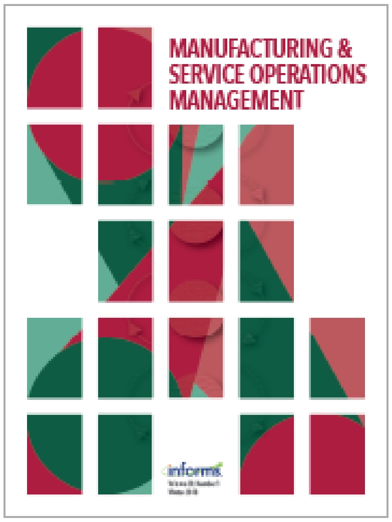


Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Optimal Electricity Imbalance Pricing for the Emerging Penetration of Renewable and Low-Cost Generators

Yashar Ghiassi-Farrokhfal, Rodrigo Belo, Mohammed Reza Hesamzadeh, Derek Bunn

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Articles in Advance, pp. 1–18
https://pubsonline.informs.org/journal/msom
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

Optimal Electricity Imbalance Pricing for the Emerging Penetration of Renewable and Low-Cost Generators

Yashar Ghiassi-Farrokhfal,[a,]* Rodrigo Belo,[b] Mohammed Reza Hesamzadeh,[c] Derek Bunn[d]

[a] Rotterdam School of Management, Erasmus University, 3062 PA Rotterdam, Netherlands; [b] Nova School of Business and Economics, Universidade NOVA de Lisboa, Campus de 2775-405 Carcavelos, Portugal; [c] KTH Royal Institute of Technology, 10044 Stockholm, Sweden; [d] London Business School, London NW1 4SA, United Kingdom

*Corresponding author

Contact: yashar@rsm.nl, https://orcid.org/0000-0001-6365-1001 (YG-F); rodrigo.belo@novasbe.pt, https://orcid.org/0000-0001-6839-5086 (RB); mrhesa@kth.se, https://orcid.org/0000-0002-9998-9773 (MRH); dbunn@london.edu, https://orcid.org/0000-0003-1770-4488 (DB)

Received: November 16, 2021
Revised: June 14, 2022; December 19, 2022; April 25, 2023
Accepted: May 13, 2023
Published Online in Articles in Advance: July 13, 2023

https://doi.org/10.1287/msom.2021.0555

Abstract. ***Problem definition*:** With the rise of renewables and the decline of fossil fuels, electricity markets are shifting toward a capacity mix in which low-cost generators (LCGs) are dominant. Within this transition, policymakers have been considering whether current market designs are still fundamentally fit for purpose. This research analyses a key aspect: the design of real-time imbalance pricing mechanisms. Currently, markets mostly use either single pricing or dual pricing as their imbalance pricing mechanisms. Single-pricing mechanisms apply identical prices for buying and selling, whereas dual-pricing mechanisms use different prices. The recent harmonization initiative in Europe sets single pricing as the default and dual pricing as the exception. This leaves open the question of when dual pricing is advantageous. We compare the economic efficiency of two dual-pricing mechanisms in current practice with that of a single-pricing design and identify conditions under which dual pricing can be beneficial. We also prove the existence of an optimal pricing mechanism. ***Methodology/results*:** We first analytically compare the economic efficiency of single-pricing and dual-pricing mechanisms. Furthermore, we formulate an optimal pricing mechanism that can deter the potential exercise of market power by LCGs. Our analytical results characterize the conditions under which a dual pricing is advantageous over a single pricing. We further compare the economic efficiency of these mechanisms with respect to our proposed optimal mechanism through simulations. We show that the proposed pricing mechanism would be the most efficient in comparison with others and discuss its practicability. ***Managerial implications*:** Our analytical comparison reveals market conditions under which each pricing mechanism is a better fit and whether there is a need for a redesign. In particular, our results suggest that existing pricing mechanisms are adequate at low/moderate market shares of LCGs but not for the high levels currently envisaged by policymakers in the transition to decarbonization, where the optimal pricing mechanism will become more attractive.

Funding: Rodrigo Belo acknowledges funding by Fundação para a Ciência e a Tecnologia (UIDB/00124/2020, UIDP/00124/2020 and Social Sciences DataLab – PINFRA/22209/2016), POR Lisboa, and POR Norte (Social Sciences DataLab, PINFRA/22209/2016).

Supplemental Material: The e-companion is available at https://doi.org/10.1287/msom.2021.0555.

Keywords: energy-related operations • environmental operations

1. Introduction

Electricity is an instantaneous commodity requiring the real-time balance of production and consumption in power networks. This is challenging, given the uncertainties of generation and consumption, the operational restrictions of generators, and network constraints. To address these challenges, trading in electricity markets consists of successive rounds of forward trading from several months, or more, to a few minutes ahead of actual delivery. The traded products in this process become progressively more granular with seasonal forwards often being simply for continuous baseload power, whereas, closer to real-time, 15-minute or even 5-minute delivery periods may be actively traded to facilitate the precise matching of demand and supply.

Within this sequence of forward trading between generators and retailers (consumers), the day-ahead market has become the most actively traded to enable production commitments and as a reference for financial hedging instruments. In the day-ahead market, market participants place their offers/bids for each hour (or half-hour) of the next day. During each period of real-time delivery (hourly or less), all market participants are likely to incur imbalance volumes to some extent, reflecting the fact that retail demand may not be exactly the same as forecast, and generators may find that their production varies from that planned. It is the aggregate of all of these participant imbalance volumes that causes the need to administer the real-time balancing to ensure the match of generation and consumption for the power network. This is settled in the so-called real-time imbalance market. Because adjustments in real time need to be administered by central control, the system operator is the counterparty for performing the real-time balancing adjustments and applying imbalance prices to market participants on their imbalance volumes. Thus, in all markets just before real-time delivery starts, there is a point called "gate closure" when the wholesale market trading between generators and retailers stops and all further transactions by either generators or retailers are with the system operator. At gate closure, the generators and retailers nominate their physical positions going into the real-time delivery periods, that is, how much they expect to produce or consume, and it is against these nominations that their actual production and consumption are compared to determine the production of the imbalance volumes. These imbalance volumes are costs to the system, and therefore, they incur imbalance prices. The revenue from these imbalance payments compensates the system operator for the costs of its real-time balancing actions. With the rise of intermittent renewable generation and greater demand-side engagement, the uncertainties in the delivery periods have increased substantially, and as a consequence, the design of the most appropriate imbalance pricing mechanism has become an important and controversial question for regulators and policymakers.

The effectiveness of a pricing strategy is greatly impacted by the composition of the supply mix and the technologies that comprise it. In electricity markets, the technological basis of market power may be changing with the energy transition. Incumbent generators with large fossil fuel facilities are seeing their coal facilities being decommissioned and replaced by large wind (or solar) farms. Thus, for example, RWE, historically one of the four dominant coal-fired generators in Germany, has faced the policy of coal eradication in Germany by developing a fleet of renewable facilities, including the world's largest offshore wind farm of 1.4 GW (www.rwe.com). In the United Kingdom, where electricity was once dominated by a duopoly of coal generators, government policy determined that the last coal power station will close before 2025, and a target of 40 GW offshore wind was set for 2030, most of which will be developed by a few major operators (HMSO 2020). In 2022, because of the geopolitical turbulence in the oil and gas supplies to Europe, The Netherlands, Denmark, Germany, and Belgium signed a deal to substantially expand the capacity of the North Sea wind farms to 65 gigawatts in 2030 and 150 gigawatts by 2050 (https://www.dutchnews.nl). Therefore, market power in the energy transition going forward may be acquired more by the low marginal cost generators (LCGs), such as renewables and nuclear, and indeed nuclear has faced accusations of market power in the past (Ofgem 2000). Furthermore, research has already started to appear on the potential for the emerging market concentration of wind and solar facilities to exercise market power; see, for example, Sunar and Birge (2019). With this technological shift in market power, we raise the question of whether the existing imbalance prices are still fit for purpose and whether there is a need for designing new pricing mechanisms. Therefore, we analyze and compare important existing imbalance pricing mechanism designs in the context of potential market power by LCGs, exercised possibly, but not exclusively, by the emerging scale of renewable generators.

Although different imbalance pricing mechanisms are used in different jurisdictions, most of them fall into two main categories known as *single-pricing mechanisms* — used, for example, in Germany, United Kingdom, and most U.S. markets — and *dual-pricing mechanisms* — used, for example, in France, Belgium, and Italy (Morales et al. 2014, EC 2016). In a single-pricing mechanism, the price applied to being out of balance in a positive or negative direction will be the same. Thus, for a generator, if it produces less (more) in real time than it nominated at gate closure, it is short (long) and will be charged (receive) the imbalance price on the imbalance volume, similarly for a retailer in the opposite direction. With the price being the same for short or long imbalances, a portfolio player can hedge quite effectively by netting its imbalance exposures across several assets and/or between its generation and retail activities. However, the single-pricing mechanism can give rise to opportunistic behavior if a market participant anticipates whether the system operator will be a net buyer or seller and thereby deliberately becomes imbalanced in the opposite direction. The participant thereby profits (Lisi and Edoli 2018, Bunn and Kermer 2021, Matsumoto et al. 2021). Furthermore, because these imbalance prices will generally differ from the prices clearing in the forward markets (the balancing function can capture a flexibility premium), the consequent price spreads between forward and imbalance prices can also lead to market players taking deliberate imbalance positions as shown both theoretically (Ito and Reguant 2016) and in practice (IRENA 2017, p. 68). Whereas some regulators (e.g., in the United Kingdom) are comfortable with market participants seeking to manage their imbalance positions in the opposite direction to that of the whole market, thereby reducing the balancing needs of the system operator, others (e.g., in Germany) take the view often expressed by system operators that this creates more uncertainty and potential instability for them to manage. One approach to making deliberate imbalance positions less attractive is to reduce the ease of hedging by having a dual-pricing mechanism. In this mechanism, participants face different imbalance prices when they are long or short. Thus, the prices are designed to be penalties for imbalance and thereby encourage the participants to manage their real-time production/consumption as close as possible to their prior gate closure nominations. However, dual pricing by itself does not preclude participants from seeking profit by being deliberately out of balance (Matsumoto et al. 2021).

Overall, single-pricing mechanisms are simpler and preferred in circumstances where the regulator sees benefits in market participants taking arbitrage positions that may help their hedging and reduce the amount of aggregate imbalance volume that the system operator needs to manage (e.g., in the United Kingdom (Ofgem 2014a)). Dual-pricing mechanisms have more flexibility in design and are generally intended to encourage participants not to have imbalances, but their advantage is unclear, especially because their particular way of distortion from the actual imbalance costs might lead to market inefficiency (IRENA 2017, p. 69). Facing this question of the most appropriate design in the context of European Union (EU)-wide market harmonization, the European Commission (ACER 2020) suggested single pricing as the basis and dual pricing only when it can be justified and on an ad hoc basis. Indeed, some markets have already shifted from a dual-pricing to a single-pricing mechanism (e.g., the United Kingdom in 2015 (Ofgem 2014b) and Nordic in 2021 (Nordic Balancing Model 2021)). Yet in practice, it is still unclear which market mechanism can best accommodate the technological and possible market power changes created by the ongoing energy transition. In particular, as the market evolves to become dominated by low-cost intermittent resources, with their consequent emergence of market power, the balancing market design not only will have more work to do in responding to the intermittency but also needs to be effective against this new source of strategic behavior.

Academic research on this particular aspect is sparse. Although researchers have studied the merits of each of these pricing mechanisms, there is no study that systematically compares and assesses the advantages of each mechanism for different market structures, namely with regard to cost, heterogeneity, and market power. Studies in this area provide little guidance and are often contradictory on whether a single pricing (Vandezande et al. 2010, van der Veen et al. 2012) or a dual pricing (Clò and Fumagalli 2019) is the better option. This apparent contradiction stems in part from the fact that the relative performance of these mechanisms is highly dependent on market characteristics such as the generators' cost structures and market power as well as supply and demand uncertainties.

Thus, in this paper, we analytically compare various pricing mechanisms and their impact on the economic efficiency and market's outcomes under different heterogeneity and market power conditions. Furthermore, we then seek to devise a new optimal imbalance pricing mechanism to face the challenge of the energy transition with the potential emerging market power of renewables. In particular, we study an electricity market with two stages (forward and real-time) and two types of generators, one (or multiple) strategic low-cost generator (LCG) and one (or multiple) flexible but fringe high-cost generator (HCG). Low-cost generators may, but not necessarily, have a large share of renewables in their portfolios. HCGs represent technologies such as gas turbines that can ramp up/down on short notice, albeit at high marginal costs. For real-time imbalance pricing, we consider the most typical form of the single-pricing (SP) as well as two important forms of dual-pricing schemes that we refer to as the *typical dual pricing* (TDP) and the *renewable-based dual pricing* (RDP). Whereas, with a single imbalance price, participants may offset positive and negative imbalances across their portfolios, in the typical dual-pricing scheme they would face different positive and negative imbalances prices, and therefore, such simple hedging would be precluded. For the renewable-based dual pricing, renewables would not be paid for any excess production when they produce more than they committed in the forward market (the argument being that their production is zero marginal cost anyway). Because regulatory policies are driven by economic efficiency, we measure the impact of market power not only on market participant conduct but also on social welfare maximization, which reduces to the total cost minimization when demand is inelastic. We characterize market behavior and economic efficiency in equilibrium as a function of the number of LCGs, their total market share, and their uncertainty, as well as demand uncertainty. We provide a comprehensive overview of the impact of different factors on the efficiency of the dual-pricing mechanisms with respect to the single-pricing mechanism.

Our results show that the performance of single-pricing versus dual-pricing mechanisms is highly dependent on market characteristics. We find that one of the dual-pricing schemes (RDP) will be attractive to regulators in markets with small to medium shares of LCGs. Counterintuitively though, as the market share of renewables increases, these mechanisms are outperformed by the typical SP mechanism. Thus, the more forward-looking regulators could be more attracted to consider other mechanisms because their market share of renewables is expected to increase above a critical point. In this regard, we propose an optimal imbalance pricing (OP) mechanism that can entirely mitigate LCGs' market power and recover the corresponding efficiency loss.

This is achieved by encouraging players to behave as if in a perfect-competition scenario. This is a new theoretical result on the impact of the real-time imbalance pricing on market power mitigation. We further complement the comparison through simulation studies for cases that reflect existing markets along with sensitivity analysis with respect to the choice of parameters. Our simulation findings consistently confirm that with a large share of LCGs, the OP mechanism can substantially reduce the deadweight loss compared with others.

The rest of the paper is organized as follows. In Section 2, we review the related research. In Section 3, we model the problem and formulate the economic efficiency in terms of the deadweight loss, given the market outcome. In Section 3, we formulate and analyze the economic efficiency of the market when the SP, the TDP, and the RDP are used in Sections 4–6, respectively. We formulate our optimal pricing mechanism and present its performance with respect to the three other pricing mechanisms in Section 7. We discuss practical concerns in Section 8. Finally, in Section 9, we provide a summary of pricing comparisons and conclude the paper.

2. Background Research

Our work relates to a large body of literature investigating the market behavior of renewables (as LCGs) in electricity markets (see, e.g., Löhndorf and Minner 2010, Zhou et al. 2015, Peura and Bunn 2021, and Sunar and Swaminathan 2021) or the market behavior in general multistage markets with similar concerns (see Anderson 1991 for a survey). Below, we classify this literature across three dimensions: problem setting, methodology, and solution.

Prior studies have considered different market setups, including Cournot competition (Allaz 1992), Bertrand competition (Mahenc and Salanié 2004), and supply function equilibrium (Anderson 2004, Al-Gwaiz et al. 2017, Sunar and Birge 2019). Different market characteristics have been studied in the literature, such as market liquidity (Hesamzadeh et al. 2020), reliability (Sunar and Birge 2019), and efficiency (Ito and Reguant 2016). Existing research also varies in terms of market interconnection assumptions. Several studies have focused on modeling the economic efficiency of an isolated electricity market (see, e.g., Zhang and Xu 2013), whereas somewhat less research has appeared on multinode interconnected electricity markets (see, e.g., Kamat and Oren 2004). In our work, we focus on equilibrium behavior, market power, and economic efficiency. We are among the few to analyze the strategic behavior of LCGs, especially the potential that renewables may pose, as indicated in Ito and Reguant (2016) and Sunar and Birge (2019).

There are various approaches used to study two-stage markets in general and electricity markets in particular. Because of the complexity of the problem, there are only a few studies that provide closed-form theoretical solutions (see, e.g., Allaz 1992, Allaz and Vila 1993, Ito and Reguant 2016, Sunar and Birge 2019). Most studies, especially those with more practical assumptions, use mathematical programming, such as a mathematical program with equilibrium constraints (MPEC) (Luo et al. 1996, Su 2007, Yao et al. 2007), regression analysis (Zarnikau et al. 2019), dynamic programming (Jiang and Fei 2011, Kim and Powell 2011), or empirical/simulation approaches (Borenstein 2002, Ito and Reguant 2016), to bring new insights to the market operation and market behaviors of sequential markets. Our work adds to the theoretical literature by providing a closed-form formulation for economic efficiency (in terms of the total cost of production and deadweight loss) and market behavior of heterogeneous suppliers (LCGs and HCGs).

The related literature can also be classified in terms of the type of solutions considered for market power mitigation. The principles of market surveillance and market rule interventions in electricity have understandably evolved to mitigate market power, mainly regarding the conduct of thermal generators and in consideration of their marginal costs. Market power analysis—as discussed by Biggar and Hesamzadeh (2014)—has generally looked at markups above short-run marginal costs (PWC 2018) and capacity withholding (Willis and Altozano 2016) as well as structure and conduct considerations (FERC 2014). Several solutions have been proposed in the literature to mitigate market power and its negative impact on the economic efficiency of sequential electricity markets. Examples are adding more forward markets (Allaz and Vila 1993), introducing virtual bidders or financial speculators (Güler et al. 2010, Ito and Reguant 2016, Jha and Wolak 2019), storage (Löhndorf and Minner 2010, Secomandi 2010, Kim and Powell 2011, Jiang and Powell 2015, Zhou et al. 2015), curtailment (Wu and Kapuscinski 2013, Al-Gwaiz et al. 2017), and enforcing price caps (Yao et al. 2007). Each of these solutions has its own strengths and limitations. We provide a detailed discussion on some of these solutions in Section 8.

Unlike these studies, we focus on real-time imbalance pricing as a solution to mitigate the adverse effects of market power. This solution has been studied recently by Sunar and Birge (2019) in a similar context. They showed that if the renewable firms with market power are charged at a market-based rate for their real-time production deviation from the day-ahead commitments, then imposing or increasing the real-time market-based penalty rates might be counterproductive by reducing reliability. Our work differs from Sunar and Birge (2019) in multiple ways. First, we are focusing on economic efficiency, whereas Sunar and Birge (2019) focused on supply reliability. Secondly, HCGs in our framework are flexible fast ramp-rate generators (e.g., gas turbines), whereas Sunar and Birge (2019) assumed inflexible generators (e.g., coal power plants). We argue that the energy transition to low-carbon technologies is leading to a replacement of the older inflexible fossil fuel generators with renewables and a contraction of the mid-merit segment of the supply function because both low-cost renewables expand and more flexible but expensive generators are needed to cope with renewables intermittencies. Finally, price formation in the second stage, in our work, is modeled differently and according to the real-time pricing mechanisms that recognize that these markets in reality often clear by facilities that are already scheduled in the first stage, and thereby, the second-stage market prices are linked to the first-stage market prices.

Our work, similar to Petruzzi and Dada (1999) and Liberopoulos and Andrianesis (2016) but in a different context, provides an analytical review and comparison of some pricing mechanisms with an extension to the optimal pricing. Thus, our work also relates to existing studies that compare the single-pricing and the dual-pricing mechanisms. To the best of our knowledge, this line of research is also rather limited. Clò and Fumagalli (2019) took an empirical approach using data from the Italian electricity market before and after it went through a transition from the single-pricing to the dual-pricing mechanism. They concluded that the dual pricing for that case outperforms the single pricing in terms of the total cost of production. van der Veen et al. (2012) took a simulation approach to compare multiple versions of single-pricing and dual-pricing mechanisms and concluded that the typical single pricing outperformed others in their study set. Vandezande et al. (2010) used simple numerical examples to compare the quantity adjustments of wind power producers under single and dual pricing and concluded that the single pricing is better. Although these studies are insightful, they are not conclusive, given the specific model setup or the market setting. In this context, our work is the first to provide a more thorough comparison between the single-pricing and the dual-pricing mechanisms as a function of market parameters, leading to an alternative optimal real-time pricing mechanism. Moreover, in these other studies, real-time prices are exogenous, and price formations are not modeled. Finally, we are focusing specifically on the impact of these real-time pricing mechanisms on market power mitigation and improving the economic efficiency (especially because the share of LCGs is increasing), none of which are the focus of these prior studies.

3. Model

3.1. Market Setup and Elements

In this section, we introduce the market setup. The reader can refer to Online Appendix A for a list of notations. To help readability, as also shown in Online Appendix A, we use regular English letters for endogenous variables and Calligraphic or Greek letters for exogenous ones.

3.1.1. Demand. We focus on the market for the delivery of electricity at a specific time period. We assume that the demand is exogenous and randomly distributed. Ahead of delivery time, the actual demand is unknown, but its expectation is common knowledge and, without loss of generality, normalized to 1. At delivery time, the actual demand is realized and revealed to all players: $1+\epsilon$. The parameter ϵ is an unbiased ($\mathbb{E}[\epsilon]=0$) adjustment to the expected demand ahead of delivery time and is assumed bounded such that $\epsilon \geq -1$. This ensures that the actual demand is never negative. We further assume that the demand is inelastic, as is roughly the case in practice, given that it is a necessary good.

3.1.2. Supply. The market is composed of two types of generators: $N \geq 1$ heterogeneous strategic low-cost generators (LCGs) and one nonstrategic (hence, not a decision-maker), high-cost generator (HCG). Although these assumptions might be simplifications for the current market setting, it is imminent that they are also practically relevant. LCGs are already becoming dominant players in some markets. For example, Germany has recorded many hours with even more than 80% of the total demand (International Energy Agency 2020). Additionally, HCGs are losing market power because of the rise of renewables (and much less leftover demand to be served by HCGs) and the replacement of large producers with many smaller ones, as already observed in longer-term forward electricity markets (see, e.g., National Grid ESO 2021).

The HCG has an unlimited production capacity and faces marginal costs of production that grow linearly in quantity (i.e., total costs are quadratic in quantity). For production contracted ahead of delivery time, the marginal costs increase at a rate of α_1. For production contracted at delivery time, the marginal costs are assumed to increase at a rate of α_2. We assume that both α_1 and α_2 are common knowledge. Typically, market players get informed about the historical values of market prices and hence, can estimate the slopes (α_1 and α_2) with high accuracy (e.g., Birge et al. 2017 and Chen et al. 2019 showed how such values can be estimated using inverse optimization techniques.). Nonetheless, in Online Appendix C, we have ensured that our conclusions are robust with respect to mispredictions of these slopes. We further assume that $\alpha_2 \geq \alpha_1 \geq 0$. This is also consistent with most energy-generating technologies; short-term adjustments result in changes in marginal costs proportional to the size of the adjustment, and they are more costly compared with scheduling them in the first stage, given the lead time.

Let us denote by H_1 the HCG's production scheduled ahead of delivery time and by H_2 the production adjustments at delivery time. Note that $H_1 \geq 0$, whereas H_2 can also take negative quantities because adjustments can also be in form of reducing the initially scheduled production. We also use $\mathcal{H}$ to refer to the total production of the HCG after adjustments; that is, $\mathcal{H} \equiv H_1 + H_2$. With this notation, the HCG's marginal costs for production scheduled ahead of delivery (MC_1) and adjusted at delivery time (MC_2) are

$$MC_1 = \alpha_1 H_1 \tag{1}$$

$$MC_2 = MC_1 + \alpha_2 H_2 = \alpha_1 H_1 + \alpha_2(\mathcal{H} - H_1). \tag{2}$$

Figure 1 depicts the functions of MC_1 and MC_2 with respect to H_1 and $\mathcal{H}$. This is due to the fact that the marginal cost of the HCG is directly proportional to its production. Therefore, the marginal cost at the delivery time MC_2 is anchored on the marginal cost with the scheduled production before delivery time MC_1, and it changes with the short-term adjustments of H_2 at the delivery time, with a faster pace compared with the one ahead of the delivery (i.e., $\alpha_2 \geq \alpha_1$). If we shrink the production in real-time $H_2 < 0$, the marginal cost will decrease, resulting in $MC_2 < MC_1$. However, this continues only until the point where MC_2 reaches zero, because the marginal cost cannot be negative. Another important point is that although we attribute HCG cost functions to one HCG, we can relax this assumption by interpreting these cost functions to be the equivalent cost function of a mixture of multiple HCGs (see Online Appendix B.1).

Using the marginal cost formulations in Equation (1) and Equation (2), the overall cost of production of the HCG, denoted by $C(H_1, H_2)$, as a function of the HCG scheduled production in Stage 1 (H_1) and its adjustment in Stage 2 (H_2) is given by the area under the curve in Figure 1:

$$\begin{aligned} C(H_1, H_2) &= \int_0^{H_1} \alpha_1 h dh + \int_{H_1}^{\mathcal{H}} [\alpha_1 H_1 + \alpha_2(h - H_1)] dh \\ &= \alpha_1 \frac{\mathcal{H}^2}{2} + (\alpha_2 - \alpha_1)\frac{H_2^2}{2}. \end{aligned} \tag{3}$$

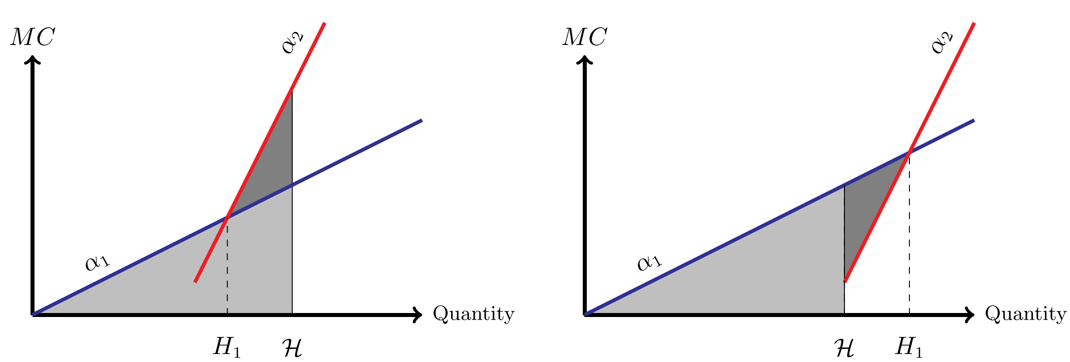
The first element in Equation (3) ($\alpha_1 \mathcal{H}^2/2$) represents the light gray area, and the second term ($(\alpha_2 - \alpha_1)H_2^2/2$) represents the dark gray area in Figure 1. Figure 1, left, represents the total cost incurred by the HCG when $H_2 > 0$, whereas Figure 1, right, represents the same costs when $H_2 < 0$.

The LCGs face a negligible marginal cost of production (assumed zero). The realized production of each LCG is randomly distributed such that, ahead of delivery time, each LCG observes a private, unbiased, but noisy forecast of their actual production: γ^i. At delivery time, the LCG's actual production is realized; $\gamma^i(1 + \nu^i)$, with ν^i representing an unbiased ($\mathbb{E}[\nu^i] = 0$) adjustment to the expected production and assumed bounded such that $\nu^i \geq -1$. Using a multiplicative adjustment is a modeling choice (similarly with ϵ for the demand), which fits practical scenarios because energy forecast errors are often reported in terms of normalized percentages. Moreover, the multiplicative error model allows us to exclude long-term errors and focus on short-term forecast errors.

We use the notation $\gamma \equiv \sum_{i=1}^{N} \gamma^i$ for the total expected production across all LCGs. Without loss of generality, we assume that γ^i is normalized to the expected demand, but with the same dimension as demand (i.e., *MWh*). As such, we also interchangeably call γ the market share of LCGs. We restrict our focus to $\gamma \leq 1$, because this represents the practical situation where there is no curtailment of excess LCGs (e.g., wind) and where uncertain LCGs must be combined with an energy backup such as HCG to cover the residual demand. We make the assumption that LCGs do not hold back their total production during the delivery period. In practice, they receive significant subsidies for each unit of production, which are usually large enough to discourage any withholding in real time. Although these subsidies do not impact our analysis, they do eliminate any potential incentives that would lead to $L_1^i + L_2^i < \gamma^i(1 + \nu^i)$. Besides subsidies, additional capacity caps can also prevent production withholding (see Online Appendix B.12).

3.1.3. Market Operator. The market operator's main goal is to guarantee that supply matches demand. To achieve this goal, the market operator defines the market rules, namely, the pricing settlement mechanism,

Figure 1. (Color online) Marginal Costs (MC_1 and MC_2, respectively, the Lines with Slope α_1 and α_2), the Total Costs (Sum of Light Gray and Dark Gray Areas), and the Deadweight Loss (Dark Gray Area) When $H_2 > 0$ (Left) and $H_2 < 0$ (Right)



that is, how prices are formed at each period of the market. All generators interact directly with the market operator and must follow the rules, at the expense of being excluded from the market if they do not do so. We discuss the properties of different price settlement mechanisms in Section 4.

3.1.4. Timing of the Model and Sequence of Decisions. The model has two stages (see Figure 2). In Stage 1—the period before the actual delivery of electricity, also known as the day-ahead market or the forward market—all suppliers receive a common signal about overall expected demand, with LCGs further receiving a private signal about their individual production capacity. With this information, LCGs decide how much production to commit to the market operator for delivery at delivery time. In Stage 2, also known as the delivery time or real-time imbalance market, the actual demand and production capacities are realized, and adjustments are made such that supply meets demand.

Stage 1: Each LCG receives a signal about their own production at delivery time, γ^i, here represented as a fraction of the expected demand. Each LCG then chooses how much of their expected production to offer to the market operator: $\gamma^i(1+d^i)$. We call the decision variable d^i the *offer quantity adjustment* because it corresponds to the deviation of the offer as a fraction of the total capacity of the LCG i. When $d^i=0$, the LCG commits its total expected energy production in Stage 1. The total quantity committed by LCGs in Stage 1 is $L_1 \equiv \sum_{i=1}^{N}\gamma^i(1+d^i)$, where $d^i \in [-1,(1/\gamma)-1]$ to guarantee that the quantity commitment of every LCG is nonnegative and that the total commitment of all LCGs does not surpass the expected demand.

The HCG is not a decision-maker in our model, acting as a backup to ensure supply-demand matching in the presence of LCGs. This means that in Stage 1, HCG schedules to produce the leftover demand, that is, $H_1 = 1 - L_1$, at a marginal cost of $MC_1 = \alpha_1 H_1$. Prices are determined according to the pricing settlement mechanism defined by the market operator, who then pays the committed quantities to the generators. Note that with these definitions, $L_1 \in [0,1], H_1 \in [0,1]$.

Stage 2: At delivery time, the actual demand $(1+\epsilon)$ and the actual production for LCGs $(\gamma^i(1+\nu^i))$ are realized. Each LCG i then faces financial adjustments for deviations between the actual realized production, $\gamma^i(1+\nu^i)$, and the committed quantity in Stage 1, $\gamma^i(1+d^i)$. The difference is $\gamma^i(\nu^i-d^i)$ because in an efficient scenario all of the LCGs' production is incorporated in the market before any HCG's production because of the lower marginal costs of the former. At this stage, the total adjustment by the LCGs is $L_2 \equiv \sum_{i=1}^{N}\gamma^i(\nu^i-d^i)$.

The LCGs that committed more than their actual production will have to further purchase quantity at the prices determined by the pricing mechanism to fulfill their commitment to the market operator, and LCGs that committed less than their actual production will have to sell their unsold production to the market operator according to the pricing mechanism. For notation simplicity, we also define L_1^i and L_2^i, respectively, as the quantity commitment of the LCG i in Stage 1 and the quantity adjustment in Stage 2, given the quantity adjustment offer. With our notation, this means that $L_1^i \equiv \gamma^i(1+d^i)$ and $L_2^i \equiv \gamma^i(\nu^i-d^i)$.

The HCG then must provide the remaining of the unfulfilled realized demand, if any, $H_2 = \epsilon - L_2$, at a marginal cost $MC_2 = \alpha_1 H_1 + \alpha_2 H_2$. The prices are determined according to the pricing settlement mechanism defined by the market operator. If the overall committed quantity by LCGs in Stage 1 was higher than the actual production, the HCG will have to adjust its production to produce more than initially planned, now with higher marginal costs. If the overall committed quantity by LCGs is lower than the actual production, LCGs will sell their extra production to the market operator, making the HCG adjust its production down, with implications on marginal costs: $MC_2 = MC_1 + \alpha_2 H_2$, in which H_2 can be negative.

3.2. Economic Efficiency

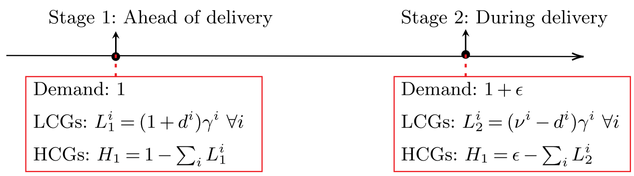
Given that demand is inelastic and the marginal costs of the HCG are given in Equations (1) and (2), the efficient outcome is the one that minimizes total costs. Thus, in the efficient scenario, all of the LCGs' production is integrated, because their marginal cost is always lower than the marginal cost of the HCG.

Because our focus is economic efficiency, we use the deadweight loss (DWL) to compare a given outcome with the first-best outcome, that is, the efficient outcome. In the case of inelastic demand, the DWL is equivalent to the additional cost of serving the demand beyond the minimum possible cost. According to Equation (3) and also Figure 1, the minimum possible cost of production happens when $H_2 = 0$ because it eliminates the dark gray area in Figure 1. Therefore, the DWL, which is the dark gray area in that figure, can be expressed as

$$DWL = C(H_1,H_2) - C(\mathcal{H},0) = (\alpha_2-\alpha_1)\frac{H_2^2}{2}. \tag{4}$$

Accounting for the uncertainty of the LCSs' production and uncertainty in demand, the expected value of DWL

Figure 2. (Color online) Market Setup and Elements



can be formulated as follows (see Online Appendix B.2 for the proof):

Lemma 1. *The expected value of the deadweight loss can be expressed as*

$$\mathbb{E}[DWL] = \frac{(\alpha_2 - \alpha_1)}{2}\mathbb{E}[(\epsilon - L_2)^2] = \frac{(\alpha_2 - \alpha_1)}{2}(VAR(\epsilon) + \mathbb{E}[L_2^2]) \tag{5}$$

or equivalently as

$$\mathbb{E}[DWL] = \frac{(\alpha_2 - \alpha_1)}{2}\left(\sum_{j=1}^{N}\sum_{i=1}^{N}\gamma^j\gamma^i d_1^j d_1^i + VAR(\epsilon) + \sum_{j=1}^{N}\sum_{i=1}^{N}\gamma^j\gamma^i Cov(\nu^j,\nu^i)\right). \tag{6}$$

Equation (5) implies that given demand forecast errors as exogenous, economic efficiency maximizes if the total that LCG offers in Stage 2 is minimized (i.e., $L_2 = 0$). Equation (6) further differentiates the role of forecast errors from strategic behavior. The first term in Equation (6) is attributed to the strategic behavior of LCGs, and the last two terms correspond to the DWL caused by the forecast errors (ϵ and ν^i's). Thus, economic efficiency can increase by improving the forecasting algorithms and/or by aligning LCG's market behavior with an efficient outcome. In this paper, we focus on the latter (actions on market behavior), treating forecast errors as exogenous. The important market behavior of any LCG i in our context is its offer quantity adjustment (d^i) because it directly affects the DWL (see Equation (6)). An immediate observation from Equation (6) is that a sufficient (not necessary) condition to remove the inefficiency caused by the strategic behavior of LCGs occurs when all LCGs offer their total predicted production in Stage 1 (i.e., $\forall i: d^i = 0$).

The market behavior of LCGs depends on the pricing mechanisms in the first and the second stages of the market. Market price mechanisms in Stage 1 are typically similar across all markets, chosen to be the marginal cost of the HCGs; that is,

$$p_1 = MC_1, \tag{7}$$

where MC_1 is given in Equation (2)), given the fringe nature of HCGs and the uniform pricing mechanism used in Stage 1 (day-ahead market). However, market price mechanisms in Stage 2 (imbalance market) differ considerably across different markets. We call the pricing mechanism in Stage 2 the imbalance pricing. In what follows, we analyze the market behavior of LCGs under different imbalance pricing mechanisms and the resulting impact on the economic efficiency.

4. Single-Pricing

In Stage 2 of the market, any LCG i trades its quantity adjustment L_2^i, acting as a seller (when $L_2^i > 0$) or a buyer (when $L_2^i < 0$). In a single-pricing (SP) mechanism, the buying and selling prices for adjustments in Stage 2 of the market are identical. In the most typical form of a SP mechanism, this price is the marginal cost of procurement by the HCG in the balancing market; that is, $p_{2,sp} = MC_2$ (see Equation (2)). This can be summarized in the following definition:

Definition 1 (SP). The single pricing is an imbalance pricing mechanism in which the price applied to the LCG i in the imbalance market is given by

$$\textbf{SP}: \quad p_{2,sp}^i = p_1 + \alpha_2 H_2. \tag{8}$$

According to the above definition and Equation (8), the imbalance price in the SP mechanism $p_{2,sp}^i$ is the same for all LCGs and is anchored on the price in Stage 1 of the market p_1 and deviates from that as a function of total quantity adjustments of the HCG in the second stage (H_2). Thus, any changes in the price of Stage 1 will also affect the price in Stage 2. Accordingly, in the SP mechanism, the LCG i chooses the offer quantity adjustment d^i in the first stage to maximize the profit, given by

$$\Pi_{sp}^i = p_1 L_1^i + p_{2,sp}^i L_2^i, \tag{9}$$

where the optimal strategy of the LCG i and the resulting DWL are characterized as below.

Theorem 1. *In a market with a SP mechanism, the offer quantity adjustment of any LCG i in an oligopoly of N heterogeneous LCGs, in equilibrium, is given by*

$$d^i = -\frac{\alpha_1}{\alpha_2}\left(1 - \frac{\gamma}{(N+1)\gamma^i}\right), \tag{10}$$

which leads to

$$\mathbb{E}[DWL_{sp}] = \frac{(\alpha_2 - \alpha_1)}{2}\left[\left(\frac{\alpha_1\gamma}{(N+1)\alpha_2}\right)^2 + Var(\epsilon) + \sum_{j=1}^{N}\sum_{i=1}^{N}\gamma^j\gamma^i Cov(\nu^j,\nu^i)\right]. \tag{11}$$

Proof. See Online Appendix B.3 for the proof.

Equation (10) implies that LCGs offer below their predicted energy production; that is, they under-commit. This is so because LCGs can take advantage of their market power to influence prices. The LCGs, by committing less quantity than they expect to produce, are actively contributing to an increase in the price of energy in Stage 1; the HCG needs to schedule a higher production, raising its marginal cost of production. This rise in price means that the LCGs sell their production in Stage 1 at a higher unit cost than if they had committed their full expected production. Moreover, they will still sell the remaining of their production in Stage 2, increasing their profit.

Several other interesting observations can be made based on Theorem 1 in terms of the behavior of LCGs, according to Equation (10), and the resulting DWL, according to Equation (11). To start with, Theorem 1 suggests that forecast errors (ϵ and ν^i) have no impact on the behavior of LCGs in a SP mechanism. This means that the strategic behavior of LCGs is not because of uncertainties but because of their market power. However, forecast errors negatively impact the DWL. Indeed, by comparing the general form of DWL in Equation (6) and the DWL in the case of a SP mechanism in Equation (11), we find that the impact of the SP mechanism is reflected in the first term of Equation (11), whereas the last two terms relate to the forecast errors and are not affected by the pricing mechanism.

Contrary to the impact of forecast errors, supply heterogeneity (γ^i across all $i \le N$) does not impact the DWL but affects the market behavior of individual LCGs. This implies that supply heterogeneity in the SP mechanism makes LCGs take opposite deviations with respect to an aggregate monopolistic counterpart in a way that the impact of heterogeneity on the overall DWL remains zero. It is also important to observe the impact of the total capacity (the market share) of LCGs (γ) as a result of capacity expansion. If we define the capacity expansion process as the process of uniformly increasing γ^i for all i (and correspondingly γ), then such expansion would not affect the market behavior of LCGs but would lead to an increase in the DWL. Therefore, as the total market share of all LCGs keeps increasing, the economic efficiency deteriorates further to the extent that it might make the SP mechanism unfit for the market. This is an important and alarming observation that the SP mechanism might not be a good choice for future electricity markets because the share of renewables (as LCGs) in the energy mix is increasing.

Finally, we investigate how introducing competition affects the market behavior and the DWL. To study the impact of competition, we fix the total share of LCGs γ and increase N. Theorem 1 shows that as the number of LCGs increases from $n = 1$ (monopoly) to $N = \infty$ (perfect competition), the optimal offer quantity adjustment moves toward the efficient offer ($d^i = 0$) asymptotically as fast as $\Theta(N^{-1})$, and the DWL decreases as fast as $\Theta(N^{-2})$. Note that the Big-Θ notation provides both lower and upper bounds on the growth rate, whereas Big-$\mathcal{O}$ provides only an upper bound (see Cormen et al. 2022, p. 51, for a more formal definition). Investigating the asymptotic trend as N increases is crucial because the number of LCGs can be significant, especially with the emergence of modular solar and wind power as the dominant LCGs in future markets. For example, by analyzing the trend, regulators can gain valuable insights into the scale at which increasing N can induce competitive behavior. For the special case of perfect competition ($N = \infty$), LCGs offer their entire capacity in Stage 1 ($d^i = 0$), and the DWL reaches its minimum value as the first term in Equation (11) disappears. In a deterministic setting, a fully competitive market reaches full economic efficiency ($DWL = 0$) in the SP mechanism. In a probabilistic setting, though, the expected DWL will never be zero because of its component associated with uncertainty. All of the above discussion is summarized in the following:

Observation 1. In a market with the SP mechanism, forecast errors affect DWL negatively but have no impact on market behavior. Supply heterogeneity affects the market behavior of LCGs but does not impact the overall DWL. Increasing the total share of LCGs (γ) through capacity expansion would not affect the market behavior of LCGs, but it would result in an increase in the DWL. Finally, increasing competition (increasing N) steers the market behavior of LCGs toward the efficient offer ($d^i = 0$) as fast as $\Theta(N^{-1})$ and decreases the DWL as fast as $\Theta(N^{-2})$.

5. Typical Dual Pricing

In this section, we study the most typical form of the dual-pricing mechanism, used in several locations such as the Iberian market, the former British market, and the current Nordpool market (EC 2016). For this reason, we call it the *typical dual pricing (TDP)*. We investigate the implications of TDP with respect to the SP mechanism and when LCGs become the dominant players.

Unlike the SP mechanism, in which all market players are exposed to a single price for their imbalances, in a TDP, selling and buying prices in the real-time imbalance market are different. A TDP mechanism works as follows (Morales et al. 2014). If $\epsilon - L_2 > 0$, then there is unserved demand in Stage 2, and the market operator needs to schedule the production shortage from the HCG ($H_2 > 0$). This means market inefficiency. In this case, according to Equation (2), the marginal cost pricing of Stage 2 is higher than that of Stage 1 (i.e., $MC_2 > MC_1$, where MC_1 and MC_2 are marginal costs from Equation (2)). Any LCG i who is also a buyer (i.e., $L_2^i < 0$) when the overall market itself is in the energy shortage mode is indeed contributing to market inefficiency. In the TDP mechanism, such a market player pays p_2^i for its purchase. However, any LCG i who is a seller in this case (i.e., $L_2^i > 0$) receives $p_1 = MC_1$ (which is lower than p_2) for its production in Stage 2. Similarly, if $L_2 - \epsilon > 0$, then there is a surplus production in Stage 2, and the market operator needs to request the HCG to adjust by reducing its Stage 1 scheduled production ($H_2 < 0$), hence market inefficiency. In this case, according to Equation

(2), the marginal cost pricing of Stage 2 is lower than that of Stage 1 (i.e., $MC_2 < MC_1$). Any LCG i who is also a seller (i.e., $L_2^i > 0$) at that time is indeed contributing to this market inefficiency. In the TDP mechanism, such a market player receives $p_2^i = MC_2$ (which is lower than $p_1 = MC_1$) for its positive adjustments in Stage 2. However, any LCG i who is a buyer in this case (i.e., $L_2^i < 0$) pays $p_1 = MC_1$ (which is higher than MC_2) for its purchase in stage 2. In summary, the TDP mechanism applies Stage 2 marginal cost prices (i.e., MC_2) only when it is less attractive to the LCGs, and otherwise applies Stage 1 marginal cost prices (i.e., $p_1 = MC_1$). This is formally defined below.

Definition 2 (TDP). The typical dual pricing is an imbalance pricing mechanism in which the price applied to the LCG i in the imbalance market is given by

$$\textbf{TDP}: \quad p^i_{2,tdp} = p_1 + \alpha_2 H_2 \mathbb{I}_{(H_2 L_2^i < 0)}, \tag{12}$$

where $\mathbb{I}_x$ is 1 if x is true and is 0 otherwise.

Accordingly, as shown in the following proposition, we can express the average profit of a given LCG i in a market with the TDP mechanism (Π^i_{tdp}) as a function of a similar scenario with the SP mechanism (Π^i_{sp}) (see Online Appendix B.4 for the proof):

Proposition 1. *The profit of any LCG i with any arbitrary offer quantity adjustment in an oligopoly of LCGs in a market with the TDP mechanism versus a similar scenario in a market with the SP mechanism can be expressed as*

$$\Pi^i_{tdp} = \Pi^i_{sp} - \alpha_2 \mathbb{E}[[H_2 L_2^i]_+], \tag{13}$$

where Π^i_{tdp} and Π^i_{sp}, respectively, represent the profit of LCG i in the TDP and the SP mechanisms with the same offer quantity adjustment (d^i).

This leads to the following lemma (see Online Appendix B.5 for the proof).

Lemma 2. *The profit of the LCG i in a market with the TDP mechanism as formulated in Equation (13) is a concave function of its offer quantity adjustment d^i.*

Using these results, we compare the efficiency of the TDP mechanism with respect to the SP mechanism. In this and the next sections, we assume that forecast errors can be characterized by uniformly distributed random variables. This keeps our derivations tractable while still being sufficiently practical. With this assumption, the probability density functions of any LCG i (denoted by f_{ν^i}) and the demand (denoted by f_ϵ) are, respectively, given by

$$f_{\nu^i}(x) \sim \mathbb{U}[-\overline{\nu}, +\overline{\nu}] \text{ and } f_\epsilon(x) \sim \mathbb{U}[-\overline{\epsilon}, +\overline{\epsilon}], \tag{14}$$

where $\overline{\nu} \geq 0$ and $\overline{\epsilon} \geq 0$ model the precision of predictions. The case of $\overline{\nu} = \overline{\epsilon} = 0$ corresponds to the perfectly predictable supply and demand. For tractability, in this section and the next section, we assume uncorrelated forecast errors. Nonetheless, as observed in Online Appendix C, our conclusions remain unchanged for correlated errors as well as the skewness of the distribution. Under a monopoly of one LCG, the following theorem compares the market power mitigation and economic efficiency under TDP and SP mechanisms (See Online Appendix B.6 for the proof):

Theorem 2. *In a monopoly of a LCG, the economic efficiency of the TDP mechanism with respect to the SP mechanism depends on the forecast errors and the market share of the LCG. We can characterize three regions: R_{sp}, R_{tdp}, and R_n, representing the regions where the SP outperforms the TDP, the TDP outperforms the SP, and the region where the SP and TDP perform identically in terms of economic efficiency, respectively. These regions are given by*

$$R_{sp} = (Z_1 \cap Z_2^c \cap Z_4) \cup (Z_2 \cap Z_3^c \cap Z_4) \tag{15}$$

$$R_{tdp} = (Z_2 \cap Z_4^c) \cup (Z_3 \cap Z_4) \tag{16}$$

$$R_n = (R_{sp} \cup R_{tdp})^c, \tag{17}$$

where 4 zones (Z_1 to Z_4) are defined as below,

$$Z_1 := \{(\overline{\nu}, \overline{\epsilon}) | \overline{\epsilon} > \gamma(\overline{\nu} + d^*)\}$$

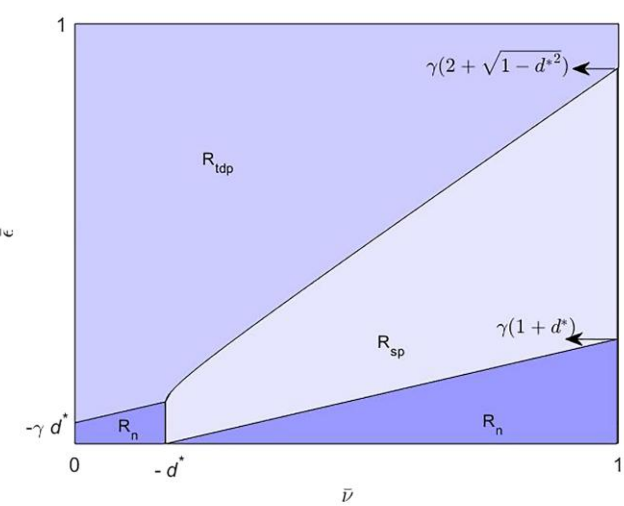
$$Z_2 := \{(\overline{\nu}, \overline{\epsilon}) | \overline{\epsilon} > \gamma(\overline{\nu} - d^*)\}$$

$$Z_3 := \{(\overline{\nu}, \overline{\epsilon}) | \overline{\epsilon} > 2\gamma\overline{\nu} + \gamma\sqrt{\overline{\nu}^2 - d^{*2}}\}$$

$$Z_4 := \{(\overline{\nu}, \overline{\epsilon}) | \overline{\nu} > -d^*\},$$

and d^* is the offer quantity adjustment of the LCG in the SP mechanism in equilibrium (as formulated in Theorem 1 for n = 1).

Figure 3. (Color online) Visualizing Theorem 2



Notes. Comparing the economic efficiency of the TDP and the SP as a function of the forecast errors of the demand (y-axis) and of the LCGs' production (x-axis). R_{sp}: the SP is better; R_{tdp}: the TDP is better; R_n: the SP and the TDP perform equally effectively.

We parametrically visualize Theorem 2 in Figure 3. Theorem 2 lays out the general efficiency comparison of SP and TDP mechanisms. Two missing points in this comparison are the impact of the number of LCGs and the impact of heterogeneity. We study the former with the following theorem (proof is in Online Appendix B.7) and discuss the latter in our simulation results in Section 7.

Theorem 3. *In an oligopoly of N homogeneous LCGs, the profit functions of any LCG I under SP and TDP mechanisms and for any identical but arbitrary offer quantity adjustment converge to each other with the rate $\mathcal{O}(N^{-1})$; that is,*

$$\Pi^i_{tdp} - \Pi^i_{sp} = \mathcal{O}\left(\frac{1}{N}\right), \tag{18}$$

where Π^i_{tdp} and Π^i_{sp}, respectively, represent the profit functions of the LCGs i in TDP and SP mechanisms. Equation (18) also infers that the market behavior, market power mitigation, and economic efficiency under these two mechanisms converge to each other as N increases.

Combining Theorems 2 and 3 and Figure 3, we reach the following (see Online Appendix B.8 for the rationale).

Observation 2. With low-demand uncertainty, the TDP mechanism is not advantageous. Increasing the demand uncertainty favors the TDP over the SP mechanism. In contrast, increasing the supply uncertainty of LCGs and increasing the total market share of LCG favor the SP over the TDP. There is, however, no general trend in the case of increasing the number of LCGs in the market. The relative performance of the TDP and the SP depends on other factors in this case.

In summary, we find that the performance of the SP mechanism compared with the TDP mechanism depends on multiple market characteristics, including LCGs' market power, demand, and supply uncertainty and market share of LCGs. Contrary to the ambitions behind the TDP mechanism designs, our results show that the TDP mechanism does not always outperform the SP mechanism. Depending on market characteristics, either of the two mechanisms can outperform the other, or they could perform identically in terms of economic efficiency. This justifies the contradictory observations in the literature (as discussed in the Introduction) about the relative performance of a single pricing with respect to a dual pricing. Figure 3 implies that the TDP mechanism mostly outperforms the SP mechanism when HCGs and LCGs are likely to take the same selling/buying position in Stage 2 of the market (for example, when the demand uncertainty increases or the share of LCGs is small). Conversely, when HCGs and LCGs have a higher chance to take opposite positions in Stage 2 (for example, when the demand uncertainty $\overline{\epsilon}$ is low and LCGs' uncertainty $\overline{\nu}$ is high), the SP mechanism leads to a higher economic efficiency.

6. Renewable-Based Dual Pricing

Variable renewable sources such as wind and solar were formerly exempted from real-time imbalance payments. However, because the share of renewables is increasing, there is a growing consensus to lift this exemption (IRENA 2017, p. 70). Indeed, the former strategy could have led to market manipulations and fewer incentives for improving the predictions of renewable generators. Apart from large uncertainties, renewable resources also differ from conventional resources in that they have negligible marginal costs. For this reason, some markets, such as the Italian electricity market, use specific real-time imbalance pricing for renewables EC (2016). A logical real-time adjustment would be not to pay them for their extra production in the imbalance market beyond their day-ahead commitments because this does not cost them, but paying for it might create an incentive for extra under-commitments in the day-ahead market.

Therefore, we study this real-time imbalance pricing mechanism, in which generators do not get paid for their overproduction (beyond their Stage 1 commitments), but they still need to buy their underproduction (below their Stage 1 commitments) in the second stage of the market. We call this mechanism the *renewable-based dual pricing (RDP)* because this is an appropriate fit for renewables and the prices differ for buying and selling, hence, dual pricing. This is formally defined below.

Definition 3 (RDP). The renewable-based dual pricing is an imbalance pricing mechanism, in which the price applied to the LCG i in the imbalance market is given by

$$\textbf{RDP}: \quad p^i_{2,rdp} = (p_1 + \alpha_2 H_2)\mathbb{I}_{L^i_2<0}. \tag{19}$$

Accordingly, the profit of LCG i, denoted by Π^i_{rdp}, under the RDP can be expressed as

$$\Pi^i_{rdp} = p_1 L^i_1 + \mathbb{E}[p^i_{2,rdp} L^i_2], \tag{20}$$

where p_1 is the marginal price in the first stage from Equation (2).

Intuitively, the RDP, compared with the SP mechanism, should make the trading in Stage 2 less attractive for LCGs because they cannot have positive revenue in that market stage. One might infer that the RDP leads to higher efficiency than the SP. We show here that this might not always be the case, depending on the market characteristics. To do so, we first formulate the offer quantity adjustment of N homogeneous LCGs in equilibrium (see Online Appendix B.9 for the proof):

Theorem 4. *Consider an oligopoly of N homogeneous LCGs in a market with the RDP mechanism. With iid supply and demand forecast errors distributed as in Equation (14), and assuming that $\alpha_2 > (2N+1)\alpha_1/(2N)$, the offer* *quantity adjustment of any LCG i in equilibrium is*

$$d^i = \begin{cases} \dfrac{-B-\sqrt{B^2-4AC}}{2A} & \dfrac{-B-\sqrt{B^2-4AC}}{2A} > -\overline{\nu} \\ \dfrac{N}{\gamma(N+1)}-1 & \dfrac{-B-\sqrt{B^2-4AC}}{2A} \le -\overline{\nu} \ \& \ \gamma \ge \dfrac{N}{N+1} \\ -\overline{\nu} & \dfrac{-B-\sqrt{B^2-4AC}}{2A} \le -\overline{\nu} \ \& \ \gamma < \dfrac{N}{N+1}, \end{cases} \tag{21}$$

where

$$A \equiv \frac{\gamma^2}{4N^2\overline{\nu}}((2N+1)\alpha_1 - 2N\alpha_2) \tag{22}$$

$$B \equiv -\left[\frac{\gamma(1-\gamma)\alpha_1}{2N\overline{\nu}} + (\alpha_1+\alpha_2)\gamma^2\frac{(N+1)}{2N^2}\right] \tag{23}$$

$$C \equiv \frac{\alpha_1\gamma}{2N^2}(N-(2N+1)\gamma) - (\alpha_2/2-\alpha_1/4)\frac{\gamma^2}{N^2}\overline{\nu}. \tag{24}$$

An interesting observation is that, unlike the SP mechanism, d^i with the RDP is a function of LCG's forecast error. Another interesting observation is the special case of $\overline{\nu} = \overline{\epsilon} = 0$ corresponding to the scenario of perfectly predictable supply and demand. In such a scenario, Equation (21) reduces to

$$d^i = \begin{cases} \dfrac{N}{\gamma(N+1)} - 1 & \text{if } \dfrac{N}{N+1} < \gamma \\ 0 & \text{if } \dfrac{N}{N+1} \ge \gamma. \end{cases} \tag{25}$$

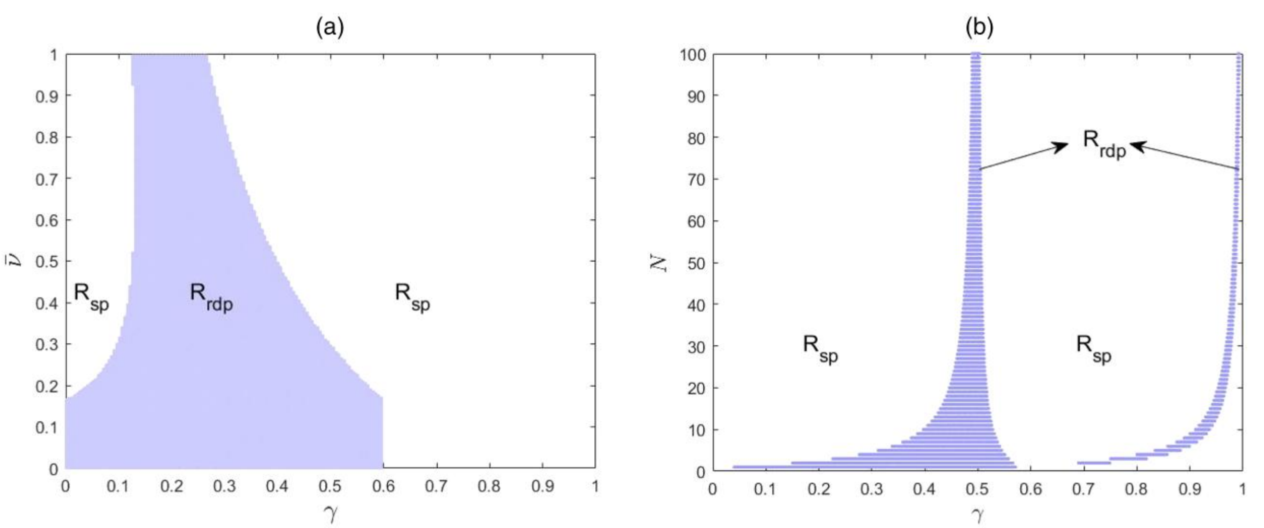
Equation (25) shows that in a deterministic setting, applying the RDP mechanism leads to the perfect competition outcome for any $\gamma \le N/(N+1)$. As N increases, the range of γ over which the perfect competition outcome is happening (i.e., $\gamma \in [0, N/(N+1)]$) increases. In the extreme case of perfect competition among LCGs, the RDP leads to maximum economic efficiency for any value of γ.

We use a numerical example to show how the impact of the RDP on market power mitigation compares to that of the SP mechanism. We compute and compare d^i from Theorem 4 with the one from Theorem 1 for identical market parameters and insert the values in Lemma 1 to obtain the corresponding DWL and compare market efficiencies. Figure 4 illustrates the regions in which each mechanism outperforms the other in terms of market power mitigation and efficiency. We set $\alpha_1 = 1$ and $\alpha_2 = 3$. In Figure 4(a), we study the case of $n = 1$ (monopoly) for any $\gamma \in [0,1]$ and $\overline{\nu} \in [0,1]$. In Figure 4(b), we fix $\overline{\nu} = 0.2$ and instead vary $N \in [1,100]$ to capture the impact of market power. Our observations are summarized as follows (see Online Appendix B.10 for supporting evidence).

Observation 3. Increasing the supply uncertainty of LCGs and increasing the total market share of LCGs favor the SP over the RDP. However, the relative performance of the RDP with respect to the SP is insensitive to the demand uncertainty, and there is no general trend with an increasing number of LCGs (depending on other factors in this case).

In summary, for a small to medium share of LCGs in the market, the RDP mechanism mitigates market power to a reasonable extent compared with the SP. Switching to the RDP in the early stages of the transition toward a high share of LCGs may be beneficial. However, with higher shares of LCGs, its efficiency degrades

Figure 4. (Color online) Economic Efficiency of the SP Versus the RDP as a Function of LCGs' Market Share (x-Axis), Forecast Errors (y-Axis, Left), and No. of LCGs (y-Axis, Right)



Notes. Shaded area: the RDP outperforms the SP (indicated by R_{rdp}). White area: the SP outperforms the RDP (indicated by R_{sp}). (a) $(\gamma, \overline{\nu})$ space, $n = 1$; (b) (γ, N) space, $\overline{\nu} = 0.2$.

quickly. From Sections 5 and 6, we find that as the market share of LCGs increases, the single-pricing mechanism will most likely outperform either of the dual-pricing mechanisms (TDP and RDP) in most cases. Thus, although these dual-pricing mechanisms have been useful for small to moderate shares of renewables when the larger share of LCGs appears, regulators would be advised to move away from them, possibly to the single-pricing mechanism. Choosing single pricing over dual pricing is aligned with the harmonization initiative (ACER 2020), which sets single pricing as the default, and dual pricing might be adopted only when justified. Importantly, our observation that none of the single- or dual-pricing mechanisms absolutely outperforms the others motivates us to develop the optimal pricing mechanism and evaluate its performance compared with others. This is done in the next section.

7. Optimal Pricing

Article 52 of the EU harmonization (EU 2017) emphasizes that imbalance prices should support competitive behavior among suppliers. Therefore, we develop an imbalance pricing mechanism that effectively removes the LCGs' incentives to exercise market power. This is defined below:

Definition 4 (Optimal Pricing). The optimal pricing (OP) is an imbalance pricing mechanism, in which the price applied to the LCG i in the imbalance market is given by

$$\textbf{OP}: \quad p^i_{2,op} = p_1 + \alpha_2 H_2 - \alpha_1 L^i_1. \tag{26}$$

The OP mechanism formulated above can be classified as a single-pricing mechanism because it does not differentiate between selling and buying imbalance prices. Upon comparing Equation (26) to Equation (8), it becomes apparent that the OP mechanism can be considered as an adjustment to the SP mechanism. The following theorem clarifies why we call it optimal (see Online Appendix B.11 for the proof).

Theorem 5. *In a two-stage electricity market described in this paper, any LCG i behaves as if in a perfect competition setting, choosing $d^i = 0$, if the OP mechanism (Definition 4) is used as the imbalance pricing mechanism. With this imbalance pricing mechanism, full economic efficiency in a deterministic setting (i.e., DWL = 0) and the maximum economic efficiency (i.e., $\mathbb{E}[DWL_{op}] = 1/2(\alpha_2 - \alpha_1)(Var(\epsilon) + \sum_{j=1}^N \sum_{i=1}^N \gamma^j \gamma^i Cov(\nu^j, \nu^i)))$ in a probabilistic setting are achieved.*

Theorem 5 implies that the imbalance pricing mechanism, if designed properly, can substantially deter the exercise of market power. As stated in Theorem 5, the OP mechanism achieves the full efficiency in a deterministic setting and the maximum possible in a probabilistic setting; the remaining DWL corresponds to forecast errors, which are exogenous in our setting.

We complement the theoretical results in Theorem 5, with simulation results. We evaluate and compare the performance of the OP mechanism with others. This allows a fair and simultaneous comparison of the efficiency and market behavior of all four mechanisms. For our numerical examples, we consider two LCGs competing in the market to serve the inelastic demand, and the leftover is served by HCGs. We choose $\alpha_1 = 1$ and $\alpha_2 = 5$, leading to the ratio of $\alpha_2/\alpha_1 = 5$, which is a practical value, used also by Ito and Reguant (2016), taken from the Iberian market. We further assume that ϵ and ν are all i.i.d. random variables with a uniform distribution in $[-0.1, 0.1]$. We include two settings, the homogeneous setting and the heterogeneous setting, both described below.

7.1. The Homogeneous Setting

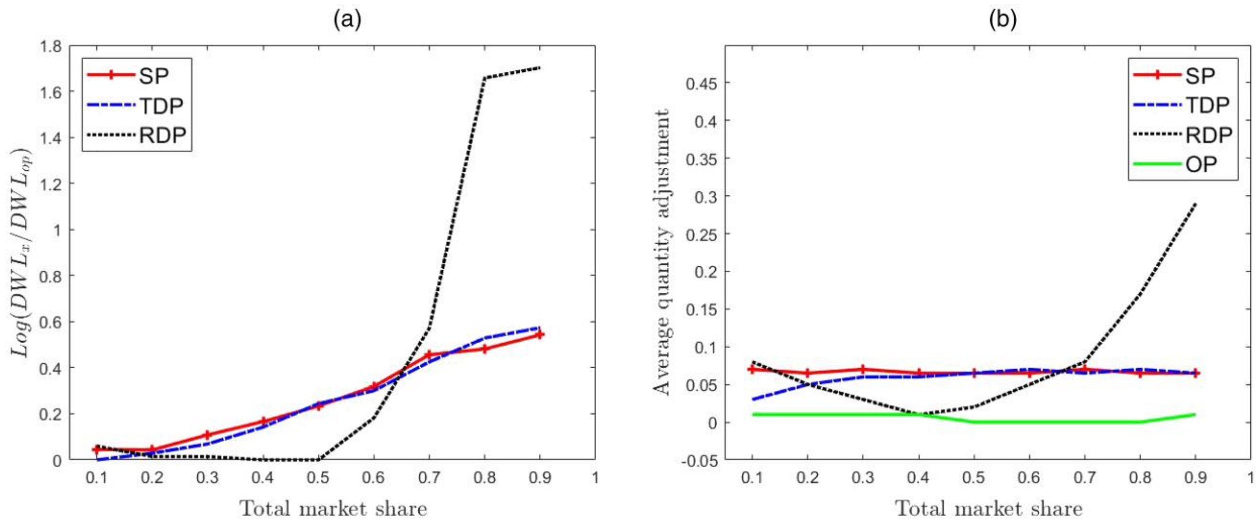
Here, we assume that LCGs are homogeneous and that the total share of LCGs in the supply mix is increasing. Figure 5(a) shows the logarithm of the ratio of the DWL in any of the three mechanisms with respect to that of the OP mechanism as a function of the total share of LCGs. Note that, unlike a deterministic setting, in a probabilistic setting the DWL of the OP mechanism is not zero. As shown in this graph, the OP mechanism substantially outperforms the others. For example, the OP mechanism reduces the DWL by almost 80% compared with the SP or the TDP in a homogeneous scenario and when the total share of LCGs is 90%. This percentage reduction is even more substantial (450%) when compared with the RDP. Another important observation is that the RDP can also lead to an efficient outcome if the market share of LCGs (γ) is low/medium and it starts to deteriorate drastically as γ keeps increasing. The SP and the TDP are outperformed by the RDP for a low/medium γ, but they outperform the RDP for a large γ.

To analyze the market behavior of LCGs in different mechanisms, we illustrate the average of the absolute quantity adjustments of both LCGs in the same scenarios in Figure 5(b). This graph corroborates our analytical observations that the market behavior under the SP mechanism in a homogeneous scenario is independent of the market share, and the TDP is not much advantageous if demand uncertainty is not large.

7.2. The Heterogeneous Setting

Here, we create heterogeneity in the production capacities of the two LCGs by varying their expected productions. Without loss of generality, we assume that LCG 1 has a smaller expected production (denoted by γ^1) than LCG 2 (denoted by γ^2). The x-axis presents the ratio between these expected productions, that is, γ^1/γ^2,

Figure 5. (Color online) Comparing the Performance of Different Mechanisms; the Homogeneous Scenario



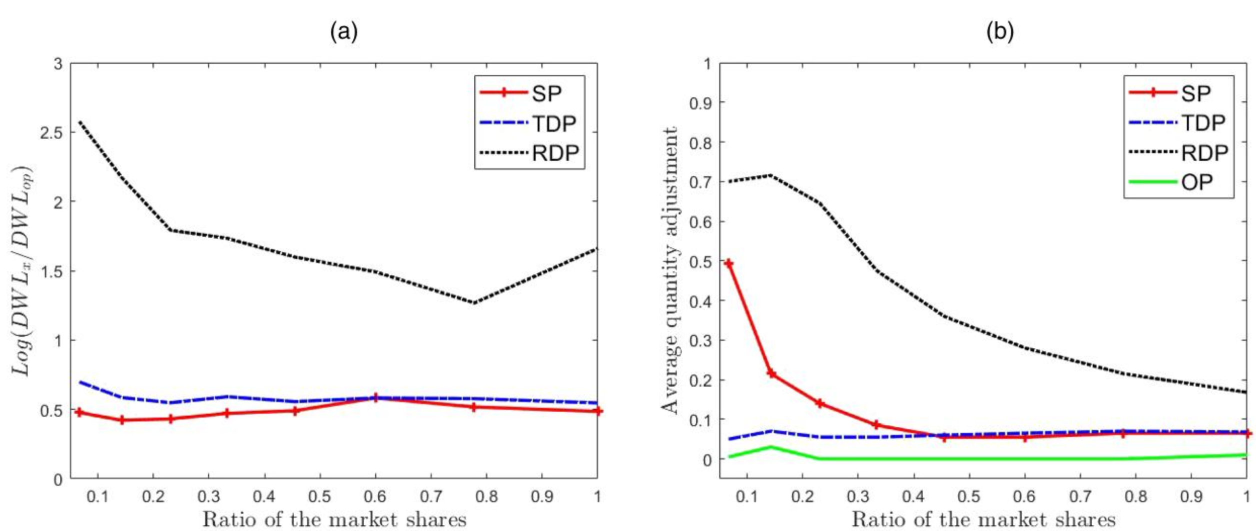
Notes. The x-axis is γ and y-axes are $Log(DWL/DWL_{op})$ (left plot) and $(|d^1| + |d^2|)/2$ (right plot). (a) DWL; (b) average deviation.

fixing their total expected production to $\gamma^1 + \gamma^2 = 0.8$. As we move on along the x-axis, the heterogeneity decreases. The extreme case of $x = 1$ represents the homogeneous scenario. We then repeat the same types of examples as in Figure 5.

Figure 6 suggests that heterogeneity does not have much effect on the DWL in the SP and the TDP, but it has major effects under the RDP. Market behavior in the optimal pricing and TDP also seems to be insensitive to heterogeneity. In contrast, even though the DWL is not affected much in the SP mechanism, we still observe considerable deviations from bidding the expected production (i.e., $d^1 = d^2 = 0$). Indeed, in the SP mechanism, in a very heterogeneous environment, market players deviate from bidding their expected production, but in the opposite directions. This is why the DWL remains unaffected by heterogeneity, unlike the market behavior.

In summary, the theoretical results and numerical examples show the optimality and the substantial efficiency gain of the OP mechanism with respect to the other pricing mechanisms. We believe that the OP mechanism is new, and as such, it is untested in practice (see Section 8). In Online Appendix C, we provide additional results to show that our conclusions are not sensitive to values of important factors such as the type of distribution, skewness, and correlations of supply or demand predictions as well as the uncertainty in slope estimations.

Figure 6. (Color online) Comparing the Performance of Different Mechanisms; the Heterogeneous Scenario



Notes. The x-axis is γ^1/γ^2, and the y-axes are $Log(DWL/DWL_{op})$ (left plot) and $(|d^1| + |d^2|)/2$ (right plot). (a) DWL; (b) average deviation.

8. Discussion

As stated in Section 2, imbalance pricing is not the only solution to improve economic efficiency in the presence of market power. Three alternatives are (1) introducing more market stages, (2) storage, and (3) virtual bidders. Each of these solutions has its own strengths and weaknesses. Here, we discuss the practical concerns of these solutions as well as ours.

The Allaz-and-Vila effect suggests that market power inefficiency can be reduced by adding more forward markets and involving market players in parallel participation. Regulators have been aware of this since the early 1990 s, attempting to mandate additional forward trading in illiquid and incomplete markets, but with limited success. However, this approach conflicts with liberalized principles of allowing market participants to hedge their risks as they see fit. It also complicates the market structure, making clearing each stage more time-consuming. Achieving full efficiency would require an infinite number of stages, which is impractical. Moreover, maintaining liquidity in all added stages is challenging, as seen in electricity markets with low liquidity in introduced intraday markets (Weber 2010).

Storage can help mitigate market power and enhance economic efficiency by shifting energy demand to more affordable periods. However, practical challenges hinder its widespread adoption. Scalability is an issue, with certain technologies like batteries being financially unattractive in electricity markets (Fares and Webber 2014, Zhou et al. 2015). Land limitations restrict the viability of technologies like pumped hydro. Additionally, integrating storage into existing market structures requires new pricing mechanisms because of the unique nature of storage technologies. Furthermore, the effectiveness of introducing storage to mitigate market power and enhance economic efficiency may not meet expectations and could potentially even reduce efficiency (Sioshansi 2014).

Virtual bidders, as used in some electricity markets such as Pennsylvania-New Jersey-Maryland (PJM), are traders with no physical assets who can arbitrage between the first and the second stage of the market. With a sufficiently large number of virtual bidders, the prices of the first and the second stage of the market converge. Multiple studies show the benefits of using virtual bidding (see, e.g., Hogan 2016 and Mather et al. 2017). In a competitive market, virtual bidding improves economic efficiency. However, virtual bidding is shown not to be a good solution to mitigate market power (Celebi et al. 2010). In the presence of market power, Ito and Reguant (2016) showed that virtual bidding is not necessarily welfare-enhancing, reducing consumer costs but increasing deadweight loss (as large as 2x). This is because generators tend to withhold production entirely in the presence of many virtual bidders. Additionally, there are several other practical concerns about virtual bidding. In 2015, PJM released a white paper (S&P Global 2018), bringing evidence that virtual bidding does not necessarily result in more efficient market operation (Hogan 2016). Multiple reasons were raised as the practical problems with virtual bidding, including the exacerbation of the congestion problem, the complexity of the market clearance with a large number of virtual bidders, and more uncertainty in real time because of the lack of generation assets of virtual bidders and creating large uncertainty in the final deliveries of electricity, which affects the system reliability. Accordingly, the Federal Energy Regulatory Commission (FERC) reduced the eligible nodes for virtual bidding by 87.9% (S&P Global 2018). Although virtual bidding has become more limited for those who have implemented it, because of these concerns, many markets such as most European markets have remained entirely uninterested in even allowing virtual bidding.

Considering the implementability of our proposed pricing mechanism, an immediate observation is that the optimal pricing (OP) with respect to dual-pricing mechanisms (here, TDP and RDP) has less complexity because the OP is a continuous mechanism and the latter involves integer variables (to distinguish selling and buying). The OP mechanism is more complex than the SP mechanism, but only with an additional continuous term. The variables involved in the optimal price formation (the price in Stage 1, the LCGs' forward quantity offer, and the HCG's quantity in the real-time market) are all already being used in single- or dual-pricing mechanisms and hence, implementable. Similarly, knowing the type of technologies (here, LCGs) needed in the proposed optimal pricing is practically sound, as is used in some types of dual pricing. In particular, in Europe at least, the transparency directives by the regulators make almost all relevant physical system data available to the market, often within minutes (e.g., see https://www.bmreports.com for the case of the United Kingdom). Moreover, because imbalance price determination and contract settlements are usually undertaken by an independent agency or system operator, they necessarily have full visibility of all contracts, including the technology and quantities for settlement. Thus, all of the information required by the balancing and settlement agency to compute the optimal price is already part of their routine data processing. It remains a question, though, whether regulators put the effort to identify the best imbalance pricing algorithm or, rather, settle with suboptimal mechanisms because of legacy.

Although we discuss the practical concerns of different solutions for market power mitigation here, we do not argue that imbalance pricing, and in particular, the optimal pricing, is the only solution. Instead, we find great potential in this solution, and we advocate embracing this solution wherever possible. In reality, we project that a combination of these solutions could

Table 1. Comparing Different Imbalance Pricing Mechanisms

	$p^i_{2,\cdot}$	Demand uncertainty ↑	Supply uncertainty ↑	LCG Heterogeneity	Low share of LCGs	high share of LCGs
SP	$p_1 + \alpha_2 H_2$	✓	✓✓	✓✓	✓	✓
TDP	$p_1 + \alpha_2 H_2 \mathbb{I}_{(H_2 L^i_2 < 0)}$	✓✓	✓	✓✓	✓	✓
RDP	$(p_1 + \alpha_2 H_2)\mathbb{I}_{L^i_2 < 0}$	✓	✓✓	✗	✓✓	✗
OP	$p_1 + \alpha_2 H_2 - \alpha_1 L^i_1$	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓

Notes. The number of ✓ per item reflects the suitability of the pricing mechanisms in the presence of that item. When ✗ is used for an item and a pricing mechanism, it means that the corresponding pricing mechanism is not a good fit in the presence of that item.

create the best solution, depending on market characteristics, technology maturity level, human interactions, etc.

9. Conclusions

This research has been motivated by one of the major challenges in electricity markets because of the energy transition. Many markets are moving toward a technology mix in which low-cost renewable generators will have predominant market shares, whereas the high-cost conventional gas generators may be left with little more than reserve and balancing services. Sunar and Birge (2019) have shown that real-time imbalance prices can have substantial and counterintuitive impacts on market behavior and the reliability of supply in electricity markets. In this paper, we focus on how real-time imbalance pricing can accommodate the fundamental changes in the energy mix.

There are two main types of pricing mechanisms used in imbalance markets, single pricing and dual pricing, where the latter is an attempt to further mitigate the strategic market behavior. Recently, the European Commission has called for more clarity on the effectiveness of the chosen pricing mechanisms and has proposed greater harmonization. Interestingly, in this call, markets are encouraged to use the single pricing and only move to dual pricing if they can justify the need. This raises several important questions. How do they perform with respect to each other as the share of renewables keeps increasing? How well can they handle the possible strategic behavior of renewables when they are dominant? Is there a better mechanism apart from the existing ones to deal with the technological and market power changes in the energy supply mix?

Motivated by these research questions, we conducted an analytical review and compared the performance of two important dual-pricing mechanisms, which we refer to as the typical dual pricing (TDP) and the renewable-based dual pricing (RDP) with respect to the single-pricing (SP) mechanism in terms of economic efficiency. Additionally, we proposed a new optimal pricing (OP) mechanism and showed for the first time that an optimal pricing mechanism for the imbalance market can be designed to create a competitive outcome in a market with dominant low-cost generators (LCG), potentially mostly renewables.

We used analytical and simulation approaches in complementary ways. Table 1 summarizes our comparison results by presenting the suitability of each market pricing mechanism as a function of different market characteristics. Some highlights of our observations are as follows. Generally speaking, we showed that our proposed OP mechanism is the most efficient one. We also found that the RDP was effective in improving efficiency for small to moderate market shares of renewables but less so, compared with the SP, as market share increases. The TDP can outperform the SP only in the presence of significant demand uncertainty. These comparative analyses suggest that for small to moderate market shares of renewables, switching to the RDP (for small demand uncertainty) and the TDP (for significant demand uncertainty) might be sufficiently effective. However, as the market share of renewables increases beyond that, there will be no incentive to switch to such heuristics because the SP might outperform them. Instead, there will be a need to consider the new optimal pricing, perhaps along the lines of the optimal pricing formulated in this work. We also found that heterogeneity among LCGs has a mixed effect. The only one that seems to be highly impacted by heterogeneity is the RDP mechanism.

We used simulations to complement our analytical findings by (1) enabling a simultaneous comparison and quantification of pricing mechanisms in one shot and under different market conditions, (2) ensuring the robustness of the results with respect to multiple factors such as the cross-correlation and the skewness of the forecast errors, and (3) further investigating the impact of heterogeneity. Most notably, our simulation results revealed that at a large share of LCGs, the OP mechanism can provide a substantial reduction in the deadweight loss compared with other mechanisms (80%, 90%, and 450% when compared with the SP, the TDP, and the RDP, respectively).

Our analytical model is very stylized, but not more so than most of the theoretical research on this theme. In practice, reality will bring many complications. Therefore, we included a detailed discussion on the practical concerns of using imbalance pricing design and those of other promising solutions, such as virtual bidding, storage, and adding more market stages. Although in this paper we focused on the impact of one solution (imbalance pricing design), we expect that a combination of multiple solutions will also soften the opportunities for price manipulation, and perhaps more significant regional interconnections will reduce market concentrations. Nevertheless, we expect the principles in this paper to motivate regulators to consider more efficient imbalance pricing mechanisms as important levers for economic efficiency in the near future.

An important consideration of the urgency of this transition is to recognize that wholesale power is a sequence of separate delivery periods (hourly, or less), which are distinct markets that clear according to the fundamentals for each period. For example, in the case of Germany with 40% of overall solar and wind power, there were about 100 hours of the year that the share of renewables in those hours was higher than 80% of the total demand (International Energy Agency 2020). Therefore, waiting for the overall market share to reach a critical point oversimplifies the situation; the critical points materialize gradually and selectively by hour and by season. This suggests a proactive approach to the implementation of optimal imbalance pricing in anticipation of it gradually becoming more appropriate for a greater proportion of the days and year.

We identify multiple directions for future work. For example, we assume a linear and symmetric marginal cost for HCGs. It would be interesting to see how the results will change in the presence of nonlinear and nonsymmetric marginal costs. Another interesting extension is to consider a combination of multiple solutions, for example, imbalance pricing and virtual bidding or storage. Along this line, the flexibility in demand, through demand response and storage and demand elasticity, is worth considering as another solution. Finally, in this study, we consider a market in isolation. It is interesting to also investigate how the imbalance pricing can affect and be affected in connected markets in the presence of congestion between them.

References

ACER (2020) Decisions of the agency for the cooperation of energy cooperation of energy regulators No. 18-2020. Accessed December 10, 2022, https://documents.acer.europa.eu.

Al-Gwaiz M, Chao X, Wu OQ (2017) Understanding how generation flexibility and renewable energy affect power market competition. *Manufacturing Service Oper. Management* 19(1):114–131.

Allaz B (1992) Oligopoly, uncertainty and strategic forward transactions. *Internat. J. Indust. Organ.* 10(2):297–308.

Allaz B, Vila JL (1993) Cournot competition, forward markets and efficiency. *J. Econom. Theory* 59(1):1–16.

Anderson RW (1991) Futures trading for imperfect cash markets: A survey. In: Phlips L, ed., *Commodity, Futures and Financial Markets. Advanced Studies in Theoretical and Applied Econometrics* (Springer, Dordrecht, Netherlands), 207–248.

Anderson SC (2004) Analyzing strategic interaction in multi-settlement electricity markets: A closed-loop supply function equilibrium model. PhD thesis, Harvard University. Accessed December 10, 2022, https://sites.hks.harvard.edu/crump/papers/Anderson_thesis.pdf.

Biggar DR, Hesamzadeh MR (2014) *The Economics of Electricity Markets* (Wiley-IEEE Press, West Sussex, UK).

Birge JR, Hortaçsu A, Pavlin JM (2017) Inverse optimization for the recovery of market structure from market outcomes: An application to the MISO electricity market. *Oper. Res.* 65(4):837–855.

Borenstein S (2002) The trouble with electricity markets: Understanding California's restructuring disaster. *J. Econom. Perspect.* 16(1):191–211.

Bunn DW, Kermer SO (2021) Statistical arbitrage and information flow in an electricity balancing market. *Energy J.* 42(5):19–40.

Celebi M, Hajos A, Hanser PQ (2010) Virtual bidding: The good, the bad and the ugly. *Electr. J.* 23(5):16–25.

Chen R, Paschalidis IC, Caramanis MC, Andrianesis P (2019) Learning from past bids to participate strategically in day-ahead electricity markets. *IEEE Trans. Smart Grid* 10(5):5794–5806.

Clò S, Fumagalli E (2019) The effect of price regulation on energy imbalances: A difference in differences design. *Energy Econom.* 81:754–764.

Cormen TH, Leiserson CE, Rivest RL, Stein C (2022) *Introduction to Algorithms*, 4th ed. (MIT Press, Cambridge, MA).

EC (2016) Overview of European electricity markets – METIS technical note T4. Accessed December 10, 2022, https://ec.europa.eu.

EU (2017) Commission regulation (EU) 2017/2195. Accessed December 10, 2022, https://eur-lex.europa.eu.

Fares RL, Webber ME (2014) A flexible model for economic operational management of grid battery energy storage. *Energy.* 78(1): 768–776.

FERC (2014) *Staff Analysis of Energy Offer Mitigation in RTO and ISO Markets*. Accessed December 10, 2022, https://www.ferc.gov.

Güler T, Gross G, Litvinov E, Coutu R (2010) On the economics of power system security in multi-settlement electricity markets. *IEEE Trans. Power Syst.* 25(1):284–295.

Hesamzadeh M, Biggar D, Bunn D, Moiseeva E (2020) The impact of generator market power on the electricity hedge market. *Energy Econom.* 86:1–15.

HMSO (2020) Energy white paper. UK government. Accessed June 7, 2023, https://www.gov.uk/government/publications/energy-white-paper-powering-our-net-zero-future.

Hogan WW (2016) Virtual bidding and electricity market design. *Electr. J.* 29(5):33–47.

International Energy Agency (2020) Global Energy Review – 2020. Accessed December 10, 2022, https://www.iea.org.

IRENA (2017) Adapting market design to high shares of variable renewable energy. Accessed December 10, 2022, https://www.irena.org.

Ito K, Reguant M (2016) Sequential markets, market power, and arbitrage. *Amer. Econom. Rev.* 106(7):1921–1957.

Jha A, Wolak FA (2019) Can financial participants improve price discovery and efficiency in multi-settlement markets with trading costs? Working Paper No. 25851, National Bureau of Economic Research.

Jiang B, Fei Y (2011) Dynamic residential demand response and distributed generation management in smart microgrid with hierarchical agents. *Energy Procedia* 12:76–90.

Jiang DR, Powell WB (2015) Optima hour-ahead bidding in the real-time electricity market with battery storage using approximate dynamic programming. *INFORMS J. Comput.* 27(3):525–543.

Kamat R, Oren SS (2004) Two-settlement systems for electricity markets under network uncertainty and market power. *J. Regul. Econom.* 25(1):5–37.

Kim JH, Powell WB (2011) Optimal energy commitments with storage and intermittent supply. *Oper. Res.* 59(6):1347–1360.

Liberopoulos G, Andrianesis P (2016) Critical review of pricing schemes in markets with non-convex costs. *Oper. Res.* 64(1):17–31.

Lisi F, Edoli E (2018) Analyzing and forecasting zonal imbalance signs in the Italian electricity market. *Energy J.* 39(5):1–20.

Löhndorf N, Minner S (2010) Optimal day-ahead trading and storage of renewable energies—An approximate dynamic programming approach. *Energy Syst.* 1(1):61–77.

Luo ZQ, Pang JS, Ralph D (1996) *Mathematical Programs with Equilibrium Constraints* (Cambridge University Press, Cambridge, UK).

Mahenc P, Salanié F (2004) Softening competition through forward trading. *J. Econom. Theory* 116(2):282–293.

Mather J, Bitar E, Poolla K (2017) Virtual bidding: Equilibrium, learning, and the wisdom of crowds. *IFAC-PapersOnLine.* 50(1):225–232.

Matsumoto T, Bunn DW, Yamada Y (2021) Mitigation of the inefficiency in imbalance settlement designs using day-ahead prices. *IEEE Transactions on Power Systems*. Accessed December 10, 2022, https://ieeexplore.ieee.org.

Morales JM, Conejo AJ, Madsen H, Pinson P, Zugno M (2014) *Balancing markets. Integrating Renewables in Electricity Markets* (Springer, New York), 101–136.

National Grid ESO (2021) *Capacity market published round results*. Accessed December 10, 2022, https://www.emrdeliverybody.com.

Nordic Balancing Model (2021) Single-price model – Timeline and milestones. Accessed December 10, 2022, https://nordicbalancingmodel.net.

Ofgem (2000) The market abuse licence condition for generators: A decision document. Accessed December 10, 2022, https://www.ofgem.gov.uk.

Ofgem (2014a) Electricity balancing significant code review. Accessed December 10, 2022, https://www.ofgem.gov.uk.

Ofgem (2014b) Electricity balancing significant code review – Final policy decision. Accessed December 10, 2022, https://www.ofgem.gov.uk.

Petruzzi NC, Dada M (1999) Pricing and the newsvendor problem: A review with extensions. *Oper. Res.* 47(2):183–194.

Peura H, Bunn DW (2021) Renewable power and electricity prices: The impact of forward markets. *Management Sci.* 67(8):4772–4788.

PWC (2018) *15th global power utilities survey*. Accessed December 10, 2022, https://www.pwc.com.

S&P Global (2018) FERC accepts PJM proposal to limit virtual transaction bidding points. Accessed December 10, 2022, https://www.spglobal.com/marketintelligence/en/news-insights/trending/5_3zsn9vjecn7diw2dmcia2.

Secomandi N (2010) Optimal commodity trading with a capacitated storage asset. *Management Sci.* 56(3):449–467.

Sioshansi R (2014) When energy storage reduces social welfare. *Energy Econom.* 41:106–116.

Su CL (2007) Analysis on the forward market equilibrium model. *Oper. Res. Lett.* 35(1):74–82.

Sunar N, Birge JR (2019) Strategic commitment to a production schedule with uncertain supply and demand: Renewable energy in day-ahead electricity markets. *Management Sci.* 65(2):714–734.

Sunar N, Swaminathan JM (2021) Net-metered distributed renewable energy: A peril for utilities? *Management Sci.* 67(11):6716–6733.

van der Veen RA, Abbasy A, Hakvoort RA (2012) Agent-based analysis of the impact of the imbalance pricing mechanism on market behavior in electricity balancing markets. *Energy Econom.* 34(4):874–881.

Vandezande L, Meeus L, Belmans R, Saguan M, Glachant JM (2010) Well-functioning balancing markets: A prerequisite for wind power integration. *Energy Policy* 38(7):3146–3154.

Weber C (2010) Adequate intraday market design to enable the integration of wind energy into the European power systems. *Energy Policy* 38(7):3155–3163.

Willis P, Altozano H (2016) Spanish authority fines Iberdrola 25m Euros in first REMIT market manipulation infringement decision. Accessed December 10, 2022, https://www.twobirds.com.

Wu OQ, Kapuscinski R (2013) Curtailing intermittent generation in electrical systems. *Manufacturing Service Oper. Management* 15(4):578–595.

Yao J, Oren S, Adler I (2007) Two-settlement electricity markets with price caps and Cournot generation firms. *European J. Oper. Res.* 181(3):1279–1296.

Zarnikau J, Woo C, Zhu S, Baldick R, Tsai CH, Meng J (2019) Electricity market prices for day-ahead ancillary services and energy: Texas. *J. Energy Markets* 12(1):43–74.

Zhang D, Xu H (2013) Two-stage stochastic equilibrium problems with equilibrium constraints: Modeling and numerical schemes. *Optimization* 62(12):1627–1650.

Zhou YH, Scheller-Wolf A, Secomandi N, Smith S (2015) Electricity trading and negative prices: Storage vs. disposal. *Management Sci.* 62(3):880–898.