*, when $\beta = 0$), the firm chooses y_T^i to exclusively maximize profit in the wholesale electricity market.

Of course, the possibility of raising p_T by some form of market manipulation would also increase type 2 firm's profits, albeit to a smaller degree. In the absence of any consequence for manipulating the pipeline tariff, one might then expect to see *any* vertically integrated firm taking actions to drive p_T up. But such manipulations run the risk of attracting attention from the FERC, which can lead to penalties. In this way, all these firms are subject to disincentives to manipulate prices. But these disincentives are likely to be similar for any firm; since the pre-penalty gain from market manipulation is larger for the type 3 firm it then follows that we would expect such behavior to be more common for firm 3. In the setting we study, the manipulation arises from scheduling deliveries that contribute to the anticipated flow in the pipeline, and hence reduce the expected unused capacity – which drives up the price of pipeline services. The volume of such phantom deliveries that a firm might have to nominate will depend on background conditions, in particular the anticipated level of demand in markets R and W. Plausibly, this combined demand would be larger under inclement weather conditions, for example during periods of particularly cold temperatures. In such a setting, the expected costs from manipulating markets would likely be smaller (as it would be harder to detect the manipulation), raising the expected profit from that behavior.

The model sketched above articulates the differentially large incentive a firm that is vertically integrated into electricity markets and local distribution gas markets can have to manipulate the price of wholesale natural gas. 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 take-off points, which we refer to as points A and B. Point B lies farther along the pipeline than point A, and so we will occasionally term these points "upstream" (point A) and "downstream" (point B). There is one seller of type 3 located at each point; to keep the notation parallel, we will call the seller located at point A firm 3_A and the seller located at point B firm 3_B. There are end users located at both nodes, so there are markets of type R at each point; we presume the market at point B is larger than the market at point A (*e.g.*, point A might be Hartford and point B might be Boston). As above, sellers in the R market are obliged to meet all demand at that node.

One important distinction to the model above is that firm 3_B has the right to sell gas in the W market at A, whereas firm 3_A cannot sell gas in the W market at B.[7] We assume the demand structure is linear, as above, with the slight variation that we suppose the difference between electricity demand at A and B arises from different intercepts, but that the slopes are identical. To minimize notation we denote the intercept in market k as $a_k, k = A, B$. We also need to distinguish between the transportation cost to points A and B. As suggested by the model above, we write the unit transportation cost to point A as $\tau_A(x_A)$, where x_A is the spare pipeline capacity on the segment that terminates at A. The unit transportation cost to point B is $\tau_A(x_A) + \tau_B(x_B)$, where x_B is the spare pipeline capacity on the segment between A and B; since some gas is extracted at A, we assume $x_B > x_A$. Also, since the second component is non-zero, it follows that the marginal costs for the two type 3

[7] One can think of these firms holding contracts for delivery, with firm 3_A's contract guaranteeing delivery to point A, and firm 3_B's contract allowing delivery to point B or points upstream – *i.e.*, point A. With this interpretation, firm 3_B could choose to withdraw gas at either point, while firm 3_A would be obliged to remove gas at point A.

firms satisfy $m_{3_A} < m_{3_B}$.

With minor adaptation, the raising rivals' costs arguments described above can be applied here. The key point such an extension would deliver is that the incentive to raise costs by manipulating the price of delivered gas is larger for firm 3_A than for firm 3_B, for two reasons: first, because $x_A < x_B$ it takes less manipulation by firm 3_A to engender any particular level of increase in delivered price. Second, because of the additional pipeline tariff firm 3_B must pay on the segment between points A and B, firm 3_A has a natural cost advantage over firm 3_B. As we noted above, any such cost advantage is the root source of motives to manipulate 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. We suppose the intercept in market k takes the form $a_k = \alpha_k \epsilon$, where α_k is the mean value of the intercept and ϵ is a multiplicative demand shock representing weather impacts (and where its mean is one). Then in weather conditions that raise demand, we expect to see a larger impact in market B than in market A. If these shocks are anticipated when flows are scheduled, as seems likely, then the disproportionate increase in downstream demand will have spillover effects in market A: because an increase in scheduled deliveries to B raises the amount of gas shipped to A they must reduce spare capacity on the segment to A. 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 costs up.

5 Detecting Withholding Behavior

In order to determine whether price manipulation has been realized in New England, we examine the scheduling patterns of all 118 delivery nodes on the Algonquin pipeline over a three year period from mid-2013 through mid-2016.[8] We observe several LDC delivery nodes engaged in a practice of consistently reducing their scheduled daily quantities in the last few hours of the gas day (hereafter referred to as "downscheduling"). This scheduling pattern enables firms to tie up pipeline capacity without actually flowing gas by signaling to the pipeline company that they are flowing at a higher rate than they actually are for the majority of the gas day, which prevents other shippers from using that capacity. The node reduces its scheduled daily quantity at the end of the gas day to match what was actually flowed which makes it possible for a firm to avoid incurring the accounting imbalance penalties described in section 3. Figure 1 illustrates the scheduling pattern of a node engaged in withholding behavior and Figure 2 shows that of a typical LDC delivery node for contrast.

Table 1 quantifies this downscheduling behavior by listing the twenty nodes on the Algonquin pipeline which have the greatest average reductions to their scheduled quantities. While most delivery nodes either downschedule or upschedule quantities in the range of a few hundred MMBtu on average, six nodes are clear outliers with average daily downscheduling in excess of 2,000 MMBtu.[9] The ten nodes which downschedule the most on average are

[8]Hourly scheduled quantities for all nodes are downloaded from the Algonquin pipeline's FERC-mandated electronic bulletin board. Note that we only observe scheduled quantities; actual flows are known only to the pipeline and individual nodal operators.

[9]At the other end of the distribution (not shown in the table), the node that upschedules the most increases its daily nomination by 652 MMbtu on average.

all operated by just two parent firms out of 28 that hold contracts on the pipeline. Six of these nodes, all operated by "Firm A," are characterized by downscheduling on both summer and winter days, with much more variation in the winter.[10] The other four nodes, operated by "Firm B," are characterized by downscheduling primarily on winter days with high variability.[11] That Firm B's nodes only engage in this behavior during the winter season suggests that the behavior is specifically motivated by an awareness of the capacity constraint. On average, aggregate scheduled reductions across all delivery nodes on the pipeline (which represents net unused capacity, as it accounts for any upscheduled quantities) averaged 49,014 MMBtu over the entire study period and 56,503 MMBtu in the winters. On 46 days out of our three-year period, these downscheduled quantities exceeded 100,000 MMBtu, which is roughly 7% of the pipeline's total capacity[12] and roughly 28% of the total supply to electricity generators on Algonquin through the wholesale gas market.

Spatially, we observe that eight of ten nodes that downschedule the most on average are located in close proximity to one another in Connecticut (see Figure 3). This section of the pipeline is downstream of its major bottleneck at the Stony Point compression station.[13] While Firm A serves heating customers

[10]The winter season is defined here as between December 1 and March 31, following the delineation in the Algonquin pipeline's tariff.

[11]Daily variation in downscheduling behavior is driven by variation in heating demand due to weather. If more capacity is needed to supply their heating customers, less capacity is available for these firms to either sell on the spot market or employ in price manipulation. This may be the only mechanism driving the high level of daily variation in downscheduling, or oligopolistic pricing may also be a factor, in which case the variation would depend on demand for gas from generators as well.

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

[13]In order to keep gas flowing at a high rate across long physical distances, interstate pipelines have compression stations every 50 to 100 miles that effectively break the pipeline

Table 1: Schedule change between final intra-day cycle (10pm) and end of gas day (9am) for 20 nodes that reduce their scheduled quantity the most on average. The winter season is delineated as December 1 through March 31, consistent with the Algonquin pipeline's tariff.

Rank	**Schedule Change**		**Schedule Change (Winter Only)**		**Node Type**	**Node Operator**
	Mean	SD	Mean	SD		
1	-18,427	7,894	-17,923	10,785	LDC	Firm A
2	-10,592	7,512	-9,279	7,274	LDC	Firm A
3	-7,105	6,234	-7,766	7,155	LDC	Firm A
4	-3,893	3,779	-2,532	4,012	LDC	Firm A
5	-3,809	5,662	-8,430	5,707	LDC	Firm B
6	-2,413	4,314	-4,971	4,354	LDC	Firm B
7	-919	2,995	-2,298	4,712	LDC	Firm B
8	-861	1,756	-286	1,088	LDC	Firm A
9	-680	5,003	-1,604	7,516	LDC	Firm B
10	-566	1,067	-59	395	LDC	Firm A
11	-351	3,635	18	1,635	Generator	Firm J
12	-245	2,469	-580	2,571	Generator	Firm H
13	-230	2,229	-220	1,925	LDC	Firm C
14	-225	1,952	-262	1,380	LDC	Firm K
15	-209	721	-434	1,002	LDC	Firm B
16	-204	911	-532	1,473	LDC	Firm B
17	-203	883	-307	1,040	LDC	Firm B
18	-195	897	-486	1377	LDC	Firm B
19	-186	1,081	-33	511	LDC	Firm B
20	-142	661	-403	1,114	LDC	Firm B
All Nodes	**-419**	**2,843**	**-483**	**3,138**	–	–
All Except Firms A & B	**26**	**1,460**	**23**	**1,723**	–	–

only in Connecticut, Firm B also has large LDC operations in Massachusetts and Rhode Island as well. Both firms' clustering of downscheduling behavior in Connecticut can be explained by a relatively weak presence of gas-fired

into a series of segments. On the Algonquin pipeline, Stony Point is the compression station that most frequently reaches its operating capacity first, and it is located East of all of Algonquin's major Western receipt points.

electricity generators along this stretch of the pipeline (see Table 2) combined with the institutional framework governing the contracts. Contracts for capacity guarantee the holder's ability to transport gas to the delivery node listed in the contract, but 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 (in the same segment) or upstream of that node. However, nominating to transport gas further downstream than the contracted delivery node will often be impossible when the pipeline is operating near its capacity in the winter. With spot market demand driven by electricity generators who must submit bids for firm capacity to ISO-NE one day in advance, this uncertainty greatly reduces the value of contracts delivering gas upstream of centers of demand. In other words, contracts held by Firms A and B delivering gas to Connecticut are significantly less valuable for selling gas on the wholesale market than contracts delivering gas further downstream to Massachusetts and Rhode Island. However, they are just as valuable as those contracts for tying up pipeline capacity at the Stony Point bottleneck. Accordingly, a straightforward linear regression at the segment level reveals a high degree of correlation between withholding behavior and the location of gas-fired generation capacity (see Table 3).

While the ten most-downscheduling nodes are all *operated* by Firms A and B, several independent marketers also manage contracts delivering gas to these locations.[14] We find indications that Firms A and B are responsible for the

[14]The node operator manages deliveries at that location and is typically responsible for the majority of the node's deliveries, but it is not necessary to operate a node to make deliveries

Table 2: There is limited demand from electricity generators in the spot market for natural gas within and upstream of the segment between Oxford and Cromwell where a majority of the withholding nodes are located.

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

Table 3: Correlation between upstream generation capacity as a proxy for demand in the spot market and (1) levels of withholding for each delivery node over entire 3-year study period with standard errors clustered at the node level and (2) average withholding for each delivery node across the study period.

	(1) Downscheduled Gas	(2) Average Downscheduled Gas
Upstream Generation Capacity	-0.171*	-0.168**
	(-1.95)	(-2.16)
Constant	1077.4**	1063.1***
	(2.24)	(3.31)
N	133397	118

t statistics in parentheses

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

withholding behavior at these nodes by examining the correlation over time between downscheduling activity and firms' holdings of no-notice contracts sourcing gas from Texas Eastern at the node level. Table 4 demonstrates a strong and significant relationship between Firm A's and Firm B's holdings

to it. Independent marketers in particular use contracts to deliver gas to the region but do not operate nodes.

of these kinds of contracts and downscheduling behavior. In our preferred specification with parent firm fixed effects, we do find a significant relationship between downscheduling and no-notice contracts from Texas Eastern for several other firms, but these coefficients are at least an order of magnitude smaller than those for Firms A and B.

Table 4: Correlation between downscheduling behavior and firms' holdings of no notice contracts delivering gas from Texas Eastern at the node level with (2) and without (1) fixed effect dummies for the node operator's parent company. Both specifications control for temperature. Firms without any no-notice contracts from Texas Eastern are excluded due to collinearity (but their fixed effects are included in the model).

No Notice Contracts from Texas Eastern held by Firm:	(1) Downscheduled Gas	(2) Downscheduled Gas
Firm A	0.545***	0.486***
	(21.81)	(9.08)
Firm B	0.146***	0.170***
	(4.56)	(5.64)
Firm C	-0.00409	0.0105***
	(-0.41)	(4.22)
Firm D	-0.0156	0.000281
	(-1.65)	(1.54)
Firm E	-0.00415	0.00582**
	(-0.67)	(2.85)
Firm F	0.0305**	0.0451***
	(3.16)	(14.17)
Firm G	-0.00711	0.000209
	(-1.64)	(0.43)
Firm H	0.0112	0.0234*
	(0.88)	(2.09)
Firm I	-0.0151	0.000110
	(-1.68)	(0.13)
Firm J	-0.295	-0.404
	(-1.76)	(-0.52)
Firm K	0.223*	0.249*
	(2.12)	(2.27)
Firm L	-0.327	-0.434
	(-1.83)	(-0.57)
Firm M	-0.00379	0.00576*
	(-0.63)	(2.03)
Firm N	-0.0132	0.00632**
	(-0.92)	(2.76)
Firm O	0.0146	0.0286**
	(1.27)	(3.34)
Firm P	-0.357	-0.460
	(-1.91)	(-0.59)
Firm Q	0.00494	0.0196*
	(0.48)	(2.39)
Heating Degree Days	11.46*	11.08*
	(2.08)	(2.05)
Heating Degree Days^2	-0.286*	-0.281*
	(-2.28)	(-2.26)
Parent FE		X
N	133397	133397

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figures 4 and 5 further explore the institutional mechanisms of no notice contracts and storage on Texas Eastern that enable withholding behavior. We confirm that Firm A's and Firm B's shippers both have significant quantities of contracts for capacity on the Algonquin pipeline that are sourced from Texas Eastern, a large portion of which are no notice. These figures also clearly show that Firm C also owns substantial contracts sourcing gas from on-demand storage on Texas Eastern for delivery to Algonquin. However, nodes operated by Firm C do not appear to engage in downscheduling. We note two major differences between Firm C and Firms A and B: Firstly, 23 of the 24 nodes operated by Firm C's shippers are located in the Massachusetts and Rhode Island, where contracts for capacity are more valuable due to greater electricity demand. Secondly, considering companies' incentives to raise wholesale electricity prices, Firm C owns negligible generation assets in New England compared to the other two (see Table 5). The other three shippers with no notice contracts sourcing gas from storage on Texas Eastern individually have small market shares, are independently owned, and also ship gas only to Rhode Island and Massachusetts.

Figures 6 through 8 show downscheduling behavior against Firm A's and Firm B's contract positions over time at the segment level.[15] For most segments of the pipeline, there appears to be no relationship between downscheduling behavior and no notice contracts sourcing gas from storage on Texas Eastern. However, for the segment between Cromwell and Chaplin, the downscheduling

[15] We aggregate by segment here for ease of presentation under the assumption that gas contracted to flow to a certain segment of the pipeline can be easily withdrawn from any other node in that segment.

Table 5: While Firm C also holds significant no notice contracts sourcing gas from on-demand storage on Texas Eastern, and thus possesses the technical capability to withhold to manipulate prices, they don't have the same incentive to do so as they do not own significant generation capacity in New England.

	Average Withholding (MMBtu)	Capacity Owned (MW)	No Notice Contracts from TE (MMBtu)
LDC Firm A	**41,535**	**232**	**37,237**
LDC Firm B	**9,490**	**1,177**	**99,941**
LDC Firm C	-70	10	158,492
LDC Firm D	221	32	0
LDC Firm E	4	0	0
LDC Firm F	-4	0	782
LDC Firm G	-6	0	0
LDC Firm H	-16	116	17,159
LDC Firm I	-19	23	4,725
LDC Firm J	-292	0	42,047
LDC Firm K	-363	1,008	0
Industrial End User A	-17	0	0
Industrial End User B	-249	0	0
14 Electricity Generating Firms	min: -399 max: 357	min: 140 max: 2903	min: 0 max: 0

behavior is of roughly the same order of magnitude as company A and B's no notice contracts for gas from Texas Eastern delivering gas to that segment. Between Oxford and Cromwell, company A and B's no notice contracts from Texas Eastern appear to be an approximate upper bound on the level of downscheduling that occurs on this segment.

5.1 Alternative hypothesis

While there is little doubt that the downscheduled quantities of gas correspond to unused pipeline capacity, there remains uncertainty as to whether the intent of this behavior is oligopolistic pricing, higher wholesale electricity prices, new pipeline development, or some other objective. One alternative hypothesis is that Firm A's and Firm B's shippers are simply exercising risk aversion by reserving capacity to ensure they will have access to it if demand turns out to be higher than expected. We reject this hypothesis for two reasons. Firstly, while all LDCs are in the very same position, the pattern of downscheduling large amounts of capacity is only exhibited by those who have both the incentives and the ability to do so without facing imbalance penalties. This explanation would require some reason why these two firms are exceptionally poor at predicting next-day heating demand compared to every other LDC using the pipeline. Secondly, Firms A and B both own significant no notice contracts that guarantee their ability to ramp up capacity usage at any point in the gas day, making it unnecessary for them to reserve capacity

6 Simulations

Next, we attempt to quantify the impact of these firms' withholding behavior on the Algonquin City Gate price series and the welfare effects of withholding on the New England wholesale electricity market, which market we believe bore the majority of the incidence of this behavior. First, we use an instrumental variables approach to estimate the elasticity of demand for natural gas, and use our estimated demand elasticity to reconstruct the counterfactual Algonquin City Gate price series that would have resulted had there been zero withholding.[16] Then, we calculate the foregone gas-market profits associated with withholding, which we later compare to the change in electricity market profits resulting from withholding. Analysis showing that the profit gained

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

from exploiting the latter margin far exceeded the opportunity cost in regards to the former margin provides further evidence that it was in the withholding firms' strategic interest to withhold capacity. Finally, we employ a dispatch model of the wholesale electricity market operated by the New England Independent System Operator (ISO-NE) to calculate the wholesale electricity prices and generation profiles resulting from our counterfactual series of gas prices, comparing those to the observed prices and generation profiles. We use these simulation results to calculate the welfare and distributional impacts of the withholding firms' behavior on producers and consumers in the NE-ISO market.

6.1 Estimating counterfactual natural gas prices

Ideally, we would directly observe the effect of withholding on the Algonquin City Gate price. Unfortunately, we do not. Constructing a counterfactual series of gas prices is not as simple as using a reduced-form model to estimate the effect of withholding (*i.e.*, regressing quantity on price) and then using that coefficient to adjust the realized gas price. Withholding behavior is correlated with temperature, which affects demand for gas on a seasonal and daily basis. As highlighted in Figures 2 and 6, these firms primarily engage in withholding on higher-priced days and during the winter months, when capacity is more likely to be constrained, due to exogenous factors (*e.g.*, weather). This means that on a seasonal basis, greater withholding will be correlated with lower temperatures, higher pipeline congestion, and higher demand for natural gas from generators. Moreover, we believe day-to-day variation in withholding during the winter is driven primarily by the quantity of excess contracts each withholding firm has available to them, after supplying demand from residential and commercial heating customers, which is likely correlated with temperature (though likely differently from the seasonal source of correlation).

Our strategy for addressing this endogeneity is to estimate the price elasticity of demand for natural gas using an instrumental variables approach. Then, using the estimates from our IV regression, we construct a counterfactual Algonquin City Gate price series by shifting supply outward by the quantity of gas withheld in each day in our time series. To begin, we specify a model of demand:

$$D_t = \alpha_0 + \beta_1 P_t^{AGT} + \beta_2 HD_t + \beta_3 HD_t^2 + \beta_4 X_t + \epsilon_t$$

D_t: Log of demand in the natural gas spot market

P_t^{AGT}: Log of Algonquin Citygate gas price

HD_t: Heating degree days

HD_t^2: Heating degree days squared

X_t: Month-of-year and weekend-day indicator variables

t: Time index

β_1, which captures the relationship between price and quantity, is our coefficient of interest. Our endogeneity concern is that P_t^{AGT} will be correlated with ϵ_t resulting in biased estimates of β_1. 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)). Price in another market is often thought to be a good candidate instrument, as it is likely to be somewhat determined by the same supply-side drivers and different demand-side drivers, making it a reliable supply shifter. One obvious candidate is the Henry Hub price. Henry Hub is the major trading point for natural gas in the United States. The Henry Hub price is likely driven by some common and some distinct factors from the Algonquin City Gate price. However, when we test the relationship between Henry Hub price and temperatures in New England (i.e., our primary driver of New England demand), we find some evidence of endogeneity, even after controlling for month-of-year and weekend-day fixed effects.

Table 6: Instrumental variables approaches using the Henry Hub price, and alternatively the portion of variation in the Henry Hub price that is orthogonal to our other covariates, yield consistent estimates of the elasticity of demand for pipeline natural gas.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Log(p_AGT)	Log(q_AGT)	Log(p_HH)	Log(p_HH)	Log(q_AGT)
Log(p_HH)	1.102***				
	(0.0385)				
HDD	0.00188	-0.00654**	0.00709*	0.00977**	-0.00654**
	(0.00429)	(0.00275)	(0.00378)	(0.00431)	(0.00275)
HDD Sq.	0.000595***	7.00e-05	6.80e-05	0.000669***	7.00e-05
	(8.26e-05)	(6.21e-05)	(7.00e-05)	(8.23e-05)	(6.21e-05)
Weekend	-0.0580**	-0.107***		-0.0703***	-0.107***
	(0.0241)	(0.0142)		(0.0240)	(0.0142)
Log(p_AGT)		-0.264***			-0.264***
		(0.0198)			(0.0198)
Res_HH				1.102***	
				(0.0385)	
Constant	-0.0310	13.26***	0.780***	0.831***	13.26***
	(0.0865)	(0.0521)	(0.0702)	(0.0830)	(0.0521)
Observations	1,117	1,117	1,117	1,117	1,117
R-squared	0.786	0.740	0.083	0.786	0.740

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6 summarizes our investigation into the use of the Henry Hub (HH) price as an instrument for the Algonquin City Gate (AGT) price. Column 1 reports the results of the first-stage regression of the AGT price on the HH price, a linear-quadratic function of Boston Heating Degree Days (HDD), plus weekend and month-of-year fixed effects. The coefficient on the HH price is significant at the 1% level, suggesting it is a strong predictor of the AGT price. Column 2 reports results for the corresponding IV regression. The coefficient on the AGT price is significant at the 1% level. Our implied demand elasticity of -0.264 falls in the squarely in the center of the range of estimates found in (**?**) (*i.e.*, -0.1 to -0.34). Column 3 tests the relationship between the HH price and HDDs in New England. If HDDs in New England are a strong predictor of HH prices, it would fail the exclusion restriction. Though the quadratic coefficient is insignificant, the coefficient on the linear term is significant at the 10% level. This is not particularly surprising, as New England is a major source of natural gas demand, and fluctuations in New England's demand could plausibly affect upstream hub prices in addition to transmission prices. To overcome the potential endogeneity, we employ an orthogonal instruments approach, using residuals from the Column 3 regression as an instrument for the AGT price. As shown in (Ackerberg *et al.* , n.d.), this approach yields consistent estimates of the price elasticity of demand, by construction. That is, using the residuals as an instrument for the AGT price will yield consistent estimates of the price elasticity of demand as long as the instrument is not correlated with the other explanatory variables, which, by construction, it is not.[17] Columns 4 and 5 report results for the first and second stages of the orthogonal IV regression. As before, both coefficients are significant at the

[17]However, in general, this approach does not guarantee consistent estimates for other regressors.

1% level. Somewhat surprisingly, using the alternative instrument left our IV estimate unchanged; the estimated elasticity of demand for natural gas remains -0.264, despite changes in the coefficients on the linear and quadratic HDD terms. This may owe to the fact that, while the HH price appeared to be weakly endogenous to New England HDDs, the magnitude of the contribution of HDDs to the HH price appears to have been quite modest.

6.2 Simulating counterfactual pipeline natural gas prices

Next we plug in our estimated price elasticity of demand into a constant-elasticity demand function: $D(p) = kp^{-0.264}$, and solve for the vector of date-specific k's corresponding to each day's observed level of pipeline gas deliveries and the Algonquin City Gate price. For example, on a day when 1,000 Mcf of gas is delivered to downstream nodes and the City Gate price is \$13.81 Mcf, the equation becomes: $1,000 = k(13,81^{-0.264})$, which reduces to $k = 2000$. Having identified the vector of ks, it is straightforward to solve for a counterfactual vector of ps, associated with a counterfactual vector of delivered quantities. Our counterfactual vector of quantities is constructed by summing all observed deliveries to downstream nodes, plus all downscheduled quantities at the nodes operated by Firm A and Firm B, where withholding appears to have occured, less the daily average fraction of downscheduled capacity observed at all other nodes where there does not appear to be systematic downscheduling. Adding downscheduled quantities back into the supply lowers the City Gate price. This vector of prices provides the primary fuel input data for gas-fired generators in our electricity dispatch model.

6.3 Simulating counterfactual wholesale electricity prices

In this section, we first describe our equilibrium dispatch model and then discuss how we apply data from various sources to arrive at our calculations.

6.3.1 Dispatch Model

We assume here that electric generators act in a manner consistent with perfect competition, with regards to the supplying electricity for the wholesale market. As such, the solution stemming from a perfectly competitive market is equivalent to the solution of a social planner's problem of maximizing total welfare.

The key variables and parameters of the model are grouped according to two indices: the origin plant and time period of production. The total production of plant p at time t is represented by $q_{p,t}$. Production costs $C_p(q_{p,t})$, vary by firm, technology, and location, and are constant for each plant and are unchanging over time.

$$C_p(q_{p,t}) = c_p q_{p,t}$$

where $q_{p,t} = \sum_j q_{p,t}$. Total emissions by firm and technology are determined by a constant emissions rate e_p and denoted $e_p(q_{p,t}) = e_p * q_{p,t}$.

Wholesale electricity is assumed to be a homogenous commodity for purposes of setting wholesale prices, although prices are assumed to vary by time. For each time period $t \in \{0, ..., T\}$, a perfectly competitive market outcome is obtained by solving the following welfare maximizing problem:

$$\int_0^{Q_t} P_t(Q)dQ - \sum_p C_p(q_{p,t}),$$

where $P_t(Q)$ gives the power prices in period t, and $Q_t = \sum_{p,i} q_{p,t}$. The output $q_{p,t}$ is further limited by its capacity: $q_{p,t} \leq \bar{Q}_p$.

6.3.2 Generation and Emissions Data

We utilize detailed hourly load and production data for all major fossil-fired generation sources in ISO-NE. Our primary source is the EPA Continuous Emission Monitoring System (CEMS) data, which provide hourly output for all major fossil-fired power plants. The CEMS data cover all major utility-scale sources of CO_2, but do not measure output from nuclear, combined-heat and power, wind, solar, or hydro sources. Therefore, one challenge we face is the lack of information about the power plants that are not required to report to CEMS. To model these units, we assign a zero emission rate to those units since historically they are dominated by renewables and hydro facilities. We further assume that the power sales of those "non-CEMS" units are not changed in response to the natural gas price and fix their sales $q_{p,t}$ at their average historical levels. These hourly data are aggregated by region to develop the "demand" in the simulation model, and are combined with cost data to produce cost and emissions estimates for each of the generation units in the CEMS database. Emissions, the resulting outputs for each *simulated* demand level was multiplied by the number of *actual* market hours used to produce the input for that simulated demand level.

In the following sub-sections, we describe further the assumptions and functional forms utilized in the simulation.

6.3.3 Market Demand

Aggregate demand is taken from FERC form 714, which provides hourly total end-use consumption by control-area. As described below a large portion of this demand is served by generation with effectively no fuel costs or CO_2 emissions, such as nuclear and hydro sources. This generation needs to be netted out from total demand to produce a residual demand to be met by fossil-fired sources.

End-use consumption is represented by the demand function:

$Q_t = \alpha_t - \beta p_t$.

The intercept of the demand function is based upon the actual production levels on each day calculated as described above. Summary statistics on demand are reported in the appendix. In other words, we model a linear demand curve that passes through the observed price-quantity pairs for each period. As electricity is an extremely inelastic product, we utilize an extremely low value for the slopes of this demand curve. For each region, the regional slope of the demand curve is set so that the median elasticity in each region is -.05. When the market is modeled as perfectly competitive, as it is here, the results are relatively insensitive to the elasticity assumption, as price is set at the marginal cost of system production and the range of prices is relatively modest.

6.3.4 Transmission Network Management

We assume that the transmission network is managed efficiently in a manner that produces results equivalent to those reached through centralized locational marginal pricing (LMP). For our purposes this means that the transmission network is utilized to efficiently arbitrage price differences across locations, subject to the limitations of the transmission network. Such arbitrage could be achieved through either bilateral transactions or a more centralized operation of the network. For now we simply assume that this arbitrage condition is achieved, and focus our analysis on the central hub price.

6.3.5 Hydro, Renewable and other Generation

Generation capacity and annual energy production is reported by technology type in Tables 7 and 8. We lack data on the hourly production quantities for the production from renewable resources, hydro-electric resources, combined heat and power, and small thermal resources that comprise the "non-CEMS" category. By construction, the aggregate production from these resources will be the difference between market demand in a given hour, and the amount of generation from large thermal (CEMS) units in that hour. In effect we are assuming that, under our counter-factual, the operations of non-modeled generation (*e.g.*, renewable and hydro) plants would not have changed. This is equivalent to assuming that reallocation of production occurs exclusively within the set of modeled plants. We believe that this is a reasonable assumption for two reasons. First the vast majority of the CO2 emissions from this sector come from these modeled resources. Indeed, data availability is tied to emissions levels since the data are reported through environmental compliance to existing regulations. Second, the total production from "clean" sources is unlikely to change in the short-run. The production of low carbon electricity is driven by natural resource availability (e.g., rain, wind, solar) or, in the case of combined heat and power (CHP), to non-electricity production decisions. The economics of production are such that these sources are already producing all the power they can, even with historically low gas price throughout much of the year; short-run production reallocation will have to come either from shifting production among conventional sources, a reduction in end-use, *i.e.*, leakage or reshuffling.

6.3.6 Fossil-Fired Generation Costs and Emissions

The purpose of these simulations is to model the effect of fuel price changes on plant-level output and profits in ISO-NE. To do this, we explicitly model the major fossil-fired thermal units in ISO-NE. Because of the legacy of cost-of-service regulation, relatively reliable data on the production costs of thermal generation units are available. The cost of fuel comprises the major component of the marginal cost of thermal generation. The marginal cost of a modeled generation unit is estimated to be the sum of its direct fuel and variable operation and maintenance (VO&M) costs. Total fuel costs are calculated by multiplying the price of fuel, which varies by period, by a unit's 'heat rate,' a measure of its fuel-efficiency.

Generation marginal costs are derived from the costs of fuel and variable operating and maintenance costs for each unit in our sample. SNL provides a unit average heat-rate for each of these units. These heat-rates are multiplied by a regional average fuel cost for each fuel and region, also taken from SNL. Marginal cost of each plant p is therefore constant:

$$C_p(q^i_{p,t}) = c_p q_{p,t}.$$

Emissions rates, measured as tons CO_2/MWh, are based upon the fuel-efficiency (*i.e.*, heat-rate) of a plant and the CO_2intensity of the fuel burned by that plant. The average emissions rates of all facilities are summarized by region in Table **??**.

6.4 Welfare consequences of withholding on electricity market participants

7 Discussion and conclusions

This paper reveals an instance where two firms participating in New England's natural gas transportation market appears to have exploited the rigidity of the pipeline infrastructure to capture large rents. Having devised a system to use their contracts to reserve pipeline capacity without actually flowing gas, these firms have been able to reduce the supply of gas to the wholesale market by up to 28%. Restricting supply enables them (and other LDCs and independent marketers) to sell gas to electricity generators at oligopolistic prices, which in turn raises the wholesale electricity price and increases revenues for baseload and renewable generation resources these companies also own. We estimate that the increased energy costs passed on to gas and electricity ratepayers due to this price manipulation resulted in a transfer from ratepayers to energy companies of $ XX over the period 2013-2016.

Harder-pressed capacity constraints yield higher rents for long-term contract holders, and New England has one of the tightest natural gas markets in the nation. To what extent similar market behavior exists in other regions is an important area for future study. LDC contract holders are generally well positioned to capture these rents because to provide reliable service, they must own sufficient contracts to supply their customers on the highest-demand days. In the long-run, the regulated utility is able to socialize the cost of these contracts onto their ratepayers, while in the short run, an affiliate shipping arm could use excess contracts to transport and market gas to obtain unregulated profits. Beyond giving them an advantaged position in the secondary capacity and spot markets, LDCs' reliability requirements make them more likely to control larger market shares in constrained regions and may also make them less likely to arouse suspicion of market manipulation than independent marketers.

References

Ackerberg, Daniel, Crawford, Gregory S., & Hahn, Jinyong. Orthogonal instruments: Estimating price elasticities in the presence of endogenous product characteristics. *University of Warwick Working Paper, year=2011.*

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.

EIA. 2013. Short-Term Energy Outlook Supplement: Constraints in New England likely to affect regional energy prices this winter. *U.S. Energy Information Administration*, **January 18, 2013**.

FERC. 2008. Order 712. Promotion of a More Efficient Capacity Release Market. *Federal Energy Regulatory Commission*.

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.

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 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.

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

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.*

Figures

Figure 1: A typical LDC delivery node, with most adjustments made shortly after the start of the gas day and some balancing either direction at the end of some days. Each line represents one gas day in the study period (Aug. 1, 2013 - Sept. 13, 2016). The X-axis covers 44-hour scheduling period and the Y-axis is total daily quantity of gas scheduled at a given time. The graph is broken into thirds by price (Algonquin Citygate basis over Henry Hub), with the right panel showing the lowest-priced days, and the right panel showing the highest-priced days. Line color represents basis of Algonquin Citygate over an average of three other Northeast prices (TE M3, Transco Z6 NY, & Transco Z6 Non-NY); redder is relatively higher New England prices (presumably due to capacity constraint). We constructed this graph for all 118 delivery nodes on the Algonquin pipeline.

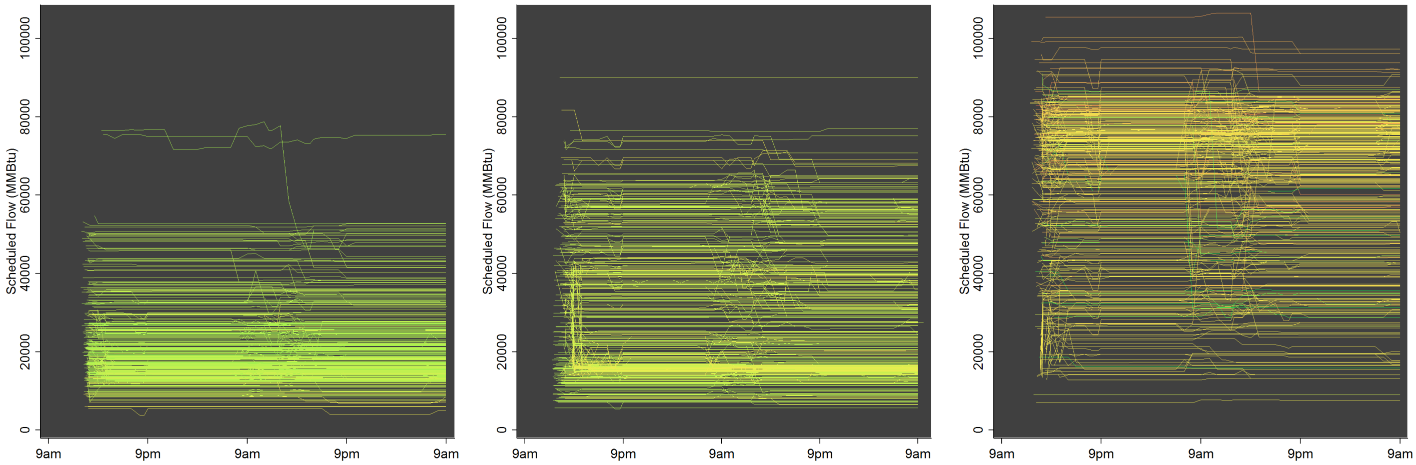

Figure 2: A node that consistently reduces its nomination in the final three hours of the gas day. This pattern is exhibited to varying degrees by ten nodes operated by Firms A and B. Reduced capacity represents unused space: For example, if the node schedules 72,000 MMBtu at the beginning of the scheduling period, they are indicating to the pipeline company that they will be flowing gas at a rate of 72,000/24=3,000 MMBtu per hour for that period and that capacity is then reserved for them. When the node reduces its scheduled quantity to 48,000 MMbtu in the last three hours 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, and is changing its schedule to match its actual flowed quantity for the day to avoid incurring an imbalance penalty. The result in this example is 24,000 MMBtu less gas entering New England for that gas day.

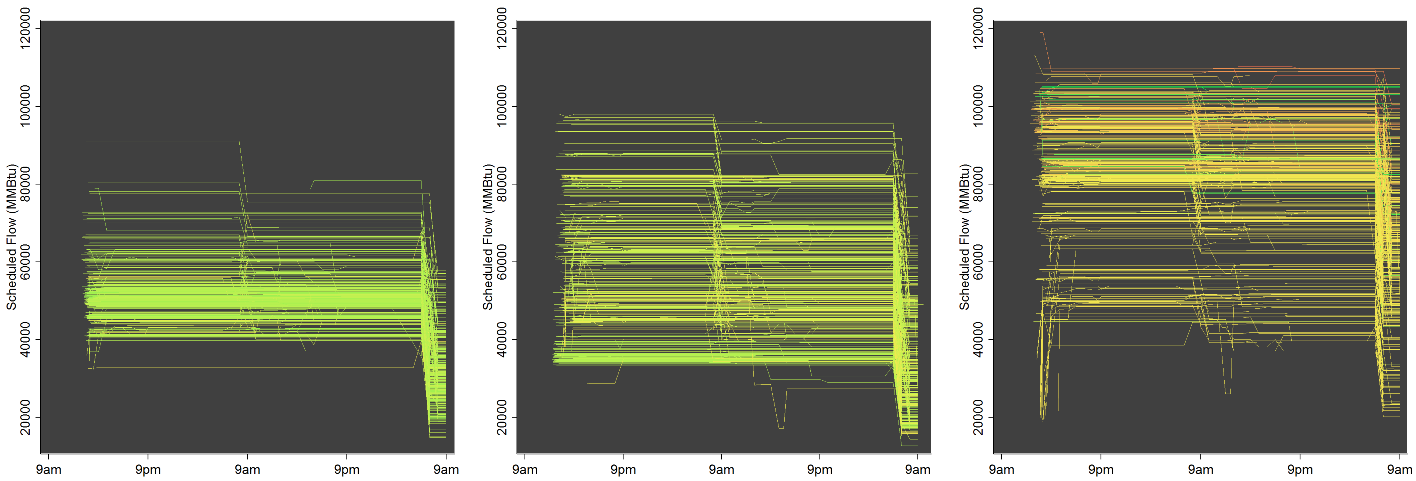

Figure 3: The locations of the 10 nodes that downschedule the most on average. Eight are located in Connecticut, which is downstream of the Stony Point compression station but upstream of most of electricity generators that demand gas on the wholesale market.

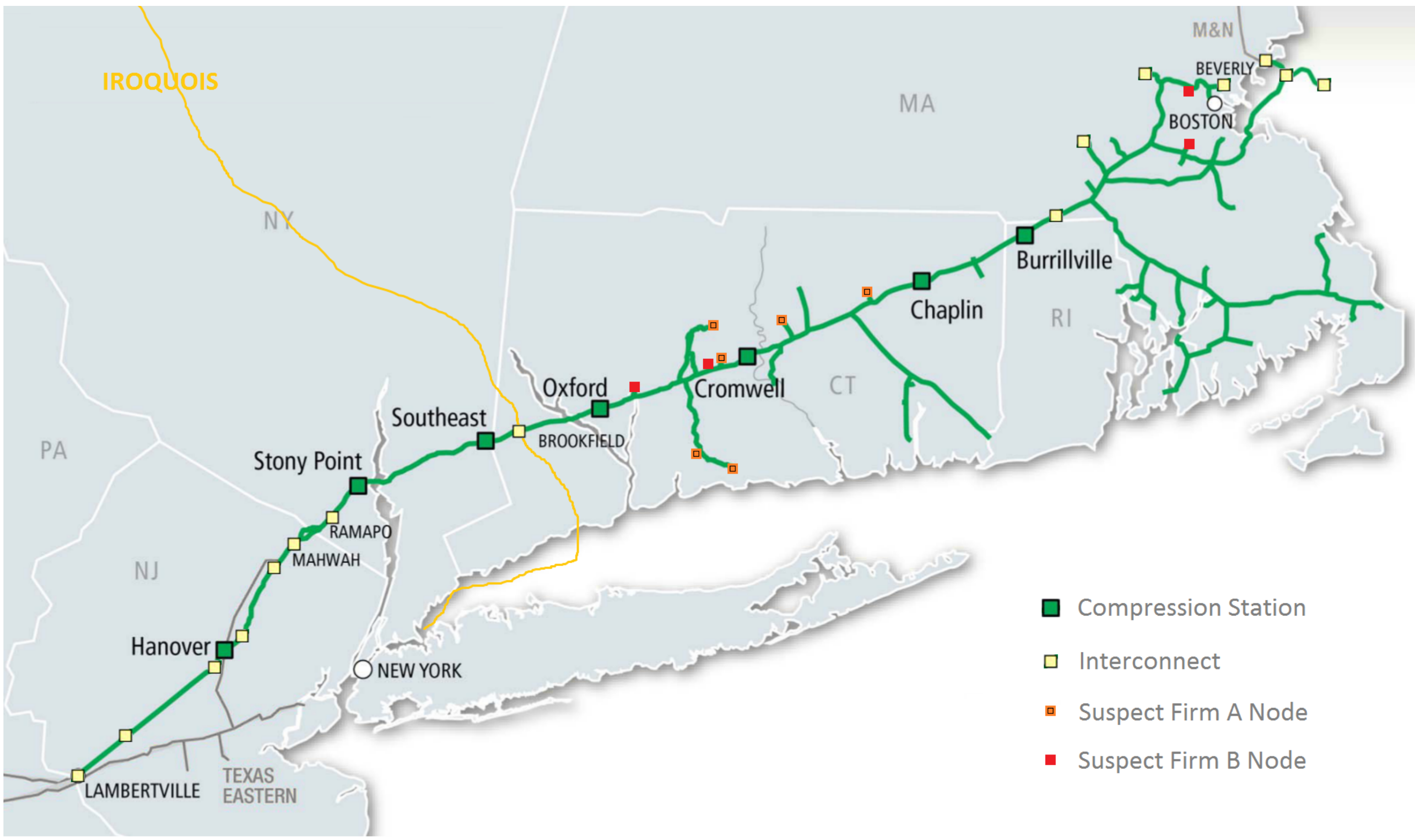

Figure 4: Average holdings of all contracts for capacity on the Algonquin pipeline from 2012 through 2016. Each bar represents a shipper; shippers are grouped by parent company for Firms A, B, and C, all of whom transport a majority of their gas through Algonquin's interconnects with the Texas Eastern Pipeline. Many of these contracts are no-notice

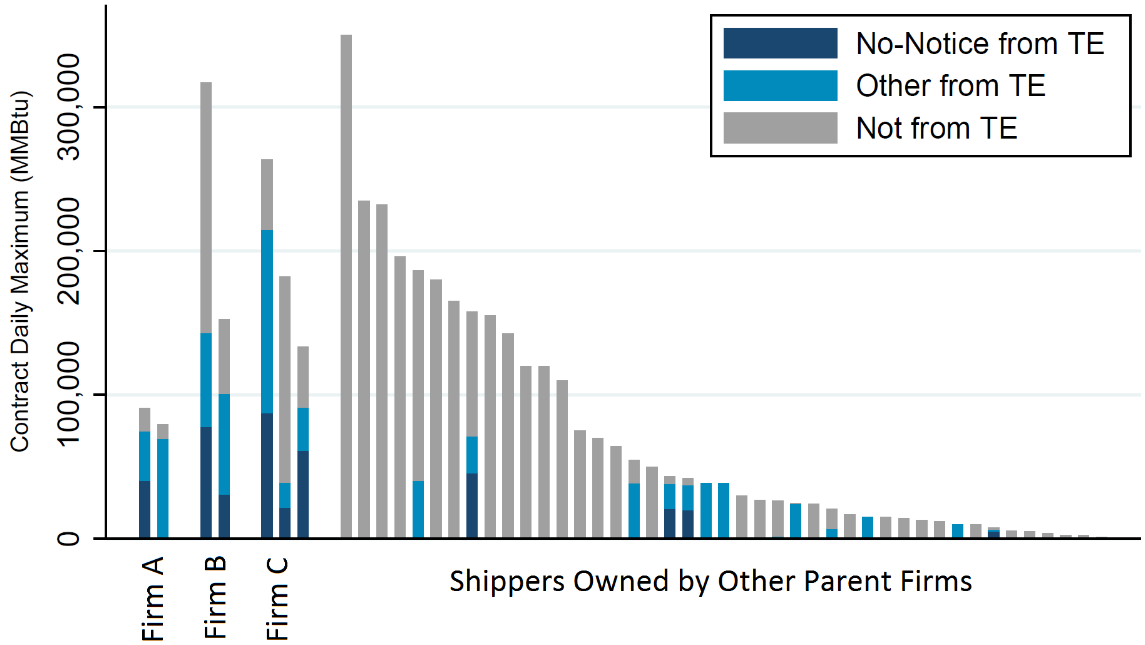

Figure 5: Average holdings of all contracts for capacity on the Texas Eastern pipeline for gas delivered to its two interconnects with the Algonquin pipeline. Of the gas sourced from Texas Eastern by Firms A, B, and C, much of it comes from storage. The upper limits of the aggregate downscheduling behavior observed (around 100,00- MMBtu) roughly match the sum of Firm A and Firm B's no-notice contracts sourcing gas from storage on Texas Eastern.

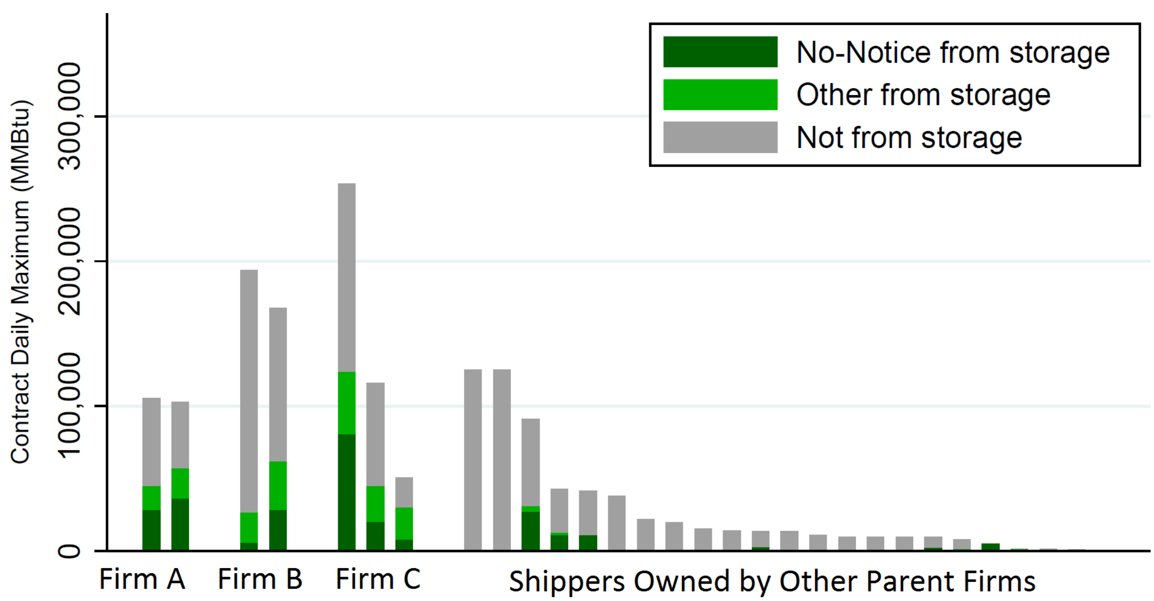

Figure 6: Aggregate downscheduling behavior compared to the contract positions of Firms A and B over time within the segment between the Oxford and Cromwell compression stations, where six regularly withholding nodes are located. These two firms' holdings of no notice contracts sourcing gas from the Texas Eastern pipeline roughly correspond to the upper bound on the amount of withholding that occurs in this segment.

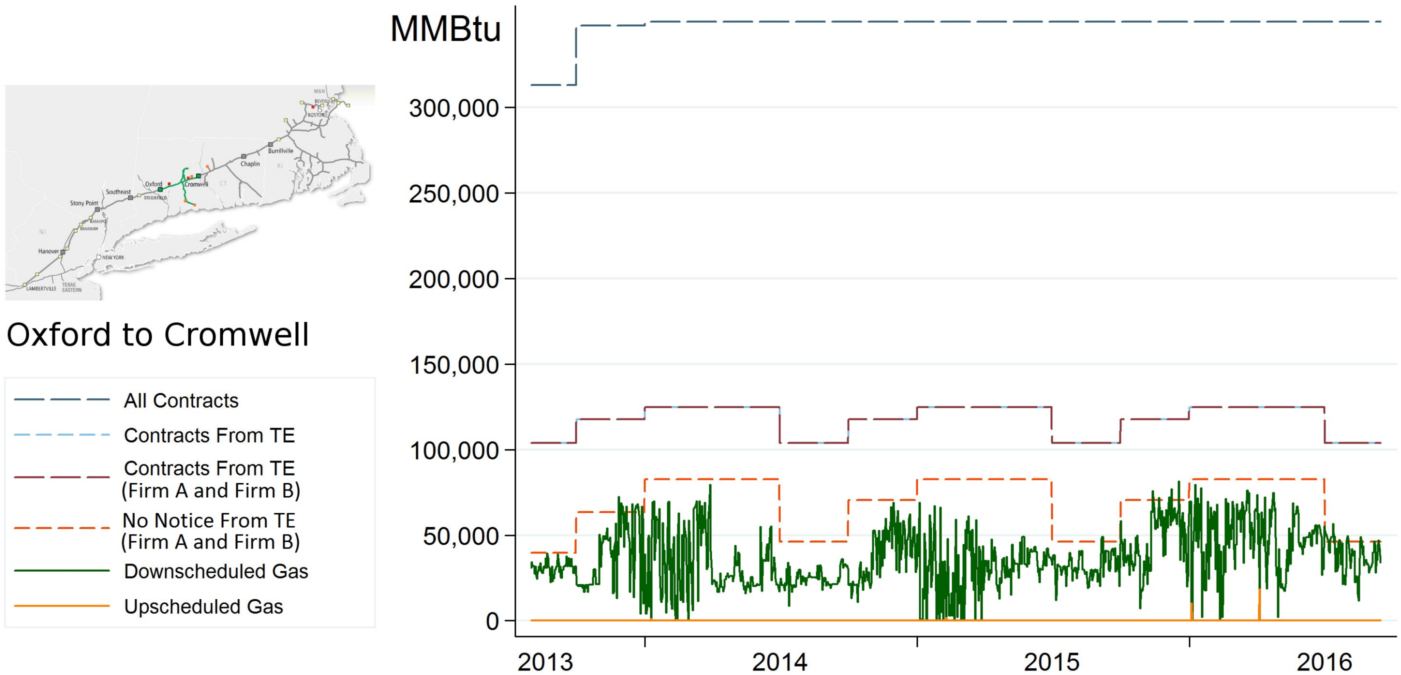

Figure 7: Aggregate downscheduling behavior compared to the contract positions of Firms A and B over time within the segment between the Cromwell and Chaplin compression stations, where two regularly withholding nodes are located. The level of downscheduling behavior is of roughly the same order of magnitude as these two firms' holdings of no notice contracts delivering gas to this segment.

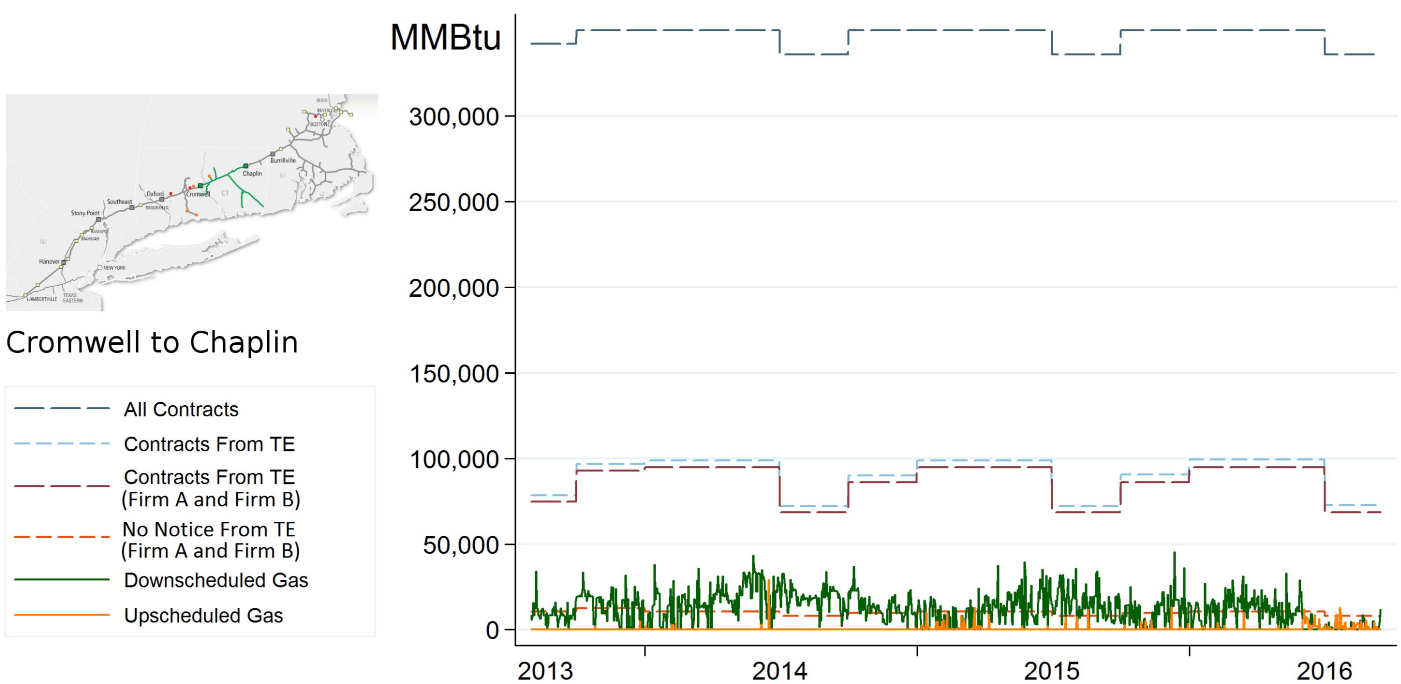

Figure 8: Aggregate downscheduling behavior compared to the contract positions of Firms A and B over time within pipeline's "J System," which serves many large electricity generators as well as a substantial heating market. Two regularly withholding nodes are located here. There is no clear correlation here between aggregate downscheduling and contracts held by Firms A and B.