Market Design Options for Scarcity Pricing in European Balancing Markets

Anthony Papavasiliou, *Senior Member, IEEE,* Gilles Bertrand, *Student Member, IEEE*

***Abstract*—The European balancing market is undergoing radical transformation through numerous market design initiatives. These initiatives aim at improving geographical coordination among European transmission system operators, and better positioning the European system for integrating renewable resources through short-term operational efficiency and long-term investment in flexible resources. However, the European design is characterized by a missing market for real-time reserve capacity, that has been inherited from a failure to recognize the central role of real-time operations as the spot market of the electric power industry. This missing market undermines the valuation of reserve capacity, and the back-propagation of price signals to forward reserve markets that can support investment in reserves. The goal of the present paper is to develop a methodology that exposes the implications of this missing market. The methodology relies on analytical insights that can be derived under an assumption of price-taking behavior. These insights are validated by a simulation model which represents the European balancing market as a Markov Decision Process. The simulation model is used for validating the analytical insights and testing the ability of various balancing market design options to back-propagate the real-time value of reserve to forward reserve markets.**

***Index Terms*—reserves, operating reserve demand curves, scarcity pricing, balancing, Markov Decision Processes**

I. Introduction

A. Trading of Energy and Reserve in European Markets

The European balancing market has been undergoing significant transformation in recent years, due to various coordinated balancing initiatives. These initiatives include integrated market clearing platforms for replacement reserve (Trans European Replacement Reserves Exchange, abbreviated "TERRE"), tertiary reserve (Manual Activated Reserves Initiative, abbreviated "MARI"), secondary reserve (Platform for the International Coordination of Automated Frequency Restoration and Stable System Operation, abbreviated "PICASSO"), and primary control (The International Grid Control Cooperation, abbreviated "IGCC").

Traditionally, European system operations have been segmented geographically and functionally. Geographical segmentation refers to the fact that each European country is commonly operated by a single, or a handful, of transmission system operators (referred to as TSOs hereafter). Functional separation refers to the fact that the trading of energy and reserve capacity[1] is not coordinated.

European TSOs are responsible for procuring reserve capacity, and for deploying reserve capacity in real time. Day-ahead procurement of reserve capacity can take place before, during, or after, the clearing of the day-ahead energy exchange, depending on the country [1], [2]. The operation of the European day-ahead and intraday market is conducted by Nominated Electricity Market Operators (NEMOs), which are separated functionally from TSOs. NEMOs are responsible for trading energy in the day-ahead and intraday time frame.

Balancing, in European parlance, refers to the trading of real-time energy. The entities that trade energy in real time are the so-called "Balancing Responsible Parties" (abbreviated BRPs hereafter) and "Balancing Service Providers" (abbreviated BSPs hereafter). BRPs are essentially portfolio owners that find themselves producing or consuming more energy in real time than they have originally traded, and are therefore essentially *price-inelastic* buyers or sellers of real-time energy. BSPs, on the other hand, refer to owners of assets that can offer reserve services. BSPs submit offers for balancing energy in the real-time balancing market, they can therefore be viewed as *price-elastic* suppliers or consumers of real-time energy. Upwards balancing refers to the selling of real-time energy by BSPs, downwards balancing refers to the procurement of real-time energy by BSPs. By selling reserve capacity in day-ahead reserve markets, BSPs essentially commit to bidding at least the amount of capacity that they have sold in the day ahead to real-time balancing markets. Each BSP must be attributed to a unique BRP portfolio, as foreseen in article 18(4).d of the European Balancing Guideline [3].

From an economic standpoint, the essential difference between BRPs and BSPs is price elasticity in the real-time energy market, and the ability of the latter to provide reserve capacity. The functional separation of BSPs and BRPs in system operations, however, has been misunderstood as a license to introduce a market distortion, whereby the two are paid differently for trading the same product of real-time energy. Concretely, BRPs are settled for their real-time energy deviations at a so-called imbalance price, whereas BSPs are settled for their real-time deviations at a so-called balancing price[2]. The two may be different, even though they apply to the same product, real-time energy. Furthermore, it is not clear that the balancing platforms mentioned in the opening paragraph of this text will be coherent in terms of setting a price for real-time energy (in the sense that balancing energy from different platforms may be priced differently).

Anthony Papavasiliou is with Center for Operations Research and Econometrics, Université catholique de Louvain; Email: anthony.papavasiliou@uclouvain.be; Phone: +32 10 474325.

[1]We ignore transmission capacity in the present paper.

[2]Although certain European national balancing markets currently rely on pay-as-bid settlements, the aforementioned integrated EU balancing platforms that are being put in place will be trading at a uniform balancing price.

Compared to US-style pools, therefore, European markets differ along the following major axes: (i) There is no co-optimization of energy and reserve in the day ahead, the two are traded in separate auctions. Energy auctions are operated by NEMOs. Reserve auctions are operated by TSOs. (ii) Energy is traded in real time by balancing platforms which are operated by TSOs. The counterparties in the trading of real-time energy are BSPs and BRPs. (iii) There is a lack of a unique price signal for real-time energy. (vi) Reserve capacity is *not* traded in real time in the European market. This creates challenges in the valuation of reserve, as we discuss next.

B. Motivation of Our Paper

The accurate valuation of energy and reserve capacity is an increasingly crucial function of real-time markets in a regime of large-scale renewable energy integration. Operating reserve demand curves (ORDCs) [4] have been proposed as a means for achieving this important goal. ORDC adders are computed on the basis of available reserve capacity in the system. As the amount of reserve capacity in the system decreases, ORDC adders increase, and reflect the value of reserve in a tight system. As the available reserve capacity increases, ORDC adders dissipate, since the system is not experiencing scarcity.

ORDC adders have been adopted in Texas [5], and their adoption is moving forward in PJM [6]. The Electricity Balancing Guideline of the European Commission, which is the reference text for European balancing legislation (and which we will refer to as "EBGL" hereafter), introduces the legal possibility of implementing ORDC adders by referring to the mechanism as a "scarcity pricing function" in article 44(3) of the legislation [3].

Belgium has made steps in advancing the implementation of scarcity pricing. A series of preliminary analyses have focused on quantifying the possible implications of the mechanism for resources that can provide reserve to the system [7], [8]. The Belgian system operator and regulator [9] has collaborated with the authors towards computing and publishing scarcity adders based on the "available reserve capacity" (ARC) of the system. These adders are computed for every quarter of the day, and published one day after operations.

In US parlance, the ORDC adder effectively sets the real-time price for reserve capacity. Since prices in energy and reserve have to be consistent in equilibrium (and this equilibrium is respected automatically in a co-optimization of energy and reserve), the ORDC adder also uplifts the real-time energy price. These first principles translate to the following market design proposals for implementing scarcity pricing in the EU market design [1]:

- **Market design proposal 1**: the introduction of a scarcity adder to the imbalance price.
- **Market design proposal 2**: the application of the same adder to the balancing energy price.
- **Market design proposal 3**: the implementation of an EU real-time market for reserve capacity (equivalently, a market for "reserve imabalances", in the same way that we operate a market for energy imbalances), which is a missing market in the existing EU balancing design.

Market design proposal 3 means that the scarcity adder should (1) apply to BSP capacity that is not activated, (2) apply to free bids that are available in real time even if they have not sold reserve capacity in the day ahead, and (3) apply for buying back reserve capacity that has been activated as upward balancing energy and is no longer available as reserve capacity in real time.

Justifying these three market design changes (especially the second and third) to stakeholders with quantitative models has been challenging, as we outline below. The present paper is an attempt to develop an analytical and simulation framework towards advancing this goal.

C. Existing Modeling Frameworks

The intuitive economic arguments of why we need the three aforementioned market design changes are the following:

- **Economic principle 1**: Real-time energy is a unique product, therefore the buyer and seller should exchange it at the same price.
- **Economic principle 2**: If we put in place a real-time market for reserve capacity, then agents will only sell reserve capacity in forward markets at the value that they would need to buy it back in real time. This second principle is especially crucial, since it allows the value of reserve capacity to back-propagate into forward reserve auctions, and send the signal to investors that the market can support investments in reserve capacity.

In previous analysis [1], stochastic equilibrium has been used as our quantitative method of choice for representing the back-propagation effect quantitatively. The stochastic equilibrium framework that we have developed, which has originally been applied in the context of investment [10], [11], reveals the strengths and weaknesses of different market design choices in back-propagating the value of reserve to forward reserve auctions. However, the stochastic equilibrium framework encountered an immediate weakness from the outset during discussions with stakeholders: it embeds economic principle 1, meaning that the model assumes a unique market for real-time energy, and therefore a unique price for real-time energy. This assumption contradicts the practice of using imbalance prices for BRP settlement that are different from balancing prices for BSP settlement. To put it differently: whereas stochastic equilibrium can be used for understanding the effect of certain market design choices on the back-propagation of reserve prices to forward markets, it cannot be used for assessing the validity of different mixtures of BSP and BRP settlement on this back-propagation.

An alternative model that is developed in this paper is the representation of the balancing market as a Markov Decision Process (MDP). Our approach is inspired by a growing body of work on the application of agent-based models to the analysis of electricity markets. In early work on this topic, Bunn and coauthors [12], [13] analyze the effect of a change of design in the England and Wales market. In recent work, with the broader use of Reinforcement Learning techniques such as Q-Learning [14], researchers have applied MDPs [15], [16] in more complex settings. However, these classical

Reinforcement Learning techniques are inefficient for high-dimensional problems because they rely on the discretization of the state and action space. This problem has been overcome recently by the development of deep-learning [17], [18]. As we discuss in section II-A, our problem is low-dimensional, and therefore we rely on the standard Q-learning algorithm [14].

In the context of our analysis, we consider BRPs and BSPs as agents that engage in trade in a balancing market, and develop trading strategies *given* different market design options. We then test the ability of agents to infer the value of the reserve capacity that they offer to the market under different market design choices, and thus the ability of different market design choices to back-propagate the value of this reserve in forward reserve markets.

The MDP framework offers powerful modeling flexibility. However, it is difficult to extract conclusions regarding first principles, since one is limited to observing the outcome of a simulation, without necessarily gaining insights about the role of a market design in driving a certain outcome. For this reason, we supplement our MDP-based market simulation framework with an analytical characterization of the best response of market agents to different balancing market design choices under an assumption of perfect competition. The MDP simulation framework is then used for providing tangible evidence for the behavior that the analytical mathematical framework predicts, which can be valuable for discussions with stakeholders.

D. Contributions and Structure

Our claimed contribution in this paper is twofold. We propose an analytical framework for analyzing European balancing markets which we supplement by an MDP-based market simulator. And we use our framework to arrive at concrete insights and recommendations regarding the design of the European balancing market. One important recommendation is to introduce a real-time reserve market in the European balancing design.

The remainder of the paper is structured as follows. In section II we describe various market design options for the European balancing market, and propose an MDP framework for simulating these different market design options. In section III we analyze these different market design options under an assumption of perfect competition, and summarize our main conclusions regarding the strengths and weaknesses of different market design proposals. In section IV we validate our theoretical results by applying the MDP simulation framework of section II in order to test the ability of different balancing market design options in back-propagating the value of reserve to forward markets. We conclude our analysis and discuss prospects for future research in section V.

II. A Model of the European Balancing Market Based on Markov Decision Processes

A. Building Up the MDP Model

In order to illustrate our full MDP model of the balancing market, we commence by the simplest possible setting and add features gradually to the model. We discuss our assumptions along the way.

As we mention in the introduction, each BSP must be attributed to a unique BRP according to article 18(4).d of the EBGL [3]. Without loss of generality, therefore, we consider a generic agent participating in the balancing market as one which owns (i) a pool of uncontrollable assets that impose a price-inelastic imbalance (positive or negative) to the system as well as (ii) a set of controllable assets with marginal cost C that is private information of the agent, and with a total upward capacity P^+ and downward capacity P^- that is common knowledge for the TSO and all market agents. The controllable set of assets can be offered to the balancing market.

1) Single-Stage MDPs: Consider an agent that wishes to decide how much balancing energy q to offer to a uniform price auction. In MDP terminology, the decision q is the action of the agent. For the moment, let us assume that the auction price is constant and equal to λ^B over episodes. The reward of the agent as a function of state and action is described as $(\lambda^B - C) \cdot q$.

This model can be enriched by introducing the possibility for the agent to submit price-quantity pairs. Concretely, the action space can be enlarged to (p, q). This would correspond to an offer of q MW at p €/MWh. Assuming that the bids of all competing agents are fixed, this bid implies a balancing price, and a quantity qa that is accepted by the auction. The reward of the agent is then expressed as $(\lambda^B - C) \cdot qa$. Note that the representation of this decision-making problem already exceeds the expressive ability of mathematical programs with equilibrium constraints [19].

The next feature that can be added to the model is uncertainty in the balancing price. This uncertainty can be represented by introducing a system-level uncertain imbalance that should be covered by the balancing offers of the agents.

2) Two-Stage MDPs: We are interested, next, in introducing a difference between the balancing price and the imbalance price to the model. This is the current practice, for example, in Belgium, where the system operator computes the imbalance price by applying a surcharge α^U whenever the system is short, or a discount α^L whenever the system is long [20]. Mathematically, the imbalance price in this setting can be expressed as:

$$\lambda^I = \lambda^B + \alpha \quad (1)$$
$$\alpha \triangleq \alpha^U \cdot \mathbb{I}[Imb^t > UI] - \alpha^L \cdot \mathbb{I}[Imb^t < LI] \quad (2)$$

The imbalance price is denoted by λ^I. Here, Imb^t corresponds to the total imbalance of the system. The parameters UI and LI represent the upper and lower imbalance thresholds at which the surcharge or discount apply, respectively.

We represent the operation of the balancing market through the following sequence of events. (1) The agent submits a price-quantity bid in the balancing platform. (2) The agent observes the imbalance Imb within its portfolio, and decides how much of it to cover. (3) The TSO observes the system imbalance, activates BSPs, and produces a uniform clearing price. (4) The TSO also computes an alpha penalty, which is added to the balancing price and is charged to BRPs.

We model this process as a two-stage MDP:

- Stage 1
  - State: a single element, the default state of the world.
  - Action: (p, q), the price-quantity offers in the balancing platform.
  - No reward is collected at this stage.
- Stage 2
  - State: (i) the bid price p, (ii) the leftover BSP capacity after some capacity has been offered to the balancing auction, and (iii) the level of imbalance Imb of an agent.
  - Action: How much of the imbalance Imb to cover (this action, denoted as ai and referred to as "active imbalance", must be limited to the leftover capacity that the BSP has not allocated to the reserve auction).
  - Reward: (i) BSP payment for upward/downward activation, expressed as $\lambda^B \cdot qa$, (ii) BRP payment for imbalance settlement, expressed as $-\lambda^I \cdot (Imb - ai)$, and (iii) fuel costs related to self-balancing and BSP activation, expressed as $-C \cdot (ai + qa)$.

Note that active imbalance, which corresponds to $ai \neq 0$, is a practice which TSOs do not necessarily encourage. Nevertheless, it is impossible to enforce $ai = 0$, since agents are in control of their private assets, and since the net demand forecast of a portfolio is private information that the TSO cannot audit [1].

3) Three-Stage MDPs: In order to model the back-propagation of the value of reserve to forward reserve capacity auctions, we introduce a uniform-price auction for reserve capacity. This corresponds, for example, to European day-ahead reserve capacity auctions for secondary or tertiary reserve [2].

The overall model can be described as the following three-stage MDP:

- Stage 1
  - State: a single element, the default state of the world.
  - Action: (p^R, q^R), the price-quantity offers in the reserve capacity auction.
  - Rewards: the payment from the reserve capacity auction.
- Stage 2
  - State: the capacity qa^R awarded in the reserve capacity auction.
  - Action: (p, q), the price-quantity offers in the balancing platform. The offered quantity can be no less than what has been cleared in the reserve auction.
  - No reward is collected at this stage.
- Stage 3: identical to the two-stage MDP.

B. Market Design Variants

Our analysis will focus on four different market design options. These options are inspired by discussions with stakeholders about different ways in which the European balancing market could be organized so as to enable a more accurate reflection of the value of reserve capacity.

1) The Vanilla European Design (D1): The default European design is the one corresponding to section II-A3, for which the imbalance penalty α of Eq. (1) is equal to zero. This implies that, in this design, the balancing price equals the imbalance price, $\lambda^I = \lambda^B$.

This design is fully compatible with the EBGL. However, as we show in the following section and verify experimentally in section IV, it fails at generating a forward reserve price signal. Inherently, therefore, this mechanism fails to value reserve capacity. The reason is that, in this design, there is a missing market for trading reserve capacity in real time.

2) Imbalance Penalties (D2): The inherent inability of design $(D1)$ to generate a forward reserve price signal that reflects the value of reserve has already been discussed based on a stochastic equilibrium framework in [1]. In response to the request of the European Commission for planned market reforms in order to implement scarcity pricing (article 20(3) of regulation 2019/943 [21]) the Belgian government [22] mentions that the imbalance penalty α of Eq. (1), "already exhibits quite some characteristics of a scarcity pricing mechanism" [22]. What we show in the sequel is that, in the case of independent imbalances and a symmetric imbalance penalty α, design $(D2)$ behaves identically to design $(D1)$.

It is important to note that design $(D2)$ relies on imbalance penalties α which depend on the level of system imbalance, which is not to be confused with the level of scarcity in the system. To clarify: a system that is exhibiting a very large positive imbalance is not experiencing scarcity if it carries abundant reserve at the moment in time when the large imbalance occurs.

In practice, the imbalance penalty in Eq. (1) depends on the imbalance of the current and previous interval (see Eq. (4) below). Therefore, the MDP model that we develop for design $(D2)$ requires an additional state variable, the imbalance of the previous balancing interval, which is is added to the state vector of stages 2 and 3.

3) Adders on Imbalance Charges (D3): Scarcity pricing, as proposed in [1] and following [23], introduces a real-time price for reserve, or ORDC adder, which is a function of the instantaneous amount of leftover capacity in the system:

$$\begin{aligned}
&\lambda^R = (VOLL - \lambda^B) \cdot \\
&LOLP(P^{+,tot} - Imb^t) \cdot \mathbb{I}[P^{+,total} - Imb^t \geq 0] + \\
&(VOLL - C^{max}) \cdot \mathbb{I}[P^{+,total} - Imb^t < 0]). \qquad (3)
\end{aligned}$$

Here, $VOLL$ is an estimate of the value of lost load in the system, $P^{+,tot}$ is the total reserve capacity that is available, $LOLP(\cdot)$ is the loss of load probability in the system as a function of available reserve capacity, and C^{max} is an estimate of the marginal cost of the most expensive unit in the system. This price signal is reflective of system scarcity, in the sense that it is adaptive to the amount of leftover reserve capacity, $P^{+,total} - Imb^t$.

The question is where this adder should be applied. It has been proposed [24] to apply this adder as an imbalance charge, as an alternative to the α penalty of Eq. (1). As we demonstrate analytically in proposition 3.3 and numerically in section IV, this market design produces a forward reserve price, however this signal is significantly weaker than the average value of reserve capacity to the system. Introducing an adder to the imbalance price does not rectify the fact that design $(D3)$, like design $(D1)$ and $(D2)$, is featuring a missing market for reserve capacity in real time.

4) Scarcity Pricing (D4): The implementation of scarcity pricing relies on a real-time market for reserve capacity. In terms of the MDP model, this implies replacing α with λ^R in Eq. (1), and introducing the following term in settlement:

$$-\lambda^R \cdot qa^R + \lambda^R \cdot (P^+ - qa - ai).$$

This term effectively implies that agents buy back their day-ahead reserve capacity at real-time reserve prices, and sell their entire real-time reserve capacity at real-time reserve prices. Introducing this settlement of real-time reserve imbalances induces agents to bid their reserve capacity in forward markets in a way that anticipates the expected price at which they would be required to buy that reserve capacity back in real time. This effect results in the back-propagation of the scarcity signal.

The mechanism amounts to introducing a real-time market for reserve capacity, and is exactly analogous to the practice of settling *energy* imbalances at prevailing real-time energy prices. Furthermore, the approach is compatible with the EBGL, since one can invoke article 18(4).d of the EBGL for attributing a BSP to an associated BRP, and article 44(3) of the EBGL for introducing an additional settlement mechanism which is separate from imbalance settlement.

Note that the representation of this design requires augmenting the MDP model of section II-A3 by adding the awarded day-ahead reserve capacity qa^R to the state of the third time step, since this quantity affects the third-stage payoff under design $(D4)$.

III. Analytical Results

This section analyzes each of the four designs that are introduced in section II under the simplifying assumption of perfect competition. Unveiling difficulties in back-propagating reserve prices in the case of perfect competition suggests fundamental market design problems, and offers insights about what to expect in the simulations of section IV-B. Our simplifying assumption can be stated as follows:

Perfect competition assumption: We consider a fringe agent, i.e. one with infinitesimal capacity.

In order to keep the development concise, we proceed by characterizing the optimal strategy of a fringe agent in section III-A. We then outline the strategy of our proofs in section III-B. The full proof for each of the following propositions is available online [25].

A. Statement of Analytical Results

Proposition 3.1: In design $(D1)$, it is always optimal for agents to bid their entire balancing capacity at the true marginal cost to the balancing auction. For agents with upward balancing capacity ($P^+ > 0$), the opportunity cost of bidding their capacity to the day-ahead reserve auction is zero. This is a pure strategy Nash equilibrium.

Proposition 3.2: Under the assumption of independent symmetric imbalances, in design $(D2)$ it is always optimal for agents to bid their entire balancing capacity at the true marginal cost to the balancing auction. For agents with upward balancing capacity ($P^+ > 0$), the opportunity cost of bidding their capacity to the day-ahead reserve auction is zero. This is a pure strategy Nash equilibrium.

Proposition 3.3: In design $(D3)$, it is sometimes, but not always, optimal for agents to bid their entire balancing capacity at the true marginal cost to the balancing auction. For agents with upward balancing capacity ($P^+ > 0$), the opportunity cost of bidding their capacity to the day-ahead reserve auction is less than or equal to the scarcity value $\mathbb{E}[\lambda^R]$. This does *not* characterize a pure strategy Nash equilibrium, since some agents find it optimal to self-balance.

Design $(D3)$ is depressing the scarcity price in two ways: (i) agents who find it optimal to self-balance face an opportunity cost which is less than the scarcity price $\mathbb{E}[\lambda^R]$, and (ii) agents who find it optimal to bid their entire capacity to the balancing auction face an opportunity cost of zero for bidding reserve in the day ahead.

Proposition 3.4: In design $(D4)$, it is always optimal for agents to bid their entire balancing capacity at the true marginal cost to the balancing auction. This is a pure strategy Nash equilibrium.

Note that design $(D4)$ emerges as the only option which back-propagates the real-time value of reserve capacity to day-ahead reserve auctions, while preserving the incentive of agents to make their balancing capacity available in the balancing market. Choosing to offer resources in the balancing auction instead of self-balancing promotes operational efficiency, since resources are pooled in the balancing auction, where price discovery and efficient allocation of resources can take place.

B. Proof Strategy

In this section, we prove the statement of proposition 3.1 for one case. This technique forms the basis for all the results of section III-A, and conveys the basic intuition of our reasoning. For a detailed proof of all the results of section III-A, the reader is referred to [25].

The first step in the proof of all the propositions is to demonstrate that there is no loss of generality in considering the case of an agent which has only downward capacity (i.e. $P^+ = 0$ and $P^- < 0$) or the case of an agent which has only upward capacity (i.e. $P^- = 0$ and $P^+ > 0$) [25].

Once this is established, we can fix the bid (p, q) in the balancing market. Under the fringe assumption, we can ignore the influence of the agent decisions ai and the agent imbalance on the expected imbalance price. In the following calculations, we denote $D \triangleq -\mathbb{E}[\lambda^B \cdot Imb]$. This is not affected by the actions of the agent, and is therefore a constant offset to the imbalance payoff of the agent.

We have two possible suppliers: (i) the ones for which $\mathbb{E}[\lambda^B] \geq C$, and (ii) the ones for which $\mathbb{E}[\lambda^B] < C$. In what follows, we limit the discussion to the the case of cheap suppliers with upward capacity ($\mathbb{E}[\lambda^B] - C \geq 0, P^+ > 0, P^- = 0$).

Our strategy is to first characterize the optimal bidding strategy in the balancing market, (p, q), by considering the effect of these decisions on imbalance settlements and balancing payments.

The imbalance payoff is computed as follows for agents with $P^+ > 0$ (and therefore $q \geq 0$):

$$\begin{aligned} &\max_{ai}(\mathbb{E}[\lambda^B] - C) \cdot ai - \mathbb{E}[\lambda^B \cdot Imb] \\ &ai + q \leq P^+ \\ &ai \geq 0 \end{aligned}$$

We have $ai^\star = P^+ - q$. The expected payoff z_I is then expressed as follows:

$$z_I = (\mathbb{E}[\lambda^B] - C) \cdot (P^+ - q) + D$$

The balancing payoff z_B can be expressed as follows:

- If $p > \lambda^B$, then $z_B(\omega) = 0$
- If $p = \lambda^B$, then $z_B(\omega) = (\lambda^B - C) \cdot qa$ for some qa which is selected by the auctioneer. We get rid of this case by assuming that the auctioneer always activates zero MW of the supplier when the bid is at the money. Since this is a fringe supplier, the auctioneer can always source the imbalance energy from alternative suppliers. Thus, we have $qa = 0$ and $z_B = 0$ in this case.
- If $p < \lambda^B$, then $z_B(\omega) = (\lambda^B - C) \cdot q$.

The realization ω corresponds to the realization of system imbalance. Note that $z_B(\omega)$ is random. In fact, the distribution of λ^B depends on the decisions of the agent, p and q. In the sequel, we denote the probability measure of the balancing price λ^B as μ.

The expected payoff can therefore be expressed as follows:

$$\begin{aligned} z_B &= \mathbb{E}[z_B(\omega)] \\ &= \int_{x>p} (x - C) \cdot q \cdot d\mu(x) \end{aligned}$$

The overall payoff of the agent can therefore be expressed as follows:

$$\begin{aligned} R(p, q) &= z_I + z_B \\ &= C_1 - C_2 \cdot q + C_3(p) \cdot q \end{aligned}$$

where the terms can be described as follows:

$$\begin{aligned} C_1 &= (\mathbb{E}[\lambda^B] - C) \cdot P^+ + D \\ C_2 &= \mathbb{E}[\lambda^B] - C \\ C_3(p) &= \int_{x>p} (x - C) \cdot d\mu(x) \end{aligned}$$

In order to determine the optimal bidding strategy, let us first fix the bid quantity q of the agent. We can express the first-order conditions with respect to p as:

$$\begin{aligned} \frac{\partial R(p, q)}{\partial p} &= C_3'(p) \cdot q \\ &= -\mu(p) \cdot (p - C) \cdot q \end{aligned}$$

We note that the payoff function $R(p, q)$ for fixed q is increasing in $(-\infty, C]$, zero at C, and decreasing in $[C, +\infty)$. Thus, for any q, an optimal strategy is to bid the true cost. And, given this strategy, the payoff becomes

$$R(C, q) = C_1 - C_2 \cdot q + C_3(C) \cdot q$$

We can show that $\frac{\partial R(C,q)}{\partial q} > 0$ [25]. Therefore, it is optimal to bid $q^\star = P^+$ in the balancing auction, and $ai^\star = 0$. This reflects the fact that, when being in active imbalance, the agent takes the risk of producing power when being out of the money. Instead, the balancing market will only activate the agent when its marginal cost is lower than the balancing price. The fact that the balancing and imbalance price are equal sends the correct incentive to the agent for bidding its entire capacity to the balancing auction.

Note that every MW cleared in a forward reserve auction comes with an obligation to bid that MW in the balancing auction, so this is profit lost in the balancing and imbalance phase. Since the optimal strategy of the agent is to anyways bid its entire capacity in the balancing auction, there is no opportunity cost for the agent, i.e. $dR^\star/dq = 0$. Thus, the reserve price at which the agent would bid in the day-ahead reserve auction is zero.

IV. Illustration on a Case Study

We now proceed to a numerical illustration in a simple case study. In section IV-A we validate the analytical results of section III by considering a single fringe agent. In section IV-B we assess the ability of the different designs to back-propagate reserve prices by considering multiple agents that compete against each other.

A. Validation of Analytical Results

Consider a system with a fringe supplier that manages a flexible upward capacity of $P^+ = 1$ MW (and downward capacity of $P^- = 0$ MW). The marginal cost of the agent is $C = 50$ €/MWh. We discretize the action space as follows: the balancing auction bid q and reserve auction bid q^R is either 0 MW or 1 MW, and the agent can bid any value p between 25 to 75 €/MWh, in increments of 5 €/MWh.

The system imbalance is assumed to be normally distributed with a mean of 0 MW and a standard deviation of 91.5 MW. The imbalance of the fringe agent is assumed normally distributed, with a mean of 0 MW and a standard deviation of 0.41 MW.

In the analytical model, the balancing supply function of the system is assumed to be affine, and is expressed mathematically as $a + b \cdot q$, where q is the amount of activated balancing capacity (with $q > 0$ corresponding to upward activation and $q < 0$ corresponding to downward activation), $a = 50$ €/MWh, and $b = 0.11$ (€/MWh)/MW. This supply function is an approximation of a balancing market with 8 agents, whose parameters are defined in Table I. The fringe agent that we are interested in is agent A5.

For the case of design $(D2)$, we use the formula proposed by ELIA [20]: $UI = LI = 150$ MW, and

$$\alpha^U = \alpha^L = \frac{200}{1 + \exp\left(\frac{450 - x}{65}\right)} \tag{4}$$

TABLE I
THE BALANCING CAPACITY AND MARGINAL COST OF DIFFERENT AGENTS FOR THE MDP CODE OF SECTION IV-A. UNITS ARE IN [MW] FOR P^+ AND P^-, AND IN [€/MWH] FOR C.

	A1	A2	A3	A4	A5	A6	A7	A8
P^+	0	0	0	0	1	100	100	100
P^-	-100	-100	-100	-50	0	0	0	0
C	20	30	40	50	50	60	70	80

where $x = \frac{|Imb^t|+|Imb^t_{t-1}|}{2}$ is the average of the absolute total system imbalances of the previous and current imbalance interval. For the case of design $(D3)$ and $(D4)$, we assume a value of $VOLL = 920$ MW.

Design	$(D1)$	$(D3)$	$(D4)$
q^* [MW]	1	0	1
p^* [€/MWh]	55	any	50
Average Profit [€]	6.34	14.43	18.85
Opportunity cost $dR^\star/dq$ [€/MWh]	0	8.11	12.71

TABLE II
RESULTS FOR $(D1)$, $(D3)$ AND $(D4)$ IN THE SINGLE-AGENT SIMULATION.

Imb^t_{t-1} [MWh]	$]\infty,-150]$	[-150,0]	[0,150]	$[150,\infty[$
q^* [MW]	1	1	1	1
p^* [€/MWh]	50	55	55	50
Average Profit [€]	6.43	6.30	6.32	6.46
$dR^\star/dq$ [€/MWh]	0	0	0	0

TABLE III
RESULTS FOR $(D2)$ FOR DIFFERENT RANGES OF Imb^t_{t-1} IN THE SINGLE-AGENT SIMULATION.

Design	$(D1)$	$(D2)$	$(D3)$	$(D4)$
q^* [MW]	1	1	0	1
p^* [€/MWh]	50	50	any	50
Average Profit [€]	4.04	4.04	12.57	16.63
Opportunity cost $dR^\star/dq$ [€]	0	0	8.53	12.59

TABLE IV
RESULTS FOR DIFFERENT MARKET DESIGNS USING THE ANALYTICAL SOLUTION.

For the single-agent simulation, we use the Q-learning algorithm [14] under a uniformly distributed policy for the purpose of learning the Q function. We use a learning rate of $\frac{1}{n(s,a)}$ for each state-action pair (s,a), where $n(s,a)$ counts the number of visits to (s,a). We run $2,000,000$ episodes for each design with the same seeds, in order to isolate the effect of the market design changes on the results.

We summarize the results of the simulation in Tables II and III, and the analytical solution in Table IV. We observe the following. (i) For every design, the bid quantity and price are equivalent for the analytical case and the MDP model[3]. (ii) The profits are in the same range for the analytical solution and the MDP model. Differences (which amount to a range of 2 €) can be expected, because the analytical model assumes a continuous supply function, which is a continuous approximation of the stepwise supply function that is used in the MDP code (see Table II). (iii) The opportunity costs are very close to each other for the analytical model and the MDP code. (iv) For design $(D2)$, the range of values in the imbalance of the previous period, Imb^t_{t-1}, does not influence the selected action or the profit, see Table III. This observation is in line with proposition 3.2.

[3]Indeed, in the MDP model, bidding at a price of 45 or 55 €/MWh is equivalent to bidding at 50 €/MWh, because there is no other producer with a marginal cost in the intervals $[45,50[$ and $]50,55]$. For design $(D3)$, the bid price does not matter, because the bid quantity is 0 MW.

B. Back-Propagation

We now concentrate on assessing experimentally the ability of the different market designs to back-propagate the real-time value of reserve to the day-ahead reserve market. For this purpose, we use our MDP model for developing a multi-agent simulation. In order to focus the analysis on the effects of the design in conditions of high competition for upward balancing capacity, we replace producers $5-8$ by 35 producers with a capacity of 10 MW and marginal costs that increase uniformly from 50 €/MWh to 84 €/MWh.

We discretize the agent action space by having agents bid in price increments of 5 €/MWh and in quantity increments of half of their capacity. Each agent is facing a portfolio imbalance which is uniformly distributed between zero, and half of its minimum and maximum capacity. There is also a system imbalance with a zero mean and a standard deviation of 21.9 MW. Agent imbalances are independent of each other and of the system imbalance. The day-ahead reserve demand curve is assumed to be identical to the real-time reserve demand curve, and based on the ORDC formula of Eq. (3).

We let every agents optimize its own policy using the Q-learning algorithm under an $\epsilon-$greedy policy. During the learning phase, ϵ_k evolves as $\frac{0.05}{N-k}$, where N is the maximum number of iterations and k is the current iteration. Since all agents are learning simultaneously, from the perspective of any single agent, the environment is non-stationary, which implies that we have no convergence guarantees. In order to cope with the non-stationarity of the environment, we use a constant learning rate [26].

We run $1,500,000$ iterations in blocks of 100. After each block of 100 iterations, we compute the outcome that we would have obtained in the reserve market if each agent were applying its policy greedily. We plot the sample average of this reserve price for the different designs in Fig. 1.

We observe the following. (i) For $(D1)$ and $(D2)$, the reserve price sample average converges to a small value. This is anticipated by the analytical results, because the opportunity cost for each agent is equal to 0. The decrease is slower for $(D2)$, because there are more states in $(D2)$ than in $(D1)$, and therefore the convergence is slower. (ii) For $(D3)$, the reserve price sample average arrives slightly above the one resulting from $(D1)$. As the analysis shows [25], under $(D3)$ certain low-cost producers may face a positive opportunity cost when bidding into the day-ahead reserve market. Nevertheless, the resulting reserve price remains close to the one of $(D1)$, because few producers are sufficiently cheap to fulfill this condition. (iii) Under design $(D4)$, the day-ahead reserve price converges to a value which is close to the average real-time scarcity adder, i.e. 9.35 €/MWh.

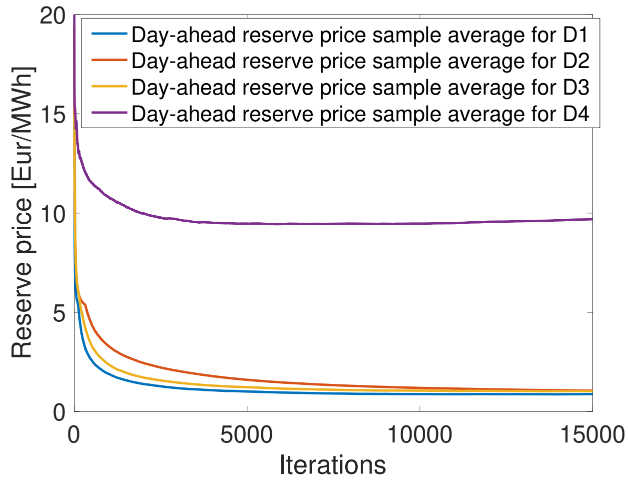

Fig. 1. The evolution of the reserve price in the simulation of section IV-B.

V. Conclusions and Perspectives

We present a methodology for analyzing the European balancing market based on an analytical derivation of optimal bidding under perfect competition assumptions, accompanied by an MDP-based simulation. The analysis exposes the inability of various market design alternatives in back-propagating the value of reserve capacity in day-ahead markets. The analysis validates the ability of a real-time market for reserve capacity [1] to back-propagate the value of reserve capacity to day-ahead markets, while also preserving the incentive of agents to make their reserve resources available in the balancing market.

The policy discussion for the implementation of scarcity pricing is advancing in Belgium. Since October 2019, the Belgian system operator publishes[4] scarcity prices one day after operations based on the available reserve capacity that has transpired during the previous day. In future research, we will further analyze numerous important aspects of the mechanism. The legal basis for the implementation of the mechanism can rely on articles 18(4) and 44(3) of the EBGL. The specific parameter choices for computing the scarcity adders, i.e. the shape of the ORDC, are currently being investigated. Finally, it is important to understand the interaction of the mechanism with neighboring energy and reserve markets that are not adopting the mechanism, and to ensure its compatibility with the legal framework of EBGL in this multi-area setting.

References

[1] A. Papavasiliou, Y. Smeers, and G. de Maere d'Aertrycke, "Study on the general design of a mechanism for the remuneration of reserves in scarcity situations," UCLouvain, Tech. Rep., 2019. [Online]. Available: https://www.creg.be/sites/default/files/assets/Publications/Notes/Z1986Annex.pdf

[2] R. Dominguez, G. Oggioni, and Y. Smeers, "Reserve procurement and flexibility services in power systems with high renewable capacity: Effects of integration on different market designs," *Electrical Power and Energy Systems*, vol. 113, pp. 1014 – 1034, 2019.

[4]https://www.elia.be/en/electricity-market-and-system/studies/scarcity-pricing-simulation

[3] European Commission, "Commission regulation (EU) 2017/2195 of 23 november 2017 establishing a guideline on electricity balancing," Tech. Rep., 2017. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R2195&from=EN

[4] W. Hogan, "On an 'energy only' electricity market design for resource adequacy," Center for Business and Government, JFK School of Government, Harvard University, Tech. Rep., September 2005.

[5] ERCOT. (2015) Ercot market training: Purpose of ORDC, methodology for implementing ORDC, settlement impacts for ORDC. [Online]. Available: http://www.ercot.com/content/wcm/training_courses/107/ordc_workshop.pdf

[6] W. W. Hogan and S. L. Pope, "PJM reserve markets: Operating reserve demand curve enhancements," Harvard University, Tech. Rep., 2019. [Online]. Available: https://sites.hks.harvard.edu/fs/whogan/Hogan_Pope_PJM_Report_032119.pdf

[7] A. Papavasiliou and Y. Smeers, "Remuneration of flexibility using operating reserve demand curves: A case study of Belgium," *The Energy Journal*, pp. 105–135, 2017.

[8] A. Papavasiliou, Y. Smeers, and G. Bertrand, "An extended analysis on the remuneration of capacity under scarcity conditions," *Economics of Energy and Environmental Policy*, vol. 7, no. 2, 2018.

[9] ELIA, "Study report on scarcity pricing in the context of the 2018 discretionary incentives," 2018.

[10] D. Ralph and Y. Smeers, "Risk trading and endogenous probabilities in investment equilibria," *SIAM Journal on Optimization*, vol. 25, no. 4, pp. 2589–2611, 2015.

[11] A. Ehrenmann and Y. Smeers, *Stochastic Equilibrium Models for Generation Capacity Expansion*, ser. Stochastic Optimization Methods in Finance and Energy, International Series in Operations Research and Management Science, Part 2. Springer, 2011, vol. 163, pp. 273–310.

[12] J. Bower and D. Bunn, "Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the england and wales electricity market," *Journal of Economic Dynamics and Control*, vol. 25, no. 3-4, pp. 561–592, Mar. 2001.

[13] D. Bunn and F. Oliveira, "Agent-based simulation-an application to the new electricity trading arrangements of england and wales," *IEEE Transactions on Evolutionary Computation*, vol. 5, no. 5, pp. 493–503, Oct. 2001.

[14] C. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, p. 279–292, May 1992.

[15] N.-P. Yu, C.-C. Liu, and J. Price, "Evaluation of market rules using a multi-agent system method," *IEEE Trans. Power Syst.*, vol. 25, no. 1, p. 470–479, Feb. 2010.

[16] V. Nanduri and T. Das, "A reinforcement learning model to assess market power under auction-based energy pricing," *IEEE Trans. Power Syst.*, vol. 22, no. 1, p. 85–95, Feb. 2007.

[17] Y. Ye, D. Qiu, J. Li, and G. Strbac, "Multi-period and multi-spatial equilibrium analysis in imperfect electricity markets: A novel multi-agent deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 130 515 – 130 529, Sep. 2019.

[18] Y. Ye, D. Qiu, M. Sun, D. Papadaskalopoulos, and G. Strbac, "Deep reinforcement learning for strategic bidding in electricity markets," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, p. 1343–1355, Mar. 2020.

[19] A. Ehrenmann, "Equilibrium problems with equilibrium constraints and their application to electricity markets," Ph.D. dissertation, 2004.

[20] ELIA, "Tariffs for maintaining and restoring the residual balance of individual access responsible parties 2020-2023," Belgian Transmission System Operator, Tech. Rep., 2019.

[21] European Commission, "Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 june 2019 on the internal market for electricity (recast)," *Official Journal of the European Union*, 2019.

[22] ELIA, "Belgian electricity market: Implementation plan." [Online]. Available: https://ec.europa.eu/energy/sites/ener/files/belgian-electricity-market-implementation-plan.pdf

[23] W. Hogan, "Electricity scarcity pricing through operating reserves," *Economics of Energy and Environmental Policy*, vol. 2, no. 2, pp. 65–86, 2013.

[24] P. Giesbertz. (2019) The power market design column - the scarcity of scarcity pricing. [Online]. Available: https://www.linkedin.com/pulse/power-market-design-column-scarcity-pricing-paul-giesbertz/

[25] A. Papavasiliou, "Analytical derivation of optimal BSP / BRP balancing market strategies," Tech. Rep., 2020. [Online]. Available: https://perso.uclouvain.be/anthony.papavasiliou/public_html/AnalyticalV3.pdf

[26] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. MIT press, 2018.