STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic Year 2018-19

The Counteraction of the Merit Order Effect: Wind Power and Strategic Withholding on the Swedish Electricity Market

Joel Blanke (23290) and Gustav Tillman (23346)

Abstract: One of the posited benefits of renewable energy in electricity markets is its potential to drive down prices - the merit order effect. However, diversified firms with both renewable and conventional energy production in their portfolios have an incentive to counteract such an effect by withholding their high cost production. Thus, it has been hypothesized that an increased share of renewable production that is owned by such firms leads to higher prices. The empirical evidence is thin, and no previous studies incorporate electricity grid congestion in their models. Using a novel approach to estimate firm wind production based on wind power plant data and wind speed measurements, we investigate both the main effect of ownership and its interaction effect with congestion on the Swedish electricity market. The results show that diversified firms counteract the merit order effect in Sweden - when the share of renewable production that is supplied by diversified firms goes from 0% to 100%, prices increase by 13.7% on average. Finally, we test several configurations of congestion, using both an instrumental variables approach and an endogenous switching regression approach to account for possible endogeneity. Our results are inconclusive but suggestive with respect to congestion, indicating that this feature is important to consider in future studies. Our results have important policy implications for electricity market monitoring and the design of renewable energy support policies.

Keywords: Electricity Markets, Market Power, Merit Order Effect, Nord Pool, Switching Regression Model
JEL: C24, L13, L94, Q40

Supervisor: Chloé Le Coq
Date submitted: May 13, 2019
Date examined: May 27, 2019
Discussant: Nicolas Leicht
Examiner: Mark Sanctuary

Acknowledgements

We would like to thank our supervisor Chloé Le Coq for helpful advice and guidance during the writing of our thesis.

Thank you also to Henrik Horn af Rantzien, Erik Lundin, Thomas Tangerås and Pär Holmberg at IFN for insightful comments.

Contents

1 Introduction

One of the claimed benefits of intermittent renewable energy such as wind and solar power is the possibility that the low marginal cost of such technologies may lower the price of electricity, an effect called the merit order effect in the literature. However, economic theory suggests that the potential benefits to consumers may be diminished due to the specific market structure of electricity markets. Specifically, firms with diversified generation portfolios consisting of both renewable and conventional technologies will have higher incentives to exercise market power by withholding expensive production to push up market prices, relative to non-diversified fringe firms that only supply renewable power at one constant marginal cost. Because wholesale electricity markets are generally characterized by a few dominant and diversified firms, this is likely to be a problem. This is formalized by Acemoglu, Kakhbod, and Ozdaglar (2017) in a symmetric Cournot representation of an electricity market.

Strategic withholding has important welfare implications since it leads to inefficient allocation on a market and thus reduces welfare. In addition, the higher prices lead to wealth transfers from consumers to firms, with distributional effects as a consequence. As the share of renewable energy increases, backed by a wide range of support policies, these issues will grow in importance for policymakers.

This thesis studies if the ownership of renewable power generation matters for market outcomes (market prices and quantities of hydro power and short-term controllable thermal power) in the Swedish wholesale electricity market. Furthermore, potential additional effects of congestion on the relationship between ownership and market outcomes are studied.

The effect of ownership of renewable energy on market outcomes, as proposed by Acemoglu et al. (2017), has been supported empirically by Genc and Reynolds (2019) and Butner (2019). However, both these studies focus on electricity markets in North America. These markets differ fundamentally from the Swedish electricity market both in terms of market design and generation technology mix. Thus, our first major contribution is that we test the effect of ownership in a market outside North America. Furthermore, neither the theoretical work nor the empirical studies done in this area have included grid congestion in the analysis. A rich literature documents the possible anti-competitive effects of congestion. Based on this, we incorporate congestion in the theoretical framework of Acemoglu et al. (2017), and test for potential interaction effects with ownership. This is our second major contribution.

With no access to firm level production data, we use a novel approach to estimate wind production on a firm level. We match detailed data on ownership and location of Swedish wind power plants with data on historical wind speeds at a granular geographical level to compute the share of total generated wind power that is provided by large, diversified firms - delta. To our knowledge, neither this approach, nor any of the data has been used in this setting before. We thus also contribute to the literature on how to test for market power in electricity markets without access to firm level data.

Because firms' investment decisions regarding the number and localization of plants are fixed in the short run, the variation in our delta variable stems from spatial and temporal variation in wind speeds, making the variable highly exogenous. Under relatively weak identifying assumptions, this allows us to estimate causal effects of wind power and delta on market outcomes. We extend the analysis by investigating potential interaction effects of delta and transmission constraints in the electricity grid.

To address potential endogeneity of such congestion we employ instrumental variables (IV) methods and a switching regressions framework for endogenously shifting congestion states. By instrumenting for congestion with planned outages on transmission lines, we estimate the marginal increase in the delta effect in areas that are import constrained.

The results confirm the existence of the merit order effect in the Swedish market: prices decrease following increases in the amount of wind power produced in the market. We also find strong evidence of a counteraction effect by delta: for a given level of wind power supplied in the market, prices are increasing in delta. The results for the effect of congestion are suggestive, but inconclusive. While it is clear that congestion matters for how firms withhold generation following changes in delta, we cannot generalize any effect across different transmission links and generation technologies.

The rest of this thesis will be organized as follows. In Section 2, we describe the design of power markets, and discuss the impact of variable renewable energy generation and potential for market power in electricity markets. In Section 3 we provide a review of the relevant literature and in Section 4 we explain the theoretical framework of Acemoglu et al. (2017), and adapt it to our setting. Section 5 contains our empirical strategy and describes our data, and Section 6 presents our results. We discuss our results and their validity in Section 7 and conclude in Section 8.

2 Background

Unlike most markets, electricity markets have not emerged from previously less centralized marketplaces. Because electrical power has been viewed as an essential service in society and because of the specific technical properties of the allocation of electricity, power markets have generally been carefully designed by regulators. Originally, the provision of electricity was not organized through markets with several vertically differentiated and competing actors, but was handled in each region by one monopoly utility. The utility would alone provide generation, transmission and distribution to households and business under a rate-of-return regulation, by which an authority would grant the utility prices to adequately recover its operational and investment costs. With the aim of increasing efficiency, electricity market deregulation began in the 1990s, vertically separating the provision of transmission from the wholesale and retail markets, and introducing competition in the two latter (Cramton, 2017).

The rest of this section provides an overview of the structure of deregulated electricity markets, with a special focus on the Swedish market, as well as an introduction to the market implications of variable renewable electricity and to market power in electricity markets.

2.1 The Swedish electricity market

As this paper focuses on the Swedish electricity market, we will here present the design of the Nordic market, of which the Swedish market is a part. In many ways, the Nordic market design is representative of the standardized design of liberalized electricity markets throughout the world, but important differences exist, reflecting both differences in local market settings, as well as political reasons.[1]

[1]See chapter 3 of Biggar and Hesamzadeh (2014) for a discussion of the key differences.

Following Norwegian electricity market liberalization in 1991, the Swedish deregulation started in 1996 by introducing competition in the generation and distribution of electricity. Transmission and access to the networks remained a monopoly under the Swedish authority Svenska Kraftnät.

Following the deregulations, significant consolidation took place in the Swedish and Norwegian markets. Between 1989 and 1996, the number of local electricity distributors declined from 290 to 250, mainly driven by forward integration of large generation companies. To address the increasing market concentration, the Swedish and Norwegian markets were integrated, thereby reducing market shares of firms. This was implemented by creating the common electricity exchange Nord Pool, which was set up in 1996 (The Swedish Competition Authority, 1996). In 1998, Finland joined and in 2000 Denmark joined, resulting in the Nordic market becoming fully integrated. From 2010 to 2013, Estonia, Latvia and Lithuania joined Nord Pool, becoming fully integrated with the Nordic markets (Nord Pool, n.d.a).[2]

The overall structure of the Nordic electricity market is illustrated in Figure 1. The market consists of a wholesale market and a retail market. In the wholesale market, generating firms compete in generating electricity and selling it to retailers and large end consumers (for instance firms in electricity-heavy manufacturing industries). In the retail market, retailing firms compete in procuring electricity at the wholesale market and selling it to end consumers (households and businesses). Vertically integrated firms may act both as producers and retailers and some firms may procure electricity on the wholesale market only to sell it again on the same market.

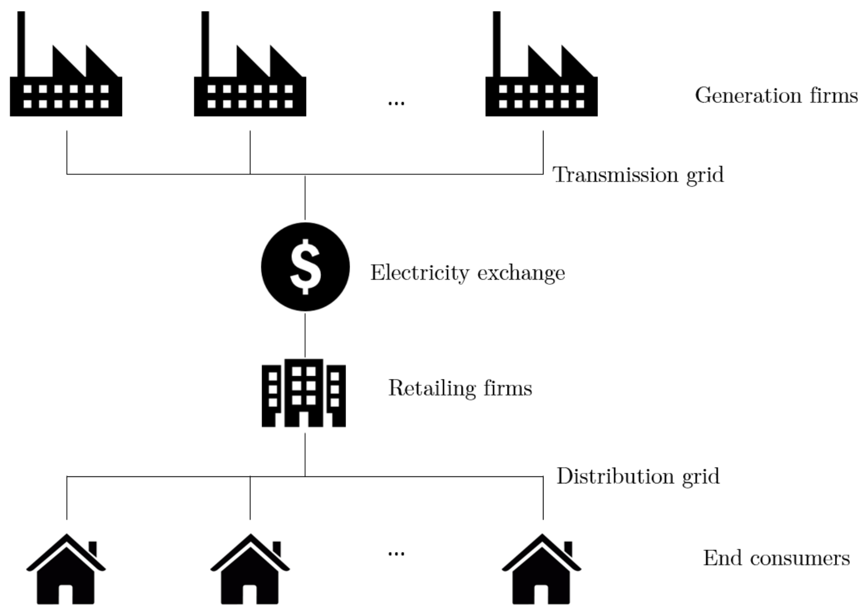

Figure 1: A stylized overview of the Nordic electricity market.

The high-voltage grid connects generation plants with retailers, and the low-voltage grid connects retailers with end customers. The high-voltage grid is owned and managed by the transmission system operator (TSO), a non-commercial actor that is usually state-owned. The low-voltage grids are owned

[2]Nord Pool offers various exchange services in several other European countries, including Germany and the UK, but these markets are not fully integrated with the Nordic-Baltic region.

and managed by local non-commercial grid operators. The TSO is also responsible for the security of supply in a country and thus also rules and controls the electricity system in that country. In Sweden, the TSO is Svenska Kraftnät (Nord Pool, n.d.b).

The majority of the wholesale trading takes place over a centralized electricity exchange, much like a stock market exchange. In the Nordics, the electricity exchange is Nord Pool, covering Denmark, Finland, Sweden, Norway, Estonia and Lithuania. Actors may also engage in bilateral trading by which contracts are signed for delivery of electricity outside of the Nord Pool market. In the Nordics, electricity traded over the Nord Pool exchange constitute 77% of total consumed electricity (Tangerås and Mauritzen, 2014).

2.1.1 The day-ahead market

The day-ahead electricity market is organized as Walrasian auctions: Market participants in advance submit bids in the electricity exchange's day-ahead market (Nord Pool Spot in the Nordics) specifying how much electricity they are willing to sell/purchase for a range of prices for each hour of delivery for the following day, yielding individual supply and demand curves of electricity for each hour. Nord Pool aggregates all submissions and creates a non-decreasing aggregate supply curve (so that cheaper supply bids always come before more expensive ones in the curve), and intersects this curve with the aggregate demand curve, resulting in the market price - the so called system price (Nord Pool, n.d.b).

Through the electricity exchange, actual deliveries between pairs of producing and procuring actors are not pinned down. Electricity is a completely homogeneous good and it is not possible to single out any specific electricity that is running through the grid. Instead, the system price together with the individual actors' posted bid curves pin down each individual actor's assigned insertion or extraction from the grid for the concerned next-day period (Nord Pool, n.d.b). The day-ahead market constitutes the vast majority of all electricity traded on Nord Pool. In 2017, the Nordic-Baltic day-ahead market had a turnover of 394 terawatt hours (TWh), making up 98% of the volume of all electricity traded on Nord Pool in the Nordics and Baltics (Nord Pool, 2018).

2.1.2 Congestion

In any grid, there are physical constraints on how much electricity can be transmitted at any one point in time. This results in transmission bottlenecks at certain points in the grid, where particularly high amounts of electricity is transmitted relative to the capacity of the transmission lines. When bottlenecks arise, electricity cannot be delivered as agreed on in the day-ahead auction. To address this issue, the grid is split according to bidding zones by the market operator.

Bidding zones are pre-determined local geographical markets whose borders are drawn first by national borders and thereafter based on the transmission bottlenecks in the grid. Each bid from a commercial actor on the day-ahead market also indicates which bidding zone that the actor is demanding or supplying electricity in. After the system price is determined, the volume of electricity to be transmitted between different bidding zones is calculated. While a bidding zone may be entirely self-sustaining, many zones will be exporting or importing electricity from another zones. This happens automatically because electricity flows freely in the grid, up to the transmission capacity constraint. If then an electricity flow reaches the transmission constraint, the importing zone will be left with excess demand, as more electricity is demanded than can flow into the market. Thus, the operator compares the implied flows of electricity based on the system price, with the transmission capacity of the grid between the zones. If a bottleneck is identified, the market area is split by the border between the two concerned zones. The operator now allows the price to increase in the import constrained zone, until local demand and supply have reached a new market equilibrium. Correspondingly, because of the excess supply in the exporting zone, that zone's price is allowed to decrease until the additional exports stops, which happens when the importing zone is not import constrained any more. This splitting process is repeated until there are no more bottlenecks in the Nordic grid, or until a maximum of fifteen bidding zones is reached. These are shown in Figure 2.

Figure 2: The Nord Pool Nordic-Baltic market: Overview of the fifteen different bidding zones.
Source: Nord Pool (n.d.b).

2.1.3 Balance of the grid

It is the responsibility of the TSO to keep the grid electrically stable at a pre-set frequency. This frequency is at 50 Hz and devices and appliances requiring electricity are designed to use electricity at this frequency. If the frequency deviates too much from this level, devices take damage and in the extreme case a complete black-out occurs. The frequency is determined by the relative insertion and extraction of electricity in the grid. If insertion and extraction are balanced, the frequency stays at 50 Hz. If more electricity is inserted into the grid than is extracted, the frequency rises, and vice versa for the case when more electricity is extracted than is inserted (Nord Pool, n.d.b).

Because day-ahead bids are based on imperfect forecasts of production (for a supply-bidding producer) and end consumer demand (for a demand-bidding retailer), commercial actors on the market will frequently be unable to extract/insert electricity from the grid as assigned in the day-ahead market. To keep the grid frequency balanced, each commercial actor is financially responsible to insert/extract as much electricity as she was assigned, given her posted day-ahead bid curve and the system price. This incentivizes each commercial actor to post bids that she can (and wants to) fulfill, thus helping the grid to stay in balance at the level pinned down by the system price.

After the day-ahead market closes, there are two additional markets that aim to help the system stay in balance - the intraday market and the balancing market. The intraday market uses continuous trading and allows market actors to fulfill their commitments on the day-ahead market in the event of unexpected changes in e.g. production capacity. The balancing market is used as a last resort to ensure balance in the grid. For a more thorough description of these, consult Appendix A.

2.2 Renewable energy in electricity markets

Owing to technological development and concerns over climate change caused by CO2 emissions, renewable electricity generation in Europe has increased rapidly in the last decade. The total installed capacity of wind and solar power in the European Union has increased from around 80 GW in 2008 to around 300 GW in 2018 (Wind Europe, 2018). Many governments actively support such expansion through the use of both investment support and subsidies in the market place. In Sweden, there is essentially no solar power, but the number of wind power plants has increased from a few hundred in 2008 to over 3,000 in 2018, while the share of total production has gone from negligible to 11 % (Energiföretagen, 2018).

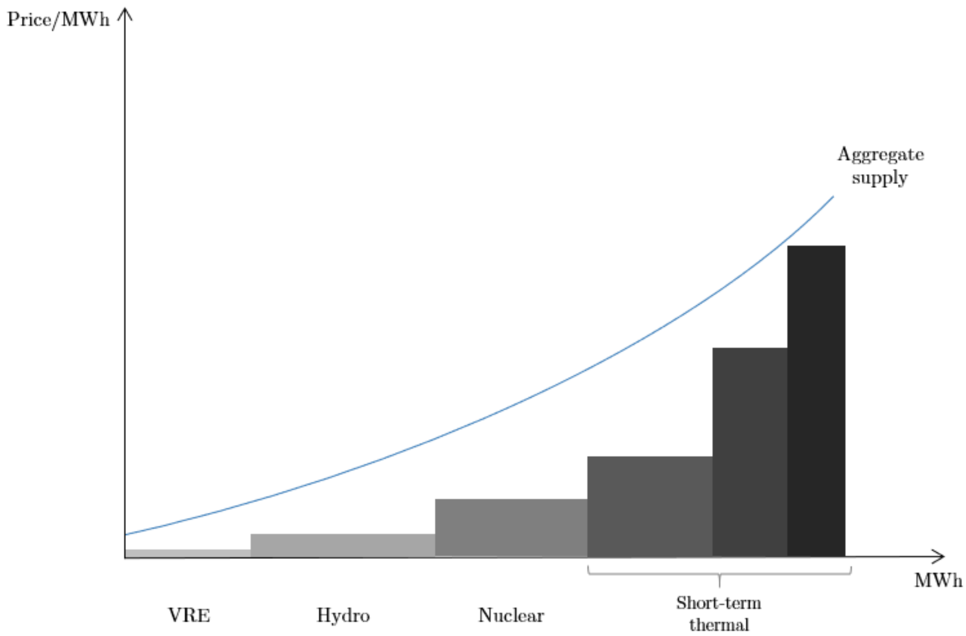

Figure 3: The merit order curve.

Beside the environmental effects of a higher renewable power penetration, its potential to reduce electricity prices has been highlighted as a benefit (Acemoglu et al., 2017). This is because of the different cost structure of renewable electricity generation, compared to conventional generation. As there is no fuel cost in renewable generation, it is characterized by very low marginal costs. This contrasts with conventional thermal sources like coal and gas that face much higher marginal costs.

As described in Section 2.1.1, when the market operator clears the day-ahead market, it aggregates the submitted supply curves of all producers in the market. In most cases, it is undesirable for an actor to bid below the marginal cost.[3] In a perfectly competitive market, the aggregate supply curve in effect reflects the marginal cost curve for the whole market. In this case, aggregate supply equals the aggregate cost curve in the market. At low quantities, only low cost technologies such as renewable and nuclear power are cleared. As the demand for production increases, increasingly expensive technologies are dispatched. This is illustrated in Figure 3.

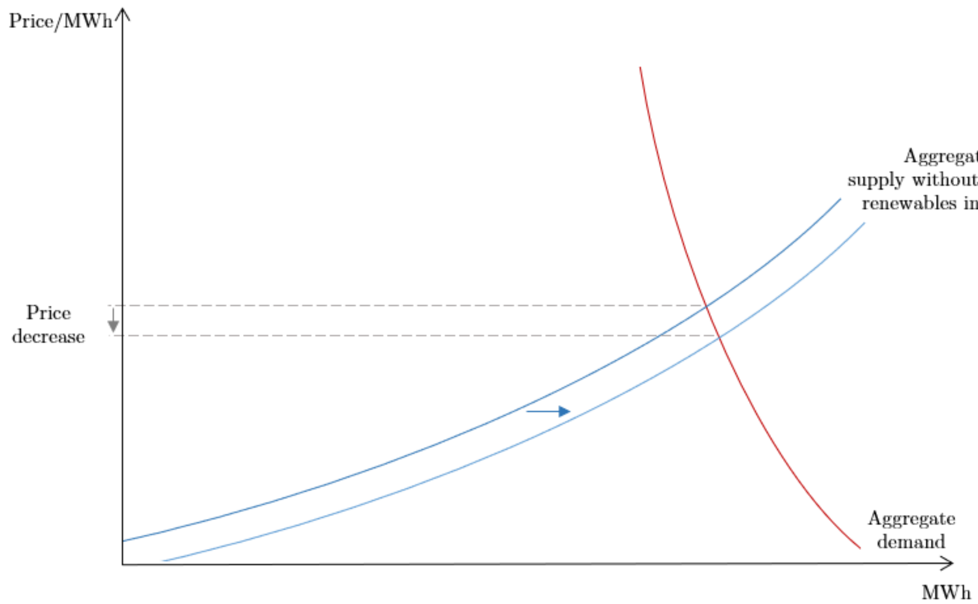

Figure 4: The merit order effect.

Besides the zero fuel cost of variable renewable energy sources (VRE), such as wind and solar power, these generation technologies also face no opportunity costs of generation since wind or sunshine cannot be stored for later generation. This translates into a marginal cost that is essentially zero. While hydro power also has no fuel cost, its opportunity cost is generally above zero since many hydro power plants store water in reservoirs and can choose when to produce by controlling the flow of water into the plant turbines. Therefore, the marginal cost of hydro power is higher than that of solar and wind power. Next comes nuclear power. After this, the aggregate supply curve is made up of conventional technologies in order of increasing fuel costs. When there is an increase in the supply of renewable electricity, this shifts the aggregate supply curve outwards. As can be seen in Figure 4, this should result in a lower

[3]An exception is if the cost of shutting down production is higher than the loss when price is lower than marginal cost.

market price, all else equal. This effect is commonly referred to as the merit order effect.

2.3 Susceptibility to market power in electricity markets

Since the deregulation of electricity markets began in the 1990s, the study of market power in the wholesale electricity market has attracted much attention both from market monitors and academic researchers. The main reason for this is that electricity markets are characterized by several attributes that increase the potential for market power compared to other markets.

Firstly, the generally high cost of storing electricity at scale limits the possibility of generation serving demand at later periods. Increased storage of electricity would have a competitive effect, as there would be more potential supply to compete for demand in any period. Secondly, hourly wholesale demand is very inelastic. Small end-users like households and small businesses generally enter flat rate electricity contracts with retailing companies and thus do not receive price signals reflecting variation in the relative supply and demand. This translates into a highly inelastic wholesale demand, increasing the potential for profitably raising prices. Lastly, electricity generation and distribution are characterized by strict capacity constraints. The cap of how much supply can be generated and distributed relative to demand in a region at a specific point in time is much lower than in other markets, and thus limits the ability of producers to compete for excess demand.[4]

These attributes combined make electricity markets very susceptible to market power, which can lead to market inefficiencies and transfers of wealth from consumers to producers. Additionally, because the electricity market infrastructure is shared between all market participants, and because the infrastructure is susceptible to congestion and blackouts due to grid frequency imbalance, the decision of one market participant to insert or extract electricity from the grid affects the infrastructure's stability for all market participants. It is therefore of interest for market monitors to align market participants' individually profitable behavior with actions that do not compromise the system security. Because of the magnitude of electricity markets and other sectors' dependence on it, the welfare implications of any kind of electricity market failure are potentially very big (Borenstein, Bushnell, and Knittel, 1999).

3 Literature review

In the following section we provide an overview of the literature related to our study. First, the general literature on market power in electricity markets is presented. Then, we cover the literature on congestion and its implications for market power. After that, we present both theoretical and empirical work on the merit order effect and how it is counteracted by diversified firms' ownership of renewable production. Lastly, we discuss the gap in the literature as we see it, motivating our study.

[4] Assume, for instance, that all but one supplier of electricity are supplying their full capacities. If only one percent of the demand is not met, the remaining firm that is not already supplying at max capacity will have monopoly power on that residual demand. Because of the uniform pricing rule that is used in electricity auctions, each generating firm will receive the monopoly price, leading to high markups over marginal cost across the market.

3.1 Market power in electricity markets

Researchers commonly differentiate between structural measures of market power, which measure the potential for profitably raising prices above marginal costs in a market, and behavioral measures of market power, which address whether such market power is actually exercised. Evidence of exercised market power has been found in several liberalized electricity markets. Using firm level data on posted bid curves, McRae and Wolak (2009) and Wolak (2003) find for New Zealand and California, respectively, that producing firms post higher bid curves when demand is less elastic. Similarly, using plant-level cost data, Wolfram (1999) find evidence of market power in the UK. When firm level data is not available, estimating market power is more difficult and researchers have usually relied on simulation or structural methods. This has largely been the case in the Nordics.

Borenstein et al. (1999) simulate several US electricity markets under Cournot competition of strategic firms and a non-strategic competitive fringe bidding their full capacity at marginal costs. Using detailed cost data, Cournot firm cost curves are constructed and market outcomes are then simulated under various demand scenarios. The results show that in high demand hours, when fringe producers are supplying at max capacity, Cournot firms strategically withhold production, leading to significantly higher prices and market power compared with the competitive outcome.

In the Nordics, Hjalmarsson (2000), Vassilopoulos (2003) and Bask et al. (2011) all study the power market using the Bresnahan-Lau conjectural variations approach. Hjalmarssson and Vassilopoulos find no statistically significant evidence of markups, while Bask does, however not at very economically significant levels. Lundin and Tangerås (2017) model the Nordic Nord Pool market as one of Cournot competition with a competitive fringe and exploit exogenous variation in the curvature of the residual demand curve to structurally estimate market power. Their results show that dominant firms in the Nordic market on average applied markups of 8%-11% during 2011-2013. Additionally, the authors find some evidence for market power increasing after the Swedish market was divided into bidding zones in 2011, indicating that the geographical configuration of the relevant market may be important for market performance.

3.1.1 Congestion

The importance of grid congestion in explaining market power has been documented in several studies. Wolak (2015) compares dominant firm bid curves with counterfactual curves capturing the firms' bidding behavior under perceived non-congestion to quantify the effect of a hypothetical expansion in transmission capacity in the Canadian Alberta electricity market, and finds that such expansion would yield strong competitive effects. Importantly, what is driving the results are dominant firms' decreased expectations of congestion in the grid, leading to submitted bid curves closer to marginal cost curves. Ryan (2017) simulates market outcomes under increased transmission capacity in the Indian Electricity Exchange, and finds that the competitive effects of grid expansion significantly affects strategic firm bidding, increasing total market surplus by up to 19%. In the Nordics, both Johnsen et al. (1999) and Steen (2005) find increased market power in congested periods in Southern Norway.

A subsection of the congestion literature concerns the endogeneity of congestion - more specifically if firms strategically induce congestion to achieve market power. Borenstein et al. (2000) show in a theoretical paper that a dominant firm may have incentives to limit generation in a zone in order to induce congestion in it. This hypothesis is investigated by Mirza and Bergland (2015), again in the Southern Norway. The authors find that firms deliberately congest transmission lines during the less elastic periods of late night and early mornings, leading to significantly higher markups during these periods. Whether congestion is exogenous or endogenous to withholding behavior is an important consideration for properly estimating causal effects of congestion. We address this issue in Section 5.

3.1.2 Congestion as a regime switch

Studying the effect of congestion on market outcomes is a potentially challenging topic for two reasons. Firstly, there may be variation in parameter estimates for different congestion configurations. Furthermore, congestion is a possibly endogenous event to market behavior, making standard approaches problematic. To properly disentangle market outcomes and congestion effects while allowing for model parameters to vary with congestion, some authors have instead suggested the use of regime-switching models. Crucial in this approach is the fact that parameter estimates are allowed to vary depending on the current regime, allowing the identification of heterogeneous effects. In electricity markets, regime switching models have mainly been applied in studying price dynamics with time series models. For instance, Weron et al. (2004) model price spikes as regime shifts with unobserved states. Congested grids may also be thought of as a changed regime, with the difference that the states are actually observed. Haldrup and Nielsen (2006) define Markov switching models allowing for three different congestion states and find that for Nordic market data, the price behaviour for single markets can vary depending on the presence or absence of bottlenecks in the electricity grid.

Our thesis differs from the above studies in that we do not employ a time-series approach. Our approach most closely resembles that of Sapio (2015), who employs an endogenous switching regressions approach following Lokshin and Sajaia (2004). In this paper, the author estimates the effects of increased renewables on price. He explicitly incorporates congestion into his model, since this changes the relevant market and thus the relevant geographical level of his variables. This framework also allows for the fact that there are two possible states - congested and uncongested - and that the realized state in each period is possibly endogenous to the model. A similar approach is used outside electricity markets by e.g. Läpple et al. (2013), who study the effect of government funded extension programs for farmers on farm profits. The framework accounts for possible self-selection, similar to how congestion may be endogenous in the electricity market context.

3.2 The merit order effect

The merit order effect has been empirically documented in several electricity markets. de Miera, del Río González, and Vizcaíno (2008) investigate increased wind power in Spain and find significant resulting price reductions, ranging between 4.75€/MWh-12.44€/MWh on average between 2005 and 2007, or 8.6%-25.1%. Sensfuß, Ragwitz, and Genoese (2008) get similar results for the German market. Importantly, they also find that while the merit order effect is sometimes completely absent in low-demand hours, in peak-demand hours it can amount to reductions of 36€/MWh. McConnell et al. (2013) find the merit order effect also for solar power in Australia, estimating the potential savings of a solar installation equivalent to the German per capita level of Solar power, to 8.6% of the total value traded through the Australian electricity pool in 2010. Würzburg, Labandeira, and Linares (2013) review previous research on the magnitude of the merit order effect and find that estimates vary both with studied regions and chosen assessment methods. However, based on their own analysis, the authors argue that the effects in different markets are less dispersed than is suggested by the literature.

3.3 Counteraction of the merit order effect

Several theoretical papers have shown that the ownership of renewable energy sources matters for market outcomes. Ben-Moshe and Rubin (2015) demonstrate that with linear inverse demand and quadratic costs in a Cournot market, firms with a diversified portfolio exercise market power by withholding their conventional production in response to their own renewable production. Building on this, Acemoglu et al. (2017) construct a simpler Cournot model with symmetric firms, but prove the results in a general setting with less strict assumptions on demand and cost functions. In their model, the oligopolistic firms own all thermal production units and some renewable units, but renewable units may also be owned by a competitive fringe. The authors show that ownership of renewable sources by oligopolistic firms counteracts the merit order effect, because the firms withhold conventional generation in response to their own renewable generation. Furthermore, when all renewable production is owned by oligopolistic firms, this counteraction is total and increased renewables production causes no price decrease in the market. The only change is that the firms make larger profits on their units sold, as the marginal cost of production is lower with the renewable units than with the thermal units.

In a recent paper, Genc and Reynolds (2019) show that with asymmetric cost functions, the above effect is stronger for a strategic firm with a steeper cost function. They also provide simulation results of their model and basic empirical tests for the Ontario electricity exchange, and find support for their theoretical predictions of a counteraction effect. Lastly, Butner (2019) creates a structural empirical model and uses individual firm bidding data from the MISO market in the US to show that diversified firms withhold conventional power in response to an increase in their own wind production also on this market.

3.4 Specification of research focus

Based on our survey of the literature, there is theoretical support for why the ownership of wind power matters for market outcomes, specifically the counteracting effect of generation portfolio diversification on the merit order effect. The empirical evidence of such an effect is to our knowledge thin but so far convincing, however limited to two electricity markets in Canada and USA. Electricity markets in North America commonly use nodal pricing, which is different from the system with zonal pricing in the Nordics. Furthermore, Butner (2019) uses individual firm data, something which is often not available to researchers in other markets. Thus, there is both a need for empirical evidence of the proposed effect in markets outside North America, and also developments in how to analyze this without firm level data.

In addition to the above, it is clear from previous work that congestion matters a great deal for market outcomes on electricity markets in general, and neglecting it in both theoretical arguments and empirical work around the counteraction of the merit order effect could be problematic. This is a gap in the literature that we intend to fill. We thus arrive at the following research questions:

- *To what extent is the merit order effect counteracted if firms with diversified production technologies supply wind power in the Swedish electricity market?*
- *Is the above described effect pronounced by congestion?*

We base our study on the theoretical model by Acemoglu et al. (2017), adopt it to the Swedish electricity market, and introduce a way of considering congestion in this framework. From this model, we derive testable predictions, and empirically test both the effect of wind power ownership and how this effect interacts with congestion.

4 Theoretical framework

In this section, we present the model of Acemoglu et al. (2017), adapt it to the Nordic market and show how to analyze network congestion with this framework. Based on this, we derive our testable predictions.

4.1 Model economy

Acemoglu et al. (2017) consider an oligopolistic energy market consisting of $n \geq 2$ diversified producers owning both thermal electricity plants and renewable plants. There is also a competitive fringe that only owns renewable production and supplies its full capacity at the marginal cost of production.[5] Each diversified actor produces q_i units thermal generation at the production cost $C(q_i)$, where C is a convex and differentiable function.

In the economy, there is a total of R units of renewable production, always supplied at zero marginal cost. Each diversified firm owns the same share δ/n of the renewable production such that the total diversified share of renewable production is $\delta \in [0, 1]$ and the fringe's share of the renewable production is $(1 - \delta)$. Total electricity production is given by $Q + R$, where $Q = \sum_{i=1}^{n} q_i$ is the total conventional output. Inverse aggregate demand is given by $P(Q + R)$, where P is a non-increasing differentiable function.

Each strategic firm engages in Cournot competition with its thermal production, choosing the quantity q_i to maximize its profit, which is given by

[5]Modeling fringe firms as non-strategic is common in the literature. While it is true in theory that even the smallest firm could be the pivotal bidder and thus unilaterally set the market price, larger firms will in almost all circumstances have stronger incentives to withhold, and their withholding diminishes the incentives for small firms to do so (see Borenstein et al., 1999, for details).

$$\Pi_i = (q_i + \delta R/n)P(Q+R) - C(q_i). \tag{1}$$

To see the merit order effect in this economy, we differentiate aggregate demand with regard to the amount of renewable production:

$$\frac{\partial P}{\partial R} = P'(Q+R)(\frac{\partial Q}{\partial R} + 1) \tag{2}$$

If the merit order effect is present, the price P should be non-increasing in the amount of renewables R, so that $\frac{\partial P}{\partial R} \leq 0$. By assumption, $P' < 0$. To see that $\frac{\partial Q}{\partial R} + 1 \geq 0$, we look at the firms' equilibrium strategies. We show the following for the duopoly case $n = 2$, but the extension to $n \geq 2$ is straightforward. Each firm chooses the profit-maximizing quantity q_i^* in equilibrium, giving us the first order condition

$$\frac{\partial \Pi_i}{\partial q_i} = P(Q^* + R) + (q_i^* + \delta R/2)P'(Q^* + R) - C'(q_i^*) = 0, \tag{3}$$

where q_i^* and Q^* denote individual and aggregate equilibrium quantities of diversified firms, respectively. By symmetry, $q_1^* = q_2^* = ... = q_n^*$ such that $q_i^* = Q^*/n$. Substituting Q^*/n for q_i^* in equation 3, differentiating w.r.t. R and rearranging[6], we get

$$\frac{\partial Q}{\partial R} = -\frac{(2+\delta)P'(Q+R) + (Q+\delta R)P''(Q+R)}{3P'(Q+R) + (Q+\delta R)P''(Q+R) - C''(\frac{Q}{2})}. \tag{4}$$

Assuming linear cost functions, it is the case that $C'' = 0$. Also, because $P' < 0$ and $P'' < 0$ by assumption, it holds that

$$-1 \leq \frac{\partial Q}{\partial R} < 0 \Rightarrow \frac{\partial Q}{\partial R} + 1 \geq 0. \tag{5}$$

Thus, $\frac{\partial P}{\partial R} \leq 0$. In other words, the price is non-increasing in the amount of renewables and there is indeed a merit order effect in the economy. The intuition for this is straightforward: a larger amount of renewable energy that is offered at zero marginal cost will push out the whole supply curve to the right. This means that absent any strategic behavior, less costly production will now satisfy the last unit of demand at each quantity level so that aggregate supply will intersect aggregate demand at a lower level, resulting in a lower market clearing price.

Next, we show the impact on prices of a higher δ, the share of the renewable generation that is supplied by the diversified firms. To investigate this effect, we differentiate the demand function $P(Q+R)$ w.r.t. δ:

$$\frac{\partial P}{\partial \delta} = (\frac{\partial Q}{\partial \delta})P'(Q+R) \tag{6}$$

Again, $P' < 0$ by assumption and we are thus interested in the sign of $(\frac{\partial Q}{\partial \delta})$ to determine the effect of δ on P. Again, substituting Q^*/n for q_i^* in equation 3, differentiating with regard to δ and rearranging,

[6] For details, see the online appendix of the Acemoglu et al. (2017) paper.

we get

$$\frac{\partial Q}{\partial \delta} = -\frac{RP'(Q+R)}{3P'(Q+R) + (Q+\delta R)P''(Q+R) - C''(\frac{Q}{2})} \leq 0, \tag{7}$$

since $R \geq 0$, $P' < 0$, $P'' < 0$ and $C'' = 0$. Thus, $\frac{\partial P}{\partial \delta} \geq 0$, and the price is increasing in the share of renewable production that is owned by the diversified firms - a counteraction of the merit order effect. The intuition for this result is that when a producer is diversified, the lower marginal cost on the renewable units makes the higher prices from withholding production increase profits more than in the case of a fringe firm. Thus, diversified firms face higher incentives to withhold conventional generation to keep prices high. It can be shown that in the extreme case of $\delta = 1$, the neutralization of the merit order effect under linear cost functions is total, so that each renewable unit increase is matched by a unit decrease of conventional production, keeping market prices unaffected.

Assuming linear demand of the form

$$P(Q+R) = \alpha - \beta(Q+R), \tag{8}$$

where α and β are intercept and slope parameters, respectively, we get the following equilibrium conditions for individual and aggregate strategic firm conventional output and price, respectively:

$$q_i^* = \frac{1}{(n+1)\beta}(\alpha - \gamma - \beta(R + \delta R/n)) \tag{9}$$

$$Q^* = \frac{n(\alpha-\gamma) - \beta(\delta R + nR)}{\beta(n+1)} \tag{10}$$

$$P^* = \frac{1}{n+1}(\alpha + \beta(-R + \delta R) + n\gamma). \tag{11}$$

The previous observations about the merit order effect and its counteraction can easily be seen in the equilibrium conditions as well. In (11), the amount of renewables R enters as $(-R + \delta R)$. When $\delta < 1$, $(-R + \delta R) < 0$ and P^* decreases in R. When $\delta = 1$, $(-R + \delta R) = 0$, meaning that the merit order effect is fully counteracted. For the diversified firms' equilibrium conventional output, δ enters negatively so that q_i^* and Q^* decrease in δ, showing the increased incentives to withhold conventional production for a more diversified firm.

4.2 Considerations for the Swedish market

Several considerations need to be made with regard to the fit of the above model to the Swedish electricity market.

First, the model assumes that the diversified actors are symmetric. Although a common assumption in Cournot models, it is clearly not realistic as firms vary in size, generation portfolios, locations of plants, and more. This assumption is, however, not crucial to the general results of the model, as shown by Genc and Reynolds (2019) who find the counteraction of the merit order effect in a setting with asymmetric firms.

Second, in the Acemoglu et al. (2017) model, the non-strategic competitive fringe is made up by all actors who are non-diversified. In reality, diversified actors may be too small to be able to impact market prices much, and actors that are both diversified and large enough to affect prices may be non-commercial, for instance because they are organized as cooperatives. This is quite often the case in Sweden, where both wind power and short-term strategic resources are regularly cooperative-owned. In our setting, we define strategic actors as the five largest and diversified electricity producers in Sweden: Vatenfall, E.ON (including Sydkraft), Fortum, Statkraft, Skellefteå Kraft, and the fringe as the rest. The production technology mix for each strategic firm is presented in Table 1. The market is dominated by Vattenfall, housing roughly a third of all installed capacity. Together with the other four largest actors, they own 70% of all installed capacity.

Company	Hydro	Nuclear	Wind	Thermal	Total
Vattenfall AB	7917	4954	303	924	14098
Sydkraft AB/E.ON Sverige AB	1794	2464	165	2036	6459
Fortum AB	3063	1553	42	627	5285
Statkraft Sverige AB	1262	0	334	1	1597
Skellefteå Kraft AB	655	64	272	54	1045
Total 5 largest firms	14646	9035	1116	3642	28484
Total Sweden	16181	9076	6520	8042	40004

Table 1: Overview of the five largest actors: Production technology mix by installed capacity in MW for the five largest actors on the Swedish electricity market as of January 2017.
Source: Authors' illustration of data from Energiföretagen (2017).

Third, in the Acemoglu et al. (2017) model there are only two technologies for generating electricity - short-term strategic thermal generation and renewable generation that is always offered at zero marginal cost. In reality, there exist both renewable and thermal generation that is short-term strategic, as well as renewable and thermal resources that are unable to change strategically in the short term. While a renewable resource, storable hydro power should for instance be considered a short-term strategic resource as dam gates can be opened and closed with short notice, and saved production has a value in that it can be used tomorrow. Similarly, nuclear power is a thermal source, but is costly to turn up and down and is therefore not fit for short-term production optimization. These two considerations are important for the present case of Sweden, where hydro power and nuclear power generally constitute a majority of total production.

Based on the above, we do the following adjustments to the Acemoglu et al. (2017) framework. The only relevant renewable power is wind power, and it is like in the original model supplied fully at zero marginal cost. Diversified firms engage in Cournot competition with their short-term strategic generation, which we define as hydro power and short-term controllable thermal power like coal, gas and biomass. Nuclear power is considered base load that is supplied non-strategically in the short term and thus does not enter into the model.

Fifth, and lastly, the framework does not include explicit ways to model congestion in the electricity grid. Congestion is crucial to understand how market power and price formation work on electricity markets and we thus argue that the exclusion of congestion in the model limits its applicability for the analysis of real electricity markets. In the next section we will show how to interpret congestion in the framework and use it to analyze the effect of congestion on market outcomes.

4.3 Congestion

When a bidding zone is congested, the relevant market of which that zone is part is effectively reduced. Thus, at each price, the demand that producing firms are facing is reduced, translating into the aggregate demand curve of the relevant market rotating inwards. This means that the demand is now less elastic, so that the demand slope is steeper. In the demand function in (8), this effect is captured by an increase in the slope parameter β.

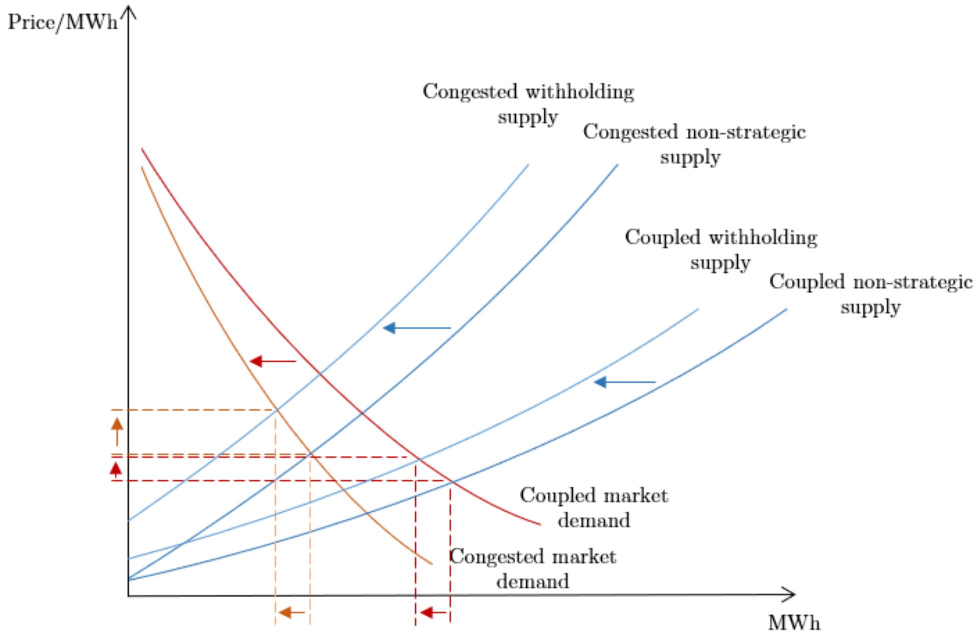

Figure 5: The effect of a decrease in market size on demand elasticity and withholding strategies.

This has two counteracting effects on the equilibrium price. Firstly, the general price reduction from renewables will be stronger since demand is less elastic, and thus price movements from supply shifts are stronger. Secondly, the incentives to withhold for diversified firms are stronger, since less elastic demand means that the reduction in quantity sold due to higher prices will be lower, and thus lead to increased profits from withholding. This can be seen in the equilibrium price equation 11, where $-\beta R$ is the first effect and $+\beta\delta R$ is the second effect. When $\delta < 1$, the first effect will dominate the second effect, making the net effect of an increase in β on prices negative.

In all three equilibrium condition equations, however, β enters multiplicatively with δ. This means that there is an interaction effect between the two, reinforcing one another such that the effect of δ

increases as β increases, and vice versa. The intuition behind this goes as follows: As δ increases, the incentives to withhold conventional sources increase because of the increasing returns of withholding, as explained previously. If the grid becomes congested so that the relevant market is decreased, β will increase, meaning that demand will decrease less in response to price increases from withholding. Therefore the benefit of withholding is pronounced, further increasing the increased incentives of withholding from a higher δ. In (9) and (10), this can be seen on the negative $-\beta\delta R$ terms, reducing equilibrium output of diversified firms. This then translates to relatively higher equilibrium prices, as seen by the positive $\beta\delta R$ term in (11). Figure 5 shows the effect of congestion on firms' general withholding incentives. Congestion makes demand less elastic, which makes withholding relatively more profitable - for the same degree of withholding, the resulting price increase is larger in the congested (decoupled) market than in the coupled market.

Importantly, when a zone is import constrained, no generation from outside the zone can enter into it, and all electricity generated in the zone will be consumed in the zone. Therefore, we expect to see pronounced withholding behavior in zones specifically in the periods that they are import constrained. While there may be strategic behavior outside of the relevant zone that is inducing or sustaining congestion, as long as the congestion is sustained such behavior is not affecting market outcomes in the congested zone on the margin. This is instead determined by the relative supply and demand *within* the zone. Thus, in periods of congestion, δ should reflect the share of renewable production owned by diverse producers within the import constrained zone, rather than in the whole country. Similarly, the relevant short-term strategic capacity is that which lies within the zone. An important assumption is thus that there still exists both wind power owned by diverse firms, as well as sufficient short run strategic resources, in the import constrained zone. If this is not the case, we would expect no strategic behavior due to δ in times of congestion. Furthermore, in our framework we are interested in how strategic behavior in a relevant market is affected by the exogenous addition of congestion, as opposed to strategic behavior that may endogenously induce and sustain congestion. In Section 5.2, we address how to make such a distinction empirically.

4.4 Testable predictions

Based on the theoretical framework of Acemoglu et al. (2017) and our discussion of how congestion can be analyzed with it, we thus argue that the increased incentives of withholding short-term strategic generation from an increased diversified share of renewable production will lead diversified firms to withhold more short-term strategic generation when they supply a higher share of the renewable generation. This will translate to higher market prices when the diversified share of renewable generation is higher.

Because of the decrease in the price elasticity of demand in a zone that is congested, the incentives to withhold are further increased and the predicted negative effect of δ on output and the positive effect on prices of the diversified share of renewable generation should be further increased. We thus arrive at the following testable predictions about the Swedish electricity market, to test our research questions:

- *For a given level of wind power production, the share of the wind production that is supplied by* *diversified firms in the Swedish market has a negative effect on the short-term strategic output by diversified firms.*

- *The above effect translates into higher prices, so that the diversified share of renewable production has a positive effect on prices.*
- *The above effects are stronger in relevant zones under periods of congestion.*

5 Method

In this section we present our strategy for testing our theoretical predictions to be able to answer our research questions. We begin by discussing considerations for identifying strategic behavior in a market based on observed aggregate outcomes. With this in mind, we then present our empirical strategy, and continue by discussing our identifying assumptions. Lastly, we present the data that we use.

5.1 Identifying strategic behavior

Under the uniform price rule which is used in the Nordic market, the marginal bid (the highest-priced dispatched supply bid) sets the price for all dispatched units. As such, firms do not unilaterally choose the prices they will receive for their production, as is common in most markets. Instead, the price that each producer receives is set by the firm that produces the last dispatched unit of electricity. Each firm whose supply is needed to satisfy total demand (taking other firms' supply as given), is said to face a residual demand curve. On this demand, the firm is a monopolist and unilaterally sets the price. Taking other firms' bids as given and assuming that bid curves are all non-decreasing (in quantity), a firm can try to exercise market power either by bidding less capacity at marginal costs, physical withholding, bidding its capacity at bids higher than marginal costs, economic withholding, or a combination of the two. Any of these strategies shifts the aggregate supply curve inwards, resulting in a higher market clearing price (Biggar and Hesamzadeh).[7]

It follows that withholding behavior can be studied straightforwardly if one has access to individual firms' posted bid curves - our theoretical prediction would then be that diversified firms' posted bid curves shift inwards when δ (hereafter delta) increases, and especially so in periods of congestion. Firm level bid data is, however, unfortunately not available for the Nordic market. Instead, we use aggregate data on short-term strategic output as well as on prices. If our hypothesized effects exist, inwards-shifted bid curves should translate into measurable effects on aggregate market outcomes. This assumption is realistic given the strong link between posted bids and assigned production quantities.

[7] It will not necessarily increase the market price because sometimes parts of the supply curve will be step-wise, so that a shifted curve may still be intersected by the demand curve at the same price level. However, as the probability of a higher market price is increased by engaging in any kind of withholding, this does not change the implications for firms' strategies.

5.2 Empirical strategy

The causal relationships proposed in our theoretical predictions can be estimated with the following empirical models, where we choose to specify all continuous variables (except delta) in logs so that coefficient estimates can be interpreted as elasticities and semi-elasticities:

$$\ln(Q_t) = \beta_0^Q + \beta_1^Q \ln(W_t) + \beta_2^Q \delta_t + \beta_3^Q C_t + \beta_4^Q (\delta_t \times C_t) + \mathbf{X}_t \beta + \mathbf{a}_t \alpha + \epsilon_t^Q \tag{12}$$

$$\ln(P_t) = \beta_0^P + \beta_1^P \ln(W_t) + \beta_2^P \delta_t + \beta_3^P C_t + \beta_4^P (\delta_t \times C_t) + \mathbf{X}_t \beta + \mathbf{a}_t \alpha + \epsilon_t^P, \tag{13}$$

where Q_t is short-term strategic output (hydro power or thermal power), P_t is the market price, W_t is wind power production, δ_t is the share of wind power production owned by diversified firms and C_t is a dummy variable equal to 1 if the market is import-congested and zero otherwise. Finally, $\mathbf{X}_t$ is a vector of controls, $\mathbf{a}_t$ is a vector of time fixed effects, and ϵ_t^Q and ϵ_t^P are error terms. The superscripts Q and P denote parameter estimates for the chosen output variable and market price as dependent variables, respectively.

Assuming proper identification, β_1^Q is the causal effect of additional wind power production on short-term strategic output, β_2^Q is the causal effect on short-term strategic output from an increase in the diversified share of renewable electricity production, and β_4^Q is the change in the latter effect of induced congestion. Lastly, β_0^Q is an intercept and β_3^Q is the effect of congestion on short-term strategic output. In (13), coefficients are interpreted equivalently, and β_1^P is the merit order effect.

We define our vector of controls as $\mathbf{X}_t = \{\ln(N_t), \ln(I_t), \ln(D_t)\}$, where N_t is nuclear power production, I_t is hydro inflow, and D_t is forecasted demand. Nuclear power is base load - it is produced at relatively constant levels over time, with little or no short-term strategic capabilities. However, any variation in its supply should have an effect on market outcomes and the potential for market power. Thus, we include it as a control. At first glance, it might seem reasonable to also control for the amount of hydro and thermal power in the regressions where they are not the dependent variables. However, they would then be what Angrist and Pischke (2008) refer to as "bad controls".[8] Instead, another way to control for hydro power conditions is to include hydro inflows as a supply shifter. This is a weekly measure of how much much water flows into the hydro reservoirs, and is standard in the literature to use as a control, since it only depends on precipitation and thus is exogenous. Furthermore, it is an important determinant for how much strategic behavior is possible (Lundin, 2017).

Our last control, forecasted demand, accounts for the size of the market. For a given aggregate supply curve, prices may increase simply due to an outward shift of the demand curve, and vice versa. Since we intend to study the effect of supply-side strategic behavior on prices and quantities cleared, we wish to remove any effects that changes in demand has on those market outcomes. One approach common in the literature is to to control for the total consumption, since this should proxy well for total demand. However, since the demand curve is not perfectly inelastic (vertical), such an approach could

[8]This is the case if a control variable actually is a causal channel through which our explanatory variable affects our dependent variable. Since our hypothesis is that diversified producers change their supply decisions of hydro and thermal power (technically their bidding decisions) in order to affect the price, this would clearly be an issue.

suffer from reverse causality issues. That is, price and supply factors depend on actual consumption, and actual consumption depends on price and supply factors. Thus, to properly isolate demand shifts, we choose to use forecasted demand. This variable is generated by Svenska Kraftnät one day ahead of delivery, and is based on factors such as temperature prognoses and other indicators exogenous to our model. Given that we control for this, we are able to isolate supply effects.

Lastly, we define our vector of fixed effects to include hour-of-day dummy variables. Thus, we account for invariant time patterns in market outcomes, for instance increasing demand in the evening when electrical lights are used more, and systematic trends in wind speed during a day.

We now go on to describe our approach in more detail, presenting both the specifications and estimation methods that we use. We begin by abstracting from import transmission constraints and directly test our baseline model. After that, we first add external congestion from other countries, and then also incorporate internal congestion between zones in Sweden.

5.2.1 Baseline model

In our baseline model we ignore potential effects of congestion by letting C_t in (12) and (13) always equal zero. This corresponds to directly testing the predictions by Acemoglu et al. (2017). Since both the amount of wind power and delta are based on wind speeds, they should be exogenous. As such, we may estimate their effects on our dependent variables with OLS. Based on our predictions from Section 4, we expect a positive coefficient on delta when price is our dependent variable. When one of the output variables is our dependent variable, we instead expect a negative coefficient, reflecting withholding behavior.

In our initial regressions, we use our full sample and price in SE3 as our dependent variable. This could be problematic, since in times of congestion the size of the relevant market changes and hence the relevant level of controls is no longer country wide. Thus, we also run the same regressions but restrict our sample to periods where there is no internal congestion in Sweden (so that the price is the same in all four zones). This ensures that our explanatory variables are at the correct geographical level.

5.2.2 External congestion

Next, we extend the models to account for the effects of transmission constraints by letting C_t take on the value of one in periods of congestion. We begin by studying congestion from other countries to Sweden. We will consider the two most frequently congested connections, those from Denmark (DK1 and DK2), and southern Norway (NO1). To allow for heterogeneous effects of congestion for the two different transmission lines, we perform the analyses separately for Denmark and Norway. Effects could be heterogeneous, for instance, because the production technology mix differs for the two countries. Thus, in one case C_t is equal to one when there is congestion from Denmark. In the other case, C_t is equal to one whenever there is congestion from Norway.

As an initial approach we estimate (12) and (13) with OLS. Based on our theoretical predictions, we expect the coefficient β_4^Q to be negative, and β_4^P to be positive. The above approach assumes that congestion is exogenous. However, while withholding may be an effect of congestion, it may also be a cause of congestion. Consider the case when a producing firm, for whatever reason, cuts back short-term strategic output in a zone. Demand levels, that are largely determined by consumer behavior and weather factors, would not go down just because supply has, and therefore more imports will be needed in the area to clear the market. Regardless if this is a strategy by the firm to induce congestion and push prices up, or because of a random non-strategic event, the risk for congestion into the zone has clearly increased because of the reduced production by the firm. Because of this, congestion may be endogenous to our models and since our OLS estimates do not account for this endogeneity, they are potentially biased.

To avoid such bias we also estimate the above models with an instrumental variables (IV) approach. To instrument for congestion, we use planned transmission line outages by the grid operator Svenska Kraftnät. Such outages occur for instance when there is a need for maintenance work on the transmission grid, and effectively reduce the available transmission capacity. Therefore, the outages should increase the probability of congestion, making it a relevant instrument. The instrument is also valid, since transmission outages are determined and implemented by the grid operator Svenska Kraftnät, which is unrelated to the supplying and demanding agents on the electricity market. Therefore, we do not expect transmission outages to affect our dependent variables, short-term strategic output and prices, in any other way than through increasing the probability of congestion.[9]

An important feature here is that our endogenous variable, C_t, is binary. A temptation would therefore be to estimate the first-stage regression with a probit estimator and then use the fitted values in the second-stage. However, this is what is called a forbidden regression, since it uses non-linearities in a 2SLS estimation process, which is not allowed. In fact, only OLS is guaranteed to produce residuals in the first-stage that are uncorrelated with fitted values and covariates (Angrist and Pischke, 2008). Thus, we follow Angrist and Pischke and estimate the first-stage with OLS, corresponding to an ordinary 2SLS approach. While this can reduce efficiency, it ensures consistent estimates. Due to the high number of observations in our sample, we consider the issue of efficiency secondary to consistency. Note that since C_t appears both as a main effect and as an interaction effect with delta in our specifications, we effectively use two instruments - planned transmission outages on its own and interacted with delta.

An important difference between the IV estimator and the OLS estimator is that when treatment effects are heterogenous, the IV estimator is the local average treatment effect (LATE) (Imbens and Angrist, 1994). In the OLS approach used above, we instead estimate average treatment effects - the average effect of all congestion periods on our dependent variable. This implicitly assumes that the treatment has the same effect, regardless of what "type" of congestion it is. The LATE allows for heterogeneity, and instead estimates the average effect for those observations whose treatment status is changed by the instrument. In our case, this translates to congestion that is induced by planned outages.

This is an important point, since we should in fact expect that the effect is larger in such periods. To see this, consider the fact that not all congestion is predictable. Non-predictable congestion can only become known *after* the day-ahead bidding stops. In that case, strategic actors have no way of changing

[9] For an in-depth discussion on relevance and validity of IV instruments, see Angrist and Pischke (2008).

their day-ahead market behavior in response to the congestion. Planed outages are usually announced before the day-ahead market closes, and given that they increase the probability of congestion, strategic actors have a possibility to change their bidding on the day-ahead market in response to this. Thus, we should see a larger effect due to congestion in the IV specification.

However, confirming that this is actually the reason for an increase in the IV coefficient is not straightforward. In general, IV estimates could become larger than OLS estimates due to a variety of reasons. It could be that the OLS estimates in fact are biased due to endogeneity, or it could be a symptom of weak or invalid instruments (where we may only test for the former). A simple approach to try to identify the source of any increase in the IV estimate is to run OLS, but change the definition of the congestion dummy to periods in which there is congestion *and* planned outages. If the estimated effect of this congestion is larger than in the regular OLS, it would indicate that there is heterogeneity in the treatment effect. Note that this estimate may still be subject to endogeneity, since the congestion variable is still used to select which variation in the planned outages variable we use in the newly created variable. Therefore, these results should not be seen as estimates of the effect of congestion. They are rather indicators of the degree to which OLS and IV estimates of our main specification differ because of heterogeneous treatment effects.[10]

5.2.3 Internal congestion

Next, we consider internal congestion. Similarly to external congestion from another country to Sweden, there may be internal congestion between price zones in Sweden. As previously argued, under congestion, the only thing that matters for market outcomes on the margin is what happens within the import constrained bidding zone. If, for instance, SE4 is import constrained from the northern zones, strategic behavior in order to affect market outcomes in SE4 on the margin can only be undertaken from *within* SE4. Thus, the relevant geographical level of our variables changes when there is internal congestion.[11]

A naive solution to address this would be to split our sample and run regressions separately for each zone when they are congested. As emphasized previously, however, congestion is possibly endogenous. Thus, splitting the data based on congestion could lead to selection bias. Instead, we use an endogenous switching regression approach similar to that used by Sapio (2015). In this approach, there are two states of the world, determined by a selection equation. For each possible state, there is one outcome equation and the variables differ for each state. Furthermore, the state of the world is allowed to depend on both exogenous factors and the variables in the outcome equations. This makes it suitable for modelling congestion as a potentially endogenous selector of the relevant geographical level of variables. Formally, consider the following model:

[10] An alternative approach could be to run OLS for periods in which planned outages are more common. If we then observe a higher effect from congestion, we could conclude that there are heterogeneous effects. However, when observing the data we find little to no pattern in which hours and days planned outages occur in. Thus, this type of analysis is not possible.

[11] Note that this is also the case for when there is external congestion. But since we lack data on markets outside Sweden, we can only distinguish between different geographical levels of our variables when there is congestion *within* Sweden.

$$
\begin{aligned}
& & C_t = 0 \quad & \text{if} \quad Z_t\gamma + u_t \leq 0 \\
& & C_t = 1 \quad & \text{if} \quad Z_t\gamma + u_t > 0 \\
\text{Regime}\,1: \ \text{SE} & \quad y_{1t} = \mathbf{X}_{1t}\beta_1 + \epsilon_{1t} & \text{if} \quad & C_t = 0 \\
\text{Regime}\,2: \ \text{Zone} & \quad y_{2t} = \mathbf{X}_{2t}\beta_2 + \epsilon_{2t} & \text{if} \quad & C_t = 1,
\end{aligned}
\tag{14}
$$

where C_t is the state variable, Z_t is a vector of variables that determine the state, y_{1t} and y_{2t} are the dependent variables of the outcome equations, and $\mathbf{X}_{1t}$, $\mathbf{X}_{2t}$ are vectors of explanatory variables that differ according to the state and u_t, ϵ_{1t} and ϵ_{2t} are error terms. For our purposes, $C_t = 1$ translates to internal congestion between two Swedish zones. As before, we may use either price, hydro power or thermal power as the dependent variable, and the relevant level of the variables will vary depending on the state.

Consider, for instance, the case of thermal power in SE4. If SE4 is congested from SE3 (so that $C_t = 1$), we use thermal power production in SE4 as the dependent variable, and in the non-congested state (when $C_t = 0$) we use national thermal production as the dependent variable. Similarly, the geographical level of the explanatory variables in $\mathbf{X}_{1t}$, which are used when there is no congestion, is nationwide. Equivalently, the level of the explanatory variables in $\mathbf{X}_{2t}$ is the import constrained zone. Specifically, $\mathbf{X}_{1t} = \{\ln(W_t^{SE}), \delta_t^{SE}, \ln(N_t^{SE}), \ln(I_t^{SE}), \ln(D_t^{SE})\}$ and $\mathbf{X}_{2t} = \{\ln(W_t^{Zone}), \delta_t^{Zone}, \ln(N_t^{Zone}), \ln(I_t^{Zone}), \ln(D_t^{Zone})\}$, where the variables are defined as before and the superscript *SE* means that the geographical level is for the whole of Sweden, while the superscript *Zone* means that the variable is on a zonal level (for instance SE4).

Let Σ be the variance-covariance matrix of the error terms ϵ_{1t}, ϵ_{2t} and u_t. We make the following assumptions:

- $\epsilon_{jt} \sim N(0, \sigma_j^2), \quad j = 1, 2$
- $u_t \sim N(0, \sigma_u^2)$
- $\sigma_{12} = 0$
- $\sigma_u^2 = 1$

The first two assumptions are the usual assumptions of maximum likelihood models. The third assumption is common practice in the literature and is needed since the outcomes of different states are never observed together and thus σ_{12} cannot be estimated. The last assumption is needed since γ in (14) can only be estimated up to a scalar factor. Putting these together, we can write:

$$
\Sigma = \begin{pmatrix} \sigma_1^2 & & \\ 0 & \sigma_2^2 & \\ \sigma_{1u} & \sigma_{2u} & 1 \end{pmatrix} \tag{15}
$$

To understand the need for the switching regressions framework in the presence of endogeneity, consider the following. In the case that C_t is exogenous, $\sigma_{1u} = \sigma_{2u} = 0$. That is, each of the equations in (14) are independent of each other. In that case, a switching regressions framework is redundant, and each of the outcome equations may be estimated separately using sample splitting. However, if C_t is endogenous to the model, $\sigma_{1u} \neq 0$ and $\sigma_{2u} \neq 0$. That is, there is not independence between the error term in the selection equation and the error terms in the outcome equations. Not accounting for this and regressing each equation independently would then bias the results. Estimating the above equations simultaneously with maximum likelihood estimation corrects for this.

Given the above assumptions, the log likelihood for the system in (14) is (Dutoit, 2007):

$$\begin{aligned} \ln L = \sum_t \Bigg(& C_t w_t \Big[\ln\Big\{ F(\eta_{1t}) \Big\} + \ln\Big\{ f(\epsilon_{1t}/\sigma_1)/\sigma_1 \Big\} \Big] + \\ & (1-C_t) w_t \Big[\ln\Big\{ 1 - F(\eta_{2t}) \Big\} + \ln\Big\{ f(\epsilon_{2t}/\sigma_2)/\sigma_2 \Big\} \Big] \Bigg) \end{aligned} \tag{16}$$

where F is a cumulative normal distribution function, f is a normal density distribution, w_t is an optimal weight for observation t, and

$$\eta_{jt} = \frac{(Z_t\gamma + \rho_j \epsilon_{jt}/\sigma_j)}{\sqrt{1-\rho_j^2}} \quad j = 1, 2 \tag{17}$$

where ρ_1 is the correlation coefficient between ϵ_{1t} and u_t, and ρ_2 is the correlation coefficient between ϵ_{2t} and u_t.

Z_t, contains instruments for the selection variable, requiring similar assumptions about relevance and validity as in an instrumental variables approach. If no instrument is specified, the model may still be identified by using non-linearities. However, since we have a suitable instrument, planned transmission outages, we use this to instrument for congestion just as in our IV approach. To do the estimation, we use the sampleSelection package in R (Toomet and Henningsen, 2008), which uses the Newton-Raphson algorithm.

In our analysis with this approach we focus on SE3 and SE4, as these are the zones where there is a meaningful amount of congested hours from the other Swedish zones. We consider these jointly (from now on referred to as SE3SE4). That is, the congested state is defined as periods where SE3 is import constrained from SE2, but SE3 and SE4 are internally unconstrained. This allows us to study the effects of internal congestion while maintaining a sufficient amount of observations for the congested state. In 1748 periods, SE3 and SE4 are not congested between each other, but import constrained from SE2. Out of those, SE3SE4 is congested from Denmark in 280 periods and from Norway in 1196 periods.

The first analysis we do with this approach is to compare estimates for the uncongested regime with estimates for the congested regime. Based on our theoretical framework, we expect stronger effects of delta under congestion since the relevant market is smaller and thus demand should be less elastic. A second approach is to always use SE3SE4 as the study object, identify periods when the number of congested transmission lines into the zone increases from one to two, and see if the estimated effect of congestion changes. For instance, a comparison of the zone SE3SE4 under congestion from SE2, with SE3SE4 under *simultaneous* congestion from SE2 *and* DK, would capture the effect of an unambiguous increase in congestion into SE3SE4. This analysis is done by running two separate switching regression estimations, and changing the definition of the congested state. The difference in the coefficient estimates for delta can then be interpreted as the interaction effect of delta and an increase in the degree of congestion. This approach should provide better identification since we can compare the effect of two different degrees of congestion for the same zone. Because we always study the same zone, this approach is not as sensitive to unobservables that may vary between the country level and zonal level.

5.3 Identifying assumptions

In order to interpret our estimates as causal effects, we need to make some identifying assumptions. We will discuss the validity of these in Section 7.

Our first identifying assumption is that wind power is always supplied non-strategically at the marginal cost of production. Thus, after wind plants are built, they generate wind power according to how the wind blows. This is a key assumption in our theoretical framework as it makes individual and aggregate supply of wind exogenous.

We calculate delta by geographically pairing local wind speeds with locations of wind plants and the owning firms, and then calculate the amount of diversified firm-owned wind power production as a share of total wind power production. If there is an error, and this systematically correlates with variables in our model, we would get biased estimates.

Our second identifying assumption is that our delta variable is exogenous to our model. There are three main ways in which this can be violated. Firstly, we could have reverse causality in which delta depends on the dependent variable. However, wind speeds are clearly exogenous, determined by the meteorological environment in the area at hand, meaning that delta should not depend on our dependent variables.[12] Secondly, we may have omitted variable bias in which unobservables in the error term of the model correlate with delta. We include relevant control variables and fixed effects to capture such effects. While many factors influence outcomes in an electricity market, we are unable to control for them all. Thirdly, we may have measurement error in the way we construct the delta estimates. We estimate firm level wind production by combining wind speed measurements with location and ownership data of wind power plants, and then calculate delta as wind power production owned by diverse firm as a share of total wind power. If there is an error, and this error systematically correlates with variables in our model, we would get biased estimates. This identifying assumption is arguably our strongest one, and will be one of our major discussion points in Section 7.

Our third identifying assumption is that our instrument in our IV specifications is valid. We have argued why this should be the case in Section 5.2.2.

5.4 Data

We would ideally have used data on firms' supply bids to investigate strategic behavior. While that data is available on aggregate level for the whole Nord Pool market, it is unfortunately not available at a zonal or firm level. Instead, we use data from Svenska Kraftnät on hourly actual production per

[12] While investment decisions and the locations of wind plants may be affected by the variables in our model, after those decisions have been made, the amount of wind plants and their location in a zone is exogenous to market outcomes in the short-run, since neither the number of plants nor their locations can be changed with short notice.

technology on zonal level, aggregating it to country level in the relevant cases (Svenska Kraftnät, 2019). As discussed in Section 5.1, the underlying assumption for this approach is that strategic behavior in day-ahead bidding translates to observable changes in market clearing quantities and prices.

Additionally, we would ideally have used firm level data on wind production to create our main variable of interest, delta. This data is not available to us, and so we use granular data on wind turbines and hourly wind speed data to estimate wind production on a firm level. We use data on all wind power turbines in Sweden that have applied for renewable energy certificates (The Swedish Energy Agency, 2019).[13] Since such certificates grant extra income at little or no cost, we assume most market actors apply for them. The data should thus contain most wind power plants built in Sweden, and information on their commission date, owner, capacity, and location. From the Swedish Meteorological and Hydrological Institute, we obtain historical hourly wind speed estimates from around 150 weather stations from across the country (The Swedish Meteorological and Hydrological Institute, 2019).

We combine the two above data sources to estimate the hourly wind production in each plant, by using the weather station closest to each plant. Both the wind power plant data and the weather data contain information on which "tätort" each observation belongs to. A tätort is defined by Statistics Sweden as a town with at least 200 inhabitants, of which there are almost 2 000 in Sweden (Statistics Sweden, 2018). We use this information to match wind plants and weather stations together in the following way. We obtain shapefiles containing GIS layers with information on the location of each tätort from Statistics Sweden (2019) and use this to find the centroid of each tätort. We make the simplifying assumption that each wind power plant is located at the centroid in its corresponding tätort, and then find the closest weather station using GIS software.

When estimating wind production, we follow The British Wind Energy Association (2005) and assume that no electricity is produced when the wind speed is less than 3 m/s, or above 25 m/s. In addition, plants produce at full capacity when the wind speed is above 15 m/s, and the amount produced increases cubically between 3 m/s and 15 m/s. Since we have data on the ownership and commission date of each plant, we may aggregate all wind production to the relevant level and for the relevant group of firms for each hour. That is, we may create both zonal and countrywide estimates of wind production by all diversified actors. We then compute our delta variable by dividing diversified production by total wind production (on either zonal or country level). Because the detailed plant level data on wind power plants is only available for Sweden, we have restricted our study to the Swedish electricity market, which is part of the integrated Nord Pool market. This could be a potential source of bias, something we will discuss further in Section 7.

We obtain bidding zone level data on hourly prices, forecasted demand, planned outages on transmission lines, and weekly hydro inflow from Nord Pool (2019).[14] Following Haldrup and Nielsen (2006), we construct congestion dummies on specific transmission lines equal to one if the zone in question has a higher price than the zone at the other end of the connection, and zero otherwise. As explained in Section 2.1.2, this means that the higher-priced zone is import constrained from the lower-priced

[13]This is a subsidy that has been implemented by the Swedish government as an instrument to encourage investment in renewable energy sources.

[14]Most of this data is available through the Nord Pool website. To obtain data for all years studied in this paper, data needs to be accessed through the Nord Pool ftp server.

zone. For example, c_NO is equal to 1 if the price in SE3 is higher than that in NO1. Similarly, we create planned outage dummies for each relevant transmission line equal to 1 if there is an implemented planned outage on that line during an hour of observation.

Variable	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
price (SEK)	307.49	124.30	0.00	239.13	355.05	2,573.41
hydro power (MWh)	7,657.09	2,597.97	1,345.72	5,715.69	9,676.03	13,694.21
thermal power (MWh)	864.76	530.73	108.80	356.80	1,298.32	3,642.32
wind power (MWh)	1,556.23	1,067.03	14.99	721.68	2,176.81	5,874.00
delta	0.10	0.10	0.00	0.05	0.12	0.71
c_DK	0.06	0.24	0	0	0	1
c_NO	0.30	0.46	0	0	1	1
pl_DK	0.04	0.19	0	0	0	1
pl_NO	0.02	0.14	0	0	0	1
nuclear power (MWh)	7,038.34	1,396.40	2,963.08	5,986.09	8,214.44	9,140.81
hydro inflow, weekly (GWh)	1,288.38	1,072.80	210.00	586.00	1,526.00	7,125.50
forecasted demand (MWh)	15,763.00	3,430.99	8,568.00	13,072.00	18,116.00	27,558.00
hydro power SE3SE4 (MWh)	1,340.45	497.19	262.21	939.63	1,723.68	2,569.01
thermal power SE3SE4 (MWh)	750.76	477.06	83.43	303.15	1,126.69	3,419.97
delta SE3SE4	0.10	0.10	0.00	0.05	0.11	0.56
c_DK SE3SE4	0.005	0.07	0	0	0	1
c_NO SE3SE4	0.02	0.14	0	0	0	1
nuclear power SE3SE4 (MWh)	7,038.34	1,396.40	2,963.08	5,986.09	8,214.44	9,140.81
forecasted demand SE3SE4 (MWh)	12,716.97	2,891.22	6,488.00	10,455.00	14,699.00	22,919.00

Table 2: Summary statistics.

We combine all the above sources to construct a dataset with hourly observations spanning from 2012 to 2018 inclusive, giving us 61 368 observations. However, some variables include a small amount of missing values, reducing the effective number of observations depending on which specification is used. All data cleaning, tidying and analysis is done in R (R Core Team, 2018).

Summary statistics for the used variables are presented in Table 2. Note that we use the same price variable, the SE3 price, both in the zonal and countrywide case. We observe fairly large variation in delta and short-term strategic generation (hydro power and thermal power) both on country- and zonal level, indicating that there is potential for our proposed effect on both geographical levels. Note also that average short-run strategic output constitutes about half of the average forecasted demand at the country level and about a sixth at the zonal level. This difference is substantial and could imply varying opportunities for short-term withholding at the two levels, but we should still be able to observe the proposed effect in SE3SE4. In fact, a majority of all thermal power in Sweden is located in SE3SE4.

6 Results

In this section we present our results. We first study our baseline specification, which directly tests the counteraction of the merit order. In the following subsection, we add external congestion for market outcomes on the country level as another dimension in our analysis. Finally, in the last subsection we employ a switching regressions approach to study our hypotheses in periods of internal congestion in Sweden.

6.1 The merit order effect and its counteraction

Table 3 presents estimates of our baseline model, which abstracts from congestion. In the first four columns, we regress log price on log wind power and delta, adding fixed effects in columns 2 and 4 and restricting the sample to periods of no internal congestion in Sweden in columns 3 and 4. That is, we restrict the sample to periods when the price is the same in all four zones. Columns 5 and 6 use our preferred specification from column 4 but for log hydro power production and log thermal power production as dependent variables, respectively.

	log price				log hydro	log thermal
	(1)	(2)	(3)	(4)	(5)	(6)
log wind power	-0.061^{***}	-0.062^{***}	-0.065^{***}	-0.067^{***}	-0.128^{***}	0.075^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
delta	0.402^{***}	0.143^{***}	0.383^{***}	0.137^{***}	-0.196^{***}	-0.029
	(0.019)	(0.022)	(0.021)	(0.024)	(0.013)	(0.019)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	-	Hour	-	Hour	Hour	Hour
Observations	61,184	61,184	54,105	54,105	54,105	54,105

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 3: Baseline delta regressions.

We begin by observing that there indeed seems to exist a merit order effect in the Swedish electricity market - price is decreasing in the amount of wind power. In our preferred specification, a 1% increase in the amount of wind power generated leads to a reduction in price by 0.067% on average. This translates roughly to a 15% increase in wind power leading to a 1% decrease in prices, an economically significant effect.

We next turn to our main variable of interest, the share of wind power production owned by diverse producers - delta. We observe a positive and strongly significant effect in the first four specifications, even though the magnitude is decreased by roughly a third when adding fixed effects. Note that delta represents a share out of total wind power. Hence, as opposed to our logged variables which can be interpreted as elasticities, we interpret our delta coefficients as the percentage change in the dependent variable as a result of a 100 percentage point increase in delta. In our preferred specification in column 4, we see that for a given level of wind power, if delta goes from 0 to 1, we expect a 13.7% higher price. That is, if we go from a situation where no wind power is owned by diverse producers to one where all wind power is owned by such actors, we would expect to see 13.7% higher prices for a given level of wind power. This is an economically significant effect, supporting our second theoretical prediction.

In column 5, hydro power is the dependent variable. We observe a negative coefficient on wind power. The interpretation is that in general, less hydro power is produced in times of high wind power production. This is to be expected due to the fact that hydro power generally is bid higher than wind power. A notable result is that this decrease in production depends on who owns the wind power production. Specifically, for a given level of wind power, hydro power production would be 19.6% lower if delta increased from 0 to 1, showing support for our first theoretical prediction. In column 6, we add thermal power as our dependent variable. Here, the effect of wind power is in fact positive. Furthermore, we find no significant effect for delta, indicating that thermal power is not used for strategic behavior as a response to delta.

Overall, we find strong support for our second prediction, that delta has a positive effect on the price in the market and therefore counteracts the merit order effect. We find indicative evidence that this happens through the withholding of hydro power.

Our decision to restrict our sample to periods where there is no internal congestion in Sweden warrants some further discussion. As argued in Section 5.2, this approach is appropriate, since only then does our dependent variable, the bidding zone price for SE3, correspond to the correct geographical level of our explanatory variables in every period. Note also that the number of observations is still very large in the restricted sample, meaning that the resulting decrease in observations should not in itself affect the results. Furthermore, the results for the two samples (in e.g. columns 2 and 4) are very similar. We will henceforth use the restricted sample as our main one. In Section 6.4, we run the same regressions with the full sample as a robustness check.

6.2 Incorporating external congestion

In this section we study the potential interaction effect between delta and external congestion. More specifically, we look at two major transmission connections: one between Denmark and Sweden, and one between Norway and Sweden. We begin by presenting the results for the specifications with price as the dependent variable, and then present the results from using hydro and thermal power as the dependent variables, respectively.

	log price					
	DK			NO		
	OLS	IV	OLS	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
log wind power	-0.069^{***}	0.055^{***}	-0.068^{***}	-0.049^{***}	-0.038^{***}	-0.067^{***}
	(0.002)	(0.021)	(0.002)	(0.002)	(0.005)	(0.002)
delta	0.134^{***}	1.059^{***}	0.133^{***}	-0.066^{**}	-0.731^{***}	0.121^{***}
	(0.024)	(0.230)	(0.024)	(0.031)	(0.163)	(0.024)
c_DK	0.070^{***}	-0.978				
	(0.014)	(1.212)				
pl_c_DK			0.099			
			(0.183)			
c_NO				0.169^{***}	0.159^{***}	
				(0.006)	(0.060)	
pl_c_NO						0.118^{***}
						(0.022)
delta × c_DK	-0.513^{***}	-18.405				
	(0.154)	(15.277)				
delta × pl_c_DK			-0.739			
			(2.797)			
delta × c_NO				0.382^{***}	1.638^{***}	
				(0.038)	(0.303)	
delta × pl_c_NO						0.567^{***}
						(0.132)
First-stage F-stat C	-	29.601	-	-	214.212	
First-stage F-stat $delta \times C$	-	23.524	-	-	384.090	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Hour	Hour	Hour	Hour	Hour	Hour
Observations	54,105	52,838	52,838	54,105	53,953	53,953

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 4: Delta and congestion, log price as dependent variable.

Table 4 shows the marginal effects that import constraints from Denmark (columns 1 to 3) and Norway (columns 4 to 6) have on the effect of delta on Swedish electricity prices. These estimates are shown in the last four rows of the main section of the table. For each transmission link, we present OLS and IV estimates for the congestion variables, denoted by c, and OLS estimates for the alternative planned outage-congestion variables, denoted by pl_c.

The results indicate that the impact of congestion is very different for the different transmission lines. According to our theoretical predictions, the interaction effect of delta and congestion should have a positive effect on price, because diversified firms withhold relatively more generation in congested periods. The reason for this is that the congestion leads to a smaller relevant market, which in turn reduces the elasticity of demand so that withholding strategies become more profitable.

For the Norwegian transmission line, the estimated marginal increase of congestion on the effect of delta on prices ranges between increases in price of 38.2% (OLS, column 4) to 163.8% (IV, column

5), depending on specification and estimation method. The strong and significant effects support our hypothesis of congestion increasing the effect positive effect of delta on prices. On the other hand, the estimated effect of congestion on the Danish line is either a decrease in price of 51.3% (OLS, column 1) or insignificant, thus contradicting our hypothesis of the effect of congestion.

The dramatically different results suggest that there may be qualitative differences to how the different transmission lines matter for the Swedish electricity market. We discuss this in further detail in Section 7. There are also dramatic differences between the OLS results and IV results of our main specification (columns 1 and 2, and columns 4 and 5, respectively). Part of this may be due to endogeneity in the OLS estimates, leading to downward bias. However, as discussed in Section 5.2.2, the IV method estimates the local average treatment effect (LATE), whereas OLS estimates average treatment effects. Since the congestion induced by planned outages is arguably more predictable, we should expect the local average treatment effect estimated to be higher than the OLS estimates, but it is hard to say by how much.

For comparison, we therefore include OLS results from the same specification but with our planned outage-congestion variables (columns 3 and 6). These are defined as being equal to 1 when there is congestion *and* planned outages on the relevant transmission line. For the Norwegian transmission line, the estimated marginal effect of this congestion variable on the effect of delta on price is an increase with 56.7%, compared to an increase with 38.2% in the original specification. This indicates that there is some heterogeneity involved, and that congestion induced by planned outages has a higher effect on prices than the average type of congestion. In turn, this provides a possible explanation for why the estimated IV coefficients are so large.

Weak instrument tests confirm that the instruments used in the IV regressions are highly relevant - F-statistics for both first-stage regressions are well above the rule of thumb of 10. Note, however, that F-statistics for the Norwegian line specification are much higher than for the Danish line specification. This could explain the high standard errors in the Danish line IV specification. The first-stage estimates for all our IV specifications are available in Table 9 in Appendix B.

While not the main focus of this section, we also provide some comments on the main effects of variables. In all specifications except that in column 2, we get negative coefficient estimates for wind power, reaffirming the existence of the merit order effect in the Swedish market. Estimates of the main effect of delta vary both with regard to magnitude and direction. While the coefficient on the main effect of the congestion dummy could be indicative of general strategic behavior in periods of congestion, it could also be a result of the nature of congestion, by which prices increase mechanically. Thus, we do not analyze the results for this variable.

	log hydro power					
	DK			NO		
	OLS	IV	OLS	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
log wind power	-0.123^{***}	-0.193^{***}	-0.126^{***}	-0.133^{***}	-0.117^{***}	-0.128^{***}
	(0.001)	(0.011)	(0.001)	(0.001)	(0.003)	(0.001)
delta	-0.172^{***}	-0.572^{***}	-0.172^{***}	-0.176^{***}	-0.101	-0.199^{***}
	(0.013)	(0.126)	(0.013)	(0.017)	(0.096)	(0.013)
c_DK	-0.149^{***}	1.157^{*}				
	(0.008)	(0.662)				
pl_c_DK			0.032			
			(0.101)			
c_NO				-0.054^{***}	0.151^{***}	
				(0.003)	(0.035)	
pl_c_NO						0.015
						(0.012)
delta × c_DK	0.547^{***}	1.733				
	(0.085)	(8.349)				
delta × pl_c_DK			0.187			
			(1.537)			
delta × c_NO				-0.037^{*}	-0.180	
				(0.021)	(0.179)	
delta × pl_c_NO						0.042
						(0.073)
First-stage F-stat C	-	29.601	-	-	214.212	
First-stage F-stat $delta \times C$	-	23.524	-	-	384.090	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Hour	Hour	Hour	Hour	Hour	Hour
Observations	54,105	52,838	52,838	54,105	53,953	53,953

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 5: Delta and congestion, hydro power production as dependent variable.

Next, we investigate if the effects on prices are driven by changes in short-term strategic output in response to delta and congestion. We begin by focusing on hydro power. Table 5 presents the results from the same specifications as in Table 4, but with the log of hydro power generation as the dependent variable.

The results indicate that congestion does not impact the marginal effect of delta on withholding of hydro power, contrary to our hypothesis. Only the main specification OLS estimates of the interaction effect of delta and congestion are significant, but in the Denmark line specification it has an unexpected sign (column 1) and in the Norway line specification (column 4) it is only significant at the 10% level. Estimated effects of wind and delta are significant with the expected sign in all specifications except in column 5, where the estimated effect of delta is not significant.

	log thermal power					
	DK			NO		
	OLS	IV	OLS	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
log wind power	0.072^{***}	0.136^{***}	0.075^{***}	0.059^{***}	0.038^{***}	0.075^{***}
	(0.002)	(0.013)	(0.002)	(0.002)	(0.004)	(0.002)
delta	-0.054^{***}	0.421^{***}	-0.034^{*}	0.090^{***}	0.792^{***}	-0.008
	(0.020)	(0.141)	(0.020)	(0.025)	(0.138)	(0.020)
c_DK	0.032^{***}	-0.483				
	(0.011)	(0.741)				
pl_c_DK			-0.001			
			(0.148)			
c_NO				-0.161^{***}	-0.260^{***}	
				(0.005)	(0.051)	
pl_c_NO						-0.041^{**}
						(0.018)
delta × c_DK	0.470^{***}	-9.032				
	(0.125)	(9.344)				
delta × pl_c_DK			0.270			
			(2.259)			
delta ×c_NO				-0.223^{***}	-1.545^{***}	
				(0.031)	(0.257)	
delta × pl_c_NO						-0.719^{***}
						(0.107)
First-stage F-stat C	-	29.601	-	-	214.212	
First-stage F-stat $delta \times C$	-	23.524	-	-	384.090	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Hour	Hour	Hour	Hour	Hour	Hour
Observations	54,105	52,838	52,838	54,105	53,953	53,953

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 6: Delta and congestion, thermal power production as dependent variable.

Table 6 presents the results from our main specifications again, but with the log of short-term strategic thermal power as the dependent variable. The results again show that the impact of congestion on the effect of delta differs substantially for the Danish and Norwegian transmission lines.

Estimates of the interaction effect of delta and congestion for the Norwegian line are all negative and strongly significant, and range between -22.3% and -154.5%. The marginal increase of congestion on the effect of a unit increase in delta is thus a reduction in the amount of thermal output by 22.3%, according to specification 4, and by 154.5%, according to specification 5. This supports our hypothesis of congestion leading to larger effects of delta on withholding behavior of thermal generation. Danish line congestion estimates either have an unexpected sign (column 1) or are insignificant. Again, OLS and IV estimates for the main specification differ widely, with the planned outage-congestion estimate in between. Just as for price, this indicates heterogeneity in the effect of congestion on thermal power output. Estimates of the effect of wind on thermal output are positive, and delta main effect estimates vary both with regard to magnitude and direction.

Overall, our prediction of congestion pronouncing effects of delta is quite consistent in the specifications concerning the Norwegian transmission line but not consistent at all for the specifications with the Danish lines. While we find support for increased withholding of thermal power with congestion from Norway, we cannot say the same for hydro power. The differing results indicate that the effects of congestion may be transmission-line specific, and we are therefore unable to draw any general conclusions about our predicted interaction effect based on these results.

6.3 Internal congestion - an endogenous switching regressions approach

We now continue by studying internal congestion in Sweden - congestion between Swedish price zones. As explained in Section 5.2.3, we do this using an endogenous switching regressions approach. Like in the IV setting, we use price, hydro power, and thermal power as dependent variables. Because the number of observations decreases dramatically with the different congestion configurations (see the bottom of Table 7), we choose to not include any fixed effects in our switching regressions specifications, as it would leave us with very few observations per possible value of the fixed effects variable.

	log price		
	(1)	(2)	(3)
c = 0			
log wind power	-0.068^{***}	-0.067^{***}	-0.069^{***}
	(0.002)	(0.002)	(0.002)
delta	0.453^{***}	0.448^{***}	0.430^{***}
	(0.021)	(0.021)	(0.021)
c = 1			
log wind power SE3SE4	-0.106^{***}	-0.186^{***}	-0.086^{***}
	(0.013)	(0.034)	(0.012)
delta SE3SE4	-0.150	-0.274	-0.167^{**}
	(0.097)	(0.272)	(0.082)
Congested from	SE2	SE2,DK	SE2,NO
Num. obs.	59,616	57,869	59,616
Num. obs. c = 0	57,868	57,589	58,420
Num. obs. c = 1	1,748	280	1,196

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 7: Switching regressions, price as dependent variable.

Table 7 reports the results from our first switching regressions specification, with the log of price as the dependent variable. The correct geographical level of our variables changes based on whether the studied zone is congested or not, as determined by the selection equation. We report estimates for the selection equations for all specifications in Appendix B. When $c = 0$, there is no internal congestion and the relevant market is the whole of Sweden. For instance, in column 1, the effect of a unit increase in delta is a 45 % increase in price. When $c = 1$, there is internal congestion, and the relevant market is SE3SE4. In column 1, we do not find any significant effect from the zonal delta for SE3SE4.

Overall, the results indicate that internal congestion does not seem to increase the marginal effect of delta on prices. For all three congestion configurations, the estimated effect of delta is lower for congested periods compared to the uncongested country level estimates. In fact, all three estimated coefficients are either insignificant or in an unexpected direction, suggesting that the predicted effect from delta does not exist on a zonal level. However, this could be due to the number of observations being low in the congested state.

As discussed in Section 5.2.3, however, the most controlled way to test how congestion influences the marginal effect of delta is to compare the effects of delta under unambiguously different degrees of congestion in SE3SE4. In our case, the relevant comparisons are between congestion from SE2 and congestion from SE2+DK, and between congestion from SE2 and congestion from SE2+NO1.[15] None of these estimates are statistically different from another, contradicting our hypothesis of a higher degree of congestion increasing the marginal effect of delta.[16]

	log hydro			log thermal		
	(1)	(2)	(3)	(4)	(5)	(6)
c = 0						
log wind power	-0.130^{***}	-0.128^{***}	-0.129^{***}	0.052^{***}	0.053^{***}	0.052^{***}
	(0.001)	(0.001)	(0.001)	(0.005)	(0.002)	(0.002)
delta	0.126^{***}	0.158^{***}	0.133^{***}	-0.903^{***}	-1.044^{***}	-0.901^{***}
	(0.012)	(0.012)	(0.012)	(0.041)	(0.019)	(0.017)
c = 1						
log wind power SE3SE4	-0.052^{***}	0.064^{**}	-0.033^{**}	-0.049^{***}	-0.105^{***}	-0.039^{***}
	(0.009)	(0.027)	(0.014)	(0.008)	(0.026)	(0.012)
delta SE3SE4	-0.214^{***}	0.045	-0.171^{*}	-0.384^{***}	-0.357^{*}	-0.516^{***}
	(0.071)	(0.220)	(0.090)	(0.056)	(0.211)	(0.085)
Congested from	SE2	SE2,DK	SE2,NO	SE2	SE2,DK	SE2,NO
Num. obs.	59,616	57,869	59,616	59,616	57,869	59,616
Num. obs. c = 0	57,868	57,589	58,420	57,868	57,589	58,420
Num. obs. c = 1	1,748	280	1,196	1,748	280	1,196

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 8: Switching regressions, hydro power or thermal power as dependent variables.

Table 8 reports results from the switching regressions specifications using hydro power and thermal power as dependent variables. For hydro power generation, the results suggest that congestion may have a negative marginal effect on the effect of delta on withholding, as delta estimates under congestion in the SE2 and SE2+NO cases are stronger than the uncongested estimates. This supports our prediction that congestion pronounces the effect of delta on withholding. Again, the most relevant comparison to make is that between estimates of differently congested states. We find that no such estimate is significantly different from another, indicating that the degree of congestion into SE3SE4 does not matter for how hydro power is strategically withheld with regard to delta, and thus contradicting our hypothesis.

For thermal power, none of the delta estimates under congestion are of higher magnitude than their uncongested counterparts. Furthermore, none of the delta estimates under congestion are statistically

[15] We cannot certainly say how the degree of congestion differs between SE2+DK and SE2+NO1. It could, however, shed light on the relative importance of each transmission line for the Swedish market.

[16] To test this, we use the test for comparing coefficients from different regression models provided by Clogg, Petkova, and Haritou (1995).

different from another, suggesting that the degree of congestion in SE3SE4 does not affect the impact of delta on thermal power withholding.

To conclude, we find little support for our theoretical prediction that the effect of delta is pronounced in periods of congestion when the analysis is done on a zonal level. We discuss possible reasons for this in Section 7.

6.4 Robustness checks

In order to understand how sensitive our results are to the specifications used, we perform several robustness checks. These are available in Appendix B. We begin by including some results that are excluded from our baseline model estimates in Table 3. Specifically, we estimate columns 5 and 6 where hydro power and thermal power are dependent variables in a similar progression to that in column 1 to 4. That is, we test how sensitive the results are to the inclusion of fixed effects and to the usage of the unconstrained sample. The results are available in Table 12 and Table 13. Results are mainly consistent with those estimated earlier. However, two notable exceptions exist. Firstly, delta in fact has a positive effect on hydro output without fixed effects, indicating that hydro output has significant time-pattern invariant effects. This is supported also by the positive delta estimates for the uncongested case in the switching regressions specification in Table 8, where we excluded fixed effects. Secondly, the coefficient on delta is significant and negative in the specifications where the full sample is used for thermal power as the dependent variable. This is consistent with the fact observed in our study of congestion, which suggests that thermal is withheld in periods of congestion.

Our second robustness check is done by including monthly fixed effects in addition to the hourly fixed effects in our main specifications. We estimate the specifications of Table 3 and 4 again with this modification, and as can be seen in Table 14 and 15, we find largely similar results. This indicates that our more parsimonious model with hourly fixed effects is sufficient.

7 Discussion

We begin by summarizing the results from the previous section. Firstly, we find strong evidence of a merit order effect in the Swedish market - as the supply of wind power increases, prices fall. Secondly, we find strong evidence of delta counteracting the merit order effect - for a given level of wind power, an increase in the diversified firms' share of wind power leads to higher prices. We find a strong negative effect of delta on hydro power generation but no significant effect of delta on thermal power generation. This indicates that the channel through which the counteracting effect happens is through diversified firms' strategic withholding of hydro power. We thus confirm our second theoretical prediction about delta counteracting the merit order effect on price, and partly confirm our first hypothesis about the withholding behavior that drives the effect: a higher delta makes diversified firms withhold hydro power but has no effect on thermal power.

Our last prediction states that congestion should pronounce the found counteraction effect. We test this in two ways: on country level for external congestion and on zonal level with internal congestion between zones within Sweden. These results are inconclusive. For the transmission line from Norway, we find that the counteraction effect is pronounced by congestion and that this seems to be driven by the withholding of thermal production. For the transmission lines from Denmark, we cannot conclude that the addition of congestion alters the effect of delta on prices or withholding behavior in any way. Similarly, we cannot conclude that increased congestion into the combined SE3SE4 zone in Sweden increases the effect of delta on market outcomes.

A plausible reason for why the effect of delta seems to be lower in congested zones than in the whole of Sweden, contrary to our predictions, is that the internal congestion reduces the amount of short-term strategic capacity in the relevant market, so that it becomes more difficult to use withholding strategies in that market. In Sweden, much of the short-term strategic capacity is made up of hydro power, which is located in the northern Swedish zones SE1 and SE2. When the whole of Sweden is still the relevant market, such capacity can be used to influence the price for the whole country. As expected, the ability to do so seems to increase as Sweden becomes congested from neighboring countries. In the case of the smaller relevant market consisting of the congested combined zone SE3SE4, however, this hydro power can no longer influence prices on the margin. The only capacity left to use strategically is then the relatively small capacity of short-term strategic thermal power in the two zones and thus the ability to drive prices up is limited.

The above discussion highlights a potential issue with out theoretical framework. It is likely that the relationship between different generation technologies changes when going going from the country level to a zonal level. This could affect our results, but it is difficult to assess to what extent or in what direction without explicitly accounting for how the technology mix changes with different configurations of congestion. This highlights the need for theoretical developments in this area that considers congestion more explicitly. Furthermore, in this thesis we focus on exogenous congestion since we are interested in the marginal effect of delta on withholding strategies in a smaller relevant market, and not in potential strategies to actually induce that smaller market per se. However, as previous studies have shown, strategic actors may have incentives to induce congestion. This adds another dimension in which we may observe strategic behavior, and one which may interact with the effect of delta on withholding behavior.

Trhee peculiarities of our results are worth discussing a bit more. Firstly, the interaction effect of delta and congestion is consistently stronger for the Norwegian transmission link than for the Danish ones. This could be due to peculiarities of our data. While our instrumental variable is highly relevant for congestion on the Danish lines, it is much more relevant for the Norwegian line, which in turn could lead to the different results in the IV estimates - this is supported by the generally lower standard errors on the Norwegian congestion variables. Furthermore, in the internal congestion setting the number of observations where SE3SE4 is congested from both SE2 and Denmark is much lower than for congestion from both SE2 and Norway - 280 observations compared to 1196. This is likely to drive part of the difference. Another possible explanation for the difference is that the transmission lines affect the Swedish market in such different ways that the implications of congestion on them are simply inherently different. Norway and Denmark have different generation portfolios and therefore some heterogeneity is expected, but it is hard to say to what extent. Extending our approach to the whole Nord Pool market, or simulating the entire market and investigating the interplay of heterogeneous generation portfolios and transmission links under different market structures, would help with understanding this.

Secondly, some of the delta main effect estimates are not very robust to the inclusion of the interaction of delta and congestion. For instance, the main effect of delta on price becomes negative when including congestion from Norway in the IV regression. A plausible explanation for this is that the congestion effect in fact drives a large part of the general effect of delta that we find when excluding congestion. Thus, when explicitly accounting for congestion in our specifications, these variables capture the majority of the effect and the general delta effect becomes negative or insignificant. As discussed in Section 3, the exercise of market power has been found to vary strongly with congestion. Thus, while negative effects of delta on price are in themselves unexpected, the fact that the main effects change quite a lot with the inclusion of the interaction effects is to be expected and emphasizes the importance of accounting for congestion to understand electricity market dynamics.

Thirdly, we find indicative evidence of heterogeneous effects of delta on hydro power and thermal power, respectively, and we find that the delta-congestion interactions affect hydro and thermal output in different ways. In the general Sweden-wide specification, we find a strong negative effect of delta on hydro while the effect of the delta-congestion interaction is inconclusive. For thermal power, we find no Sweden-wide effect but fairly consistent negative effects of the delta-congestion interaction. This indicates that hydro power is used more for withholding in uncongested periods, and thermal power in congested periods, following increases in delta. This is an important result that warrants closer studies.

7.1 Internal and external validity

In this section, we discuss the internal and external validity of our results. We begin with the former, where we assess if the methods we have used to produce evidence allow us to make causal inference in an appropriate way. We will do this by both discussing the validity of our identifying assumptions and the appropriateness of the methods chosen.

We begin by discussing our first identifying assumption, which is that wind power is always supplied non-strategically. In principle, there could be situations in which not all potential capacity for wind power is used. If an actor is sufficiently large, there could be potential for driving up market prices by withholding wind power. However, if there are other technologies in that firm's portfolio, it would be rational to withhold such output before any wind power, since they have higher marginal costs. Thus, we should only expect this type of withholding from firms that solely own wind. To our knowledge, there is no firm in Sweden with such a portfolio that is large enough to be thought of as a strategic actor.

Next, we turn to our second identifying assumption which is that our delta variable is exogenous. The important potential sources of endogeneity for our delta estimates are omitted variable bias, reverse causality and measurement error, and we now discuss them in turn. While we include relevant control variables and fixed effects that capture unobserved time-pattern invariant effects, we are unable to control for all possible confounders. One potential major concern that we have abstracted from in our analysis is that the Swedish electricity market is a part of the integrated Nord Pool market. If what happens in these markets affects market outcomes in Sweden and is correlated with delta, we would have omitted variable bias. For instance, consider the scenario where wind speeds are weak in Sweden in general but strong in one area where diversified firms own a lot of the wind plants. Then the value of delta in that region would largely determine the value of delta of the whole country. If wind speeds in that area are correlated with wind speeds in some other unobserved market in Nord Pool, a potential merit order effect in this outside market could drive system prices down, making our delta capture that effect. If this scenario happens systematically in some way, our estimates for delta could then be biased. Note however, that if there is indeed a counteraction of the merit order effect by diversified firms, we would expect this to happen in other markets outside Sweden as well, limiting the amount of bias such a scenario could cause. Furthermore, in periods with congestion the market effectively becomes smaller. This enables more precise identification since markets outside the congested zone become irrelevant, limiting the potential bias from their omission in our model.

Another possible issue is that of reverse causality between delta and market outcomes. While this should not be the case in the short run when investment is taken as given, in the longer run it could be that past market outcomes are driving future investment behavior. This is a complex issue that we have abstracted from and that could potentially bias our results. Some previous work has characterized electricity markets as a two-stage game where investment behavior affects spot market behavior and vice versa (see e.g. Bushnell and Ishii, 2007). Future research should investigate incentives to investing in wind power, and how these differ for different types of firms.

The last part of the exogeneity assumption is that delta does not contain any systematic measurement error. In our estimation of the wind production on plant level we indirectly assume that the only thing that matters for production output is wind speed. However, in reality several other factors may affect it. One potential issue is that the weather station data is measured at a height of 10 meters above ground level, while many wind turbines are located up to a 100 meters above ground. However, adjusting for this would just consist of multiplying our estimates by a constant. This is true since we lack height data on the individual plants (meaning that we would need to use the height of a standard wind power plant as a proxy). This would not affect any results in our analysis. Another issue is that there will inevitably be some loss due to inefficiencies in the turbine or in the transmission lines. What matters is that any such omitted variable is not systematically correlated with our main signal, wind speed. This would be the case, for instance, if the loss factor is increasing in the wind speed. Analysing this further is beyond the scope of this paper, but we argue that for our purposes, the signal obtained from our delta variable is satisfactory. Future studies on this topic should explore using more sophisticated estimation methods. If actual production is available for some wind power plants, this can be used as training data for usage in advanced prediction methods.

Our third identifying assumption is that the instrument in our IV specifications is valid. We have argued that planned outages should affect our outcome variables through no other channel than the fact that it increases the probability of congestion. In general, planned outages really only have an effect on market outcomes if they induce congestion - if not all transmission capacity is used, the reduction in transmission capacity will have no effect. While this should hold most of the time, one could imagine instances where this assumption could be violated. It could be the case that actors adjust their bidding behavior in response to announcements about planned outages to actively *induce* congestion by withholding production. However, given that one of the main goals of the system operator is to ensure system stability and minimize costs due to congestion, they should take the risk of such behavior into account when deciding appropriate periods to perform maintenance work, e.g. by avoiding peak hours of demand. Since planned outages are evenly spread across hours and weekdays, this indicates that they do not take such behavior into account. Hence, this supports our exclusion restriction argument.

One potential issue with our methods is the fact that we use aggregate market outcomes. Many previous studies in the field use individual firm data, either bid curves or cost data, to make inference about strategic behavior. Instead, we use market prices and aggregate production of certain technologies as dependent variables. One possible problem is that such variables also contain the output of small fringe firms with no potential for market power. We make two points in response to this. Firstly, a large portion of total capacity is owned by our diverse firms, giving us a good enough signal. Secondly, assuming that no fringe actors would try to exercise market power, we would expect them to produce at a greater percentage of their installed capacity than strategic firms do. Thus, this issue should make our estimates conservative, reducing the risk that our results are false positives. Similarly, we have abstracted from the other markets that open after the day-ahead market (see Appendix A for details). While all markets could matter for firm behavior on the day-ahead market, the fact that the day-ahead market constitutes 98% of the total volume of electricity traded on Nord Pool means that our results should not suffer in any significant way from this simplification.

Next, we discuss external validity, which is the generalizability of our results. The results for our main variable of interest, delta, should be expected to exist in other markets with a similar structure to the Swedish market. That is, the effect should exist in markets characterized by large actors with the potential for market power, that also own diverse portfolios with both renewable energy and short-term strategic production. Similar effects have been found by Butner (2019) and Genc and Reynolds (2019) in two markets in North America.

However, when it comes to the interaction effect between delta and congestion, we argue that our results have less external validity. This is to be expected, since each electricity market is vastly different in terms of bidding zone configurations, production technology mix in each such zone, and the transmission capabilities between them. This is exemplified by the results in our paper, which are vastly different for congestion from Denmark and Norway. However, we do believe that our results indicate that congestion in general matters when studying the effect of ownership of renewables on market outcomes.

8 Conclusion

In this thesis, we have analyzed whether the ownership of renewable power generation matters for market outcomes. Specifically, we test the hypothesis that if firms with diversified generation portfolios own wind power generation, they withhold generation from higher-cost technologies to counteract the negative effect of renewable generation on prices. Furthermore, we test if such behavior is pronounced in periods of congestion in the transmission grid. We test this empirically on a novel dataset combining detailed wind power plant data, local wind speeds and market outcome data.

Our results indicate that an increase in the share of wind power that is supplied by diverse firms leads to an increase in strategic withholding of short-term strategic generation and an increase in prices, supporting our main hypotheses. Regarding congestion, we cannot conclude that it changes the marginal effect of delta in general. However, we do find such evidence for certain transmission lines and generation technologies, indicating that congestion in some circumstances may increase the potential for the strategic firm behavior that is proposed in this paper.

Our results have two major limitations. Firstly, due to data availability we are only able to study the Swedish electricity market, which in fact is a part of the integrated Nordic market. This may potentially bias our results. Secondly, we use aggregate market outcomes to study the effect of *firm level* strategic behavior. This should result in more conservative estimates and may also cause us to miss important nuances in firm behavior.

Our general results on the effect of delta have important policy implications. While renewable sources of electricity indeed seem to reduce electricity prices, induced withholding behavior of diversified generation firms may lead both to market inefficiencies and wealth transfers from consumers to firms. This has direct implications both for welfare and distributional effects of the electricity market. Additionally, withholding behavior may compromise the stability of the electricity grid, the dependability of which is crucial for all sectors of the economy. Given the goal of increasing the share of renewable power in both Sweden and around the world, these issues are likely to grow in importance. Our results have implications for the optimal design of renewable electricity support schemes. Should, for instance, all firms receive the same degree of support, regardless of their existing generation portfolio? With regard to congestion, our inconclusive but suggestive findings highlight the importance of market operators and authorities to properly understand how the effects of transmission constraints on market outcomes may change with new generation technologies.

Our thesis highlights the need for further research in several directions. Firstly, we suggest our approach is used with actual bidding data. While we confirm our main hypotheses for aggregate level market outcomes, understanding exactly which types of diversified firms that strategize in their bidding, and how they do so, is important for appropriate policy-making. We therefore suggest our approach is adapted to markets and periods for which firm level bidding data is available. Similarly, exact measures of wind power generation per firm would yield more accurate delta figures, which would improve the reliability of the results. Secondly, we suggest our approach is used for studying an entire integrated electricity market. While our geographical limitation to Sweden is necessary due data limitations, studying the whole relevant market is important to fully understand the mechanisms driving market outcomes. This is especially true regarding the dynamics of congestion, as emphasized by our inconclusive results about its effects. Studies of renewable power ownership that estimate the effects of congestion in an entire integrated market, or explicitly model its impact under different market structures, are interesting directions for future research.

References

Acemoglu, D., A. Kakhbod, and A. Ozdaglar (2017). Competition in electricity markets with renewable energy sources. *Energy Journal 38*, 137–156.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bask, M., J. Lundgren, and N. Rudholm (2011). Market power in the expanding Nordic power market. *Applied Economics 43*(9), 1035–1043.

Ben-Moshe, O. and O. D. Rubin (2015). Does wind energy mitigate market power in deregulated electricity markets? *Energy 85*, 511–521.

Biggar, D. R. and M. R. Hesamzadeh (2014). *The Economics of Electricity Markets*. John Wiley & Sons.

Borenstein, S., J. Bushnell, and C. R. Knittel (1999). Market power in electricity markets: Beyond concentration measures. *The Energy Journal 20*(4), 65–88.

Borenstein, S., J. Bushnell, and S. Stoft (2000). The competitive effects of transmission capacity in a deregulated electricity industry. *RAND Journal of Economics 31*(2), 294–325.

Bushnell, J. B. and J. Ishii (2007). An equilibrium model of investment in restructured electricity markets. CSEM Working Paper No. 164, University of California.

Butner, M. (2019). Gone with the wind: Consumer surplus from renewable generation. Unpublished.

Clogg, C. C., E. Petkova, and A. Haritou (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology 100*(5), 1261–1293.

Cramton, P. (2017). Electricity market design. *Oxford Review of Economic Policy 33*(4), 589–612.

de Miera, G. S., P. del Río González, and I. Vizcaíno (2008). Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. *Energy Policy 36*(9), 3345–3359.

Dutoit, L. C. (2007). Heckman's selection model, endogenous and exogenous switching models: A survey. Unpublished.

Energiföretagen (2017). Energiåret 2016. `https://www.energiforetagen.se/globalassets/energiforetagen/statistik/energiaret/energiaret2016_elproduktion_19-mars-2018.pdf?v=luM_jHmMkGP-Ac09E6nOogQLDRs` (accessed April 22, 2019).

Energiföretagen (2018). Vindkraft. `https://www.energiforetagen.se/sa-fungerar-det/elsystemet/produktion/vindkraft` (accessed May 6, 2019).

Genc, T. S. and S. S. Reynolds (2019). Who should own a renewable technology? Ownership theory and an application. *International Journal of Industrial Organization 63*, 213–238.

Haldrup, N. and M. Ø. Nielsen (2006). A regime switching long memory model for electricity prices. *Journal of Econometrics 135*(1-2), 349–376.

Hjalmarsson, E. (2000). Nord pool: A power market without market power. Working Papers in Economics nr. 28, Goteborg University.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Johnsen, T. A., S. K. Verma, and C. D. Wolfram (1999). Zonal pricing and demand-side bidding in the norwegian electricity market. Unpublished.

Läpple, D., T. Hennessy, and C. Newman (2013). Quantifying the economic return to participatory extension programmes in Ireland: an endogenous switching regression analysis. *Journal of Agricultural Economics 64*(2), 467–482.

Lokshin, M. and Z. Sajaia (2004). Maximum likelihood estimation of endogenous switching regression models. *The Stata Journal 4*(3), 282–289.

Lundin, E. (2017). Market power and joint ownership: Evidence from nuclear plants in Sweden. IFN Working Paper No. 1113.

Lundin, E. and T. Tangerås (2017). Cournot competition in wholesale electricity markets: The Nordic power exchange, Nord Pool. IFN Working Paper No. 1191.

McConnell, D., P. Hearps, D. Eales, M. Sandiford, R. Dunn, M. Wright, and L. Bateman (2013). Retrospective modeling of the merit-order effect on wholesale electricity prices from distributed photovoltaic generation in the australian national electricity market. *Energy Policy 58*, 17–27.

McRae, S. D. and F. A. Wolak (2009). How do firms exercise unilateral market power? Evidence from a bid-based wholesale electricity market. Unpublished.

Mirza, F. M. and O. Bergland (2015). Market power in the Norwegian electricity market: Are the transmission bottlenecks truly exogenous? *The Energy Journal*, 313–330.

Nord Pool (2018). Annual report 2017. `https://www.nordpoolgroup.com/globalassets/download-center/annual-report/annual-report-nord-pool_2017.pdf` (accessed April 15, 2019).

Nord Pool (2019). Historical market data. `https://www.nordpoolgroup.com/historical-market-data/` (accessed March 14, 2019).

Nord Pool (n.d.a). History. `https://www.nordpoolgroup.com/About-us/History` (accessed April 14, 2019).

Nord Pool (n.d.b). The Nordic electricity exchange and the Nordic model for a liberalized electricity market. `https://www.nordpoolgroup.com/globalassets/download-center/rules-and-regulations/the-nordic-electricity-exchange-and-the-nordic-model-for-a-liberalized-electricity-market.pdf` (accessed April 29, 2019).

R Core Team (2018). R: A language and environment for statistical computing.

Ryan, N. (2017). The competitive effects of transmission infrastructure in the Indian electricity market. NBER Working Paper No. 23106.

Sapio, A. (2015). The effects of renewables in space and time: A regime switching model of the Italian power price. *Energy Policy 85*, 487–499.

Sensfuß, F., M. Ragwitz, and M. Genoese (2008). The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in germany. *Energy Policy 36*(8), 3086–3094.

Statistics Sweden (2018). Tätorter. https://www.scb.se/hitta-statistik/statistik-efter-amne/miljo/markanvandning/tatorter/ (accessed May 7, 2019).

Statistics Sweden (2019). Open geodata on tätorter. https://www.scb.se/hitta-statistik/regional-statistik-och-kartor/geodata/oppna-geodata/tatorter/ (accessed April 14, 2019).

Steen, F. (2005). Do bottlenecks generate market power? An empirical study of the Norwegian electricity market. Unpublished.

Svenska Kraftnät (2019). Förbrukning och tillförsel per timme. https://www.svk.se/aktorsportalen/elmarknad/statistik (accessed April 2, 2019).

Tangerås, T. and J. Mauritzen (2014). Real-time versus day-ahead market power in a hydro-based electricity market. IFN Working Paper No. 1009.

The British Wind Energy Association (2005). Wind turbine technology. https://www.nottingham.ac.uk/renewableenergyproject/documents/windturbinetechnology.pdf (accessed April 4, 2019).

The Swedish Competition Authority (1996). Deregulation of the Swedish electricity market. Report No. 1996:3. http://www.konkurrensverket.se/globalassets/english/publications-and-decisions/deregulation-of-the-swedish-electricity-market.pdf (accessed April 12, 2019).

The Swedish Energy Agency (2019). Marknadsstatistik: Godkända ansökningar. http://www.energimyndigheten.se/fornybart/elcertifikatsystemet/marknadsstatistik/?currentTab=0#mainheading (accessed April 1, 2019).

The Swedish Meteorological and Hydrological Institute (2019). Historical meteorological observations. https://www.smhi.se/klimatdata/meteorologi/ladda-ner-meteorologiska-observationer/ (accessed April 14, 2019).

Toomet, O. and A. Henningsen (2008). Sample selection models in R: Package sampleSelection. *Journal of Statistical Software 27*(7).

Vassilopoulos, P. (2003). Models for the identification of market power in wholesale electricity markets. *UFR Sciences of Organizations, DEA 129.*

Weron, R., M. Bierbrauer, and S. Trück (2004). Modeling electricity prices: Jump diffusion and regime switching. *Physica A: Statistical Mechanics and its Applications 336*(1-2), 39–48.

Wind Europe (2018). Wind energy in Europe in 2018. `https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Statistics-2018.pdf` (accessed May 6, 2019).

Wolak, F. A. (2003). Measuring unilateral market power in wholesale electricity markets: the California market, 1998-2000. *American Economic Review 93*(2), 425–430.

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

Wolfram, C. D. (1999). Measuring duopoly power in the British electricity spot market. *American Economic Review 89*(4), 805–826.

Würzburg, K., X. Labandeira, and P. Linares (2013). Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria. *Energy Economics 40*, 159–171.

A Electricity market details

A.1 The intraday market

The intraday market opens the same day as, and closes one hour prior to the delivery of the electricity is due. Trading on the intraday market is continuous, much like on a stock exchange. It is used primarily by commercial actors to balance new, better predictions of their possibility to deliver as assigned by the day-ahead market. For instance, a plant may break down and require the firm to procure electricity on the intraday market by a producer with spare capacity. Similarly, the wind may blow stronger than expected and may leave a producer with excess capacity that it can sell on the intraday market. For retailers, the opposite reasoning holds. If a retailing firm realizes that temperatures will be higher than expected and thus less electricity will be needed than it asked for on the day-ahead market, it can sell off electricity on the intraday market. And if the retailing firms realize it needs more generation, it can buy from a producer with spare capacity (Nord Pool, n.d.).

A.2 The balancing market

After the intraday market closes, the balancing market opens. All transactions in the balancing market consists of the TSO on one side and a commercial actor on the other side. Each commercial actor that is involved in a transaction on the balancing market (with the TSO) either cannot deliver electricity as assigned by the day-ahead market, or has short-term flexibility in their insertion or extraction of electricity in the grid (either they have spare capacity or are able to quickly reduce their electricity consumption). If an actor cannot deliver as assigned in the day-ahead market, they will have to either buy their missing electricity from the TSO or sell their excess generation to the TSO. The TSO then sells/buys this electricity to another actor, the one with short-term flexibility. This means that even though one or many actors may deviate from their assigned levels of extraction/insertion, the TSO will be able to keep the grid stable by balancing the over-/underproduction by a corresponding adjustment made by another firm. Such firms in advance submit bids about their short-term flexibility to the TSO and are called upon when balancing is needed. The deviating firm is paid/have to pay a less favorable price than the balancing firm is paid/have to pay, thus firms are incentivized to balance their insertion/extraction into the grid, but also to provide short-term flexibility if they are able to do so.

B Additional estimates

B.1 First-stage estimates

	c_DK	delta × c_DK	c_NO	delta × c_NO
	(1)	(2)	(3)	(4)
log wind power	0.052^{***}	0.004^{***}	-0.086^{***}	-0.009^{***}
	(0.001)	(0.0001)	(0.002)	(0.0004)
delta	0.287^{***}	0.035^{***}	-0.018	0.520^{***}
	(0.015)	(0.001)	(0.026)	(0.004)
pl_DK	-0.057^{***}	-0.003^{***}		
	(0.009)	(0.001)		
pl_NO			0.217^{***}	-0.024^{***}
			(0.020)	(0.003)
delta × pl_DK	0.168^{**}	-0.0004		
	(0.069)	(0.006)		
delta × pl_NO			0.294^{**}	0.446^{***}
			(0.133)	(0.020)
Controls	Yes	Yes	Yes	Yes
FE	Hour	Hour	Hour	Hour
Observations	52,838	52,838	53,953	53,953

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 9: First-stage estimates from instrumental variable approach.

B.2 Selection equation estimates

	log price		
	(1)	(2)	(3)
pl_SE2	0.304***	−0.133**	0.496***
	(0.022)	(0.062)	(0.026)
pl_DK		0.615***	
		(0.070)	
pl_NO			0.344***
			(0.061)
sigma1	0.402***	0.404***	0.403***
	(0.001)	(0.001)	(0.001)
sigma2	0.743***	0.359***	0.287***
	(0.034)	(0.045)	(0.011)
rho1	−0.161***	−0.136***	−0.196***
	(0.022)	(0.039)	(0.020)
rho2	−0.917***	−0.504**	−0.260*
	(0.013)	(0.201)	(0.143)
Congested from	SE2	SE2,DK	SE2,NO
Num. obs.	59,568	57,869	59,568
Num. obs. c = 0	57,822	57,589	58,374
Num. obs. c = 1	1,746	280	1,194

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 10: Selection equation estimates for endogenous switching regression approach with price as dependent variable.

	log hydro			log thermal		
	(1)	(2)	(3)	(4)	(5)	(6)
pl_SE2	0.309***	−0.292***	0.435***	0.308***	−0.429***	0.460***
	(0.022)	(0.058)	(0.025)	(0.022)	(0.064)	(0.020)
pl_DK		0.702***			0.590***	
		(0.069)			(0.076)	
pl_NO			0.316***			0.123***
			(0.063)			(0.037)
sigma1	0.225***	0.227***	0.225***	0.794	0.359***	0.375***
	(0.001)	(0.001)	(0.001)	(0.534)	(0.001)	(0.001)
sigma2	0.284***	0.248***	0.306***	0.719***	0.307***	0.662***
	(0.011)	(0.014)	(0.007)	(0.013)	(0.041)	(0.027)
rho1	−0.021	0.674***	−0.064	0.990	−0.295***	0.925***
	(0.074)	(0.028)	(0.054)	(10.161)	(0.041)	(0.005)
rho2	0.239	−0.130	0.071	0.990***	−0.669***	0.952***
	(0.150)	(0.336)	(0.189)	(0.001)	(0.127)	(0.007)
Congested from	SE2	SE2,DK	SE2,NO	SE2	SE2,DK	SE2,NO
Num. obs.	59,616	57,869	59,616	59,616	57,869	59,616
Num. obs. c = 0	57,868	57,589	58,420	57,868	57,589	58,420
Num. obs. c = 1	1,748	280	1,196	1,748	280	1,196

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 11: Selection equation estimates for endogenous switching regression approach with quantity as dependent variable.

B.3 Robustness checks

	log hydro power			
	(1)	(2)	(3)	(4)
log wind power	-0.122^{***}	-0.122^{***}	-0.128^{***}	-0.128^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
delta	0.058^{***}	-0.165^{***}	0.041^{***}	-0.196^{***}
	(0.011)	(0.012)	(0.012)	(0.013)
Controls	Yes	Yes	Yes	Yes
FE	-	Hour	-	Hour
Observations	61,184	61,184	54,105	54,105

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 12: Baseline model, different modifications with hydro power as dependent variable.

	log thermal power			
	(1)	(2)	(3)	(4)
log wind power	0.065^{***}	0.077^{***}	0.061^{***}	0.075^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
delta	-1.064^{***}	-0.111^{***}	-1.069^{***}	-0.029
	(0.018)	(0.018)	(0.020)	(0.019)
Controls	Yes	Yes	Yes	Yes
FE	-	Hour	-	Hour
Observations	61,184	61,184	54,105	54,105

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 13: Baseline model, different modifications with thermal power as dependent variable.

	log price				log hydro	log thermal
	(1)	(2)	(3)	(4)	(5)	(6)
log wind power	-0.061^{***}	-0.053^{***}	-0.065^{***}	-0.056^{***}	-0.121^{***}	0.014^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
delta	0.402^{***}	0.090^{***}	0.383^{***}	0.081^{***}	-0.242^{***}	-0.048^{***}
	(0.019)	(0.020)	(0.021)	(0.022)	(0.013)	(0.013)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Hour, Month	Hour, Month	Hour, Month	Hour, Month	Hour, Month	Hour, Month
Observations	61,184	61,184	54,105	54,105	54,105	54,105

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 14: Baseline model with month and hour fixed effects.

	log price					
	OLS	IV	OLS	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
log wind power	-0.058^{***}	0.030^{*}	-0.059^{***}	-0.044^{***}	-0.025^{***}	-0.057^{***}
	(0.002)	(0.017)	(0.002)	(0.002)	(0.008)	(0.002)
delta	0.070^{***}	0.978^{***}	0.070^{***}	-0.135^{***}	-1.006^{***}	0.062^{***}
	(0.022)	(0.188)	(0.023)	(0.029)	(0.156)	(0.023)
c_DK	0.060^{***}	-0.571				
	(0.013)	(1.173)				
pl_c_DK			0.161			
			(0.168)			
c_NO				0.136^{***}	0.243^{**}	
				(0.005)	(0.111)	
pl_c_NO						0.080^{***}
						(0.020)
delta × c_DK	-0.225	-19.450				
	(0.142)	(13.472)				
delta×pl_c_DK			-0.456			
			(2.564)			
delta×c_NO				0.355^{***}	1.917^{***}	
				(0.035)	(0.290)	
delta×pl_c_NO						0.533^{***}
						(0.121)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	Hour, Month	Hour, Month	Hour, Month	Hour, Month	Hour, Month	Hour, Month
Observations	54,105	52,838	52,838	54,105	53,953	53,953

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses

Table 15: External congestion model with month and hour fixed effects.