Abstract

Wholesale electricity markets use different market designs to handle congestion in the transmission network. We compare nodal, zonal and discriminatory pricing in general networks with transmission constraints and loop flows. We conclude that in large games with many producers who are allowed to participate in the real-time market the three market designs result in the same efficient dispatch. However, zonal pricing with counter-trading results in additional payments to producers in export-constrained nodes.

Keywords

Congestion management, wholesale electricity market, transmission network, nodal pricing, zonal pricing with countertrading, discriminatory pricing, large game

JEL Classification

C72, D44, D61, L13, L94
Abstract
Wholesale electricity markets use different market designs to handle congestion in the transmission network. We compare nodal, zonal and discriminatory pricing in general networks with transmission constraints and loop flows. We conclude that in large games with many producers who are allowed to participate in the real-time market the three market designs result in the same efficient dispatch. However, zonal pricing with counter-trading results in additional payments to producers in export-constrained nodes.

Keywords: Congestion management, wholesale electricity market, transmission network, nodal pricing, zonal pricing with counter-trading, discriminatory pricing, large game

JEL codes: C72, D44, D61, L13, L94

---

1 We are grateful for very helpful comments from Richard Friberg, Sven-Olof Fridolfsson, Jenny Fridström, Thomas Tangerås, an anonymous referee, seminar participants at Stockholm School of Economics and Research Institute of Industrial Economics (IFN), and conference participants at IAEE 2011 in Stockholm. We also want to thank Erik Lundin for research assistance and Christina Lönnblad for proof-reading our paper. The work has been financially supported by Jan Wallander’s and Tom Hedelius’ Research Foundation and the Research Program The Economics of Electricity Markets.
2 Research Institute of Industrial Economics (IFN), P.O. Box 55665, SE-102 15 Stockholm, Sweden, phone +46 8 665 45 59, fax + 46 8 665 4599. E-mail: Par.Holmberg@ifn.se. Associate Researcher of Electricity Policy Research Group (EPRG), University of Cambridge.
3 Research Institute of Industrial Economics (IFN). P.O. Box 55665, SE-102 15 Stockholm, Sweden, fax + 46 8 665 4599, Stockholm School of Economics, Department of Economics, Sveavägen 65, Stockholm, Sweden. E-mail ewa.lazarczyk@hhs.se
1. Introduction

Storage possibilities are negligible in most electric power networks, so demand and supply must be instantly balanced. The consequence is that transmission constraints and how they are managed often have a large influence on market prices. The European Union's regulation 1228/2003 (amended in 2006) sets out guidelines for how congestion should be managed in Europe. System operators should coordinate their decisions and choose designs that are secure, efficient, transparent and market based. Regulatory authorities shall regularly evaluate the used congestion management methods, with respect to compliance with the principles and rules established in the regulation and guidelines. In this paper, we compare the efficiency of three market designs that are in operation in real-time electricity markets: nodal, zonal and discriminatory pricing.

Real-time markets are open for offers from producers just before electricity is going to be produced and delivered. During the delivery period, the system operator accepts offers in order to clear the real-market, taking transmission constraints into account. The auction design influences which offers are accepted and their payments. Nodal pricing or locational marginal pricing (LMP) is used in Argentina, Chile, Ireland, New Zealand, Russia, Singapore, in several US states (e.g. California, New England, New York, PJM and Texas), and Poland is on the way to implement it as well. This design acknowledges that location is an important aspect of electricity which should be reflected in its price, so all accepted offers are paid a local uniform-price associated with each node of the electricity network.

There is no uniform market price under discriminatory pricing, where accepted offers are paid as bid. Still the system operator considers all transmission constraints when accepting offers, so similar to nodal pricing, discriminatory pricing allows production in import-constrained nodes to be accepted at a higher price than production in export constrained nodes. Discriminatory pricing is used in Britain and Iran, and Italy has decided to implement it as well. One (somewhat naïve) motivation for this auction format has been that low cost production is supposed to bid low and accordingly to be paid a low price, which would increase consumer's and the auctioneer's welfare at producers' expense.
Real-time markets with zonal pricing consider inter-zonal congestion, but have a uniform market price inside each region, typically a country or a state, regardless of transmission congestion inside the region. Originally this design was thought to minimize the complexity of the pricing settlement and politically it is sometimes more acceptable with one price in a country/state. This is why zonal pricing was adopted by Australia and by most European countries. Originally, zonal pricing was also used in most deregulated electricity markets in the US, but they have now switched to nodal pricing, at least for generation. One reason for this change in US is that zonal pricing is, contrary to its purpose, actually quite complex and pricing is not very transparent behind the hood. The problem is that the system operator needs to order redispatches after the zones of the real-time market have been cleared if transmission lines inside a zone would otherwise be overloaded. Such a redispatch increases supply in import constrained nodes and reduces it in export constrained nodes in order to relax inter-zonal transmission congestion. There are alternative ways of compensating producers for their costs associated with these adjustments. The compensation schemes have no direct influence on the cleared zonal prices, but indirectly the details of the design may influence how producers make offers into the real-time market. We consider the market oriented redispatch, which is called countertrading, and where countertraded offers are paid as bid. This design is used in the real-time market of the Nordic countries (Nord Pool) and was used in the old Texas design. The three market designs that we consider are summarized in the table below.

**Table 1: A summary of the three designs of real-time markets.**

<table>
<thead>
<tr>
<th>Congestion technique</th>
<th>Considered transmission constraints</th>
<th>Auction format</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Uniform-price</td>
</tr>
<tr>
<td>Nodal</td>
<td>All</td>
<td>x</td>
</tr>
<tr>
<td>Discriminatory</td>
<td>All</td>
<td></td>
</tr>
<tr>
<td>Zonal –stage 1</td>
<td>Inter-zonal</td>
<td>x</td>
</tr>
<tr>
<td>Zonal –stage 2</td>
<td>Intra-zonal</td>
<td></td>
</tr>
</tbody>
</table>
We compare the three market designs by means of a game-theoretical model. Our analysis considers a general electricity network where nodes are connected by capacity constrained transmission lines. Producers' costs are assumed to be common knowledge, and demand is certain and inelastic. There is a continuum of infinitesimally small producers in the market that choose their offers in order to maximize their individual payoffs. Subject to the transmission constraints, the system operator accepts offers to minimize total stated production costs, i.e. it clears the market under the assumption that offers reflect true costs. We consider one-shot games and characterize the Nash equilibrium of each market design. We compare prices, payoffs and efficiencies for the three designs. As far as we know there is only one previous game-theoretical comparison of congestion management techniques. But this study by Willems and Dijk (2011) is limited to two-node networks with constant marginal costs, and they do not consider the discriminatory pricing design. 4

In the nodal pricing design, producers maximize their payoffs by simply bidding their marginal costs. Thus, in this case, the accepted offers do maximize short-run social welfare. We refer to these accepted equilibrium offers as the efficient dispatch and we call the clearing prices the network's competitive nodal prices. We compare this outcome with equilibria in the alternative market designs.

For fixed offers, the system operator would increase its profit at producers' expense by switching from nodal to discriminatory pricing. But we show that discriminatory pricing encourages strategic bidding even if there are infinitely many producers in the market and this will exactly offset the system operator's increased payoff. Hence, in the Nash equilibrium of the pay-as-bid design, all accepted offers are at the network's competitive nodal prices. Moreover, accepted production is the same as in the efficient dispatch. Thus, market efficiency and payoffs to producers and the system operator are the same as for nodal pricing. As payoffs are identical in all circumstances, this also implies that the long-run effects are the same in terms of investment incentives.

4 Dijk and Willems (2011) also extend their model to consider cases with a finite number of producers and market power.
Equilibrium offers are also similar in the zonal real-time market with counter-trading, and the dispatch is the same as for the two other market designs. But the producers' payoffs are larger under zonal pricing at the system operator's expense. The reason is that the two-stage clearing of the real-time market gives producers the opportunity to either sell at the uniform zonal price or at the discriminatory equilibrium price in stage 2 (which equals the competitive nodal price of the producer's node), whichever is higher. In addition, even when they are not producing any energy, production units in export-constrained nodes can make money by selling at the uniform zonal price and then to buy back the same amount at the discriminatory price, which is lower, in the countertrading stage. This type of strategic behavior has also been observed in practice, and the unnecessarily high payoffs for strategic producers at the system operators' expense is another reason why US markets have chosen to switch from zonal to nodal pricing. Our results are the same regardless of whether producers are allowed to make new offers in the counter-trading stage or whether the rules are such that the same offers are used in the two stages.

As in Dijk and Willems' (2011) two-node model, the additional payments to producers in the zonal market cause long-run inefficiencies; producers overinvest in export-constrained nodes. We attribute this deficiency to a suboptimal combination of the two different auction formats, uniform- and pay-as-bid pricing. If one wants to get the investment incentives right, it is better to stick to a pure format, such as nodal pricing or a pure pay-as-bid auction.

In section 2 we discuss congestion management and the literature on this topic in more detail. Section 3 presents our model and section 4 an analysis of different congestion management designs. In section 5, market equilibria are discussed for a simple two node example. The paper is concluded in section 6 and a brief discussion of the results is presented in section 7.

---

5 In a model with infinitesimally small firms, this simply follows from excessive payments to generators in export-constrained nodes. Dijk and Willems (2011) use a two-stage game with entry and Cournot competition to prove this for their setting with market power.
2. Literature review

2.1 Nodal pricing

The early literature on issues of electricity pricing in networks concentrates on the calculation of nodal prices. Schwepe et al. (1988) and Hsu (1997) present a model where the spot price is derived from a social welfare maximization problem subject to a number of constraints and where the difference in spot prices between any two locations corresponds to a price of transmission costs between those two nodes.

Electric power networks normally have alternating currents (AC), which results in a non-linear model of the network. Hence, this model is often simplified by a linear approximation called the DC load flow approximation. It is, for example, used by Hogan (1992) and Chao and Peck (1996). They present a version of nodal pricing that incorporates the technological externalities associated with transmission congestion and transmission losses, adopt tradable transmission capacity rights and introduce a trading rule that specifies the transmission loss compensation required for power transfers. They demonstrate that a competitive equilibrium with property rights and their trading rule can replicate a social optimum.

2.1.1 Equilibrium Analysis of Electricity Networks

The early literature on locational marginal pricing (LMP) or nodal market design focuses on the development of computational methods to determine the competitive market prices in electricity networks. Models analyzed in this strand of literature are based on detailed systems of equations stipulating the mechanics of electrical power flow and the physical constraints associated with transmission networks, and not much space is left to analyze whether producers and consumers have incentives to bid strategically. However, there are exceptions; Escobar and Jofre (2008) prove the existence of Nash equilibria (NE) in electricity auctions for a network with one generator per node and a central agent. Wilson (2008) characterizes the supply function equilibrium (SFE) in a multi-unit auction constrained by limited

2.2 Zonal pricing

Although providing correct short-term incentives, the nodal prices design is often criticized to be too complex and in need of a large degree of coordination. This “fault” is often dealt with by bundling particular nodes together into areas with one price which is referred to as zonal pricing or market splitting. In many European countries, the whole country constitutes a zone, so there is one price in the whole country (this is called uniform pricing). In Australia each state constitutes a zone. Denmark and Norway are also divided into several zones. The first design of zonal pricing in Norway was based on flexible boundaries but this created some uneasiness among market participants, since the procedure was not sufficiently transparent. Thus, the regime was changed in 2000 into a system where zone-boundaries are predetermined once or twice a year. The Swedish government has introduced four zones in Sweden from November 2011.

2.2.1 Zonal pricing models without internal congestion

Advisors of zonal pricing argue that this approach does not only limit the complexity of nodal pricing by breaking up the system into just a few zones, but some also claim that the system in question divides naturally into a few uniformly priced areas. In some cases, zonal division

---

7 Advantages with the zonal design are for example discussed by De Vries et al. (2009), Leuthold et al. (2008) and Stoft (1997).
can be quite straightforward, corresponds well with nodal pricing and can thus be considered as an effective simplification of the LMP approach. However, the grid is normally more complex, and as shown by Stoft (1997) in his examples, the zonal approach distorts the nodal prices and assigns uniform prices to nodes that in reality bear different costs and should thus be priced differently.

While zonal pricing is supposed to be a simplification of nodal pricing, it trades simplicity for efficiency. Björndal and Jörnsten (B&J) (2001) show that the zonal approach is just a second-best solution. They discuss some problems with the system and point out difficulties in defining the zones and redistribution effects. They show that a zone allocation system based on the absolute values of optimal nodal price differences does not necessarily lead to a zonal system with maximal social surplus and they present an example of a small network where they identify a multitude of possible zone constructions. Their concerns about zone composition are confirmed by Ehrenmann and Smeers (2005). Using a six-node model in which two lines have capacity restrictions, they show that in a meshed network, the number of possible zone decompositions can be large and thus, the selection of the nodes into “right” areas is not trivial. Moreover, freezing of zones is not always a good solution, as the characteristics of a “good” partitioning may change with time.

Zonal pricing also appears to be inferior to nodal pricing when it comes to market power. Hogan (1999), Harvey and Hogan (2000) and Green (2007) present a set of examples where they show that LMP is better suited to prevent market power when compared to the zonal approach (although it does not eliminate it and additional measures are necessary in order to mitigate market power).

2.2.2. Zonal pricing with counter trading

Zones are supposed to be chosen such that intra-zonal congestion is limited. However, the first clearing stage, where the zonal price is calculated, only considers inter-area/cross-border congestion. Intra-zonal congestion is dealt with by means of a redispatch in the
second stage. The simplest redispatch “is exercised as a command and control scheme” (Krause, 2005); the system operator orders adjustments without referring to the market and all agents are compensated for the estimated cost associated with their adjustments. When the redispatch is market based as in the Nordic countries and the old zonal market in Texas, it is called countertrading. In this case, all changes after the first clearing are compensated as in a pay-as-bid auction, so that all agents are paid the stated costs associated with their adjustments.

Although internal congestion is quite important, it is normally omitted from the analysis of zonal markets. Dijk and Willems (2011) is the only exception that we know of. They argue that counter-trading leads to strategic bidding by generators who have an incentive to bid a very low price in the energy market in order to create congestion and be paid for not producing. They show that under perfect competition, entry is efficient with nodal pricing, but inefficient with counter-trading and they derive an NE of counter-trading bids for the market with an incumbent and a finite number of entrants. They conclude that compared to nodal pricing, counter-trading can be seen as a subsidy to entrants in the export constrained area whenever there is congestion.

Countertrading is costly for the system operator. To avoid this cost, a network operator sometimes has incentives to resolve intra-zonal congestion by imposing fictive inter-zonal constraints. Björndal et al. (2003) illustrate this by numerical examples. The authors also conclude that zonal pricing is not just a “mere simplification of nodal pricing” and the particular aggregation of nodes into a zone can change the allocation of the social surplus. Similar distortions of the congestion signal resulting in “perverse” incentives to the system operators (SO) were identified by Glachant and Pignon (2005) who analyze congestion techniques and their impact on grid users on an example of a stylized Norwegian and Swedish interconnected grid.
2.3 Quantitative comparison of market designs

Another strand of literature compares different pricing strategies for real markets; often the system in place with an optimal electricity dispatch based on LMP. Bernard and Guertin (2002) simulate a three-node model of Hydro-Quebec’s electric network. The simulated nodal prices differ by 18% between Montréal's (load) and the hydro generator's node, which is significantly larger than the 5.2% flat rate for transmission losses that was imposed at the time of the study. The large difference indicates that investors are provided with erroneous price signals when choosing the site for new generations.

Leuthodl et al. (2008) analyze the impact of increased wind power production on the German power grid. They show that changing from a uniform national price to nodal pricing in Germany has a negligible influence on welfare, but the social welfare rises by 0.8% if seven of its neighbours also introduce nodal pricing, which would make cross-border flows more efficient.

Green (2007) analyzed three different transmission pricing schemes: uniform – with one national price for both generation and demand; nodal price for generators and uniform price for demand; and nodal – single price for each node for both generation and demand. All three pricing schemes were set up for a 13-node network in England and Wales. His results show that LMP raises welfare by 1.5% in comparison with a uniform approach (where demand cannot actively bid into the market). He calculates the welfare for different elasticities and shows that it increases with a larger elasticity value. Moreover, a system with nodal prices also provides correct investment signals. In a later article (Green, 2010), he once again stresses the importance and “usefulness” of a market design that accommodates spatial prices. He argues that the current uniform pricing of electricity is not well-suited for accommodating intermittent wind generation that is often located far from demand.
2.4 Pay-as-bid auctions

The debate between proponents of uniform-price and proponents of pay-as-bid auctions has a long history, mainly in the literature on treasury auctions. For unchanged offers, discriminatory pricing would lead to lower electricity prices. But in theory and practice, agents compensate for this by adjusting their offers upwards and, in the end, the design's influence on the spot price is small. Electricity prices did go down when Britain changed from uniform to pay-as-bid pricing. But as shown by Evans and Green (2004), this can mainly be explained by a simultaneous change in the ownership structure. Pay-as-bid auctions have been studied both theoretically and empirically; see, for example, the survey by Holmberg and Newbery (2010). But these studies have not taken the network into account.

2.4.1 Model

The model described in this section is used to evaluate and compare three market oriented congestion management techniques: pay-as-bid, nodal pricing and zonal pricing with counter-trading. For the three designs we compare the Nash equilibrium of a one-shot game with a continuum of infinitesimally small and perfectly informed producers.8

We study an electricity network (possibly meshed) with \( n \) nodes that are connected by capacity constrained transmission lines. In each node there is a continuum of infinitesimal producers. Each producer in the continuum of node \( i \) can be indexed by the variable \( q_i \). \( C'_i(q_i) \) is the marginal cost of the producer \( q_i \) and the producer submits an offer price \( o_i(q_i) \) for its production to the real-time market. Without loss of generality, we assume that producers are sorted with respect to their marginal cost and that the index \( q_i \) is scaled such that if producer \( q_i \) is the marginal producer in node \( i \), then the local output is \( q_i \) and the total production cost in node \( i \) is given by \( C_i(q_i) \) (provided offers follow the merit-order). We assume that the marginal cost is continuous and strictly increasing in \( q_i \). Maximum total production in node \( i \) is denoted by \( \bar{q}_i > 0 \). Demand in each node is given by \( D_i \), which is certain and inelastic. It is

---

8 The idea to calculate Nash equilibria for a continuum of agents was first introduced by Aumann (1964).
assumed that costs are common knowledge among producers; they have perfect information. For simplicity, we assume that each producer is only active in one node of the network.

The system operator’s clearing of the real-time market must be such that local net-supply equals local net-exports in each node and such that the physical constraints of the transmission network are not violated. Any set \( \{q_i\}_{i=1}^n \) of nodal production that satisfies these feasibility constraints is referred to as a feasible dispatch. The system operator chooses a dispatch among the feasible dispatches in order to minimize the stated production cost or equivalently to maximize

\[
W = -\sum_{i=1}^n \int_0^{q_i} o_i(y) dy,
\]

which maximizes social welfare if offers reflect the true costs. Thus, we say that the system operator acts in order to maximize the stated social welfare subject to the feasibility constraints.

In a market with nodal pricing, all accepted offers in the same node are paid the same nodal price. The nodal price is determined by the node’s marginal offer. In a pay-as-bid auction, all accepted offers are paid according to their offer price. In the zonal pricing design with counter-trading, the real-time auction is cleared in two stages. First the system operator clears the market without regard for the intra-zonal transmission constraints (constraints inside zones). The zonal clearing prices are chosen such that welfare is maximized (costs are minimized) and total net-supply in the zone equals total net-exports from the zone. Similar to a uniform-price auction, the zonal price \( \Pi_k \) is paid to all accepted production in zone \( k \). In case intra-zonal transmission-lines are overloaded after the first clearing, there is a second clearing stage where the system operator increases accepted production in import constrained nodes and reduces it in export constrained nodes. This is called counter-trading. All deviations from the first-clearing are settled on a pay-as-bid basis. We consider two version of the zonal design: one where the same offers are used in the two stages of the real-time market and one version where agents are allowed to make new offers in the countertrading stage.
We consider networks where a Nash equilibrium exists for a nodal pricing design with a continuum of infinitesimal producers, and where the market outcome is unique in the sense that any Nash equilibrium has the same nodal prices and dispatches. Our analysis applies to general networks with possible loop flows. In principle it could be a lossy AC network. But to ensure a unique outcome we restrict the analysis to cases where the feasible set of dispatches is convex, i.e. if two dispatches are possible, then any weighted combination of the two dispatches is also feasible. This is, for example, the case under the DC load flow approximation of general networks with alternating current (AC) (Chao and Peck, 1996). This is a simplified, linear form of modeling an AC system, which is normally used in economic studies of complicated networks (Green, 2007; Björndal and Jörnsten, 2001, 2005, 2007; Adler et al., 2008; Schweppe et al., 1988; Glanchant and Pignon, 2005).

2.4.2 Analysis

We start our game-theoretical analysis of the three market designs by means of a technical result that we will use in the proofs that follow.

**Lemma 1.** Assume that offers are shifted upwards (more expensive) in some nodes and shifted downwards (cheaper) in others, then the accepted supply is weakly lower in at least one node with more expensive offers or weakly higher in at least one node with cheaper supply.

**Proof:** We let the old dispatch refer to the feasible dispatch \( \{q_i^{\text{old}}\}_{i=1}^n \) that maximized stated social welfare at old offers when supply in node \( i \) is given by \( o_i(q_i) \). Let \( \Delta o_i(q_i) \) denote the shift of the supply curve, so that \( o_i(q_i) + \Delta o_i(q_i) \) is the new supply curve in node \( i \). The new dispatch refers to the feasible dispatch \( \{q_i^{\text{new}}\}_{i=1}^n \) that maximizes stated social welfare for new offers. Thus we have

---

9Generally, it is only marginal offers that are uniquely determined in an equilibrium of an auction with nodal or uniform pricing and no uncertainties. Technically, offers above and below the margin, which are not price-setting, may differ from the marginal cost. So technically these non-marginal offers are not uniquely determined in equilibrium, but on the other hand they do not influence the market outcome.
Now, make the contradictory assumption that in comparison to the old dispatch, the new dispatch has a strictly higher accepted supply in all nodes where offers have been shifted upwards (more expensive) and strictly lower accepted supply in all nodes where offers have been shifted downwards (cheaper). Thus \( q_i^{\text{new}} > q_i^{\text{old}} \) when \( \Delta o_i(q) > 0 \) and \( q_i^{\text{new}} < q_i^{\text{old}} \) when \( \Delta o_i(q) < 0 \), so that

\[
\sum_{i=1}^{\infty} \int_0^{q_i^{\text{new}}} (o_i(x) + \Delta o_i(x)) dx \geq \sum_{i=1}^{\infty} \int_0^{q_i^{\text{old}}} (o_i(x) + \Delta o_i(x)) dx. \quad (2)
\]

But summing (2) and (3) yields

\[
-\sum_{i=1}^{\infty} \int_0^{q_i^{\text{new}}} o_i(x) dx > -\sum_{i=1}^{\infty} \int_0^{q_i^{\text{old}}} o_i(x) dx, \quad (4)
\]

which is a contradiction since, by definition, the old dispatch \( \{q_i^{\text{old}}\}_{i=1}^{\infty} \) is supposed to maximize stated welfare at old offers.

One immediate implication of this lemma is that

**Corollary 2.** If one producer unilaterally increases/decreases its offer price, then accepted sales in its node cannot increase/decrease.

### 2.5 Nodal pricing

Next, we characterize the equilibrium in the nodal pricing design. The lemma below proves that the nodal pricing design has an NE where firms offer at their marginal cost. We get this competitive market outcome as an infinitesimally small producer has too little capacity to influence the market price, so they are price takers. The corresponding equilibrium has been used in previous studies by for example Hogan (1992), Joskow and Tirole (2000) and Green (2007).
Lemma 3. In a market with nodal pricing, the large game with a continuum of producers has at least one NE where producers offer at their marginal cost.

Proof: We note that the objective function (stated welfare) in (1) is continuous in the supply \( q_i \) for the considered offers. Moreover, the feasible set (the set of possible dispatches) is closed, bounded (because of capacity constraints) and non-empty. Thus, it follows that there always exists an optimal dispatch when offers reflect true costs (Gravelle and Rees, 1992).

Next, we note that no producer has a profitable deviation from the competitive outcome. Marginal costs are continuous and strictly increasing and from Corollary 2, we know that if a firm unilaterally deviates and increases its offer in node \( i \), the accepted supply in that node will not increase. Hence, no producer with an accepted offer can increase its offer price above the marginal offer of the node and still be accepted, as its offer price would then be above one of the previously rejected offers in the same node. No producer with a rejected offer would gain by undercutting the marginal offer, as the changed offer would then be accepted at a price below its marginal cost. Thus, there must exist an NE where all firms offer to produce at their marginal cost.

As the system operator clears the market in order to maximize social welfare when offers reveal true costs, we note that this equilibrium dispatch must be efficient. Offers above and below the margin are not uniquely determined in the Nash equilibrium of a market with nodal pricing, but it can be shown that the dispatches and nodal prices are the same in all Nash equilibria.

Proposition 4. The NE of a nodal pricing market has the following properties:

1) Marginal offers are at the marginal cost.

2) Dispatches \( \{q_i^N\}_{i=1}^w \) and nodal prices \( p_i^N = C_i'(q_i^N) \) are the same in every NE.

Proof: We first realize that offers cannot be accepted below their marginal cost in equilibrium. Moreover, marginal offers must be at the marginal cost in each node and all offers from production units with a marginal cost below the marginal offer in the same node must be accepted. Otherwise, there would exist some infinitesimally small producer in the node with a marginal cost below the marginal offer, but whose offer is not accepted. Thus, it
would be a profitable deviation for such a producer to slightly undercut the marginal offer and we know from Corollary 2 that such a unilateral deviation will not decrease the accepted supply in the node.

Now, consider the case when offers are strictly increasing in output. In this case, the objective function (stated welfare) is strictly concave in the supply, \( q_i \). Moreover, the set of feasible dispatches is by assumption convex in our model. Thus, it follows that the objective function has a unique local (and global) maximum (Gravelle and Rees, 1992) and the system operator’s dispatch can be uniquely determined. From the above, it follows that the unique dispatch of an NE must be such that marginal offers are at the marginal cost. Actually, the dispatch is not influenced by changes in offers below and above the margin. Thus, we realize that any NE with strictly increasing offers must result in the optimal dispatch \( \{ q_i \}_{i=1}^{N} \) with nodal prices \( \{ p_i \}_{i=1}^{N} \) as in Lemma 3.

Finally, we argue that perfectly elastic segments in the offer curves would not change the result above. One can always construct strictly increasing offers that are arbitrarily close to such offers, and the system operator’s objective function is continuous in offers. Thus, we can use the same argument as above with the difference that the system operator may sometimes have multiple optimal dispatches for a given set of offer curves, but it is only one of them, \( \{ q_i \}_{i=1}^{N} \), where the marginal offers in each node are at the marginal cost, a necessary condition for an NE.

Next, we will analyze the other two market designs. In these calculations, we will refer to the nodal pricing equilibrium. From the analysis above, it follows that we can calculate a unique market outcome for every network with nodal pricing that we consider and we refer to it as the network’s efficient dispatch \( \{ q_i \}_{i=1}^{N} \) and the network’s competitive nodal prices \( \{ p_i \}_{i=1}^{N} \).

Schweppe et al. (1988), Chao and Peck (1996) and Hsu (1997) and others outline methods that can be used to calculate these efficient dispatches for real networks.

\[10\] To verify this statement, one can for example write down the Lagrange condition of the system operator’s optimization problem.
2.6 Discriminatory pricing

The continuum of producers that we consider are price takers in nodal markets where all agents in the same node are paid the same market price. Thus under nodal pricing we find NE where each agent offers its supply at marginal cost. But even if the market outcome is the same, we see below that equilibria where producers bid their marginal cost do not exist in the discriminatory design. The reason is that each agent is then paid its individual offer price. Thus, even if agents are infinitesimally small, they can still influence their own offer price, so they are no longer price takers.

**Proposition 5.** The Nash equilibrium of a discriminatory market has the following properties:

1) The dispatched production is identical to the network's efficient dispatch, \( q_i^N \), in each node.

2) All production in node \( i \) with a marginal cost at or below \( C'_i(q_i^N) \) is offered at the network's competitive nodal price \( p_i^N = C'_i(q_i^N) \).

3) Other offers are not accepted and are not uniquely determined in equilibrium. However, it can, for example, be assumed that they offer at their marginal cost.

**Proof:** We first realize that offers cannot be accepted below their marginal cost in equilibrium. Moreover, marginal offers must be at the marginal cost in each node and all offers from production units with a marginal cost below the marginal offer in the same node must be accepted. Otherwise, there would exist some infinitesimally small producer in the node with a marginal cost below the marginal offer, but whose offer is not accepted. Thus, it would be a profitable deviation for such a producer to slightly undercut the marginal offer and we know from Corollary 2 that such a unilateral deviation will not decrease the accepted supply in the node. We also note that in a discriminatory auction, it is profitable for a producer to increase the price of any accepted offer until it reaches the marginal offer of the node.

We have assumed that marginal costs are strictly increasing in output and we have established that accepted offers must be at the marginal cost of the marginal unit in
equilibrium. Thus, in comparison to the nodal pricing dispatch, the accepted supply must be higher and supply is shifted upwards (more expensive) in all nodes $i$ with a marginal offer higher than $p_i^N$ and the accepted supply must be lower and supply is shifted downwards (cheaper) in all nodes $i$ with a marginal offer lower than $p_i^N$. However, this would violate Lemma 1. Thus, if an equilibrium exists, the optimal dispatch must be the same as for nodal pricing. From the proof of Proposition 4, we realize that increasing offers below the marginal offer does not change the optimal dispatch. Thus, equilibrium offers must be as stated and the dispatch must be the same as under nodal pricing. Finally, we realize that there are no profitable deviations from the stated equilibrium if accepted offers are at the marginal cost of the marginal unit in the node, and non-accepted offers are at their marginal cost.

Thus, the discriminatory auction is identical to nodal pricing in terms of payoffs, efficiency, social welfare and the dispatch. As payoffs are identical for all circumstances, this also implies that the long-run effects are the same in terms of investment incentives etc. Finally we show that physical forward positions (such as day-ahead sales) do not influence the equilibrium outcome.

**Proposition 6.** The equilibrium dispatch and equilibrium prices in a market with nodal or discriminatory pricing is $\{q_i^N\}_{i=1}^n$ and $\{p_i^N\}_{i=1}^n$, respectively, for any physical forward position.

**Proof:** Assume that producers $q_i \in [0, f_i]$ in node $i$ have together sold $f_i$ physical forward contracts. In the real-time market the system operator accepts additional offers on top of physical forward trading and it maximizes stated social welfare for these additional offers, which is given by:

$$\Delta W = -\sum_{i=1}^{n} \int_{f_i}^{q_i} o_i(y)dy = -\sum_{i=1}^{n} \int_{0}^{q_i} o_i(y)dy + \sum_{i=1}^{n} \int_{0}^{f_i} o_i(y)dy, \quad (5)$$

but $\sum_{i=1}^{n} \int_{0}^{f_i} o_i(y)dy$ is a constant that is not influenced by the system operator’s dispatch. Thus it can be disregarded from the optimal dispatch problem, so maximizing $\Delta W$ is equivalent to
maximizing $W$. This implies that we can go through the same arguments used in Lemma 3 and Propositions 4-5 to prove that these statements also apply with contracting. Thus we can conclude that the equilibrium outcome is not influenced by physical forward trading. ■

We will use this result in our analysis of zonal pricing, where the first-stage clearing of the zonal market can be regarded as physical forward sales.

2.7 Zonal pricing with counter-trading

Zonal pricing with counter-trading is more complicated than the other two designs and we need to introduce some additional notation to analyze it. The network is divided into zones, such that each node belongs to some zone $k$. We let $Z_k$ be a set with all nodes belonging to zone $k$ and each node is given a number $i \in \{1, \ldots, n\}$. To simplify our equations, we number the nodes in a special order. We start with all nodes in zone 1, and then proceed with all nodes in zone 2 etc. Thus, for each zone $k$, nodes are given numbers in some range $n_k$ to $\bar{n}_k$. Moreover, inside each zone, nodes are sorted with respect to the network’s competitive nodal prices $p_i^N$, which can be calculated for the nodal pricing design, as shown in Section 4.1. Thus, the cheapest node in zone $k$ is given the number $n_k$ and the most expensive node in zone $k$ is assigned the number $\bar{n}_k$. Power flows between zones are determined and announced by the system operator before offers are submitted to the real-time market. Total net-imports to zone $k$ are denoted by $I_k^N$. We make the following assumption for these flows, as our analysis shows that it leads to an efficient outcome:

**Assumption 1:** The system operator sets inter-zonal flows equal to the inter-zonal flows that would occur for the network’s efficient dispatch $\left\{q_i^N\right\}_{i=1}^n$. These inter-zonal flows are announced by the system operator before offers are submitted to the real-time market.

---

11 In the Nordic countries inter-zonal flows are determined and announced by the system operator before offers are submitted.
The equilibrium in a zonal market with counter-trading has some similarities with the discriminatory auction. But the zonal case is more complicated, as the two stages imply that in equilibrium some producers can arbitrage between their zonal price and their individual counter-trading price, which we below show is equal to the network’s competitive nodal price $p_i^N$. Thus it is profitable for producers in export-constrained nodes, where the network’s competitive nodal price $p_i^N$ is below the zonal price, to offer production at $p_i^N$, even if this price is below the marginal cost of the unit. Such offers are accepted in the zonal market at the zonal price, but due to intra-zonal congestion, the offers are then bought back from the system operator in the counter-trade stage at the lower offer price, $p_i^N$. Hence, the offered supply in nodes where the network’s competitive nodal price is below the zonal price is given by the total production capacity in the node. As the real-time market is physical, producers in import-constrained nodes, where the network’s competitive nodal price $p_i^N$ is above the zonal price, are not allowed to first buy power at a low price in the zonal market and then sell power at $p_i^N$ in the counter-trading stage. Thus they neither buy or sell any power in the zonal market. Thus, the zonal price $\Pi_k$ in zone $k$ is set by the network’s competitive nodal price in node $m(k)$, which we define by:

$$m(k) = n \in \{p_1, \ldots, n_k\}; \quad I_k^N + \sum_{i=2}^{n-1} q_i \leq \sum_{i=2}^{n_k} D_i \quad \text{and} \quad I_k^N + \sum_{i=2}^{n-1} q_i \geq \sum_{i=2}^{n_k} D_i .$$  \tag{6}

We will consider two types of zonal markets. In the first case we assume that producers cannot make new offers to the counter-trading stage; the same offers are used in the two stages of the zonal market.

**Proposition 7.** Under Assumption 1, the Nash equilibrium of a zonal market with counter-trading and the same offers in the zonal and countertrading stages has the following properties:

1) The zonal price in zone $k$ is given by $\Pi_k = p_{m(k)}^N$, where $m(k)$ is defined in (6).
2) In nodes \(i \in Z_k\), such that \(p_i^N < \Pi_k\), the marginal offer and production with marginal costs above \(p_i^N\) are offered at the network’s competitive nodal price \(p_i^N = C'_i(q_i^N)\). For nodes in zone \(k\) where \(p_i^N > \Pi_k\), all production with a marginal cost at or below \(C'_i(q_i^N)\) is offered at \(p_i^N = C'_i(q_i^N)\).

3) Other offers are rejected in both stages and not uniquely determined in equilibrium. However, it can be assumed that they offer at their marginal cost.

4) As in the nodal pricing and pay-as-bid designs, the dispatched production in each node is given by the network’s efficient dispatch, \(q_i^N\).

**Proof:** Offers above the zonal price are never accepted in the first stage of the zonal market. For these nodes, it is the rules of the counter-trading stage that determine optimal offer strategies. Thus, the auction works as a discriminatory auction, and we can use the same arguments as in Proposition 5. The remainder of this proof deals with offers in nodes such that \(p_i^N < \Pi_k\).

We let the marginal offer of the node be the last offer in a node that is dispatched. It follows from Proposition 6 that the zonal clearing does not influence the final dispatch, so Lemma 1 and Corollary 2 are applicable. Thus the marginal offer must be at the marginal cost and all offers in the same node with a lower marginal cost must be dispatched. Otherwise, there would exist some infinitesimally small producer in the node with a marginal cost below the marginal offer, but whose offer is not accepted. Thus, it would be a profitable deviation for such a producer to slightly undercut the marginal offer and we know from Corollary 2 that such a unilateral deviation will not decrease the accepted offer in the node. Units with a higher marginal cost than the marginal offer can still sell their power in the zonal market at the zonal price and then buy it back at a lower offer price in the counter-trade stage. Thus, to maximize profits this power is offered at the lowest possible price, for which offers are not dispatched, i.e. at the marginal offer of the node.

We have now established that marginal offers are at the marginal cost of the node. Thus, in comparison to the nodal pricing dispatch, dispatched supply must be higher in nodes \(i\) with
a marginal offer higher than $p^N_i$ and dispatched supply must be lower in nodes $i$ with a marginal offer lower than $p^N_i$. The efficient dispatch is feasible under Assumption 1 and it follows from Proposition 6 that the zonal clearing does not influence the final dispatch. Thus any deviation from the efficient dispatch would violate Lemma 1. From the proof of Proposition 4, we realize that increasing offers below the marginal offer does not change the optimal dispatch. Thus, as long as offers that are marginal for nodal pricing are unchanged, the optimal dispatch will be the same as for nodal pricing.

We have already verified that non-dispatched production units would not gain by undercutting the marginal offer. Offers that are dispatched in nodes with $p^N_i < \Pi_k$ are paid the zonal price. It is not possible for one of these units to increase its offer price above $p^N_i < \Pi_k$ and still be dispatched, as non-dispatched units in such nodes offer at $p^N_i$. Moreover, it is weakly cheaper for dispatched units to produce instead of buying back power at $p^N_i$. Thus, they do not have any profitable deviations. Accordingly, the stated offers must constitute a Nash equilibrium.

The next result shows that our conclusions for markets with zonal pricing will still hold even if producers are allowed to up-date their offers in the counter-trading stage.\textsuperscript{12}

**Proposition 8.** Under Assumption 1, it does not matter for payoffs and the equilibrium outcome of the zonal market whether producers are allowed to up-date their offers in the counter-trading stage.

\textsuperscript{12}This two-stage model could also represent congestion management in the Nordic market, where the system operator does not accept offers in the zonal clearing of the real-time market if these offers will cause intra-zonal congestion that needs to be countertraded in the second-stage. This is to avoid unnecessary costs for the system operator and unnecessary payments to producers. In our model where there is no uncertainty, the zonal day-ahead market then takes the role of the first-stage of the real-time market. The latter becomes obsolete as without uncertainty, the day-ahead market has already made the zonal clearing. In this case offers to the real-time market, which are allowed to differ from day-ahead offers, are only used in the discriminatory counter-trading stage. Proposition 8 shows that under our assumptions, switching to the Nordic version of zonal congestion management is in vain, producers still get the same payoffs and the system operator’s counter-trading costs are unchanged.
**Proof:** We solve the two-stage game by backward induction. Thus we start by analyzing the countertrading stage. Accepted offers in the first stage of the real-time market are equivalent to physical forward trading. Thus it follows from Proposition 6 that the dispatch after the countertrading stage is the same efficient dispatch as in Proposition 5.

We calculate a subgame perfect Nash equilibrium of the game, so rational agents realize what the outcome of the second-stage is going to be, and make offers to the zonal market in order to maximize profits. Thus all production in nodes $i \in Z_k$, such that $p_i^N < \Pi_k$ is sold at the zonal price, and the zonal price in zone $k$ is set by the node $m(k)$ as defined in (5). Thus all agents get the same payoffs as the game in Proposition 7, where the same offers were used in the zonal and countertrading stages.

We can now conclude that the dispatch for zonal pricing with counter-trading is the same as for nodal pricing and discriminatory pricing. Thus, in the short run, the designs efficiencies are equivalent. This also confirms that the system operator should set inter-zonal flows equal to the corresponding flows in the competitive nodal market, as assumed in Assumption 1, if it wants to maximize social welfare. However, it directly follows from (5) and Propositions 7 and 8 that some producers get unnecessary high payments in a zonal pricing design:

**Corollary 9.** In comparison to nodal pricing, the total extra payoff from the system operator to producers in zone $k$ equals: $$\sum_{i \in m(k)}^{m(k)-1} (p_{m(k)}^N - p_i^N) \Pi_i$$ under Assumption 1.

Even if zonal pricing is as efficient as nodal pricing in the short run, the extra payoffs will cause welfare losses in the long run. Production investments will be too high in nodes where $p_i^N < \Pi_k$. In addition, inflexible production that cannot take part in the real-time market are paid the zonal price in the day-ahead market. Thus, the accepted inflexible supply in this market is going to be too high in nodes with $p_i^N < \Pi_k$ and too low in nodes with $p_i^N > \Pi_k$. 


3. Example

In the following section, we present equilibria for the three market designs that we have been analyzing.

![Diagram of a two-node network with one constrained transmission-line in-between. In both nodes producers are infinitesimally small and demand is perfectly inelastic. For simplicity, we make the following assumptions for each node: the marginal cost is equal to local output and the production capacity is 15 MW. In node 1, demand is at 5 MW; in node 2 demand is at 18 MW. The transmission line between these nodes is constrained and can carry only 4 MW. Demand in node 2 exceeds its generation possibilities so the missing electricity must be imported from the other node.

With nodal pricing, the equilibrium offers will be as follows:

Figure 1.
In the first node infinitesimally small producers make offers at their marginal cost (as in Lemma 3). In order to satisfy local demand and export, 9 MW are going to be dispatched. Out of these, 5 will be consumed locally and 4 will be exported; the highest possible export level that the transmission line allows for. The marginal cost and nodal price is equal to 9, which corresponds to the total production of this node. In the second node, the nodal price is 14 as there are 14 MW that have to be produced in the second node in order to satisfy demand and the transmission constraint. Production above those marginal costs (9 in node 1 and 14 in node 2) will not be dispatched. All accepted production will be paid the nodal price of the node.

The pay-as-bid design will result in the equilibrium offers presented in Fig. 3. In this design, generators are paid according to their bid. Knowing this and having perfect information, producers who want to be dispatched will bid the nodal price of their node, to ensure that they will be dispatched at the highest possible price. Thus, in node 1, they will bid 9 and in
node 2 they will bid 14. Producers who do not want to be dispatched may, for example, bid their marginal costs, which are higher than the nodal prices of the respective nodes. The dispatch will be the same as under nodal pricing design. Although producers will have different bidding strategies in both designs, the overall result will be the same. Accepted production will be paid 9 in node 1 and 14 in node 2.

Figure 3.
In the zonal design with counter-trading, producers will offer as follows:

Figure 4.

Node 1:
Producers with a marginal cost at or below the nodal price 9 may, for example, offer at their marginal cost; they will be paid the zonal price which is 14. No infinitesimally small producer in node 1 can unilaterally increase the zonal price at stage 1, as the system operator would then just accept more production from node 2. 6 units with a marginal cost above the nodal price will bid low in order to be dispatched in the first stage and be paid the zonal price of 14, but as there is a transmission constraint in the second round, they will have to buy back their supply at the bidding price. As they are interested in maximizing their profit, they want this difference in prices to be as large as possible but, at the same time, they want to ensure that they will not be chosen to produce. Therefore, they bid the nodal price 9

13 They might as well bid at the nodal price or anything below; it does not matter as they will anyway be dispatched and paid the zonal price.
so that finally, they will be “paid” not to produce and get $14 - 9 = 5$ (the gray area in the figure 5). There are no profitable deviations from these bids for producers from node 1.

Node 2:
Due to the transmission constraint, producers in node 2 know that the system operator needs to accept at least 14 units of electricity in their node after the two stages. Thus, all low-cost generators who want to be dispatched know that all offers at or below 14, the nodal price of node 2, will be accepted. But 6 units of the dispatched production in node 2 are accepted in the counter-trade stage. These units are paid as bid and accordingly, they maximize their profit by offering their supply at 14, the highest possible price for which they are going to be accepted. Producers that do not want to be dispatched at all will bid either their marginal cost or a higher price. In this way, 14 units will be produced in node 2. There are no profitable deviations from these strategies for producers in node 2.
A comparison of these three examples illustrates that although the bidding strategies are different, the dispatch is the same in all scenarios. However, the last design – zonal pricing with counter-trading – results in additional payments that affect e.g. the long-term investment incentives.

4. Conclusions

We consider a general electricity network (possibly meshed), where nodes are connected by capacity constrained transmission lines. In our game-theoretical model producers are infinitesimally small and demand is certain and inelastic. We find that the three designs, nodal, zonal with countertrading and discriminatory pricing, leads to the same socially efficient dispatch. In addition, payoffs are identical in the pay-as-bid and nodal pricing designs. However, in the design with zonal pricing and countertrading, there are additional payments from the system operator to producers. It does not matter for our results whether producers are allowed to update their offer curves in the counter-trading stage. Similar to Dijk and Willems’ (2011) two-node model, this implies that producers overinvest in export-constrained nodes.

Another result from our analysis is that there is a significant amount of firms that make offers exactly at the marginal prices of the nodes in the zonal and pay-as-bid designs, which is not necessarily the case under nodal pricing. This suggests that prices in especially the pay-as-bid design, but also in the zonal design, are less sensitive to shocks in comparison to nodal pricing. This supports the common view that the zonal design should improve liquidity in the market compared to nodal pricing. Still it is known from PJM that it is possible to have a liquid market also with nodal pricing (Neuhoff and Boyd, 2011).

4.1 Discussion

Our model is idealized in many ways: producers’ costs are assumed to be common knowledge and demand is certain and inelastic. Moreover, all agents can trade in real-time and they
choose offers to maximize their profits. A benevolent system operator accepts offers to maximize the stated social welfare, i.e. it assumes that offers reflect true costs. Finally we want to briefly discuss how our results would change if these assumptions are relaxed. In this case, the zonal design gets additional problems. Inflexible plants with long ramp-rates are often not allowed to trade in the real-time market; they have to sell at the zonal price in the day-ahead market. This imperfection will result in too much inflexible production in export constrained nodes and too little inflexible production in import constrained nodes. Related issues are analyzed by Green (2007). Hogan (1999) and Harvey and Hogan (2000) show that nodal pricing is better suited to prevent market power as compared to zonal pricing. Björndal et al. (2003) and Glachant and Pignon (2005) show that network operators may not have any incentives to maximize social welfare in markets with zonal pricing. Moreover, Björndal and Jörnsten (2001) and Ehrenmann and Smeers (2005) discuss problems with optimally choosing zones in large networks. In our analysis we assume that the system-operator sets inter-zonal flow efficiently as under nodal pricing, but in reality this may not be the case, especially for cross-border flows (Neuhoff, et al., 2011). Green (2010) points to the problems of accommodating uncertainties from intermittent power within a market design where spatial prices do not exist. In particular, as for example illustrated by Anderson et al. (2009), the elastic offers in especially the pay-as-bid design, but also in the zonal design, means that if a firm gets its offer slightly wrong, then it can have a huge influence on its dispatch. This increases the chances of getting inefficient dispatches when demand or competitors’ output is uncertain, while the nodal pricing design seems to be more robust to these uncertainties. Thus from our analysis and this discussion it can be concluded that it is difficult to find academic support for the zonal design.
References


Evans, J.E. and Green, R.J., (2005), Why Did British Electricity Prices Fall After 1998? Working paper 05-13, Department of Economics, University of Birmingham.


