Power Generation Investment under Zonal Electricity Pricing with Market-Based Re-dispatch

Quentin Lété[1,2,*], *Yves Smeers*[1], *and Anthony Papavasiliou*[3]

January 22, 2025

Abstract

There is currently an intense debate in Europe on the allocation of transmission capacity in the internal electricity market. In this context, the concept of market-based re-dispatch has regained considerable attention, as a candidate design to reconcile the zonal organization of the market with a pricing system that supports efficient real-time dispatch. In this paper, we study the problem of capacity expansion under zonal pricing with market-based re-dispatch. We propose a formulation of the problem as a mixed linear complementarity problem and analyze the existence and efficiency of solutions, both from a theoretical and empirical perspective. We find that zonal pricing with market-based re-dispatch leads to large efficiency losses. Moreover, although there exist theoretical conditions under which the efficiency of the design could be improved by means of regulatory interventions, the strictness of these conditions renders their practical implementation difficult and, therefore, the full recovery of the efficiency unlikely.

Keywords: Zonal pricing, capacity expansion, congestion management

[1]LIDAM/CORE, UCLouvain, Louvain-la-Neuve, Belgium.
[2]Power Systems Laboratory, ETH Zurich, Switzerland.
[3]School of Electrical and Computer Engineering, NTUA, Greece.
[*]Corresponding author. E-mail: quentin.lete@uclouvain.be.

1 Introduction

1.1 Motivation

The debate between nodal and zonal pricing is an old story in the evolution of the US power system. In their three-phase review of the history of the process, Hobbs and Oren (2019) note, almost in passing, that the subject had been settled in the first phase, namely after the California experience: the zonal system had fundamental flaws that the nodal system avoided. The Texas experience came a bit later but confirmed the observation (Triolo and Wolak 2021). The situation is quite different in Europe: a recent report (ACER 2022) issued in response to criticisms against the high prices observed in the power market in the end of 2021 (that is, well before the perturbations on the gas market induced by the events in Ukraine) argues that the current European (zonal) system is functioning properly, but that one might consider some improvements, among which a reduction of the size of the zones. In particular, the nodal system is mentioned as an example of this reduction (page 27 of ACER (2022)). This difference between the US and EU points of view after so many years of theory and experience on a subject that looks purely technical is strange and must have some fundamental explanation. It cannot be a misunderstanding of the nodal system, which had been described in the French literature as early as Boiteux and Stasi (1952). But institutions suggest a possible reason for divergent points of view. The restructuring of the power system in Europe was part of the integration of sectoral markets that created the internal market in 1992, with delays expected in "difficult sectors", electricity being one of them. It suffices to note, here, that a nodal model in an integrated European market would have merged national systems and interconnections into a new supranational model where all components of national networks would be placed on the same footing. Conversely, the coupling of the zonal national systems maintained their individuality and developed a framework facilitating exchanges between them through interconnections. In other words, the nodal system in the internal EU market would have had a supranational flavor that the zonal system was thought to be able to bypass.

As a consequence of the institutional barriers, the EU internal electricity market developed in general at a slower pace than their US counterparts. This can be observed for instance in how the

EU market has been organized originally around the day-ahead market, which has, until recently, been viewed as the spot market. Re-dispatching and balancing are instead seen as services used to maintain as much as possible the day-ahead positions. Continuous efforts have been made to gradually improve on this basic design with, for instance, the development of platforms to coordinate the activation of balancing capacity (MARI and PICASSO) or the efforts to integrate more precise scarcity pricing. There is, however, an intrinsic limit to these improvements, which is that the exact conditions under which electricity trading takes place are known only in real time. Therefore, as argued in Harvey and Hogan (2022), "Any pricing system that sets prices that are inconsistent with the real-time dispatch necessarily requires discriminatory pricing, limits on access, constrained-on and –off payments, or reliance on command and control to maintain reliability". With the increasing integration of renewable resources and distributed energy resources, which bring more variation and uncertainty in the grid, the reliance on these ad-hoc measures becomes increasingly complicated. In this context, market-based re-dispatch represents a seemingly appealing modification to the European design, introducing a pricing system that is closer to the real-time dispatch without completely questioning the zonal organization. Market-based re-dispatch draws indeed more and more interest among EU market design discussions, as witnessed by the fact that it is now the default option in the regulation (European Parliament, Council of the European Union 2019).

At the same time, there has recently been an increasing interest in local flexibility markets, for which a certain number of pilots are emerging all over the continent (Frontier Economics and Entso-e 2021). It is interesting to note that the primary use of these local flexibility markets is congestion management, as revealed in a poll of Nordic DSOs (Hackett et al. 2021). Therefore, most of these initiatives work on the principle of locational pricing, which implies that opportunities exist for inc-dec gaming.

In times of energy transition, it is therefore crucial to understand the relationship between power generation investment and a short-term market design based on zonal pricing followed by market-based re-dispatch. Our interest in this paper is to analyze this question from a computational perspective, by proposing a model of a long-term competitive equilibrium under this design.

Along with the development of the model, we propose a solution methodology that enables the computation of an equilibrium on large-scale instances, and apply it on a realistic instance tailored to the Central Western European network.

1.2 Related literature

There is an existing stream of literature centered on the quantitative analysis of zonal pricing with market-based re-dispatch. In the early days of the discussions on market-based congestion management in Europe, De Vries and Hakvoort (2002) analyzed five different designs among which uniform pricing followed by market-based re-dispatch, which was then referred to as *countertrading*. The authors proposed a stylized two-node one-zone example and concluded that, in theory, countertrading is efficient in the short term. The model does not account for inc-dec gaming, which is an important aspect in market-based re-dispatch. Potential inefficiencies stemming from deviations from this idealized setting analyzed in the paper are discussed qualitatively in the work of De Vries and Hakvoort (2002).

One situation that could be associated to inefficiencies in market-based re-dispatch in the short term takes place when the TSO does not act as a price-taker in the re-dispatch market. This situation is analyzed by Grimm et al. (2018) under the assumption of absence of inc-dec gaming. The authors show, on an illustrative example, that, in this case, the outcome could be inefficient.

An important drawback of re-dispatch markets is that they offer the possibility to market participants to engage in arbitrage between the zonal and re-dispatch price, the so-called inc-dec game. This leads to a distortion of long-run investment incentives as well as the extraction of short-term rents by those exercising the gaming strategy for essentially offering nothing to the system. This situation is analyzed quantitatively in Holmberg and Lazarczyk (2015). The authors show that inc-dec gaming is an arbitrage strategy that also occurs under perfect competition. They conclude that, under certain restrictive assumptions, zonal pricing followed by market-based re-dispatch is efficient in the short term, and the outcomes only differ by a redistribution of the welfare.

More recently, Hirth and Schlecht (2020) proposed a simple model on a two-node one-zone example in order to analyze inc-dec gaming. The authors emphasize that the design leads to undue arbitrage opportunities for market participants, even when they cannot exercise market power. The authors further discuss the conditions under which this arbitrage is exacerbated or can be mitigated.

Our work also relates to an emerging stream of literature on the study of the impact of locational instruments in zonal markets from a quantitative perspective. The first paper that tackles this problem quantitatively is Grimm et al. (2019). The authors propose a tri-level model of the long-run equilibrium of zonal pricing with cost-based re-dispatch. The tri-level structure aims at representing the sequential nature of the problem: grid investment by the TSO, capacity investment by private firms and re-dispatching in the short term by the TSO. A simplified spatially differentiated capacity signal based on average nodal prices is then added to the model. The authors find that, although it influences the location of investment, the introduction of a well-calibrated capacity signal only slightly improves welfare, as the signal fails to also impact the operational efficiency of the system. Schmidt and Zinke (2020) study the impact of uniform pricing in Germany on the siting of wind generation. They find that nodal pricing, by incentivizing the siting of investment closer to load centers and thereby reducing wind curtailment, has a positive effect on welfare. The paper also investigates the restoration of a locational component to the uniform price through capacity-based latitude-dependent connection charges. However, the authors find that these simplified locational instruments are not adequate for mitigating the inefficiencies associated to uniform pricing. Finally, Eicke (2021) proposes a model for quantifying the optimal capacity-based locational signal for restoring the efficiency of investment in zonal pricing with cost-based re-dispatch. Unlike the two papers previously cited, this paper is the first to consider a technology-differentiated signal. The author shows that differentiating by technology is a necessary condition for restoring the efficiency of nodal pricing. They find that, although locational capacity signals have a significant cost-saving potential, they do not lead to a full recovery of the efficiency of nodal pricing.

1.3 Contributions

This paper contributes to the existing literature in the following ways:

- Our main contribution is to propose a model that internalizes the investment decisions of private firms in generating capacity in a zonal market design with market-based re-dispatch. We show how the loss of efficiency of this design originates from its mathematical formulation and formally establish the existence of a resulting equilibrium.
- Based on this modeling framework, we discuss the theoretical conditions under which the long-run efficiency of nodal pricing can be recovered in a zonal pricing market with market-based re-dispatch.
- We show how to integrate market monitoring considerations in the analysis of the long-term equilibrium by imposing limitations on the deviation from truthful bidding for market participants.
- Finally, we provide simulations of the long-term equilibrium in zonal pricing followed by market-based re-dispatch on a large-scale instance of the Central Western European (CWE) network area and compare its efficiency to nodal and zonal pricing benchmark designs.

1.4 Organization of the paper

The paper is organized as follows: we start, in section 2, by presenting our model of zonal pricing with market-based re-dispatch in the short term. Then, section 3 extends the model to the case of the long-term equilibrium and discusses how the integration of long-term considerations leads to a loss of efficiency of the design. In section 4, we analyze how efficiency can be restored by means of additional market instruments. Section 5 extends our model to incorporate an endogenous representation of market monitoring that limits zonal-nodal arbitrage. Finally, in section 6, we present the results of our case-study before discussing some concluding remarks in section 7.

2 Short-term equilibrium

2.1 Model

We start by presenting a simple model of zonal pricing followed by market-based re-dispatch (abbreviated MBR hereafter) in the short term, i.e. when investment decisions in generation capacity are ignored. The re-dispatch problem that we consider in this paper is an economic dispatch with nodal transmission constraints and pay-as-cleared settlements. We model the economic equilibrium on the electricity market as a Nash equilibrium between three types of agents: a single TSO, the producers, and a Walrasian auctioneer that enforces market clearing conditions (i.e. a power exchange). In order to focus on the impacts of MBR on the short and long-term equilibria, we adopt a set of simplifying assumptions that allow us to isolate the effects related to this congestion management policy. In particular, we assume that the market is perfectly competitive and we model the agents as price takers. This implies that we ignore market power. Although real markets deviate from the situation of perfect competition, our view is that it remains essential to understand the performance of market designs under perfect competition, as it is unlikely that identified inefficiencies will disappear in imperfectly competitive markets. The assumption of perfect competition allows us to keep the models transparent and tractable.

Note that absence of market power also forces us to question the terminology of "inc-dec gaming". Indeed, the market design that we analyze in this paper modifies the opportunity cost of participants in the zonal market with the introduction of the subsequent re-dispatch market. This means that their cost is now an opportunity cost which differs from their marginal production cost. In our setting, participants would thus still bid at cost, whereas inc-dec gaming suggest some kind of strategic behavior due to market power. Therefore, in order to avoid any confusion, we refrain from using the terminology "inc-dec gaming" and, instead, refer to **zonal-nodal arbitrage** throughout the paper to describe the behavior of market participants.

In addition, we assume that the profit-maximizing problems of all agents are linear and that there are no intertemporal operating constraints: in the short run, all periods are independent. We consider that producers can invest in generation capacity in a continuous way, and we ignore

transmission capacity expansion, for which the continuity assumption would be unrealistic.

Electricity demand is assumed to be known and inelastic. We make this assumption in order to keep the equations and the theoretical analysis as simple as possible, but we note that our model, as well as the theoretical results, naturally extend to the case of active demand-side participation. The logic is that consumers can also engage in price arbitrage between the two sequential markets, in a symmetrical way as the behavior of the producers. This is formally shown in Appendix B, where we discuss extensions of our model to different types of market participants, including active consumers. We also show that the results of the paper naturally extend to these different cases.

Finally, our model is deterministic and we assume perfect foresight. This implies in particular that market participants are able to perfectly anticipate the subsequent re-dispatch price when bidding in the zonal market. We note that this is a strong assumption that is not likely to be verified in practice. As noted in Hirth and Schlecht (2020), exercising price arbitrage would expose market participants to risk, which would probably limit the degree to which participants would be willing to exercise this arbitrage if risk aversion is taken into account. In the present paper, we abstract from uncertainty considerations to keep the model transparent, computationally tractable and absent of additional considerations regarding risk aversion and the forecasting abilities of agents. The endogenous modeling of uncertainty in the context of market-based re-dispatch is relegated to future research. Note that the perfect foresight assumption also allows us to omit N-1 security constraints in the representation of transmission capacity.

We now proceed to the sequential description of the profit-maximizing problems of each agent under our set of assumptions.

Producers. The producers with technology $i \in I$ at node $n \in N$ of the transmission network maximize their profit from the sale of electricity given the prices at their node. The pair (i, n) is assumed to uniquely identify a generator in our model. Generators incur a linear cost with marginal cost MC_i, and face capacity limits X_{in} on their production. In zonal pricing followed by MBR, they have the opportunity to participate in two different markets at every time period $t \in T$: the zonal market with zonal price $\rho_{Z(n)t}$ and the re-dispatch market with nodal price $\tilde{\rho}_{nt}$. We

denote by y_{int} the quantity cleared in the zonal market and by $\tilde{y}_{int}$ the re-dispatch quantity cleared in the re-dispatch market at period t, which is positive (negative) in case of upward (downward) re-dispatch. Their problem at each time period t can be written as follows:

$$\begin{aligned} \max_{y_{int}, \tilde{y}_{int}} \quad & \rho_{Z(n)t} y_{int} + \tilde{\rho}_{nt} \tilde{y}_{int} - MC_i(y_{int} + \tilde{y}_{int}) \\ \text{s.t.} \quad & X_{in} - y_{int} \geq 0 \\ & X_{in} - y_{int} - \tilde{y}_{int} \geq 0 \\ & y_{int} + \tilde{y}_{int} \geq 0 \\ & y_{int} \geq 0 \end{aligned} \tag{1}$$

The constraints of problem (1) ensures that the quantities cleared in both the zonal and re-dispatch markets are physically feasible[1]. Note that here, again, we use a basic producer model to keep the equations and the theoretical analysis as simple as possible. However, it should be noted that the results that we present in this paper are robust to the specific representation of market participants. In particular, under the assumptions of absence of market power and perfect information, the results naturally extend to renewable resources and storage units. The reader is referred to Appendix B for a formal discussion of the extension of our model to the case of renewable and storage resources.

We split the description of the TSO problems in the zonal and re-dispatch markets in the interest of clarity. As the two problems are completely independent, this can be done without loss of generality.

TSO in the zonal market. The TSO maximizes its profit from the transmission of electricity given the zonal prices by controlling the net position of each zone (p_{zt}) while making sure that the net positions belong to a certain predefined set of feasible net positions. This set, that we denote

[1]Note that this prevents financial arbitrage, following the assumption of Hirth and Schlecht (2020). Using the terminology of Hirth and Schlecht (2020), our model is thus based on asset-backed arbitrage which implies that participants can only bid quantities that they would be able to physically produce in both the zonal and the re-dispatch markets.

by $\mathcal{P}$, thus encodes the zonal transmission constraints[2]:

$$\begin{aligned} &\max_{p_{zt}} -\sum_z p_{zt}\rho_{zt} \\ &\text{s.t. } p_{:t} \in \mathcal{P} \end{aligned} \tag{2}$$

where $p_{:t}$ denotes the vector of net positions at a given time period. There exist different ways for defining set $\mathcal{P}$, but we keep it general in our discussion and only assume that it includes the power balance constraint ($\sum_z p_{zt} = 0$) and that it depends solely on grid quantities.

TSO in re-dispatch market. In the re-dispatch market, the TSO has to buy re-dispatch resources ($\tilde{r}_{nt}$, negative or positive) in order to recover a nodal dispatch (r_{nt}) that is feasible for the full set of constraints of the transmission grid, that we denote by $\mathcal{R}$:

$$\begin{aligned} &\max_{r_{nt},\tilde{r}_{nt}} -\sum_n \tilde{r}_{nt}\tilde{\rho}_{nt} \\ &\text{s.t. } r_{nt} - \sum_i y_{int} + D_{nt} - \tilde{r}_{nt} = 0 \\ &\quad r_{:t} \in \mathcal{R} \end{aligned} \tag{3}$$

Note the fundamental difference between the formulations of the zonal market and nodal re-dispatch market: whereas the zonal market only imposes constraints on the zonal net positions $p_{:t}$, in the re-dispatch market the TSO needs to ensure that the net injections of every node $r_{:t}$ respect the nodal transmission constraints. Assuming that these constraints can be represented by the DC approximation, the feasible set of nodal net injections $\mathcal{R}$ can be formulated without any ambiguity as follows:

$$\begin{aligned} \mathcal{R} = \Big\{ r \in \mathbb{R}^{|N|} \,\Big|\, \exists f \in \mathbb{R}^{|K|} \;: \\ f_k = \sum_n PTDF_{kn} \cdot r_n, k \in K \\ \sum_{n \in N} r_n = 0, -TC_k \leq f_k \leq TC_k, k \in K \Big\} \end{aligned} \tag{4}$$

[2]The separation between the TSO and the market operator follows the organization of the European market. In practice, the optimality conditions of (2) are directly added to the market clearing algorithm. In the case of the European electricity market, therefore, KKT conditions of (2) correspond to the network equilibrium conditions that are part of Euphemia, the algorithm that clears the zonal day-ahead market.

where f_k is the power flow on line $k \in K$, r_n is the net injection at node $n \in N$, $PTDF_{kn}$ is the power transfer distribution factor of line k and node n, and TC_k is the thermal limit of line k.

Auctioneer in the zonal market. In the zonal market, the auctioneer determines zonal prices while ensuring that the market clears in every zone of the network. This function can be represented as a profit-maximizing problem for zone z as follows:

$$\max_{\rho_{zt}} \ \rho_z(p_{zt} - \sum_{i,n \in N(z)} y_{int} + D_{zt}) \tag{5}$$

Auctioneer in the re-dispatch market. Similarly, the auctioneer of the re-dispatch market ensures the TSO's demand for re-dispatch resources matches its supply from market participants.

$$\max_{\tilde{\rho}_{nt}} \ \tilde{\rho}_{nt}(\tilde{r}_{nt} - \sum_{i} \tilde{y}_{int}) \tag{6}$$

It is important to note that the right solution concept for this game is that of a generalized Nash equilibrium (GNE). Indeed, the feasible set of the TSO in the re-dispatch market, as described in Problem (3), is based on the nodal net injections in the physical dispatch that depend on the production in the zonal market. This implies that it is not equivalent to a single optimization problem. Nevertheless, the equilibrium can still be formulated as a single problem in the form of a mixed linear complementarity problem (MLCP).

Equivalent MLCP. The equivalent MLCP can be obtained by aggregating the KKT optimality conditions (which are necessary and sufficient for linear programs) of the profit-maximizing problem of every agent. Using Greek letters for denoting dual variables for each constraint, we obtain the

following MLCP:

$$0 \leq y_{int} \perp MC_i + \mu_{int} + \tilde{\mu}_{int} - \rho_{Z(n)t} - \delta_{int} \geq 0 \tag{7a}$$

$$\tilde{y}_{int} \text{ free } \perp MC_i + \tilde{\mu}_{int} - \tilde{\rho}_{nt} - \delta_{int} = 0 \tag{7b}$$

$$0 \leq \mu_{int} \perp X_{in} - y_{int} \geq 0 \tag{7c}$$

$$0 \leq \tilde{\mu}_{int} \perp X_{int} - y_{int} - \tilde{y}_{int} \geq 0 \tag{7d}$$

$$0 \leq \delta_{int} \perp y_{int} + \tilde{y}_{int} \geq 0 \tag{7e}$$

$$p_{zt} \text{ free } \perp \rho_{zt} + \sum_m V_{mz}\gamma_{mt} = 0 \tag{7f}$$

$$0 \leq \gamma_{mt} \perp W_m - \sum_z V_{mz}p_{zt} \geq 0 \tag{7g}$$

$$\tilde{r}_{nt} \text{ free } \perp \tilde{\rho}_{nt} - \nu_{nt} = 0 \tag{7h}$$

$$r_{nt} \text{ free } \perp \nu_{nt} + \sum_m \tilde{V}_{mn}\tilde{\gamma}_{mt} = 0 \tag{7i}$$

$$\nu_{nt} \text{ free } \perp -r_{nt} + \sum_i y_{int} - D_{nt} + \tilde{r}_{nt} = 0 \tag{7j}$$

$$0 \leq \tilde{\gamma}_{mt} \perp \tilde{W}_m - \sum_{nt} \tilde{V}_m r_{nt} \geq 0 \tag{7k}$$

$$\rho_{zt} \text{ free } \perp -p_{zt} + \sum_{i,n \in N(z)} y_{int} - D_{zt} = 0 \tag{7l}$$

$$\tilde{\rho}_{nt} \text{ free } \perp -\tilde{r}_{nt} + \sum_{in} \tilde{y}_{int} = 0 \tag{7m}$$

where we define $V \in \mathbb{R}^{|M|\times|Z|}$, $W \in \mathbb{R}^{|M|}$ and M, a set of indices, such that

$$p_{:t} \in \mathcal{P} \Leftrightarrow W_m - \sum_z V_{mz}p_{zt} \geq 0, m \in M$$

and $\tilde{V} \in \mathbb{R}^{|\tilde{M}|\times|N|}$, $\tilde{W} \in \mathbb{R}^{|\tilde{M}|}$ and $\tilde{M}$, a set of indices, such that

$$r_{:t} \in \mathcal{R} \Leftrightarrow \tilde{W}_m - \sum_n \tilde{V}_{mn}r_{nt} \geq 0, m \in \tilde{M}$$

which are well-defined as both $\mathcal{P}$ and $\mathcal{R}$ are polytopes.

This MLCP formalizes the short-term equilibrium in markets with zonal pricing and MBR, such as the one that is described on a simple one-zone two-node network in Hirth and Schlecht (2020). It can thus be used to characterize and represent quantitatively the so-called inc-dec game in perfect competition, which is in reality an arbitrage between the zonal and re-dispatch market.

In the short term, although the equilibrium cannot be formulated as a single LP, a further simplification of the MLCP can be obtained by observing that the part of the problem that relates to the re-dispatch market is independent of the part that relates to the zonal market. More precisely, let us consider the MLCP obtained by isolating the equations related to the re-dispatch problem only, i.e. the MLCP made of equations (7b), (7d), (7e), (7h), (7i), (7j), (7k) and (7m). If we denote by $\bar{y}_{in} = y_{in} + \tilde{y}_{in}$ the physical dispatch of technology i in node n, one can simplify this MLCP by eliminating variables $\delta_{in}, \nu_n, \tilde{r}_n$. These variables were indeed useful to describe the profit maximization problem of each agent, but become superfluous to analyze the equilibrium. We get:

$$0 \leq \bar{y}_{int} \perp MC_i + \tilde{\mu}_{int} - \tilde{\rho}_{nt} \geq 0 \tag{8a}$$

$$0 \leq \tilde{\mu}_{int} \perp X_{in} - \bar{y}_{int} \geq 0 \tag{8b}$$

$$r_{nt} \text{ free } \perp \tilde{\rho}_{nt} + \sum_m \tilde{V}_{mn}\tilde{\gamma}_{mt} = 0 \tag{8c}$$

$$\tilde{\rho}_{nt} \text{ free } \perp r_{nt} - \sum_i \bar{y}_{int} + D_{nt} = 0 \tag{8d}$$

$$0 \leq \tilde{\gamma}_{mt} \perp \tilde{W}_m - \sum_n \tilde{V}_{mn} r_{nt} \geq 0 \tag{8e}$$

Problem (8) fully characterizes the solution to problem (7) for the variables associated to the re-dispatch market, i.e. the variables denoted with the tilde. Observe that MLCP (8) is exactly the

set of KKT conditions of the nodal economic dispatch problem:

$$\begin{aligned}
\min_{\bar{y}_{int}\geq 0, r_{nt}} \quad & \sum_{in} MC_i \bar{y}_{int} \\
\text{s.t.} \quad & X_{in} - \bar{y}_{int} \geq 0, i \in I, n \in N \\
& r_{nt} - \sum_{in} \bar{y}_{int} + D_{nt} = 0, n \in N \\
& r_{:t} \in \mathcal{R}
\end{aligned} \tag{9}$$

This means that the physical dispatch, obtained after the re-dispatch market, corresponds to the optimal solution of the nodal economic dispatch.

Now, let us look in further detail at the MLCP obtained by isolating the equations related to the zonal market, i.e. the MLCP made of equations (7a), (7c), (7f), (7g) and (7l). Equation (7b) implies that

$$MC_i + \tilde{\mu}_{int} - \delta_{int} = \tilde{\rho}_{nt}$$

Accounting for this simplification, we get that MLCP (7a), (7c), (7f), (7g), (7l) is exactly the set of KKT conditions of the following optimization problem:

$$\begin{aligned}
\min_{y_{int}\geq 0, p_{zt}} \quad & \sum_{in} \tilde{\rho}_{nt} y_{in} \\
\text{s.t.} \quad & X_{in} - y_{int} \geq 0, i \in I, n \in N \\
& p_{zt} - \sum_{i,n\in N(z)} y_{int} + D_{zt} = 0, z \in Z \\
& p_{:t} \in \mathcal{P}
\end{aligned} \tag{10}$$

That is, the zonal part of the problem corresponds to a zonal economic dispatch problem where the marginal cost of technology i has been replaced by the opportunity cost $\tilde{\rho}_{nt}$ at that location[3].

The decomposition between the zonal and re-dispatch parts of the problem that we have just

[3]This observation is important, as it reveals the bidding behavior that a rational fringe agent would have in this setting. Indeed, the interest of an agent in this market is to bid $\tilde{\rho}_{nt}$, for the same reason why rational fringe agents bid at their marginal cost in the usual economic dispatch problem, where the coefficients in the objective function are the agents' marginal costs MC_i instead of $\tilde{\rho}_{nt}$.

highlighted has two important implications in the short term regarding (i) the solution methodology for solving the equilibrium, that we discuss in section 2.2 and (ii) the efficiency of the equilibrium, that we discuss in section 2.3.

2.2 Solution methodology

The decomposition between the zonal and re-dispatch parts reveals a simple two-step solution methodology for obtaining an equilibrium of zonal pricing followed by MBR in the short term, that we present in Algorithm 1. Note that this is implicitly the methodology that is used by Hirth

Algorithm 1 Short-Term Equilibrium Solution

1. Solve the nodal economic dispatch problem, i.e. model (9), in order to obtain the optimal values of the variables associated to the re-dispatch market as the primal-dual optimal solutions. Denote by $\tilde{\rho}^*_{nt}$ the optimal value of the re-dispatch price.
2. Solve the zonal economic dispatch with opportunity cost, i.e. model (10) with $\tilde{\rho}_{nt}$ fixed to $\tilde{\rho}^*_{nt}$. This gives the optimal values of the remaining variables, i.e. the variables associated to the zonal part of the problem.

and Schlecht (2020) to identify an equilibrium in their illustrative example: they assume that the opportunity cost of bidding in the zonal market corresponds to the nodal prices and then check that the assumed prices on the re-dispatch market turn out to be correct.

2.3 Efficiency

A second implication of the decomposition of the problem between the zonal and re-dispatch parts is the efficiency of the design in the short term. A design is efficient if it leads to the same total cost as the centralized optimization problem, which, in our case, corresponds to the nodal economic dispatch problem introduced in equation (9). The efficiency result is formalized in Proposition 1.

Proposition 1. Zonal pricing followed by MBR, as formulated in MLCP (7), is efficient in the short run.

The proofs of all propositions of this paper are provided in Appendix A.

Our model assumes implicitly that there is no irrevocable decision with respect to the outcome

Period	Duration (h)	D_{1t} (MW)	D_{2t} (MW)
1	1500	10,000	1,170
2	5500	4,502	4,502
3	1760	1,000	6,086

Table 1: Level of demand in each node in the illustrative example.

of the zonal pricing auction. This is a crucial assumption for Proposition 1 to hold. In particular, we assume that there is no unit commitment decision made based on the zonal auction and that the zonal net positions cleared in the zonal auction are not firm and can be freely modified during the re-dispatch phase, which implies that the TSOs coordinate perfectly, which includes resorting to cross-border re-dispatch[4].

Proposition 1 should be seen as closely related to Proposition 4 of Holmberg and Lazarczyk (2015) which also states that zonal pricing with MBR is efficient in the short run under a slightly different framework and assumptions. In particular, Holmberg and Lazarczyk (2015) assume that the TSO sets the inter-zonal flows to their level in the efficient dispatch. This assumption is related to our assumption of perfect TSO coordination in the re-dispatch stage: both assumptions imply that the efficient dispatch can be recovered during the re-dispatch stage.

2.4 Results on an illustrative example

In order to illustrate our model, we present in this section its results on a small example. We use the simplest possible network configuration that can be used to illustrate a zonal design: two nodes linked by a fictitious line of 0MW of capacity, grouped into a single zone. We assume that the one-year time horizon is divided into three periods of different duration in which the demand is constant in both nodes. The duration of the period as well as the level of the demand in each period and in each node is displayed in Table 1. Finally, we assume that there are four technologies available for serving the demand, with constant marginal and investment costs, as

[4]Note that although this condition is not fully respected in the current European practice, there is an ongoing effort towards improving cross-border coordination. Indeed, the European Commission guidelines now state that TSOs should develop a common methodology for coordinated re-dispatch and countertrading (European Commission (2015), Art. 35, §1).

Technology	MC (€/MWh)	IC (€/MWh)	X_{N_1} (MW)	X_{N_2} (MW)
Coal	25	16	3,502	4,916
Gas	80	5	5,498	0
Nuclear	6.5	32	1,000	1,170
Oil	160	2	0	0

Table 2: The set of options for serving demand and the optimal invested capacity.

shown in Table 2. In the short-term equilibrium, the capacities need to be fixed. We assume that they are fixed at the level that is indicated in Table 2[5]. As the transmission line in our example is purely fictitious, the two nodes can be seen as completely independent and the nodal solution can be easily obtained using, for instance, the screening curve method (Stoft 2002) on both nodes. The optimal investment is also shown in Table 2. Using these levels of capacity, one can apply Algorithm 1 to solve for the short-term equilibrium in zonal pricing with MBR.

From a qualitative point of view, there are two types of zonal-nodal arbitrage behaviors that can be exercised by producers, depending on whether they expect to be re-dispatched up or down.

The units that will be re-dispatched up in the nodal market (also called constrained-on units) anticipate a higher re-dispatch price than the zonal price. For these resources, it is preferable not to be cleared in the zonal market at all, and sell all their production in the re-dispatch market. As there exists a subsequent market on which they can trade, their opportunity cost is the re-dispatch price that they anticipate, and they have no incentive to bid at a different price than the re-dispatch price in the zonal market.

In a symmetric way, the units that will be re-dispatched down (the constrained-off units) anticipate a lower price in the re-dispatch market. For them, it is preferable to sell all their production in the zonal market at a high price and buy it back in the re-dispatch market at a lower price. Again, knowing that a subsequent market with a different price is coming, their opportunity cost becomes the re-dispatch price, and bidding at the anticipated re-dispatch price becomes their

[5]The values correspond to the optimal capacity in the nodal capacity expansion problem, which is defined formally later in section 3.1, equation (11).

Variable	N_1	N_2
y_{Coal}	3,502	4,916
y_{Gas}	582	0
y_{Nuclear}	1,000	1,170
y_{Oil}	0	0
$\rho_{Z(n)}$	109.2	

Variable	N_1	N_2
$\tilde{y}_{\text{Coal}}$	0	-4,916
$\tilde{y}_{\text{Gas}}$	4,916	0
$\tilde{y}_{\text{Nuclear}}$	0	0
$\tilde{y}_{\text{Oil}}$	0	0
$\tilde{\rho}_n$	109.2	10.4

Table 3: Optimal production levels and prices at the short-term equilibrium in period 1.

optimal strategy.

Our illustrative example is a prime example of these two types of arbitrage behaviors. Let us first focus on the results for period 1. The optimal levels of production and prices for this period are shown in Table 3. Here, the Coal unit in N_2 anticipates that the re-dispatch price will be lower than the zonal price, and that it will be re-dispatched down. Indeed, in period 1, most of the demand is located in N_1 and the local demand in N_2 can be entirely covered by the cheaper Nuclear unit, which implies that the Coal unit does not belong to the optimal mix. It can take advantage of this situation by bidding below its marginal cost in the zonal market in order to be cleared and earn the zonal price and then be re-dispatched down in the re-dispatch market at a lower price (remember that the bidding strategy of the agents is implicit in our model, see Footnote 3). In our example, the RD price is 10.4€/MWh in N_2, which is a solution of the nodal economic dispatch, i.e. Step 1 in Algorithm 1. This corresponds to the opportunity cost of the N_2 Coal unit being cleared in the zonal market and will thus correspond to its bidding price in perfect competition, which is below its marginal cost of 25€/MWh. In total, the N_2 Coal unit earns 109.2€/MWh for being cleared in the zonal market and pays 10.4€/MWh for being re-dispatched down in the re-dispatch market, which is a net profit of 98.8€/MWh, even though it ends up not producing anything.

Let us now have a look at the results in period 3, shown in Table 4. Here, the Coal and Nuclear units in Node N_2 provide a good example of units exercising the second type of inc-dec strategy that corresponds to constrained-on units. Indeed, they anticipate that the re-dispatch price is

Variable	N_1	N_2
y_{Coal}	1,570	0
y_{Gas}	5,498	0
y_{Nuclear}	0	0
y_{Oil}	0	0
$\rho_{Z(n)}$	12.56	

Variable	N_1	N_2
$\tilde{y}_{\text{Coal}}$	-1,570	4,916
$\tilde{y}_{\text{Gas}}$	-5,498	0
$\tilde{y}_{\text{Nuclear}}$	1,000	1,170
$\tilde{y}_{\text{Oil}}$	0	0
$\tilde{\rho}_n$	12.56	104.64

Table 4: Optimal production levels and prices at the short-term equilibrium in period 3.

going to be higher than the zonal price. They thus bid at the re-dispatch price, which is higher than their marginal cost, so as not to be cleared in the zonal market and sell all their production in the re-dispatch market. Note that both types of inc-dec strategies tend to aggravate congestion.

After the re-dispatch market, nevertheless, the optimal production mix is recovered. The short-term welfare is thus equal to that of the nodal design and the two policies differ only in terms of the allocation of surplus between the different agents. The situation is different in the long term, which is the focus of the next section of this paper.

3 Long-term equilibrium

3.1 Model

As discussed in past literature (Holmberg and Lazarczyk 2015, Hirth and Schlecht 2020), although efficient under certain conditions in the short term, zonal pricing with MBR distorts the investment signal and, therefore, leads to inefficiencies in the long term. Our goal in this section is to capture these ineffeciencies in our model by formulating the long-term equilibrium of zonal pricing with MBR. In order to account for investment in generation capacity, we introduce variable x_{in}, which corresponds to the capacity invested in technology i and in node n. This capacity is associated to a linear investment cost with coefficient IC_{in} that depends on the technology and the location. Before describing the modifications to the zonal and market-based re-dispatch model, one also

needs to define the corresponding long-term centralized problem that characterizes the efficient solution. In our case, it corresponds to the capacity expansion problem with nodal transmission constraints, and can be written as follows:

$$\begin{aligned} \min_{x_{in},\bar{y}_{int}\geq 0, r_{nt}} \quad & \sum_{in} IC_{in}x_{in} + MC_i\bar{y}_{int} \\ \text{s.t.} \quad & X_{in} + x_{in} - \bar{y}_{int} \geq 0, i \in I, n \in N, t \in T \\ & r_{nt} - \sum_{in} \bar{y}_{int} + D_{nt} = 0, n \in N, t \in T \\ & r_{:t} \in \mathcal{R}, t \in T \end{aligned} \tag{11}$$

This problem is the counterpart of equation (9) in the long-term case.

We now turn to the long-term model of zonal pricing with MBR. In perfect competition, investment will be made in technology i in node n if the investment cost is recovered exactly through the sale of electricity and re-dispatch resources in the predefined set of time periods T. This condition can be formulated mathematically through the following complementarity condition:

$$0 \leq x_{in} \perp IC_{in} - \sum_{t\in T} \mu_{int} - \sum_{t\in T} \tilde{\mu}_{int} \geq 0 \tag{12}$$

Moreover, the investment in capacity modifies the capacity limits in the zonal and re-dispatch markets in the following way:

$$0 \leq \mu_{int} \perp X_{in} + x_{in} - y_{int} \geq 0 \tag{13}$$

$$0 \leq \tilde{\mu}_{int} \perp X_{in} + x_{in} - y_{int} - \tilde{y}_{int} \geq 0 \tag{14}$$

The MLCP consisting of equations (12), (13), (14), (7a), (7b), (7e)-(7m), with all short-term variables indexed by time period, models the long-term equilibrium of zonal pricing with MBR.

Crucially, the investment condition represented in equation (12) now links together the two parts of the problem related to the zonal market (through variables μ_{int}) and the re-dispatch market (through variables $\tilde{\mu}_{int}$). This implies that the decomposition between the two parts that

we discussed in section 2 is not possible anymore. As a consequence, the efficiency of the design, that resulted from the decomposition in the short-term model, is not guaranteed in the long-term model. Nevertheless, the existence of an equilibrium can still be guaranteed under mild conditions, as we show in Proposition 2.

Proposition 2. If the marginal costs, the investment costs and the demand in all nodes are positive, and if the set $\mathcal{P}$ is such that $W_m \geq 0 \ \ \forall m \in M$, the MLCP that represents the long-term equilibrium has a solution.

Note that condition $W_m \geq 0 \ \ \forall m \in M$ in Proposition 2 is relatively weak. Indeed, by definition of W_m, $W_m \geq 0 \ \ \forall m \in M \Leftrightarrow 0 \in \mathcal{P}$. That is, it is equivalent to assuming that the situation without any exchange between the zones is feasible for the zonal network constraints.

Another important question is that of the uniqueness of solutions. However, the loss of equivalence with a single optimization problem implies that less can be guaranteed. One can nevertheless show that for almost all parameter data, the long-run equilibrium has a finite and odd number of solutions. This claim can be proved using the formalism of homotopy theory, following Bidard (2012). The reader is referred to Appendix C for a more thorough discussion of the uniqueness of solutions to the investment problem.

3.2 Results on the illustrative example

In this section, we illustrate the results of our long-term model on the same example as the one used for the short-term model. Contrary to the short-term case, however, we do not assume any existing capacity. We solve the MLCP using a simple splitting method, that is also used for the case study presented in section 6, and refer the reader to section 6.2 for more details on the solution methodology. Figure 1 shows the optimal capacity mix in both nodal pricing and zonal pricing with MBR. We observe that the optimal mixes are almost perfectly equivalent, with the notable exception that zonal pricing with MBR leads to an additional investment of around 5GW of Oil capacity in node N_2. This result confirms the intuition of Hirth and Schlecht (2020), who predict that MBR incentivizes investment in units with low investment costs and high variable

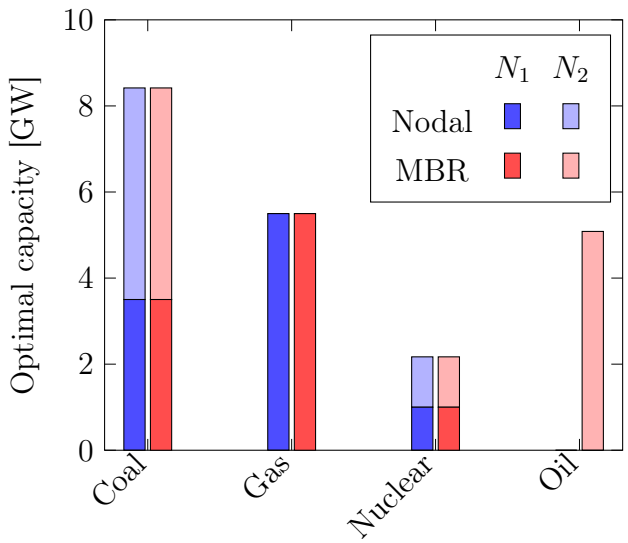

Figure 1: Optimal capacity mix in the long-term equilibrium of the illustrative example for both nodal pricing and zonal pricing with MBR.

costs, that they call "ghost plant". The Oil capacity investment that we observe in the results of our illustrative example is a prime example of such a ghost plant, Oil being indeed the technology with the lowest investment cost in our setting. This unit never produces any energy and earns its profit uniquely from the extraction of an arbitrage rent, such as the one obtained by the N_2 Coal unit that we describe in detail in section 2.4. The results of the illustrative example also confirm that, in the long term, re-dispatch markets are associated to inefficiencies. The total investment and operational costs are indeed increased by 10,168€/h in zonal pricing with MBR compared to that of nodal pricing, which corresponds precisely to the investment cost of the Oil capacity. This additional cost corresponds to a loss of efficiency of 2.1%, although this number should not be considered representative, as the example is purely illustrative. It is important to stress that this situation is obtained under the assumption of perfect competition. The inefficiencies are thus due to a market imperfection that leads to a zonal-nodal arbitrage similar to inc-dec gaming, even though market participants have no market power.

In the next section, we discuss whether zonal pricing with MBR can be modified in order to restore its efficiency.

4 Restoring efficiency

In the previous section, we present a model that captures the inefficiencies of zonal pricing with MBR and illustrates these inefficiencies on a simple example. We now investigate whether these inefficiencies can be corrected by modifying the design.

One might first think that removing the asset-backed arbitrage constraint could improve efficiency, as this would eliminate the incentive to invest in ghost plants. They are indeed the source of inefficiencies, as discussed in section 3.2. This would, however, not solve the inc-dec problem. Removing the asset-backed arbitrage constraint is equivalent to allowing pure virtual trade, i.e., allowing the participation of purely financial agents who exploit arbitrage opportunities. In this case, there would not exist any equilibrium, as zonal and re-dispatch prices could never perfectly converge. Virtual bidders would thus theoretically be able to extract infinite arbitrage rents.

Instead, we argue that, from a technical point of view, there are two ways to understand the loss of efficiency, and each is associated to a possible modification of the design for restoring efficiency. The first way to understand the loss of efficiency is to observe that the production decision of the producers imposes an externality on the decision set of the TSO, which leads to a market failure. We can therefore resort to the classical theory of Pigouvian taxation in order to understand how efficiency can be restored. Another viewpoint that can be adopted is one in which the loss of efficiency is seen as resulting from the interdependence of the zonal and the re-dispatch markets, as we discuss in section 3.1. A disconnection between the two can be imposed through an additional market instrument. These two ways of recovering efficiency are discussed successively in the remainder of this section.

4.1 Pigouvian taxation

A Pigouvian tax, named after Pigou (1932), is a well-known way to correct market failures due to externalities. The idea is to force the agent that causes an externality to internalize it, by imposing a tax exactly equal to the marginal cost of the externality at the optimal solution. In our problem,

the externality comes from the following constraint in the TSO's problem:

$$r_{nt} - \sum_i y_{int} + D_{nt} - \tilde{r}_{nt} = 0 \ \ [\nu_{nt}]$$

Following the Pigouvian theory, efficiency can thus be restored by adding $\nu_{nt} y_{int}$ to the remuneration of the producers, where ν_{nt} is the dual variable of the constraint. According to equation (7h), $\nu_{nt} = \tilde{\rho}_{nt}$, thus the objective function of the producers becomes:

$$\max \ \sum_{t \in T} (\tilde{\rho}_{nt} - MC_i)(y_{int} + \tilde{y}_{int}) + \sum_{t \in T} \rho_{Z(n)t} y_{int} \tag{15}$$

The first term corresponds exactly to the nodal profit for the nodal physical dispatch $y_{int} + \tilde{y}_{int}$. The second term is an additional zonal payment for the production in the zonal market. With this modification, it can now be checked that the long-term equilibrium is equivalent to a single optimization problem, that, after some simplifications, can be formulated as follows:

$$\min_{\substack{x_{in}, \bar{y}_{int} \\ y_{int}, r_{nt}, p_{zt}}} \sum_{int} MC_i \bar{y}_{int} + \sum_{in} IC_{in} x_{in} \tag{16a}$$

$$\text{s.t. } X_{in} + x_{in} - \bar{y}_{int} \geq 0, \quad X_{in} + x_{in} - y_{int} \geq 0, \qquad \forall int \tag{16b}$$

$$-r_{nt} + \sum_i \bar{y}_{int} - D_{nt} = 0, \quad -p_{zt} + \sum_{i \in I, n \in N(z)} y_{int} - D_{zt} = 0, \qquad \forall nzt \tag{16c}$$

$$r_{:t} \in \mathcal{R}, \quad p_{:t} \in \mathcal{P}, \qquad \forall t \tag{16d}$$

where $\bar{y}_{int} = y_{int} + \tilde{y}_{int}$. Problem (16) is, indeed, equivalent to the nodal capacity expansion problem under a mild regularity assumption on the feasible set of net positions $\mathcal{P}$, that we call *nodal consistency* and that we define below.

Definition 1. The feasible set of zonal net positions $\mathcal{P}$ is said to be **nodal consistent** if there exists a solution to the nodal capacity expansion problem, with r^* the vector of values of the nodal net injections in that solution, such that

$$p^*_{:t} \in \mathcal{P} \ \ \forall t \in T$$

with

$$p^*_{zt} = \sum_{n \in N(z)} r^*_{nt} \quad \forall z \in Z, t \in T$$

In words, $\mathcal{P}$ is nodal consistent if it contains, at each time period, a vector of zonal net positions that aggregates a vector of nodal net injections that solves the nodal capacity expansion problem.

In that case, however, one can also observe that the zonal market loses its meaning: there is always an equilibrium solution with a zonal price of 0€/MWh. Indeed, the second term of Equation (15) then vanishes, and the total profit of the producers is exactly the nodal profit for the nodal physical dispatch.

4.2 Locational capacity market

As we mention in section 3.1, the loss of efficiency can also be viewed as resulting from the interdependence between the zonal and re-dispatch markets. This is done through the investment condition, that we recall here:

$$0 \leq x_{in} \perp IC_{in} - \sum_{t \in T} \mu_{int} - \sum_{t \in T} \tilde{\mu}_{int} \geq 0$$

In this section, we argue that the efficiency of nodal pricing can be recovered by introduction a locational capacity market. The introduction of a capacity price π_{in} modifies the investment condition in the following way:

$$0 \leq x_{in} \perp IC_{in} - \sum_{t \in T} \mu_{int} - \sum_{t \in T} \tilde{\mu}_{int} + \pi_{in} \geq 0 \tag{17}$$

The capacity market clearing condition reads

$$\pi_{in} \text{ free } \perp \bar{x}_{in} - x_{in} = 0 \tag{18}$$

where $\bar{x}_{in}$ is the optimal capacity investment, i.e. the solution of the nodal capacity expansion problem. This solution should be interpreted as a target capacity computed beforehand by the

TSO through a process that is comparable to existing initiatives, such as ENTSO-E's Mid-term Adequacy Forecast in Europe (ENTSO-E 2020). Such a mechanism would indeed lead to a recovery of the efficiency, as condition (18) ensures that the investment is optimal, and we have already shown in Proposition 1 that the design is efficient in the short run. However, note that the conditions are very strong: π_{in} can be either positive or negative and is differentiated both per location and per technology.

In Proposition 3, we show that these conditions can be considerably weakened: the efficiency is guaranteed to be recovered even in the case where the locational capacity price is not differentiated per technology and is forced to be positive[6].

Proposition 3. If zonal pricing followed by MBR is augmented with a locational capacity market that determines a non-negative capacity charge, uniform for all technologies of the given location, then it is efficient, i.e. it leads to the same welfare as nodal pricing in the long run.

The intuition behind Proposition 3 is that, as all market participants have the same opportunity cost when participating to the zonal market, they will earn the same margin. Consequently, there exists a non-negative capacity charge, uniform for all units at the same location, that will cancel out the profits made by the participants through the zonal-nodal arbitrage, which restores efficiency.

The crucial point for this result to hold is thus the fundamental observation that all market participants have the incentive to bid the anticipated re-dispatch price in the zonal market. This is what we prove formally in Proposition 3. As it turns out, under the assumptions of absence of market power and perfect information, this is a robust result that does not depend on the specific representation of market participants used in this paper. We discuss this further in Appendix B.

[6]The weakening of the conditions for recovery of efficiency obtained in Proposition 3 is particularly important as it allows to connect the design to existing initiatives in Europe. As discussed in details in Eicke et al. (2020), Great Britain, Ireland and Sweden all implement capacity-based instruments that are differentiated among locations, but none of them are differentiated among technologies. These instruments are not market-based. There is a capacity market in France, but it has no spatial differentiation. However, as discussed in Eicke et al. (2020), France has recently implemented a specific tender to build new capacity in Brittany that was facing adequacy issues. This particular capacity tender thus had explicit spatial differentiation. Moreover, many countries impose a connection charge to new capacities. This shows that a locational capacity market that determines a capacity charge would share many characteristics with existing instruments in different European countries.

5 Market monitoring via price cap and price floor

As we discuss in section 2, zonal pricing with MBR incentivizes market participants to deviate from truthful bidding in the zonal market, even in the absence of market power. We have seen that the results of the zonal market can be found by solving the modified zonal economic dispatch problem given by model (10), where the marginal cost of participants has been replaced by the expected re-dispatch price at their location, which corresponds to their opportunity cost. In this case, their opportunity cost can be interpreted as the price bid that rational price-taking agents would submit in the market, as we argue in Footnote 3. In practice, however, such a deviation from truthful bidding could catch the attention of the regulator, which would have the opportunity to intervene to prevent such a behavior. Closer market monitoring is indeed sometimes suggested as a solution for improved efficiency in this context (Cramton 2022).

Our interest in the present section is to integrate market monitoring in our model of zonal pricing with MBR. For this purpose, we assume that market participants submit a bid in the zonal market that corresponds to a weighted average of their marginal cost and the anticipated re-dispatch price, i.e. $\alpha\tilde{\rho}_n + (1-\alpha)MC_i$ with $\alpha \in [0,1]$. The case with $\alpha = 1$ represents the situation discussed so far in the paper, in which there is no monitoring. Market participants are free to deviate completely from bidding at marginal cost, and they therefore bid at their opportunity cost, which is the anticipated re-dispatch price. On the contrary, the case of perfect market monitoring corresponds to $\alpha = 0$. That is, market participants cannot deviate from bidding at marginal cost without catching the attention of the regulator and being sanctioned. This is of course an idealized situation that cannot be realistically implemented in practice, as regulators do not have perfect knowledge of market participants' costs or the ability to perfectly enforce such measures. This situation corresponds to the case of *re-dispatch markets without anticipation*, as discussed in Hirth and Schlecht (2020).

Following the developments in section 2, we know that the MLCP (7), corresponding to the short-term equilibrium of zonal pricing followed by MBR, can equivalently be written, by means

of a change of variables, as:

$$
\begin{aligned}
&(8a) - (8e) \\
&0 \leq y_{in} \perp \tilde{\rho}_n + \mu_{in} - \rho_{Z(n)} \geq 0 && (19a)\\
&0 \leq \mu_{in} \perp X_{in} - y_{in} \geq 0 && (19b)\\
&p_z \text{ free } \perp \rho_z + \sum_m V_{mz}\gamma_m = 0 && (19c)\\
&\rho_z \text{ free } \perp -p_z + \sum_{i,n\in N(z)} y_{in} - D_z = 0 && (19d)\\
&0 \leq \gamma_m \perp W_m - \sum_z V_{mz}p_z \geq 0 && (19e)
\end{aligned}
$$

where equations (8a) - (8e) correspond to the complementarity conditions of the nodal economic dispatch, that define the re-dispatch price $\tilde{\rho}_n$. In order to introduce the modified bidding from market participants in the zonal market due to market monitoring from the regulator, it suffices to modify equation (19a) by

$$0 \leq y_{in} \perp \alpha\tilde{\rho}_n + (1-\alpha)MC_i + \mu_{in} - \rho_{Z(n)} \geq 0 \tag{20}$$

Interestingly, letting $\alpha = 0$ restores the block skew-symmetric property of the MLCP matrix, which implies that, in this case, the problem can be solved as a single optimization problem. That is, the long-term equilibrium of zonal pricing with market-based re-dispatch without anticipation is the solution to the following modified capacity expansion problem:

$$
\begin{aligned}
\min \quad & (16a) + \sum_{int} MC_i y_{int} \\
\text{s.t.} \quad & (16b) - (16d)
\end{aligned}
\tag{21}
$$

Note that problem (21) is not equivalent to the nodal capacity expansion problem, due to the additional term in the objective, which leads to double counting the zonal production cost. MBR without anticipation is thus in general associated to a loss of efficiency in the long term, as we also verify empirically in the case study, presented in the next section.

6 Case Study: Central Western Europe

In this section, we present simulation results for a reduced instance of the Central Western European network area.

6.1 Data

We use the same dataset as in Lété et al. (2022). The area consists of 6 countries (Austria, Belgium, Germany, France, Luxembourg and the Netherlands) grouped into 5 bidding zones, Germany and Luxembourg forming one single zone. Our network is a reduced version of the CWE transmission grid, originally based on the European grid model of Hutcheon and Bialek (2013). Our time series data (hourly demand, solar and wind production in each country) are obtained from the ENSTO-E Transparency Platform for the year 2018. The generation data is obtained from Open Power System Data (2020). Our dataset includes a total of 892 non-renewable units. The models that we use for the case study are generalized versions of the models presented in sections 2 and 3. In the models of the CWE case study, we also consider revenues from reserve provision, where reserve is assumed to be cleared simultaneously with energy. We also consider fixed operating and maintenance costs. Units that cannot cover their fixed costs are decommissioned. We assume that investment is possible in 3 different technologies, similarly to Ambrosius et al. (2020): CCGT units, OCGT units and Combined Heat and Power CCGT units. Our cost data is sourced from Ambrosius et al. (2020). Wind and solar expansion are accounted for in an exogenous way. Based on the exogenous renewable production, we build a net load duration curve, that we discretize into 20 time periods, in order to obtain a tractable model. The capacity expansion model is then run on a single representative year with 20 types of hours. Load, renewable capacity and existing generators are calibrated for their expected value on future year 2035. The reader is referred to Lété et al. (2022) for more details on the dataset, including specific values for the different costs used in the case study.

6.2 Algorithmic remarks

As we mention in section 3, the long-term equilibrium of zonal pricing with MBR corresponds to a GNE. The GNE property originates in the imperfect nature of the market as the impacts of the production decisions in the zonal market on the feasibility set of the TSO are not internalized by the producers. From a mathematical point of view, this results in the breaking of the equivalence between the MLCP that aggregates the KKT optimality conditions at equilibrium of each type of agent and the LP that maximizes welfare. However, although the equivalence does not hold, this reasoning suggests a way to leverage this structure algorithmically by resorting to a splitting algorithm. We describe below a basic splitting algorithm for solving the $LCP(q, M)$ that corresponds to our MCP (Cottle et al. 2009):

Step 0. *Initialization.* Let z_0 be an arbitrary non-negative vector, set $\nu = 0$.

Step 1. *General iteration.* Given $z^\nu \geq 0$, solve the $LCP(q^\nu, B)$ where

$$q^\nu = q + Cz^\nu$$

and let $z^{\nu+1}$ be an arbitrary solution.

Step 2. *Test for termination.* If $z^{\nu+1}$ satisfies a prescribed stopping rule, terminate. Otherwise, return to Step 1 with ν replaced by $\nu + 1$.

In our case, a natural choice for matrix B is that associated to the perfect version of the market, i.e. to the welfare-maximizing LP. The C matrix therefore captures the deviation between the Jacobian of the MLCP and the matrix of the welfare maximizing problem. For this reason, the term Cz^ν has the same analytical formulation as the Pigouvian tax that would restore efficiency, that we discussed in section 4.1. In our case, it corresponds to the lack of impact of the re-dispatch price on the zonal production decisions. The algorithm has thus the effect of adding $\tilde{\rho}_n^* y_{int}$ to the remuneration of the producers, where $\tilde{\rho}_n^*$ is the value of the re-dispatch price at the previous iteration, similarly to the Pigouvian taxation modification, that adds $\tilde{\rho}_n y_{int}$ to the remuneration of the producers.

Policy	Op. cost	Inv. cost [M€/yr]	Total cost	Efficiency losses [%]
Nodal	15,810	10,433	26,243	-
Cost-based re-dispatch	16,835	10,909	27,744	5.4
MBR with $\alpha = 0.0$	15,646	10,745	26,391	0.56
MBR with $\alpha = 0.2$	15,679	10,856	26,535	1.1
MBR with $\alpha = 0.4$	15,913	11,661	27,574	4.8
MBR with $\alpha = 1$	19,057	15,867	34,924	24.9
MBR with capacity market	15,810	10,433	26,243	0.0

Table 5: Performance comparison of the different policies.

This is the methodology that we use for solving for the GNE corresponding to the long-term MLCP. We note however that we have no guarantee of convergence for this algorithm on our problem and that we observe empirically that convergence is sensitive to the starting point.

6.3 Analysis of the results

We benchmark the results of zonal pricing with MBR against two other policies: nodal pricing and zonal pricing with cost-based re-dispatch, which ensures that the re-dispatch phase is revenue-neutral for the producers. This policy also leads to inefficiencies in the long term due to the incorrect investment signal that zonal prices represent, as discussed in further detail in Lété et al. (2022). Table 5 presents the performance in terms of total operational and investment costs of the different policies. We observe that, in the long run, zonal pricing followed by MBR is significantly more costly than the nodal pricing benchmark[7]. When $\alpha = 40\%$, we arrive at a loss of efficiency compared to nodal pricing of 4.8% of the total cost. This rise in the total cost is due almost exclusively to the investment cost, which is to be expected in light of the analysis of section 3. Indeed, the arbitrage between the zonal electricity market and nodal re-dispatch market allows producers to extract a rent that translates to significantly more investment in the long run that

[7]Note that caution must be exercised when interpreting the efficiency results, in view of the potential non-uniqueness of equilibria that is discussed in Appendix C. The results presented here relate to the equilibrium found by the splitting algorithm, but we cannot exclude the existence of other, perhaps more efficient, equilibria.

does not improve the operational cost. The operational costs between the nodal policy and the MBR policy, on the contrary, are comparable.

Still in the case with $\alpha = 40\%$, we observe that zonal pricing with MBR and cost-based re-dispatch have similar efficiency, but that they differ in the distribution of the total costs. With cost-based re-dispatch, most of the losses result from operational efficiency, which is significantly degraded.

According to our results, market monitoring has a very significant influence on the efficiency of zonal pricing with MBR. In the theoretical case of perfect market monitoring, i.e. in the case of no anticipation of the re-dispatch by market participants, the efficiency of nodal pricing is almost recovered. The efficiency gap remains small with a close monitoring corresponding to $\alpha = 0.2$, evaluated at around 1.1% in our results. Then, the loss in efficiency is significantly deteriorated when α reaches 0.4. The case $\alpha = 1$, although not very realistic in view of the large deviation between implicitly submitted bids by market participants and marginal cost, gives an interesting upper bound on how much things can go wrong. This result also confirms the super-linear trend already observed for the cases $\alpha = [0, 0.2, 0.4]$. We have added a detailed comparison of the different market designs on the illustrative example in Appendix D, to which we refer the reader for a more detailed discussion of the mechanics behind the efficiency differences among the different capacity allocation methods.

Note that these results have been obtained under the assumption of inelastic demand. As we note in section 2.1, in practice, consumers could also engage in price arbitrage between the zonal and re-dispatch market. Empirically, the results that we obtain here are thus conservative, and active consumers engaging in price arbitrage can only make things worse. This reinforces our conclusion regarding the inefficiency of the design.

Figure 2 presents the optimal investment in the three candidate technologies for the five CWE bidding zones. One observes that the main difference between the nodal pricing and MBR policies is that zonal pricing with MBR tends to over-invest, especially in technologies with lower investment cost like CCGT and OCGT. This is particularly visible in the case with $\alpha = 0.4$, where zonal-nodal

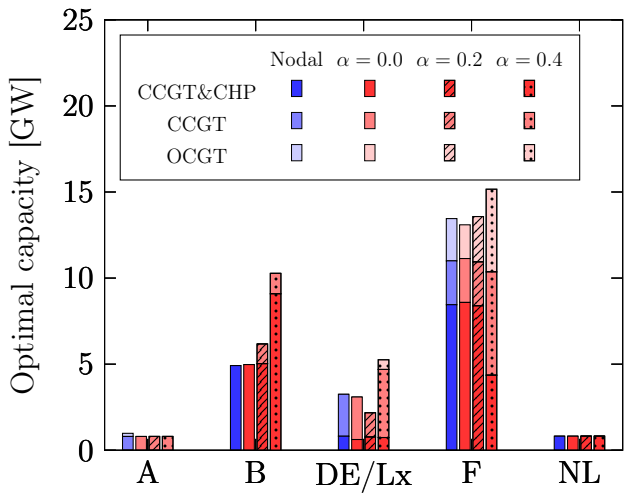

Figure 2: Optimal capacity in the long-term equilibrium in nodal pricing and zonal pricing with MBR, decomposed by bidding zone.

arbitrage is more pronounced. This confirms the results obtained with the illustrative example.

Although their total costs are similar with intermediate monitoring, the MBR policy has a notable advantage over zonal pricing with cost-based re-dispatch: its efficiency can be restored when it is combined with a locational capacity market, as discussed theoretically in section 4.2 and confirmed here empirically. In this section, we are also interested in analyzing the conditions under which the efficiency of zonal pricing with MBR could be restored. Figure 3 presents the magnitude of the optimal capacity price under this policy for each node of the network.

We observe large differences between the different values of the optimal capacity price. In particular, although the large majority of the charges are below 50k€/MW, one node at the border between Belgium and France receives a value of more than 150k€/MW. This large value can be explained by the fact that it is only connected to nodes in France while it belongs to the Belgian bidding zone. For this reason, it exhibits a large difference between its zonal and nodal price which makes it particularly prone to arbitrage.

To confirm this intuition, we plot, in Figure 4, the difference between average zonal and nodal prices as a function of the optimal capacity price. Clearly, there is a high correlation between the difference of average nodal and zonal prices in a location and the value of the optimal capacity price needed to restore the efficiency of MBR, which suggests that this difference is the main driver

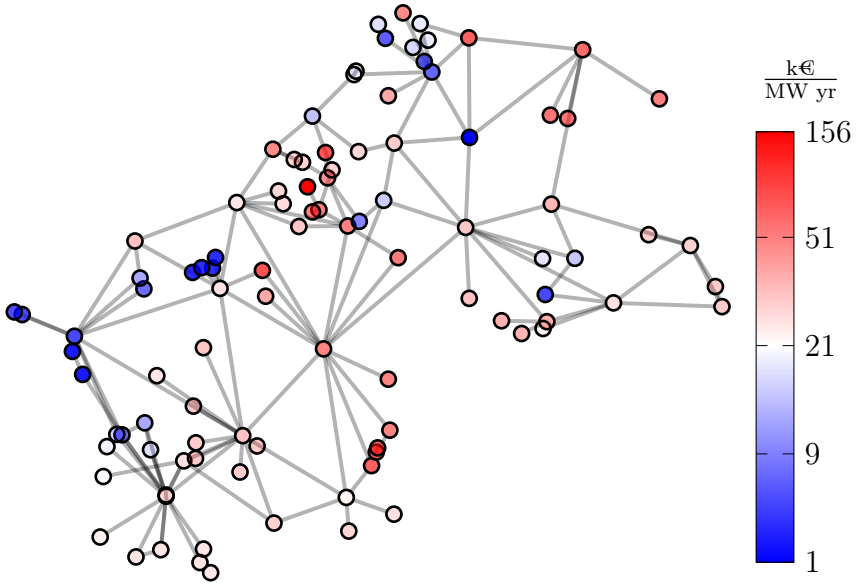

Figure 3: Spatial distribution of the optimal capacity price.

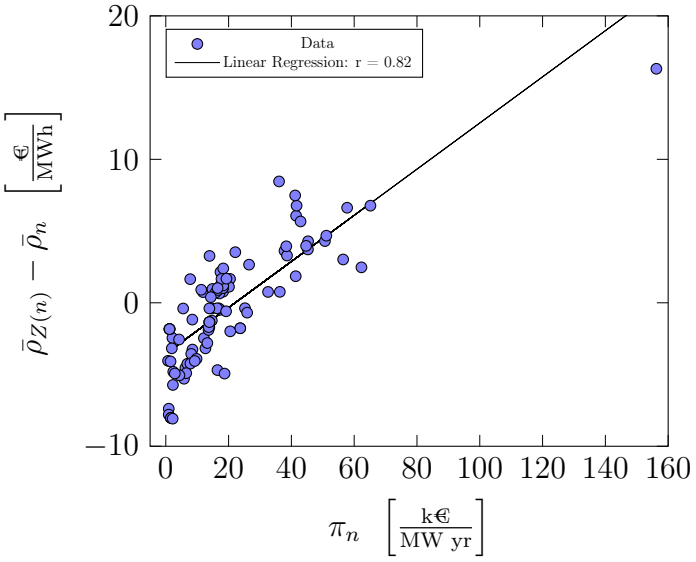

Figure 4: Difference between the average zonal and nodal prices as a function of the optimal capacity price.

for the inefficiency.

7 Conclusion

As the integration of renewable energy and distributed energy resources increases uncertainty and variability in the electricity grid, it becomes more and more important to evolve towards a pricing system that is consistent with real-time dispatch. For different reasons driven by institutional concerns, the European market has been organized around the concept of zonal pricing. Instead, locational marginal pricing, which proved efficient in various markets around the world, has so far been dismissed. In this context, it appears tempting for EU stakeholders and legislators to rely increasingly on market-based re-dispatch, which is seen as being able to reconcile a pricing system that supports real-time dispatch with the zonal organization of the market. This is for instance witnessed in the recast of the electricity regulation which makes market-based re-dispatch the default rule (European Parliament, Council of the European Union 2019). It is well known that zonal pricing followed by MBR incentivizes market participants to deviate from bidding at marginal cost, even in the absence of market power. Existing literature has studied the problem in the short term (De Vries and Hakvoort 2002, Holmberg and Lazarczyk 2015, Hirth and Schlecht 2020). In this paper, we prove formally that under a series of strong assumptions (perfect foresight and absence of irrevocable decisions based on the zonal market), zonal pricing with MBR is efficient in the short term. These assumptions, however, are known not to hold in practice. Indeed, many resources must be committed in the day-ahead stage, based on the results of the zonal market, which leads to inefficiencies in the short run (Aravena et al. 2021).

In the long run, new types of inefficiencies manifest, even under the series of strong assumptions mentioned above. In times of energy transition, it is therefore crucial to understand how MBR impacts the long-term equilibrium, both theoretically and quantitatively.

In this paper, we propose a long-term equilibrium model of zonal pricing followed by MBR. We show how inefficiencies in the long run originate from the mathematical structure of the problem, with investment conditions coupling the zonal and nodal parts of the problem. We then discuss the

conditions under which efficiency can be restored under this design as well as the impact of market monitoring on efficiency. Finally, we propose a case study on a large-scale instance representing Central Western Europe and compare MBR with nodal and zonal pricing benchmarks.

We find that zonal pricing followed by MBR is associated with significant efficiency losses, due to the fact that it incentivizes over-investment in technologies with low investment cost. This confirms the intuition discussed in Hirth and Schlecht (2020), that predicts investment in so-called "ghost plants" under this design. Our results also confirm that efficiency can be restored by introducing an additional locational capacity market, as proven theoretically in the paper. We find, however, that recovering efficiency is only possible under large deviations in capacity charges among locations, hindering the practicability of an efficient implementation. Finally, the case study highlights that market monitoring has an important impact on the results. Although efficiency can be kept close to that of nodal pricing under close monitoring, it quickly deteriorates when monitoring is relaxed.

The present work has focused on the efficiency of MBR in the best case, and has omitted the impact of uncertainty. In particular, we have assumed that market participants are able to perfectly predict re-dispatch when participating in the day-ahead market. Based on our modeling framework, it would be possible to relax this assumption by considering uncertainty in the re-dispatch price. This would make for an interesting extension to the present work, which would allow us to account for risk considerations. Moreover, an important argument for evolving towards MBR relates to the difficulty of evaluating correct compensation for re-dispatched resources in cost-based re-dispatch in a system with significant penetration of demand response and storage. Future work will build on the modeling framework presented in this paper in order to assess the impacts of MBR in systems with an increasing deployment of these types of resources.

References

Abada I, de Maere d'Aertrycke G, Smeers Y (2017) On the multiplicity of solutions in generation capacity investment models with incomplete markets: a risk-averse stochastic equilibrium approach. Math. Program. 165, 5–69 (2017). URL https://doi.org/10.1007/s10107-017-1185-9.

ACER (2022) Final assessment of the eu wholesale electricity market design. Technical report, URL https://documents.acer.europa.eu/Official_documents/Acts_of_the_Agency/Publication/ACER's%20Final%20Assessment%20of%20the%20EU%20Wholesale%20Electricity%20Market%20Design.pdf.

Ambrosius M, Grimm V, Kleinert T, Liers F, Schmidt M, Zöttl G (2020) Endogenous price zones and investment incentives in electricity markets: An application of multilevel optimization with graph partitioning. *Energy Economics* 92:104879, ISSN 0140-9883, URL https://doi.org/10.1016/j.eneco.2020.104879.

Aravena I, Lété Q, Smeers Y, Papavasiliou A (2021) Transmission Capacity Allocation in Zonal Electricity Markets. *Operations Research* 69(4):1240-1255. URL https://doi.org/10.1287/opre.2020.2082.

Bidard C (2012) The Ricardo-Lemke parametric algorithm on oddity and uniqueness. Working paper, URL https://hal.science/hal-04141043.

Boiteux M, Stasi P (1952) Sur la détermination des prix de revient de dévelopment dans un système interconnecté de production-distribution. Reprinted in G. Morlat et F. Bessière (eds.), *Vingt-cinq ans d'économie électrique*, Paris Dunod, 361-400.

Cottle R, Pang JS, Stone R (2009) *The Linear Complementarity Problem* (SIAM), URL http://dx.doi.org/10.1137/1.9780898719000.

Cramton P (2022) Local Flexibility Market. ECONtribute Discussion Papers Series 182, University of Bonn and University of Cologne, Germany, URL https://ideas.repec.org/p/ajk/ajkdps/182.html.

De Vries LJ, Hakvoort RA (2002) An Economic Assessment of Congestion Management Methods for Electricity Transmission Networks. *Competition and Regulation in Network Industries* 3(4):425–466, ISSN 1783-5917, URL http://dx.doi.org/10.1177/178359170200300403.

Eicke A (2021) Optimal generation investment under suboptimal dispatch. Working paper, Internationale Energiewirtschaftstagung an der TU Wien, URL https://iewt2021.eeg.tuwien.ac.at/download/contribution/fullpaper/176/176_fullpaper_20210831_082627.pdf.

Eicke A, Khanna T, Hirth L (2020) Locational investment signals: How to steer the siting of new generation capacity in power systems? *The Energy Journal* 41(1):281–304, ISSN 01956574, URL http://dx.doi.org/10.5547/01956574.41.6.aeic.

ENTSO-E (2020) Mid-term Adequacy Forecast Executive Summary 2020 Edition. URL

https://eepublicdownloads.entsoe.eu/clean-documents/sdc-documents/MAF/2020/MAF_2020_Executive_Summary.pdf.

European Commission (2015) Commission Regulation (EU) 2015/1222 of 24 July 2015 establishing a guideline on capacity allocation and congestion management.

European Parliament, Council of the European Union (2019) Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity.

Frontier Economics and Entso-e (2021) Review of flexibility platforms, URL https://eepublicdownloads.azureedge.net/clean-documents/SOC%20documents/SOC%20Reports/210957_entso-e_report_neutral_design_flexibility_platforms_04.pdf.

Grimm V, Martin A, Sölch C, Weibelzahl M, Zöttl G (2018) Market-Based Redispatch May Result in Inefficient Dispatch. *The Energy Journal* 43(5), ISSN 19449089, URL http://dx.doi.org/10.5547/01956574.43.5.csol.

Grimm V, Rückel B, Sölch C, Zöttl G (2019) Regionally differentiated network fees to affect incentives for generation investment. *Energy* 177:487–502, ISSN 03605442, URL http://dx.doi.org/10.1016/j.energy.2019.04.035.

Hackett S, Ahoniemi H, Goldstein H, Døvle E (2021) *Market design options for procurement of flexibility* (Nordic Energy Research), URL http://dx.doi.org/10.6027/ner2021-04.

Harvey S, Hogan W (2022) Locational marginal prices and electricity markets URL https://lmpmarketdesign.com/papers/locational_marginal_prices_and_electricity_markets_hogan_and_harvey_paper_101722.pdf.

Hirth L, Schlecht I (2020) Market-Based Redispatch in Zonal Electricity Markets: The Preconditions for and Consequence of Inc-Dec Gaming. Working paper, ZBW – Leibniz Information Centre for Economics, Kiel, Hamburg, URL http://hdl.handle.net/10419/222925.

Hobbs BF, Oren SS (2019) Three waves of u.s. reforms: Following the path of wholesale electricity market restructuring. *IEEE Power and Energy Magazine* 17(1):73–81, URL http://dx.doi.org/10.1109/MPE.2018.2873952.

Holmberg P, Lazarczyk E (2015) Comparison of congestion management techniques: Nodal, zonal and discriminatory pricing. *The Energy Journal* 36(2):145–166, ISSN 01956574, URL http://dx.doi.org/10.5547/01956574.36.2.7.

Hutcheon N, Bialek JW (2013) Updated and validated power flow model of the main continental european transmission network. *2013 IEEE Grenoble Conference*, 1–5.

Lété Q, Smeers Y, Papavasiliou A (2022) An analysis of zonal electricity pricing from a long-term perspective. *Energy Economics* 107:105853, ISSN 0140-9883, URL `http://dx.doi.org/https://doi.org/10.1016/j.eneco.2022.105853`.

Open Power System Data (2020) Data package conventional power plants. URL `http://dx.doi.org/https://doi.org/10.25832/conventional_power_plants/2020-10-01`, version 2020-10-01.

Pigou AC (1932) The Economics of Welfare. London: Macmillan.

Schmidt L, Zinke J (2020) One price fits all? Wind power expansion under uniform and nodal pricing in Germany, URL `http://hdl.handle.net/10419/227510`, EWI Working Paper, No. 20/06.

Stoft S (2002) Power system economics. Wiley-Interscience, Hoboken.

Triolo R, Wolak F (2021) Quantifying the benefits of nodal market design in the texas electricity market. Working paper, Program on Energy and Sustainable Development, URL `https://fsi.stanford.edu/publication/quantifying-bene-ts-nodal-market-design-texas-electricity-market`.

Appendix A Proofs of the Propositions

Proposition 1. Zonal pricing followed by MBR, as formulated in MLCP (7), is efficient in the short run.

Proof of Proposition 1. The result follows from the observation that, in any solution of MLCP (7), the final physical dispatch $\bar{y}_{in}$ must solve the nodal economic dispatch problem (9), as we discuss in section 2.

□

Proposition 2. If the marginal costs, the investment costs and the demand in all nodes are positive, and if the set $\mathcal{P}$ is such that $W_m \geq 0 \;\; \forall m \in M$, the MLCP that represents the long-term equilibrium has a solution.

Proof of Proposition 2. We first note that, if $MC, IC, D > 0$, then the re-dispatch prices $\tilde{\rho}_{nt}$ are always non-negative and the balance constraints can be transformed into inequalities without affecting the solution. Indeed, assume by contradiction that $\exists n, t$ s.t. $\tilde{\rho}_{nt} < 0$. Then, by equation (7d), we get that $\delta_{in} > 0 \;\; \forall i \in I$, which implies by (7e) that $y_{int} + \tilde{y}_{int} = 0 \;\; \forall i$. Then, combining (7j), (7m) and the fact that $\sum_n r_{nt} = 0 \;\; \forall t$, we get

$$\sum_{in} \left(y_{int} + \tilde{y}_{int} \right) = \sum_n D_{nt}, \quad \forall t \in T$$

which is a contradiction.

Equation (7j) can thus be transformed into

$$0 \leq \nu_{nt} \perp -r_{nt} + \sum_i y_{int} - D_{nt} + \tilde{r}_{nt} \geq 0$$

Now, note that any MLCP can be transformed into an equivalent LCP by duplicating all free variables into a positive and negative part, and duplicating all equality constraints into two inequality constraints. Let us denote by M and q respectively the matrix and the vector of independent terms associated to the LCP derived in such way from our MLCP. By Theorem 3.8.6 of Cottle

et al. (2009), if M is copositive and if, for all solutions v^* of the homogeneous LCP, it holds that $q^\top v^* \geq 0$, then there exists a solution to $LCP(q, M)$. We use this theorem to prove existence.

Let us first show that M is copositive. To do this, we note that M is the sum of two matrices:

$$M = \tilde{M} + \tilde{N}$$

where $\tilde{M}$ is skew-symmetric (it is the matrix associated to the equivalent LCP of the centralized problem, i.e. the welfare-maximizing problem) and where $\tilde{N}$ is of the following form (in block formulation):

$$\begin{array}{cc} & \nu_{nt} \end{array} \tag{22}$$

$$\tilde{N} = \begin{array}{c} y_{int} \\ \\ \\ \\ \end{array} \begin{pmatrix} 0 & \dots & I & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix} \tag{23}$$

Here, I is the rectangular identity matrix, i.e. a matrix with 1 in the entries associated to line y_{int} and column ν_{nt}, and 0 otherwise. This implies that

$$v^\top M v = v^\top \tilde{M} v + v^\top \tilde{N} v = v^\top \tilde{N} v = \sum_{int} y_{int} \nu_{nt}$$

where v is the full vector of variables of the LCP. This expression is indeed non-negative if each y_{int} and ν_{nt} are non-negative.

Now, let $v^* = (x^*_{in}, y^*_{int}, \dots, \rho^{*,+}_{zt}, \rho^{*,-}_{zt})$ be a solution to the homogeneous version of the LCP, where the free variables such as ρ^*_{zt} have been transformed into $\rho^{*,+}_{zt} - \rho^{*,-}_{zt}$ to obtain the LCP from

the MLCP. We have

$$
\begin{aligned}
q^\top v^* = & \sum_{in} IC_{in} x^*_{in} + \sum_{int} MC_i y^*_{int} + \sum_{int} X_{in} \mu^*_{int} + \sum_{int} X_{in} \tilde{\mu}^*_{int} + \\
& \sum_{mt} W_m \gamma^*_{mt} + \sum_{mt} \tilde{W}_m \tilde{\gamma}^*_{mt} + \\
& \sum_{int} MC_i \tilde{y}^*_{int} - \sum_{nt} D_{nt} (\rho^{*,+}_{zt} - \rho^{*,-}_{zt} + \nu^*_{nt})
\end{aligned}
$$

All the terms of the first and second line are non-negative, as they correspond to the product of non-negative quantities. From equation (7l) and the fact that $\sum_z p_{zt} = 0$, we deduce that $y^*_{int} = 0$. This implies from equation (7e) that $\tilde{y}^*_{int} \geq 0$ and thus that the first term of the third line is also non-negative. From equations (7e), (7h) and (7j), we get that $0 \leq \nu^*_{nt} = \tilde{\rho}^*_{nt} = -\delta^*_{nt} \leq 0$, which implies that these three quantities must be equal to 0. We then deduce from (7a) that $\rho^{*,+}_{zt} - \rho^{*,-}_{zt} \leq 0$, which, together with $\nu^*_{nt} = 0$, yields the non-negativity of the second term of the third line and concludes the proof.

□

Proposition 3. If zonal pricing followed by MBR is augmented with a locational capacity market that determines a non-negative capacity charge, uniform for all technologies of the given location, then it is efficient, i.e. it leads to the same welfare as nodal pricing in the long run.

Proof of Proposition 3. First, note that if the capacity price is forced to be positive and uniform among technologies, then the investment condition and the capacity market clearing condition become respectively as follows:

$$0 \leq x_{in} \perp IC_i - \sum_{t \in T} \mu_{int} - \sum_{t \in T} \tilde{\mu}_{int} + \pi_n \geq 0 \tag{24}$$

$$0 \leq \pi_n \perp \sum_{i \in I} \bar{x}_{in} - \sum_{i \in I} x_{in} \geq 0 \tag{25}$$

We need to show that all solutions of MLCP (24), (25), (7a) - (7m) are as efficient as the nodal solution. Let $(\hat{\cdot})$ be an arbitrary solution of the MLCP and let $(\bar{\cdot})$ be a solution of the nodal

capacity expansion problem. Assume by contradiction that

$$\sum_{in} IC_i \hat{x}_{in} + \sum_{int} MC_i(\hat{y}_{int} + \hat{\tilde{y}}_{int}) > \sum_{in} IC_i \bar{x}_{in} + \sum_{int} MC_i \bar{y}_{int} \tag{26}$$

We start by showing that, in this case, $\sum_{i \in I} \hat{x}_{in} = \sum_{i \in I} \bar{x}_{in}$. We will show this by contradiction, using thus a second embedded contradiction assumption: assume, by contradiction, that the set of nodes for which the total nodal capacity differs between the solution is non-empty, and denote this set by Q. That is, Q is the subset of N of maximum cardinality which is such that

$$\sum_{i \in I} \hat{x}_{in} < \sum_{i \in I} \bar{x}_{in} \ \ \forall n \in Q \tag{27}$$

Let us denote by ν_n the arbitrage rent in this solution, i.e. $\nu_n = \sum_{i \in I} \hat{\mu}_{int}$. Now, observe that the corresponding variables of the solution $(\hat{\cdot})$ must also be a solution of the following optimization problem:

$$\begin{aligned}
\min \ & \sum_{in} (IC_{in} - \nu_n + \hat{\pi}_n) x_{in} + \sum_{int} MC_i (y_{int} + \tilde{y}_{int}) \\
\text{s.t. } & X_{in} + x_{in} - y_{int} - \tilde{y}_{int} \geq 0, i \in I, n \in N, t \in T \ \ [\tilde{\mu}_{int}] \\
& r_{nt} - \sum_{int} (y_{int} + \tilde{y}_{int}) + D_{nt} = 0, n \in N, t \in T \ \ [\tilde{\rho}_{nt}] \\
& \tilde{W}_m - \sum_n \tilde{V}_{mn} r_{nt}, m \in \tilde{M}, t \in T \ \ [\tilde{\gamma}_{mt}] \\
& y_{int} + \tilde{y}_{int} \geq 0, i \in I, n \in N, t \in T \ \ [\delta_{int}]
\end{aligned} \tag{28}$$

By definition, the solution $(\bar{\cdot})$ is the solution of the nodal capacity expansion problem, i.e. it is a solution of:

$$\begin{aligned}
\min \ & \sum_{in} IC_{in} x_{in} + \sum_{int} MC_i y_{int} \\
\text{s.t. } & X_{in} + x_{in} - y_{int} \geq 0, i \in I, n \in N, t \in T \ \ [\tilde{\mu}_{int}] \\
& r_{nt} - \sum_{int} y_{int} + D_{nt} = 0, n \in N, t \in T \ \ [\tilde{\rho}_{nt}] \\
& \tilde{W}_m - \sum_n \tilde{V}_{mn} r_{nt}, m \in \tilde{M}, t \in T \ \ [\tilde{\gamma}_{mt}] \\
& y_{int} \geq 0, i \in I, n \in N, t \in T \ \ [\delta_{int}]
\end{aligned} \tag{29}$$

We deduce, as the feasible sets of both problems are the same, that the following should hold:

$$\begin{aligned}&\sum_{in}(IC_i - \nu_n + \hat{\pi}_n)\bar{x}_{in} + \sum_{int} MC_i\bar{y}_{int} \geq \\ &\sum_{in}(IC_i - \nu_n + \hat{\pi}_n)\hat{x}_{in} + \sum_{int} MC_i(\hat{y}_{int} + \hat{\bar{y}}_{int})\end{aligned} \tag{30}$$

By reorganizing the terms, we get

$$\begin{aligned}&\sum_{in} IC_i\bar{x}_{in} + \sum_{int} MC_i\bar{y}_{int} - \sum_{in} IC_i\hat{x}_{in} - \sum_{int} MC_i(\hat{y}_{int} + \hat{\bar{y}}_{int}) \geq \\ &\sum_{n\in Q}(-\nu_n + \hat{\pi}_n)\big(\sum_{i\in I}(\bar{x}_{in} - \hat{x}_{in})\big) + \sum_{n\in N\setminus Q}(-\nu_n + \hat{\pi}_n)\big(\sum_{i\in I}(\bar{x}_{in} - \hat{x}_{in})\big)\end{aligned} \tag{31}$$

Note that, as Q is the set of maximum cardinality among all for which (27) holds, we have that

$$\sum_{n\in N\setminus Q}\sum_{i\in I}(\bar{x}_{in} - \hat{x}_{in}) = 0$$

Moreover, for all $n \in Q$, we have $\hat{\pi}_n = 0$ by the complementarity condition (25). Therefore, equation (31) simplifies to

$$\begin{aligned}&\sum_{in} IC_i\bar{x}_{in} + \sum_{int} MC_i\bar{y}_{int} - \sum_{in} IC_i\hat{x}_{in} - \sum_{int} MC_i(\hat{y}_{int} + \hat{\bar{y}}_{int}) \geq \\ &\sum_{n\in Q}(-\nu_n)\big(\sum_{i\in I}(\hat{x}_{in} - \bar{x}_{in})\big)\end{aligned} \tag{32}$$

The left-hand side in (32) is strictly negative by our first contradiction assumption, whereas the right-hand side in (32) is non-negative, which leads to a contradiction of our second assumption. This implies that, in any solution ($\hat{\cdot}$) of the MLCP (24), (25), (7a) - (7m) such that (26) holds, $\sum_{i\in I}\hat{x}_{in} = \sum_{i\in I}\bar{x}_{in}$ for all $n \in N$.

Now, by the optimality of $(\hat{\cdot})$ for the perturbed problem (i.e. problem (28)), we have

$$\sum_{in} IC_i \hat{x}_{in} + \sum_{n \in N} (-\nu_n + \hat{\pi}_n) \sum_{i \in I} \hat{x}_{in} + \sum_{int} MC_i (\hat{y}_{int} + \hat{\tilde{y}}_{int}) \leq$$
$$\sum_{in} IC_i \bar{x}_{in} + \sum_{n \in N} (-\nu_n + \hat{\pi}_n) \sum_{i \in I} \bar{x}_{in} + \sum_{int} MC_i \bar{y}_{int}$$

which implies

$$\sum_{in} IC_i \hat{x}_{in} + \sum_{int} MC_i (\hat{y}_{int} + \hat{\tilde{y}}_{int}) \leq \sum_{in} IC_i \bar{x}_{in} + \sum_{int} MC_i \bar{y}_{int}$$

which is a contradiction to our first contradiction assumption, and, therefore concludes the proof.

□

Appendix B Extensions of the model to other types of market participants

The goal of this Appendix is to extend the model to account for other types of market participants, namely consumers, renewable resources and storage. For each of these additional types of market participants, we show that the fundamental results of the paper remain true.

Consumers. The consumers of type $j \in J$ located at node $n \in N$ with willingness to pay V_j and electricity demand D_{jn} maximize their utility for consuming while buying electricity at price $\rho_{Z(n)}$ in the zonal market and/or at price $\tilde{\rho}_n$ in the re-dispatch market. Their surplus maximization problem can be written as follows:

$$\begin{aligned}
\max_{w_{jn}, \tilde{w}_{jn}} \quad & V_j(w_{jn} + \tilde{w}_{jn}) - \rho_{Z(n)} w_{jn} - \tilde{\rho}_n \tilde{w}_{jn} \\
\text{s.t.} \quad & D_{jn} - w_{jn} \geq 0 \\
& D_{jn} - w_{jn} - \tilde{w}_{jn} \geq 0 \\
& w_{jn} + \tilde{w}_{jn} \geq 0 \\
& w_{jn} \geq 0
\end{aligned}$$

In this model, w_{jn} is the power consumed in the zonal market and $\tilde{w}_{jn}$ is the re-dispatch quantity of consumers of type j in node n. The KKT conditions of this consumer surplus maximization problem can be expressed as:

$$0 \leq w_{jn} \perp -V_j + \nu_{jn} + \tilde{\nu}_{jn} + \rho_{Z(n)} - \zeta_{jn} \geq 0$$
$$\tilde{w}_{jn} \text{ free } \perp -V_j + \tilde{\nu}_{jn} + \tilde{\rho}_n - \zeta_{jn} = 0$$

Similarly to the case of power production, we can combine the two equations in order to isolate the following equivalent optimality condition for consumers in the zonal market:

$$0 \leq w_{jn} \perp -\tilde{\rho}_n + \nu_{jn} + \rho_{Z(n)} \geq 0$$

This equation is exactly the optimality condition of consumption in a zonal economic dispatch problem where the valuation V_j would have been replaced by the re-dispatch price $\tilde{\rho}_n$. This highlights that the outcome of the zonal market in the setting where it would be followed by a re-dispatch market can be obtained by solving an economic dispatch problem where both the marginal cost of producers and the willingness to pay of consumers would be replaced by the anticipated re-dispatch price $\tilde{\rho}_n$ at their location.

Renewable producers. Modeling renewable energy sources is generally more complicated than conventional resources, as their available capacity is highly variable and uncertain. The perfect foresight assumption, however, greatly simplifies the analysis. It suffices to assume that, for every renewable resource, there is a known capacity factor for every time period of the year that defines the fraction of the total installed capacity that is available for production at time $t \in T$. We denote the capacity factor of renewable resource i in node n and time t by CF_{int}. For renewable resources, an equivalent assumption to the assumption of asset-backed arbitrage made in the main model of this paper is that operators can only bid up to the forecast capacity in the zonal and re-dispatch market, and not up to their nameplate capacity. Additionally, the marginal cost of renewable energy resources can be assumed to be zero. With this notation, the profit maximizing

problem for renewable resources can be written as:

$$
\begin{aligned}
\max_{y_{in}^{\mathrm{RE}}, \tilde{y}_{in}^{\mathrm{RE}}} \quad & \rho_{Z(n)} y_{in}^{\mathrm{RE}} + \tilde{\rho}_n \tilde{y}_{in}^{\mathrm{RE}} \\
\text{s.t.} \quad & CF_{int} X_{in} - y_{in}^{\mathrm{RE}} \geq 0 \\
& CF_{int} X_{in} - y_{in}^{\mathrm{RE}} - \tilde{y}_{in}^{\mathrm{RE}} \geq 0 \\
& y_{in}^{\mathrm{RE}} + \tilde{y}_{in}^{\mathrm{RE}} \geq 0 \\
& y_{in}^{\mathrm{RE}} \geq 0
\end{aligned}
\tag{33}
$$

Seeing the similarity with the conventional producer model, it is clear that the fundamental observation regarding the equivalence to the zonal economic dispatch with marginal cost replaced by the re-dispatch price extends to the case of dispatchable renewable resources.

Storage. Storage resources can gain profit in an electricity market by performing classical intertemporal arbitrage. Under zonal pricing with market-based re-dispatch, in addition to the intertemporal arbitrage, they can also perform the zonal-nodal arbitrage that we discuss in this paper. Fundamentally, the same logic that applies to classical electricity producers also applies to storage resources, as we discuss formally in this section. Let us denote by $p_{nt}^{\mathrm{in}}, p_{nt}^{\mathrm{out}}$ respectively the charging and discharging power cleared in the zonal market of the storage resource located in node n at time t, $\tilde{p}_{nt}^{\mathrm{in}}, \tilde{p}_{nt}^{\mathrm{out}}$ the re-dispatch quantities cleared, e_{nt} the energy level in the storage measured at time t and k_n the energy capacity of the battery. Then, the profit maximizing problem of the

storage over the time horizon T can be written as:

$$\begin{aligned}
\max_{p_{nt},e_{nt},k_n} \quad & \sum_{t\in T} \Big(\rho_{Z(n)t}(p_{nt}^{\text{out}} - p_{nt}^{\text{in}}) + \tilde{\rho}_{nt}(\tilde{p}_{nt}^{\text{out}} - \tilde{p}_{nt}^{\text{in}})\Big) \\
\text{s.t. } & e_t = \sum_{i=1}^{t} \Big(\sqrt{\eta}(p_{nt}^{\text{in}} + \tilde{p}_{nt}^{\text{in}}) - \frac{1}{\sqrt{\eta}}(p_{nt}^{\text{out}} + \tilde{p}_{nt}^{\text{out}})\Big) \quad [\mu_{nt}] \\
& 0 \le e_t \le k_n \quad [\epsilon_{nt}] \\
& 0 \le p_{nt}^{\text{in}} \le k_n/\gamma \quad [\pi_{nt}^{\text{in}}] \\
& 0 \le p_{nt}^{\text{out}} \le k_n/\gamma \quad [\pi_{nt}^{\text{out}}] \\
& 0 \le p_{nt}^{\text{in}} + \tilde{p}_{nt}^{\text{in}} \le k_n/\gamma \quad [\zeta_{nt}^{\text{in}}, \tilde{\pi}_{nt}^{\text{in}}] \\
& 0 \le p_{nt}^{\text{out}} + \tilde{p}_{nt}^{\text{out}} \le k_n/\gamma \quad [\zeta_{nt}^{\text{out}}, \tilde{\pi}_{nt}^{\text{out}}]
\end{aligned} \tag{34}$$

where η is the round-trip efficiency and γ is the discharging time at full capacity. Note that, in this model, we ignore the constraint preventing charging and discharging at the same time, in order to keep the model linear.

Let us look at the KKT conditions of charging power in both the zonal and re-dispatch market:

$$0 \le p_{nt}^{\text{in}} \perp \sum_{i=1}^{t} \sqrt{\eta}\mu_{ni} + \rho_{Z(n)t} + \pi_{nt}^{\text{in}} + \tilde{\pi}_{nt}^{\text{in}} - \zeta_{nt}^{\text{in}} \ge 0 \tag{35}$$

$$\tilde{p}_{nt}^{\text{in}} \text{ free } \perp \sum_{i=1}^{t} \sqrt{\eta}\mu_{ni} + \tilde{\rho}_{nt} + \tilde{\pi}_{nt}^{\text{in}} - \zeta_{nt}^{\text{in}} = 0 \tag{36}$$

In these equations, the term $\sum_{i=1}^{t} \sqrt{\eta}\mu_{ni}$ takes the role of the marginal cost in the case of classical thermal units. It represents the temporal arbitrage rent that the storage resource can earn and will characterize the bidding behavior of a price-taking storage unit under perfect information. Now, combining the two equations, Equation (35) can be simplified as:

$$0 \le p_{nt}^{\text{in}} \perp \rho_{Z(n)t} - \tilde{\rho}_{nt} + \pi_{nt}^{\text{in}} \ge 0 \tag{37}$$

This equation is equivalent to the optimality condition for production in the zonal market for classical production units. It indicates that the unit has an incentive to bid at the anticipated re-

dispatch price in the zonal market, regardless of the opportunity cost related to temporal arbitrage. In particular, the term $\sum_{i=1}^{t} \sqrt{\eta}\mu_{ni}$ does not appear in Equation (37). In the zonal market, the storage operator only needs to consider the zonal-nodal arbitrage rent when bidding in the zonal market, and only considers the inter-temporal arbitrage rent when bidding in the re-dispatch market. Note that a similar development can be obtained for the discharging power.

Appendix C Uniqueness of solutions

Uniqueness of equilibria is a desirable property for an economic system. It increases confidence that the system will evolve towards that particular equilibrium and enables a reliable quantitative comparison between different scenarios or market designs. It is well known that in a perfect market (convexity, completeness, perfect competition, absence of externalities), the equilibrium is unique except in degenerate cases that form a set of measure zero. This follows from the fact that the equilibrium can be described as the solution of a single linear program.

When we deviate from this perfect setting, however, such as under the market design that we analyze in this paper, uniqueness cannot be guaranteed in general. We can, nevertheless, characterize the solution set in our analysis further by stating that, for almost all parameter data, the number of solutions is odd and finite. In order to show this, we would like to apply the reasoning of Bidard (2012) who shows that for certain types of LCPs this property can be inferred from the working of the parametric Lemke algorithm. The reasoning applies for the LCP(q^*, M) with the two following properties:

1. The matrix M has zero blocks on the diagonal.
2. One can define a homotopy $q(\gamma) : [0, 1] \to \mathbb{R}^n$ such that LCP$(q(0), M)$ is known and unique; $\exists \gamma \in (0, 1) : q(\gamma) = q^*$; and there are no new solutions to the LCP for $\gamma > 1$.

The former condition is typical for LCPs that correspond to an economic equilibrium and applies naturally to our problem. The latter is less obvious in our setting. One would be tempted to take the solution of the centralized problem, which is known and unique for almost all parameter data,

as the starting point and build the homotopy in such a way to deviate from the centralized to the decentralized formulation. This is a plausible strategy, and corresponds to the reasoning applied in Abada et al. (2017). However, this type of homotopy involves modifying the matrix M with parameter γ as the matrices from the centralized and decentralized formulation differ from each other; therefore, it does not fulfill the second condition stated above.

One can, however, apply a different strategy by using the result obtained in Proposition 3. We have shown that the zonal pricing equilibrium with market-based re-dispatch is equivalent to the centralized formulation when it is augmented by a capacity market with positive capacity price. This implies that there exists an optimal positive capacity price π_n^* that leads to the equivalence between the centralized and decentralized solution. One can thus define the homotopy as the LCP(q, M) that corresponds to the decentralized equilibrium, where condition (12) is replaced by:

$$0 \leq x_{in} \perp IC_{in} + (1-\gamma)\pi_n^* - \sum_{t \in T} \mu_{int} - \sum_{t \in T} \tilde{\mu}_{int} \geq 0 \tag{38}$$

When $\gamma = 0$, the LCP is equivalent to the centralized formulation and the solution is known and unique. When $\gamma = 1$, we recover the decentralized formulation. When γ becomes large enough, the problem becomes infeasible. With this homotopy defined, we are now in the conditions of Bidard (2012) and we can apply the results of the paper to conclude that the number of solutions is finite and odd for almost all parameter data (i.e., except in degenerate instances).

Bidard (2012) shows that the property is related to the parametric Lemke algorithm. This algorithm could thus be applied to discover multiple solutions or certify uniqueness. In order to further investigate the uniqueness property for our problem, we have implemented the parametric Lemke algorithm (Cottle et al. 2009) in our illustrative example that is described in section 3.2. Figure 5 presents the evolution of the total cost of the solution with parameter γ. The parametric Lemke algorithm obtains the same solution as the splitting algorithm discussed in section 3.2, which validates the implementation. The evolution is monotone and terminates at a value of around 1.11, where the problem becomes infeasible. This numerically confirms the uniqueness of the solution for the illustrative example.

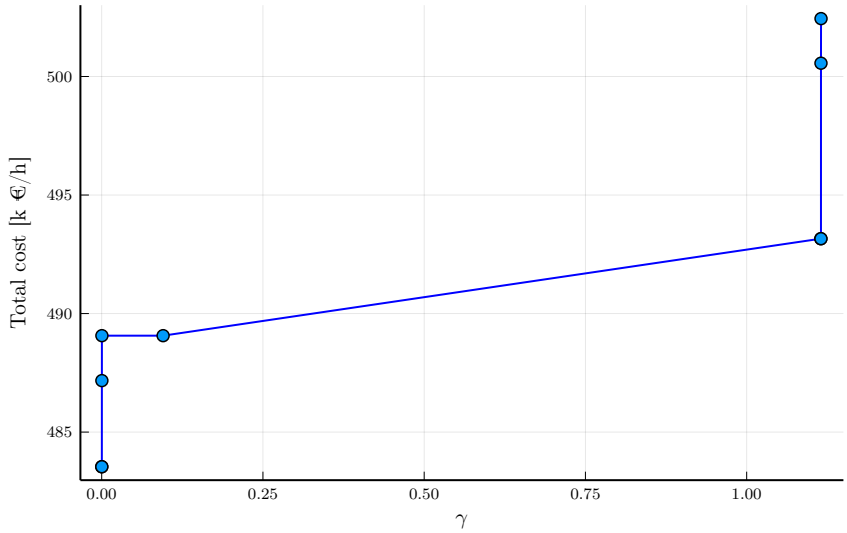

Figure 5: Evolution of the total cost of the equilibrium with the homotopy parameter γ in the implementation of the paramteric Lemke algorithm in the illustrative example.

The application of the parametric Lemke algorithm to our problem is particularly interesting in that it has an economic interpretation. It computes the value of the total cost of the equilibrium resulting from the imposition of a capacity charge that is not set at its optimal value. Often, fully pricing externalities is not possible or not desirable for different reasons, and the continuum of solutions for γ going from 0 to 1 gives a continuous of theoretically reachable equilibria. The inflection points or the potential bifurcations obtained with the parametric Lemke algorithm provides valuable information for pricing externalities when tradeoffs have to be made.

Note that the parametric Lemke algorithm is more challenging to implement than the splitting algorithm as it can be prone to numerical instabilities. Its application to the large-scale case study is thus outside the scope of this paper.

Appendix D Comparison of the capacity allocation methods on the illustrative example

In this Appendix, our goal is to compare the results of the different capacity allocation methods on the illustrative two-node one-zone example that we introduced in section 2.4. We compare the profits earned by the different units in the short term and discuss the differences in profits under

	Nodal		CBR		MBR no anticipation		MBR anticipation	
Unit	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2
Coal	16	16	9.42	9.42	10.13	11.92	16	34.5
Gas	5	4.95	0	0	4.47	0	5	23.45
Nuclear	32	32	27.92	27.92	27.92	27.92	32	50.5
Oil	0	0	0	0	0	0	0	18.5

Table 6: Marginal profit of the different units under nodal pricing [€/MWh].

the following designs:

1. Nodal pricing
2. Zonal pricing with cost-based re-dispatch (CBR)
3. Zonal pricing with MBR without anticipation
4. Zonal pricing with MBR with anticipation

Nodal pricing Nodal pricing is the first-best design in our setting. It consists of a single market with locational marginal prices. The profits of the different units, along with those under the other designs, are presented in Table 6. We observe that the marginal profits are always exactly equal to investment costs for the units that have a nonzero capacity. For the units that have zero capacity, the profit that they would earn, should they join the market, would be below their investment cost. This shows that the efficient level of investment is supported by the nodal market, which holds by design in our setting.

Zonal pricing with CBR In zonal pricing with CBR, the units are re-dispatched at cost and, therefore, only earn profits in the zonal market. The market completely omits the transmission constraint and units of a given technology earn the same profits in both nodes. As the congestion in the transmission line is not priced, electricity prices as well as profits decrease compared to the nodal pricing case. In particular, none of the units would recover their investment cost. This explains why zonal pricing tends to lead to less investment in capacity than the efficient level. This

case is analyzed in more detail in Lété et al. (2022).

Zonal pricing and MBR without anticipation In this case, there are two sequential markets, a zonal market and a re-dispatch market, but units are assumed to bid at marginal cost in the zonal market. As we discuss in the main body of the paper, this is not an optimal strategy, even in the absence of market power. It corresponds to the case with full market monitoring ($\alpha = 0$) that we introduce in section 5. In terms of profit to market participants, this case yields results that are between those of nodal pricing and zonal pricing with CBR. Indeed, in the zonal market, the producers have exactly the same profit as in the case of zonal pricing with CBR. However, they are paid at the nodal price for their re-dispatch quantities. As as result, this improves on the zonal pricing design with CBR by restoring part of the profit of producers based on nodal prices. This is the fundamental justification behind the introduction of re-dispatch markets. However, these results are only theoretical as they correspond to a suboptimal behavior of market participants, even in the absence of market power.

Zonal pricing and MBR with anticipation This is the design that we have analyzed in this paper. In terms of profits in the short term, it leads to inflated profits for the units located in Node N_2, as we discuss in section 2.4. This translates, in the long run, into inefficient investment in Oil capacity at node N_2, as we show in section 3.2.