Electricity Auctions with Integer Decisions: Pricing Rules to Signal Optimal Investment

Miguel Vazquez [*]
Michelle Hallack [**]
Carlos Vazquez [***]

TD 314
Julho/2015

© [*] Professor Adjunto da Faculdade Economia da UFF. E-mail: miguel.vazquez.martinez@gmail.com.

[**] Professor Adjunto da Faculdade Economia da UFF. E-mail: michellehallack@gmail.com.

[***] School of Mines and Energy, Universidad Politécnica de Madrid, and Gas Natural Fenosa.

ABSTRACT

Several power systems around the world have opted for using combinatorial auctions for their day-ahead markets. In them, market operators receive bids for different concepts, such as variable cost, start-up cost, no-load cost, etc., and some kind of optimization procedure determines the accepted bids. Unfortunately, this frequently implies the use of discrete variables, and there is no consensus on how to define marginal costs in presence of such integer decisions. This problem is exacerbated by the increasing need of back-up generation, which is characterized by frequent start-ups, and therefore makes integer decisions more relevant. We show that the criterion that prices should support a competitive short-term equilibrium is too loose to define the prices: many different price solutions could be adopted if only this condition were considered. We then incorporate long-term competitive equilibrium requirements, considering explicitly that the infra-marginal rents that base-load generators receive have an influence on their investment decisions, and looking for the prices that induce generation equipment to be optimal. Imposing this, we derive an optimal pricing rule, which now turns out to be a single solution. These prices represent an optimal combination of marginal price and side payment that is supported by a sound theoretical justification.

KEY WORDS: Electricity auctions; Investment signals; Side payments; Integer decisions; Marginal cost

1. Introduction[1]

Several power systems around the world –as is the case, for instance, of PJM– have opted for using a combinatorial auction for their day-ahead market. Their market operator receives bids for different concepts, such as variable cost, start-up cost, no-load cost, etc., and some kind of optimization procedure determines the accepted bids. Well-known marginal pricing principles allow the auctioneer to compute prices for each hour using the results of this optimization tool, if the model is linear. However, when the model is not convex, the previous results cannot be extended in a straightforward way. Unfortunately, the discrete variables related, among others, to start-ups introduce non-convexities in the auction problem and, therefore, there is no consensus on how to compute prices in a power auction that explicitly represents unit commitment decisions.

We first set up the frame for our analysis, in section **Erro! Fonte de referência não encontrada.**, focusing on the price in its role as a tool to clear the auction. We describe the logic for price computation in optimization-cleared markets and show that the price is obtained as the marginal cost for supplying an additional megawatt of demand. The difficulty appears when integer variables are included in the optimization. The problem becomes non-convex, hence non-derivable, so the marginal cost is ill-defined. This is the case when start-up and shut-down variables are considered explicitly, as both of them are discrete decisions which can be either one or zero, but not any intermediate value. We describe the problems that arise in this context, as well as some solutions adopted in practice or proposed in the literature, (O'Neill et al., 2005), (Hogan and Ring, 2003). The main conclusion is that there is no linear price that efficiently clears the market and if we allow some flexibility to cope with this problem, then almost any price can be a reasonable solution to the short-term equilibrium.

The rest of the paper adopts another point of view and focuses on the price in its role as long-term signal for investment in new generation capacity, pursuing the strategy developed in (Vazquez, 2003). The different pricing alternatives that are feasible under the short-term criteria of section **Erro! Fonte de referência não encontrada.** provide, all of them, the same income to the marginal generator; they only differ from each other in the remuneration that infra-marginal units receive. This is the operational profit that should pay for the investment costs. Therefore, the question that the paper answers is which of the many possible pricing schemes provides incentives to the producers to build the optimal (cheapest) generation mix. We will compare the ideal (convex) case to the non-convex one, and will show that, when considering long-term signals, it is possible to derive a single optimal pricing rule, which is more efficient that any alternative design, under the frame considered.

To that end, we continue by placing combinatorial auctions into context. One may observe several design strategies for day-ahead auctions, each of them representing solutions to the problem of an ill-defined marginal cost. Section **Erro! Fonte de referência não encontrada.** analyzes the short-term problem, showing that many different prices could be defined if only the condition that they support the short-term equilibrium were considered. Section 4 develops our proposal to include long-term considerations in the definition of the auction pricing rule. Section 5 shows the importance of our proposal in a numerical example, and section 6 collects some concluding remarks.

2. Elements of electricity auction design

One of the elements that have captured substantial amounts of interest by regulators during the process of liberalization of the power industry that has taken place in the last thirty years is the design of the spot market. The most common solution is organized around a day-ahead auction, where producers place bids to sell the energy demanded in the system, the auction mechanism selects the cheapest ones, and a generation dispatch arises. The characteristics of this auction are, however, far from being standard and go beyond the theoretical results obtained by auction theory, see for instance (Milgrom, 2004). Day-ahead electricity auctions can be described as multi-unit (many different megawatts are purchased in each hour in each auction) and multi-product (the energy corresponding to the demand of several different hours are purchased at the same time in each auction, all of them different from the rest, but all of them inter-related).

A first alternative for implementing these day-ahead auctions is to design spot markets as simple auctions, ignoring the multi-product feature, so power producers would send bids specifying the price required for selling each possible quantity. In this context, there would be one simple auction to allocate the electricity produced at each hour, and all these auctions would be independent from each other. But a simple auction only allows for the specification of a cost proportional to the units output and a maximum output constraint, and the real conditions of generators in a unit commitment problem include several technical constraints (ramping limits, minimum output...) and non-linear costs (start-ups, shut-downs...), which make the cost structure more complex than the bids allowed in a simple auction. Besides, a number of these technical conditions have the effect of interlinking the different time periods, making the results of the auction in one certain hour depend on the results of the rest of them. Having a simple-auction market design results in the need for market players to internalize into their bids and offers the part of the costs that cannot be specified in the auction bids. For instance, they might place a bid for the quantity corresponding to their minimum output with a very low price, so that this quantity is always accepted in the simple auction and therefore always dispatched.

Bidding this way is quite a difficult task. For instance, in order to internalize minimum output constraints, producers must know when their plant will be online. This depends on the auction results, so it is subject to uncertainty when producers send their bids. Thus, the design of the bids requires not only the use of the technical data of the plants, but in addition the estimation of the auction results. If such estimations are not accurate enough, the bids will not be correct and the auction dispatch will not be satisfactory for

power producers –e.g. the final unit commitment for some certain generator might not fulfill the ramping constraint and thus be technically infeasible. Hence, when a simple auction is implemented, there is a risk factor directly associated with the market clearing mechanism design, which translates into an additional production cost[2].

An answer to this problem is to allow for the creation of a number of additional markets, before and after the day-ahead auction, where generators can re-negotiate their positions and adjust their schedules, correcting the possible internalization errors (see, for instance, (Wilson, 2002). If arbitrage among the different markets works efficiently, the set of consecutive markets is equivalent to a single market with no internalization problems. This is the predominant scheme in Europe. Unfortunately, for adjustments that take place in time horizons shorter than the day-ahead market this arbitrage might be limited, so the consecutive markets solution allows for a mitigation of the bidding risk induced by the simple auction, but only partially.

An alternative solution is to facilitate the internalization process by means of an iterative auction. (Wilson, 1997) iterative mechanisms tackle the internalization problem by means of a sequence of iterations where market players reveal their preferences over the auctioned products. Such information is received by the rest of competitors and used to improve their estimations on the market clearing results. Hence, the objective is designing an auction that facilitates the internalization of the generators' characteristics, rather than reducing the need for such internalization. The process may be visualized as a sequence of simple auctions: after each iteration, market participants analyze the tentative market clearing corresponding to each iteration, and modify their bids accordingly. The process stops when no participant is willing to change her bids and the final prices and quantities are calculated using the last iteration. Nonetheless, the process tends to be very time consuming, and almost impossible to implement for the day-ahead auctions, at least without automated bidding, so its use in the electricity market has been limited to certain long-term auctions.

A third solution is the complex auction, where players are allowed to place bids that specify additional conditions to the price-quantity pairs of the simple auctions[3]. The corresponding market clearing, thus, must be found by solving some optimization problem. The pure complex auction is essentially a traditional unit commitment model, which is applied to clearing power markets, (Hobbs, 2001). This is the predominant approach in the US nowadays. The immediate advantage of these mechanisms, as of the iterative ones, is that they capture the inter-relation of the different hours and eliminates the need for internalization. On the negative side, their complexity makes their results difficult to explain and this may raise some credibility problems. More importantly,

---

[2] Note that the problem is similar to the one described in (Vickrey, 1961). In the (Vickrey, 1961) context, the English and the Dutch auctions are equivalent under perfect information, but the English one is superior since it allows for truthful bidding and therefore avoids the need for internalization. In the day-ahead electricity market, the simple and complex auctions yield the same results when the information is perfect, but the complex one is superior since it does not need any information about the rest of the competitors to get to the optimal result. However, while in the (Vickrey, 1961) case the information is needed to optimize the strategic behavior of players, which might be questioned especially in a multi-unit context, in the electricity market the information is required to incorporate the cost conditions of the bidders, which would had also happened under perfect competition.

[3] The block bids adopted in NordPool and EEX or the minimum-income conditions of the Spanish pool are examples of these rules.

pure combinatorial auctions are typically based on mixed-integer optimization, which are non-convex problems and hence have difficulties to determine the market price.

We will study the design of combinatorial auctions for electricity day-ahead markets. In that view, the rest of the paper will deal with the challenges associated with the integer non-convexities.

3. Optimal short-term signals

In this section, we analyze the problem of defining a pricing rule in combinatorial power auctions from a short-term point of view. To that end, we begin by describing the ideal situation, where no integer variables exist. Then we tackle the problem with integer decisions and show that there is a wide range of possible candidates to represent the short-term marginal cost.

a. Ideal conditions

This section is aimed at describing, in a relatively simple setting, the fundamental reasoning that we will use in the non-convex case. In particular, this section summarizes the main ideas developed in (Schweppe et al., 1988). In our context, the problem can be thought of as the definition of optimal prices for complex electricity auctions. To that end, we assume that both the producers' cost functions and the clearing algorithm fulfill the properties required to obtain a convex problem.

The basic scheme of the reasoning used in this paper is based on considering the cost-minimization problem as the reference to define an efficient market behavior; i.e. we use the centralized optimization of the system, under perfect information, to define the efficient unit commitment. On the other hand, we analyze the decision-making process of market players under perfect competition. The efficient market behavior is the one that obtains the same results as the centralized model, and the optimal price is the one that incentivizes market players to obtain such results.

Consider that the centralized power system operation can be described by the following program:

$$
\begin{aligned}
\min_g \quad & \sum_{t,i} c_{t,i} g_{t,i} \\
s.t. \quad & g_i^{min} \leq g_{t,i} \leq g_i^{max} \quad : \lambda_{t,i}^{min}, \lambda_{t,i}^{max} \\
& \sum_i g_{t,i} = d_t \quad\quad\quad : \lambda_t^{demand}
\end{aligned}
\tag{1}
$$

where $c_{t,i}$ is the variable cost of the plant i at time t, $g_{t,i}$ is its production at the same period, g_i^{min} and g_i^{max} are its production limits and d_t is the system demand at time t. Thus, the optimality conditions for $g_{t,i}$ can be written as

$$
\lambda_t^{demand} = c_{t,i} + \left(\lambda_{t,i}^{max} - \lambda_{t,i}^{min} \right)
\tag{2}
$$

On the other hand, under a market environment, power producers decide on their output by maximizing the profits associated with their electricity sales at the market price. Thus, the corresponding profit-maximization problem can be represented by

$$\min_g \quad \sum_{t,i}(\pi_t - c_{t,i})g_{t,i}$$
$$s.t. \quad g_i^{min} \leq g_{t,i} \leq g_i^{max} \quad : \lambda_{t,i}^{min}, \lambda_{t,i}^{max} \tag{3}$$

where π_t is the market price at time t. In this case, the optimality conditions of the problem for $g_{t,i}$ can be written as

$$\pi_t = c_{t,i} + \left(\lambda_{t,i}^{max} - \lambda_{t,i}^{min}\right) \tag{4}$$

Comparing (4)4 to the optimality of the centralized problem, defined by (2), the solutions of both problems are the same when the price is identified with the shadow price of the balance constraint:

$$\pi_t = \lambda_t^{demand} = c_{t,i} + \left(\lambda_{t,i}^{max} - \lambda_{t,i}^{min}\right) \tag{5}$$

Therefore, when system prices are defined as $\pi_t = \lambda_t^{demand}$; i.e., the price is the shadow price of the balance equation in each hour, the producers' output decisions obtained from the individual problems are the same than the ones resulting from a centralized dispatch. In other words, with this price definition, all agents perceive the output decisions that are efficient from the global point of view to be also their optimal responses to the market price. The efficient solution is also a market equilibrium, fulfilling the equilibrium condition: once the solution is achieved, no player has an incentive to deviate from the optimal dispatch.

b. Price signals under integer non-convexities

The main idea behind the reasoning described in the previous section is to identify the market price with the short-term marginal cost, i. e. the cost of producing an additional megawatt. However, when integer non-convexities are taken into account, the definition of such cost is considerably difficult. The following integer program represents the analogue to problem (1) when integer decisions are considered:

$$
\begin{aligned}
\min_{g,u,v,w} \quad & \sum_{t,i}\left(c_{t,i}\,g_{t,i} + cb_i\,u_{t,i} + ca_i\,v_{t,i}\right) \\
s.t. \quad & u_{t,i}\,g_i^{min} \le g_{t,i} \le u_{t,i}\,g_i^{max} \\
& \sum_i g_{t,i} = d_t \\
& u_{t,i} = u_{t-1,i} + v_{t,i} - w_{t,i} \\
& u_{t,i}, v_{t,i}, w_{t,i} \in \{0,1\}
\end{aligned}
\tag{6}
$$

where the new variables $u_{t,i}$, $v_{t,i}$ and $w_{t,i}$ represent commitment, start-up and shut-down decisions, respectively, of unit i at time t. Parameters cb_i and ca_i are the no-load and start-up costs of unit i, respectively (we assume shut-down cost is zero). The third constraint represents a continuity constraint: if the power plant is on-line at time t, then it was either on-line at time $t-1$ or it is started up at time t. Conversely, if it is not on-line, then it was either off-line at time $t-1$ or it is shut down at time t.

Since this is a mixed-integer optimization, first-order optimality conditions cannot be defined as in the previous subsection, as they are only meaningful for linear problems. The optimization algorithms perform different kind of searching techniques, which go beyond the pure marginal analysis, to reach the best solution. Therefore, before we formalize the solution of this problem and compare it with the solution of the market players' problems, let us start with a simple example to identify the most relevant issues in the discussion.

Illustrative example

We will assume a single-hour problem, where demand is d and there are three generators $i = \{i_1, i_2, i_3\}$, with maximum output g_i^{max} and a cost function that only involves a start-up cost ca_i and a variable cost cv_i for each of the generators, being $ca_i < ca_{i+1}$ and $cv_i < cv_{i+1}$. This is a very simple example, but it keeps the essential feature of including discrete variables, which are the source of non-convexities. The optimization of the centralized problem is rather easy in this case: if demand is smaller than the capacity of the cheapest generator, then only this one should produce; if demand is higher than the maximum output of the cheapest generator, but smaller than the aggregate capacity of the first and second generators, then the first one should produce at its maximum and the rest of the demand should be provided by the second unit; if demand is larger than the capacity of the first two generators, then both of them should operate at their maximum and the rest of the demand should be produced by the third generator. It results in the curve, represented in Figure 1 of total production cost as generation increases. The optimal solution is the point where generation is equal to demand.
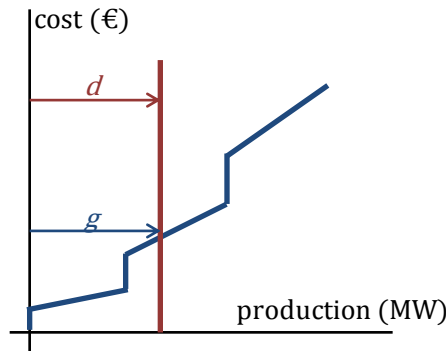
FIGURE 1. PRODUCTION COSTS AS A FUNCTION OF ELECTRICITY PRODUCED.

In a perfect market context, each generator decides its output maximizing income from market sales minus operating cost. If the prices were calculated from the direct application of the results of the convex case described in section a, then the price would be equal to the pure marginal cost: in this example, to the variable cost of unit i_2. The decision process of market players for this price is represented in Figure 2, where solid lines show their operating costs and the dotted ones represent their market income.
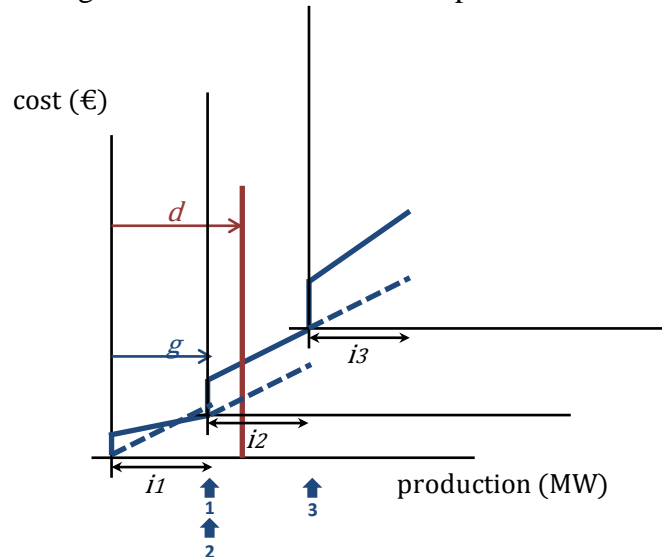

FIGURE 2. REPRESENTATION OF PRODUCERS' DECISION-MAKING PROCESS.

For generator i_1, operating margin is maximized when producing at its full capacity, since the price is higher than its marginal cost, and the higher the production the higher the net income. For generators i_2 and i_3, on the contrary, the best decision is not to produce at all, since the price is not enough to compensate for their full operating costs. Total generation under this price is lower than demand, and different from the centralized solution. Therefore, we can conclude that this price does not support the global optimum to be a competitive equilibrium.

The same result could have been obtained by noticing that the net profits associated with any plant can be written as $R = \pi g - C$, where C represents the production costs, g is the plant output and π is the market price. Thus, the net profits can be recast as $C = -R + \pi g$, so they constitute a set of straight lines with slope π and intercept $-R$. The individual decision of any certain market player, given the market price, is to choose the straight line with a slope equal to the market price and with the lowest

possible intercept (i.e., maximum possible operating profit), taking into account that only the lines with at least one point in common with the cost curve are feasible.

Therefore, a straight line with slope equal to the market price and tangent to the cost curve represents the optimal response of the system generation portfolio to a certain price. The optimal production is given by the intersection of both curves. In Figure 3 this reasoning is applied to the previous market price, the pure marginal cost, with the dotted line being the lowest line tangent to the cost curve and with the same slope than the variable costs of generator i_2, and it yields the same results as in Figure 2.
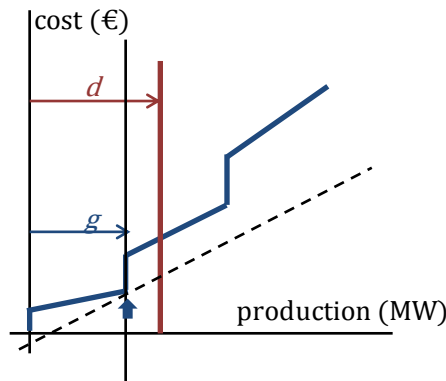

FIGURE 3. RESULTS ANALYZING NET PROFITS.

Again, generator i_2, which is the marginal plant in the optimal solution, does not recover its start-up costs, and as a consequence its optimal response to this price is to stop producing. The equilibrium implied by such price is different from the optimal dispatch. Since the price is calculated as the pure marginal cost, any cost that is related to the discrete decisions is not incorporated into the prices, so the marginal unit does not recover its full operating costs, and therefore it is not willing to produce. This price does not attract enough generation to fulfill demand.

An alternative candidate for the market price is the average cost of the marginal plant, which should ensure that the price pays for all of the operating costs of this generator, so it will come on line. This is represented in Figure 4.

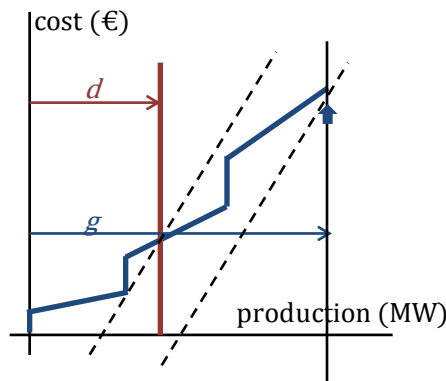**FIGURE 4. DEFINITION OF PRICE AS THE AVERAGE COST.**

This price is in fact enough to make generator i_2 willing to produce. However, i_2 does not maximize its profits by operating at the partial level determined by the global optimization, but rather by producing at its full capacity. Furthermore, the price is high enough to make plant i_3 willing to produce at its maximum too. Then, the total

generation induced by this price is much larger than demand. The equilibrium implied by these prices is not the optimal dispatch either.

A possible compromise solution is represented in Figure 5. In this case, the market price is defined as the average cost at its maximum output of the marginal plant.
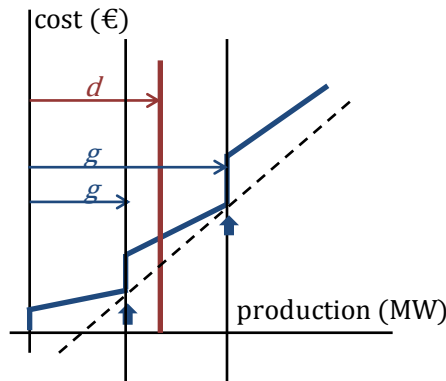

FIGURE 5. A COMPROMISE SOLUTION: PRICE AS THE AVERAGE COST AT MAXIMUM OUTPUT.

Under this solution, all the units that were not dispatched in the optimal solution are not willing to produce, and no infra-marginal unit has incentives to stop producing. This means that, for all units except the marginal one, this price supports the global optimum as equilibrium. However, the optimal response of the marginal plant is either to stop producing or to produce at its maximum output. Any generation level different than its maximum production would reduce the income of unit i_2, making it lower than its operating costs, and thus leading the plant to stop producing. The price described in Figure 5 does not create incentives for generator i_2 to produce at the level required to fulfill demand as in the optimal dispatch.

This example illustrates that often there is no price that supports the short-term competitive equilibrium in the market. If the system operates under perfect competition (i.e. each generator decides his production in order to maximize his own profits, considering that the price does not change upon his actions), then the equilibrium price should provide incentives to ensure that total production is equal to demand. In other words, the competitive equilibrium price should provide incentives for all the generators that have been dispatched in the centralized model to keep up with their production, and to all of the generators that have not been dispatched in the centralized model not to enter the market. However, in the example presented above any price slightly higher than the one described in Figure 5 would lead market players to produce more than demand; any price slightly lower than this one would lead market players to generate less than demand; and the actual price in Figure 5 would lead either to more or less generation than required, but not the exact quantity demanded. There is no price that supports the short-term competitive equilibrium.

General characterization: the duality gap

The lack of short-term equilibrium is not restricted to this example, it is rather a general result of the non-convexities of the cost functions. The following optimization describes the decision process of each of the producers, where they decide their production once the market price is known. It is the analogue to problem (3) when discrete variables are included.

$$\max_{g,u,v,w} \quad \sum_{t,i}\left(\pi_t \; g_{t,i}\right) - \sum_{i,j}\left(c_{t,i} g_{t,i} + cb_i \; u_{t,i} + ca_i \; v_{t,i}\right)$$

$$s.t. \quad u_{t,i} g_i^{min} \leq g_{t,i} \leq u_{t,i} g_i^{max}$$

$$u_{t,i} = u_{t-1,i} + v_{t,i} - w_{t,i}$$

$$u_{t,i}, v_{t,i}, w_{t,i} \in \{0,1\}$$

(7)

where π_t is the market price at time t. Compared to the centralized optimization described in problem (6), the generation-demand balance equation $\sum_i g_{t,i} = d_t$ has disappeared in problem (7) and, instead, a price element $\sum_{t,i}\left(\pi_t \; g_{t,i}\right)$ has been included in the objective function.

In fact, one could describe the market as a process where a central problem (the day-ahead auction) sets the price and each producer resolves his own subproblem to determine his production. In other words, this is a way of solving problem (6) through a decomposition technique, allowing for a distributed optimization (each generator decides his own production) and leaving price as a control variable that creates economic signals to ensure that the generation-demand balance is maintained. This kind of decomposition, where a certain constraint is replaced by a price signal, is known as Lagrangian relaxation, see for instance (Minoux and Vajda, 1986).

A common phenomenon when applying Lagrangian relaxation to non-convex problems is the existence of a duality gap. This is a difference between the solution of the primal problem (in our case, the centralized optimization) and dual problem (in our case, the market), which is zero when the problem is convex but not in the presence of non-convexities. More precisely, in the areas where the problem is not convex, the solution of the dual problem change discretely with small changes in the price signal, making it impossible to satisfy the relaxed equation. As a result, the solution of the primal problem, which satisfy exactly all of the constraints, differs from the solution of the dual problem.

This is the case with the day-ahead electricity market. There is no price that creates incentives for the generators to produce the same amounts that they were dispatched in the centralized model: in the market, all of them will produce either their full capacity or zero. Thus, the solution of the market is different than the global optimum. In the presence of non-convexities, there is no price that satisfies the conditions to be a short term competitive equilibrium.

c. Pricing under integer non-convexities: the need for side payments

The previous analysis shows that no linear price constitutes a signal for the optimal system operation. In that view, there is a need to add some specific mechanism to cope with the problem.

The most used approach is to add a side payment to the clearing price (following the original proposal of the UK, this term is often referred to as uplift). For instance, all previous proposals would require an additional payment aimed at compensating for each unit's incentives to deviate from the optimal dispatch. The first case of the example (see Figure 2), where the price is set to the variable cost of the marginal plant, requires defining an extra payment to all infra-marginal units that are not recovering their operation costs. The second case (see Figure 4), where the price is equal to the average cost of the marginal plant, requires a negative side payment to the marginal plant, and to all units above the margin with incentives to start producing. Finally, the third case (see Figure 5) requires an additional payment to the marginal plant. These additional payments are unavoidable, as the existence of integer variables precludes the possibility that some linear price supports the optimal dispatch.

The characterization of the previous side payments has been studied in a number of works. (O'Neill et al., 2005) proposes a pricing scheme based on integer duality. The pricing scheme proposed in (O'Neill et al., 2005) consists in setting the market price equal to the variable cost of the marginal plant. However, the income associated with the short-term market is not only the shadow price of the balance equation, which is mainly related to the variable cost of the most expensive plant producing in the optimal dispatch. These prices alone are not enough for all units to recover their operation costs: some of them will not perceive the fixed part of such costs. Actually, these units are the ones partially online in the linear relaxation, and thus the ones that require the additional constraint to achieve an integer solution. Hence, these plants earn an additional payment, the uplift, which compensates for the missing income. On the other hand, (Hogan and Ring, 2003) proposes a solution based on minimizing the uplifts. This pricing scheme corresponds to the third case discussed in the single-period example. Formally, it suggests the use of the convex hull slope to define the market price. In this case, the only plant of the system that requires uplift is the marginal plant. The convex hull can be calculated by the Lagrangian relaxation of the problem. Close to that solution, an alternative is proposed in (Bjørndal and Jörnsten, 2008), where a decomposition based on the Benders Decomposition, (Benders, 1962) is used.

Note, however, that the use of side payments implies a wide range of possible solutions. Actually, we might force income for all producing units to be equal or greater than their cost, and lower or equal to their costs if they are not producing. In the limit, hence, we might set the price to zero and pay all costs through side payments, or the other way around, set a very high price and use negative side payments for non-producing units.

The conclusion, from a market design standpoint, is that all prices have a complementary set of side payments that make them equilibrium prices. Under the criteria used in this section, it is not possible to choose among them, so that some additional element is required. This is the aim of the next section.

4. Optimal long-term signals

The previous section has analyzed several pricing solutions with regard to their ability to induce market players to choose the optimal dispatch, i. e. to behave optimally in the short run. Nonetheless, as shown in (Joskow and Schmalensee, 1988) or (Caramanis, 1982), among others, market prices should also provide optimal long-term signals, so

that they induce producers to build the optimal generation portfolio. In fact, a considerable part of the advantages associated with the use of markets depends on whether the market model can achieve optimal investment decisions. In this section, we will use this additional criterion to analyze pricing mechanisms for combinatorial electricity auctions.

a. Investment problem without integer variables

The objective of centralized planning is the minimization of the sum of investment and operation costs. Thus, a new investment is undertaken as long as the fixed cost of the new plant is lower than the operation savings associated with the investment. The problem can be represented by the following program:

$$
\begin{aligned}
min \quad & \sum_i ic_i \; g_i^{max} + \sum_{t,i} c_{t,i} g_{t,i} \\
s.t. \quad & 0 \le g_{t,i} \le g_i^{max} && : \lambda_{t,i}^{min}, \lambda_{t,i}^{max} \\
& \sum_i g_{t,i} = d_t && : \lambda_t^{demand}
\end{aligned}
\tag{8}
$$

The optimality conditions of the problem are

$$
ic_i \; = \sum_t \lambda_{t,i}^{max} \tag{9}
$$

$$
\lambda_t^{demand} = c_{t,i} + \left(\lambda_{t,i}^{max} - \lambda_{t,i}^{min} \right) \tag{10}
$$

On the other hand, under a market environment, the objective is the maximization of firms' total profits. Hence,

$$
\begin{aligned}
max \quad & -\sum_i ic_i \; g_i^{max} + \sum_{t,i} \left(\pi_t - c_{t,i} \right) g_{t,i} \\
s.t. \quad & u_{t,i} g_i^{min} \le g_{t,i} \le u_{t,i} g_i^{max} && : \lambda_{t,i}^{min}, \lambda_{t,i}^{max}
\end{aligned}
\tag{11}
$$

Along the same lines as in the centralized problem, the optimality conditions can be written as

$$
ic_i \; = \sum_t \lambda_{t,i}^{max} \tag{12}
$$

$$
\pi_t \; = c_{t,i} + \left(\lambda_{t,i}^{max} - \lambda_{t,i}^{min} \right) \tag{13}
$$

15

That is, the new investment is undertaken as long as the profits associated with the new plant's production are equal to or greater than the investment cost. When the market price is set to λ_t^{demand}, the results obtained with both models are the same. Therefore, the generation portfolio obtained under a market environment is optimal. Equivalently, defining market prices as the shadow price of the balance equation gives optimal long-term signals.

It is worth to illustrate the problem by means of a simple example, where the possible investments are restricted to three technologies:

- Tech1 has no investment costs and a high operation cost (these are the typical characteristics of non-served energy)

- Tech2 has fixed costs defined by IC2, and intermediate operation costs

- Tech3 has the highest investment cost and the lowest operation cost

Figure 6 shows the efficient solution for the use of available technologies. For instance, if some megawatt will be used for a number of hours higher than H2, the best solution is to build that megawatt of Tech3, whereas within H2 and H1, the best investment is of Tech2.
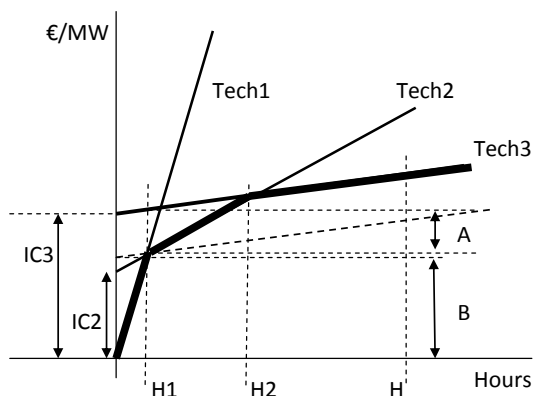

FIGURE 6. ADAPTED GENERATION PORTFOLIO.

Once H1, H2 and H3 are obtained, the optimal investment can be calculated as in Figure 7.
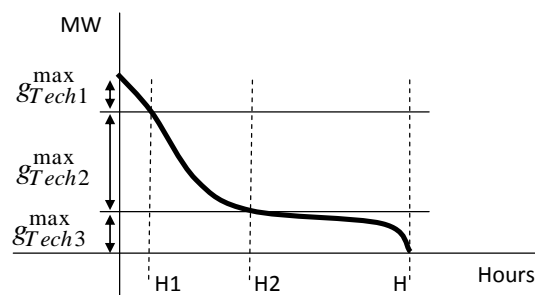

FIGURE 7. INVESTMENT FOR THE ADAPTED GENERATION PORTFOLIO.

Let us analyze, as an example, the way in which market prices allow producers recovering their investment costs in Tech3. In the period from H2 to H, Tech3 sets the

16

price, and thus it is equal to Tech3's variable cost. This price is enough to pay for Tech3's operation cost, but it does not give any payment to compensate for the investment cost. From H1 to H2, the price is equal to Tech2's variable cost, and thus Tech3 receives operation profits equal to the difference between Tech2's variable cost and its own variable cost (represented by the segment A of Figure 6). In the hours from zero to H1, the segment B of Figure 6 represents Tech3's profits. By construction, A + B = IC3.

It is important to highlight that the central idea behind the previous analysis is the description of the decision-making process to obtain the adapted generation portfolio, or equivalently, the generation portfolio that minimizes the system cost. The question is whether it is possible to design a pricing mechanism that plays the same role when integer variables are considered in the operation problem.

b.   Long-term signals under integer non-convexities

Under perfect competition, a certain firm decides to build a new power plant as long as the operation revenues that the plant can obtain from selling energy in the spot market are higher than the corresponding investment costs. If a change in the spot market rules results in lower operation profits, this will surely affect future investment decisions. One of the primary features of the pricing scheme analyzed above, in the context of continuous operation, is that spot prices are used to cover investment costs in infra-marginal units. From the example represented in Figure 6 and Figure 7, investing in Tech3 is justified as long as more expensive technologies set the spot price during hours H2. Formally, this property is represented by equation of (12).

In this view, spot prices are not only signals for the optimal operation of the system, but also for the optimal investment. To put it another way, when analyzing different pricing schemes with regard to operation incentives, the question was whether the price incentivizes supra-marginal units to be willing to enter into the spot market, or conversely, it represents an incentive to shut down plants that should be producing. Hence, it might seem that base-load plants –in the sense of plants producing under any pricing scheme– would not be affected by a change of pricing rules. However, if the long-term problem is taken into account, lower spot prices makes them to earn less operation profits, and thus to deviate from optimal investment decisions. This is the central idea that we will exploit in this section, with the aim of designing a pricing scheme that incentivizes not only the optimal operation, but also the optimal investment. From the cost minimization point of view, optimal investment decisions are taken to maximize savings associated with an additional megawatt of a certain technology. To gain some intuition on the problem, consider the case of the base-load technology. The idea of the investment problem is that, in a centralized problem, an investment in one megawatt of the base-load technology is undertaken as long as its investment cost is less than the operation savings associated with producing one megawatt with the base-load technology. In a problem without non-convexities, this savings can be viewed as the difference between the variable cost of the marginal plant and the base-load technology variable cost. The previous problem, on the other hand, implies the same condition in the investment problem. However, an additional megawatt of base-load technology also represents savings with regard to fixed costs –startups…– of marginal power plants. When the market pays for fixed operation costs in a discriminatory manner, infra-

marginal units do not receive signals associated with the savings corresponding to marginal plants' fixed costs.

To illustrate the previous idea, it is possible to consider the following simple example, represented in Table 1. There are two available technologies, just one hour, and one additional megawatt is necessary in the system. Note that both the start-up and the investment costs are normalized to a unit investment, that is, the total cost of divided by the maximum output of the plant.

	Tech1	Tech2
Variable Cost (€/MW)	5	3
Unit Start-up Cost (€/MW)	4	2
Unit Investment Cost (€/MW)	0	2

TABLE 1. TWO-TECHNOLOGY EXAMPLE.

In addition, let us assume that both technologies are required to start up. Whether to invest in Tech1 or Tech2 is decided based on:

$$Total\ Cost\ (Tech1) = g_{Tech1}^{max}(5+4) > Total\ Cost(Tech2) = g_{Tech2}^{max}(3+2+2)$$

That is, Tech2 is cheaper than Tech1. However, under a pricing proposal based on a price equal to the variable cost, plus a discriminatory payment, the price signal received is

$$Total\ Cost\ (Tech2) = g_{Tech2}^{max}(3+2+2) > Total\ Income(Tech2)$$
$$= g_{Tech2}^{max}VarCost(Tech1)$$

and hence the signal sent by the spot market is that the optimal investment is to build one extra megawatt of Tech1. The intuition behind the problem with the previous pricing scheme is that the start-up cost of Tech1 is covered by a side payment, and hence it is not a signal for the investment in Tech2. The central idea of the following reasoning is to show how that signal can be included in the market prices.

As pointed out in (Kahn, 1988), one of the first steps for marginal cost pricing is to define the unit investment. To put it another way, we are interested in defining the cost of an incremental unit of output. The most usual way of doing so is to consider that a unit investment in any technology, say in CCGT, allows producing one extra megawatt, which has a cost defined by the variable cost associated with the technology. Implicitly, this model assumes that short-term fixed costs (startup and shutdown costs, etc.) are the same as the ones corresponding to a large power plant. That is, a unit investment would be able to produce only one megawatt, but for doing so, it would be forced to pay for the same fixed costs as a large power plant. In contrast, we will assume that there are no economies of scale, so that the small power plant representing the incremental block of production will have, in addition to small capacity, small short-term fixed costs. This approach is aimed at representing that the relevant cost of an additional unit is not only

the corresponding variable cost, but also the associated part of short-term fixed costs. Therefore, we will model investments using a variable representing the small power plant as a whole, which will be termed as x_i. Furthermore, the absence of economies of scale implies that x_i can take any value. In this view, the small power plant will be defined by some standardized capacity $x_i g_i^{\max}$, the corresponding variable cost $c_{t,i}$, and the unit fixed costs at maximum capacity: the online cost $co_i x_i$, and the startup cost $cs_i x_i$. The logic for that definition is that, although in the short term startup costs and the like are fixed costs, when deciding on investments, power producers can decide on the size of the plants, and thus on the relative value of the fixed costs. In this view, considering no economies of scale is equivalent to consider that the power plant can be made of any size. This will be discussed below.

With that in mind, we can write the analogue to investment problem (8) using the following model:

$$
\begin{aligned}
max \quad & -\sum_i ic_i x_i + \sum_{t,i}(\pi_t - c_{i,t})g_{t,i} - co_i x_i u_{t,i} - cs_i x_i v_{t,i} \\
s.t. \quad & u_{t,i} x_i\ g_i^{min} \le g_{t,i} \le u_{t,i} x_i\ g_i^{max} \\
& u_{t,i} = u_{t-1,i} + v_{t,i} - w_{t,i} \\
& u_{t,i}, v_{t,i}, w_{t,i} \in \{0,1\}
\end{aligned}
\tag{14}
$$

In addition, the problem can be reformulated as

$$
\begin{aligned}
max \quad & -\sum_i ic_i x_i + \sum_{t,i}(\pi_t - c_{i,t})g_{t,i} - co_i u_{t,i} - cs_i v_{t,i} \\
s.t. \quad & u_{t,i} g_i^{min} \le g_{t,i} \le u_{t,i} g_i^{max} \\
& u_{t,i} = u_{t-1,i} + v_{t,i} - w_{t,i} \\
& u_{t,i}, v_{t,i}, w_{t,i} \in \{0, x_i\}
\end{aligned}
\tag{15}
$$

where x_i can take any value. In principle, the problem is still a nonconvex one, so the definition of a linear pricing scheme is still impossible. However, it is possible to exploit the fact that x_i is a continuous variable. As power producers can build infinitely small plants of any technology, i. e. x_i can be as small as required, any cost related to the duality gap can be eliminated without cost, by substituting one unit by several smaller units. For instance, consider that a certain marginal plant is producing at its maximum capacity in one hour of the time scope, below its maximum output in another hour, and not producing the rest of the hours. It is always possible to substitute it by two smaller plants of the same technology, with added capacity equal to the large one, so that one of them is producing in the first hour and both of them are producing in the second hour. Though there are no economies of scale and the substitution is costless, it allows eliminating the duality gap.

The idea behind the following reasoning is that such problem is equivalent to a linear problem. In fact, the duality gap will be way from zero when the startup and online

variables are below their limits in the linear relaxation of the problem. However, this will only happen when the amount of plants in the problem, or equivalently, the subscript i is fixed at some relatively low value. But, from the investment standpoint, this is not necessary because power producers can undertake, in principle, any profitable investment. Therefore, if the number of possible plants is high enough (higher than the highest demand in the scope of the problem), the optimal solution will only contain variables at their integer limits. This is the definition of the integrality property, so the solution of the linear relaxation is the same as the integer solution. The property that we have exploited above is that it is possible to consider an ideal generation portfolio made up of many small power plants. And that is possible as long as there are no economies of scale. The linear relaxation of the problem is:

$$
\begin{aligned}
max \quad & -\sum_i ic_i x_i \; + \sum_{t,i}(\pi_t - c_{i,t})g_{t,i} - co_i u_{t,i} - cs_i v_{t,i} \\
s.t. \quad & u_{t,i}g_i^{min} \le g_{t,i} \le u_{t,i}g_i^{max} && : \lambda_{t,i}^{min}, \lambda_{t,i}^{max} \\
& u_{t,i} = u_{t-1,i} + v_{t,i} - w_{t,i} && : \lambda_t^{online} \\
& 0 \le u_{t,i}, v_{t,i} \le x_i && : \lambda_{t,i}^{u}, \lambda_{t,i}^{v}
\end{aligned}
\tag{16}
$$

and the optimality conditions of this problem, with respect to investment decisions, is

$$
ic_i = \sum_t \lambda_t^u
\tag{17}
$$

where the Lagrange multiplier is defined by

$$
co_i + \left(g_i^{min}\lambda_{t,i}^{min} - g_i^{max}\lambda_{t,i}^{max}\right) + \left(\lambda_{t+1}^{online} - \lambda_t^{online}\right) + \lambda_{t,i}^u = 0
\tag{18}
$$

The investment signal in this problem takes into account fixed operation costs. Note also that, following the same reasoning (modeling investments in small power plants), the inclusion of additional technical constraints, such as ramps or reserve constraints, does not change the previous results.

c.   Analyzing the "small plants" assumption

In practice, economies of scale often exist. For instance, a CCGT cannot be found in the market with a maximum output of 10 MW. This minimum-size requirement would prevent the investment variable x_i to take any value, but it will be an integer variable. Consequently, there would be non-convexities associated with the investment variable. That problem was pointed out in (Scarf, 1994), where it was shown, using a simple example, that some technologies, in spite of having more expensive average costs, led to lower total costs in the system. The idea behind the example is that, when the size of the smallest plant available in the system is relatively large with respect to the system demand, the duality gap (the cost associated with the minimum-size requirement) could be significant and it is worth to invest in plants that are more expensive, in order to

minimize it. The example points out that, if the duality gap is significant in a certain industry, linear prices may be misleading signals. When the number of required plants is low and the demand is inelastic, the size effect represents the savings related to reducing the duality gap. Linear prices cannot capture this effect, as they do not represent the duality gap, and thus give incomplete signals.

Nonetheless, if the required number of plants are high enough, or equivalently, if the minimum size of the plants is small enough, the duality gap will be low. Moreover, in practice the system demand is subject to great uncertainty, especially with horizons as the ones required to deciding on investments. When the demand uncertainty is higher than the minimum size of the power plants, the economies of scale cannot be certainly defined, and they tend to disappear. Both features, which are associated with developed power systems, allow disregarding such economies of scale. When there are significant economies of scale (as it could be the case in a small power system), it is not clear whether competition is practicable. A small number of plants, together with large economies of scale, would suggest the possibility of a natural monopoly. In that case, the industry will likely benefit from organizing production under a centralized planning, and hence the problem of defining the pricing scheme will be irrelevant.

The previous problem, however, does not take into account the duality gap. In this regard, any pricing scheme providing signals according to this model will not represent the cost associated with the duality gap. To put it another way, any investment that is profitable to lower the duality gap, but that it is not without considering the integrality of the problem, will not be undertaken.

5. Numerical example

This example is aimed at illustrating the long-term effects of different pricing schemes, using the same reasoning as the previous section. The main idea behind the methodology is to analyze the revenues associated with the investments under two pricing schemes, so that the long-term signals given by each pricing mechanism determines the generation portfolio. To do so, we will use the static planning criteria described above, so that each plant will modify its capacity until operation profits are equal to investment costs. The example considers a time scope of seven hours. In this regard, the fundamental mechanism for producers to balance investment costs and operation profits are rather inflexible in this simplified example with only seven periods. Thus, some input data of the example –e. g. variable costs– are defined to avoid numerical problems.

With regard to the system demand definition, in addition to each hour demand, it is imposed that the total capacity of the system must exceed a certain value, in order to represent the economic impact of peak periods, which have not been explicitly modeled.

	h1	h2	h3	h4	h5	h6	h7	Total capacity
Demand (MW)	300	400	515	850	600	220	120	1000

TABLE 2. DEMAND DATA.

	Variable cost (€/MWh)	Online cost (€/h)	Start-up cost (€)	Investment cost (€/MWh)	Capacity (MW)
g1	2.5	33.33	42.5	3.5083	100
g2	3	16.67	50	1.7667	100
g3	3.73	12	89.33	0	100

TABLE 3. GENERATION DATA.

a. Prices obtained from the linear relaxation

Firstly, we will analyze the results corresponding to defining the spot price equal to the linear relaxation (note that this pricing rule is related to some examples in the US[4]), as proposed in section 4.b. These results are represented in Table 4 and Table 5.

	Capacity	Operation cost	Investment cost	Income	Net profits
g1	250	4665	877	5542	0
g2	330	3815	583	4398	0
g3	420	1559	0	1559	0

TABLE 4. GENERATION PORTFOLIO ADAPTED TO THE PRICING SCHEME PROPOSED IN THIS PAPER

	h1	h2	h3	h4	h5	h6	h7	SP
g1	250	250	250	250	250	220	120	53
g2	50	150	265	330	330	0	0	88
g3	0	0	0	270	20	0	0	200
Price	3.2	3.2	3.2	4.75	3.85	2.83	2.83	
Total generation cost: 10038 €								

TABLE 5. PRODUCTION AND PRICES CORRESPONDING TO THE PRICING SCHEME IN THIS PAPER.

Under this pricing scheme, spot prices in all periods include the impact of online costs. For instance, the price set by g2 in h1, h2 and h3 is not only its variable cost, but also its online cost divided by its capacity.

On the other hand, with regard to the part of revenues that is not associated to spot prices, g1 receives a side payment equal to part of its online costs associated with production in h6 and h7, which is proportional to the capacity that is online but not producing in these hours. Thus, the unit earns a side payment equal to 0.33x(250-220) in h6, and 0.33x(250-220) in h7. Analogously, g2 earns a compensation for part of its online costs equal to 0.17x(330-50) in h1, to 0.17x(330-150) in h2 and to 0.17x(330-265) in h3. Finally, g3 earns a side payment compensating for not only its online costs,

---

[4] "Real-time prices are set by the ideal dispatch pass, in which inflexible (i.e., they must operate at zero or their maximum output) gas turbines are dispatched economically over their entire operating range, even if they are not actually capable of running at anything other than zero or their maximum output." Federal Energy Regulatory Commission Order, FERC Docket No. ER05-1123-000, July 19, 2005.

but also for its start-up costs. Thus, g3 earns 0.12x(420-20) corresponding to online costs in h5, and 0.12x(420-270) plus 0.89x(420-270) in h4, the first term representing online costs and the second start-up costs.

b. Prices equal to variable cost of marginal plant

Firstly, we analyze the generation portfolio resulting from the pricing scheme proposed in (O'Neill et al., 2003), where spot price is the variable cost of the marginal unit in each period. The mechanism also defines side payments to the power plants that do not cover their operation costs. The corresponding results are represented in Table 6 and Table 7.

	Capacity	Operation cost	Investment cost	Income	Net profits
g1	200	3852	702	4553	0
g2	190	2925	336	3261	0
g3	610	3843	0	3843	0

TABLE 6. GENERATION PORTFOLIO ADAPTED TO THE PRICING SCHEME PROPOSED IN (O'NEILL ET AL., 2005).

	h1	h2	h3	h4	h5	h6	h7	SP
g1	200	200	200	200	200	200	120	67
g2	100	190	190	190	190	20	0	63
g3	0	10	125	460	210	0	0	838
Price	3	3.73	3.73	3.73	3.73	3	2.5	
Total generation cost: 10620 €								

TABLE 7. PRODUCTION AND PRICES CORRESPONDING TO THE PRICING SCHEME IN (O'NEILL ET AL., 2003).

In this case, the price set by g3 in h4 is lower than the one obtained in the previous case, as it does not include any part of the fixed operation costs. This makes the income of g2 lower and thus the investment associated with g2 decreases. Consequently, g3 sets the price in two additional hours, h2 and h3, with respect to the linear price.

Spot prices set by g2, the ones corresponding to h1 and h6, do not include either fixed operation costs. Hence, the income corresponding to g1 is lower than in the previous case and the installed capacity of this technology decreases. In turn, this effect makes g1 to stop setting the price in h6.

Compared to the linear price case, it can be observed that the operation costs are higher in this case, and that the increase of operation costs is higher than the savings in investment costs. Therefore, some investments that could result in operation savings greater than their investment costs have not been undertaken. This implies that spot prices are too low to provide with enough long term incentives for the investment problem. When producers are deciding on the amount of new investments, they do not perceive –in terms of expected operation profits– the savings in fixed operation costs that would result from the new investments. Hence, producers build less base-load units than the ones required for an optimal generation portfolio. Such under-investment is compensated for with peaking plants.

6.   Concluding remarks

We have shown that there is no unique definition for prices in a power auction with integer decisions that support a short-term competitive equilibrium. There are no linear prices that send the optimal short-term signals alone, without a side payment, due to duality gap, and one can find a wide range of combinations of marginal prices and side payments that represent optimal short-term signals. We have also shown that, when one considers long-term competitive equilibrium requirements, the range of possible solutions is significantly restricted. To show that, we consider a situation where plants are small compared to size of the system and to the magnitude of the uncertainty affecting long-term decisions. Under those hypotheses, if we consider explicitly that the infra-marginal rents that base-load generators receive have an influence on their investment decisions, and look for the prices that induce generation equipment to be optimal, we find that the prices defined by a very specific linear relaxation of the long-term problem turns out to be a single solution.

From the market design point of view, the results obtained in this paper allow pointing at a direct problem of day-ahead auction design. The increasing need of backup generation complicates significantly the operation of power systems. In particular, the amount of start-ups that are needed under considerable amount of renewable generation (typically non-dispatchable technologies) puts the representation of those integer decisions at center stage.

Moreover, we have shown the importance of the pricing methodology to choose the production technology portfolio. In a situation where technology choice is one of the most important challenges of power system planning, short-term prices may be an extremely useful mechanism to coordinate long-term decisions. But that requires that those short-term prices include relevant cost signals.

7. References

Benders, J.F., 1962. Partitioning procedures for solving mixed-variables programming problems. Numer. Math. 4, 238–252. doi:10.1007/BF01386316

Bjørndal, M., Jörnsten, K., 2008. Equilibrium prices supported by dual price functions in markets with non-convexities. European Journal of Operational Research 190, 768–789. doi:10.1016/j.ejor.2007.06.050

Caramanis, M., 1982. Investment decisions and long-term planning under electricity spot pricing. Power Apparatus and Systems, IEEE Transactions on 4640–4648.

Hobbs, B.F., 2001. The Next Generation of Electric Power Unit Commitment Models. Springer Science & Business Media.

Hogan, W.W., Ring, B.J., 2003. On minimum-uplift pricing for electricity markets. Working Paper, Harvard Electricity Policy Group.

Joskow, P.L., Schmalensee, R., 1988. Markets for power: an analysis of electrical utility deregulation. MIT Press Books 1.

Kahn, A.E., 1988. The economics of regulation: principles and institutions. Mit Press.

Milgrom, P., 2004. Putting Auction Theory to Work, Cambridge University Press.

Minoux, M., Vajda, S., 1986. Mathematical programming: theory and algorithms. Wiley New York.

O'Neill, R.P., Sotkiewicz, P.M., Hobbs, B.F., Rothkopf, M.H., Stewart Jr., W.R., 2005. Efficient market-clearing prices in markets with nonconvexities. European Journal of Operational Research 164, 269–285.

Scarf, H.E., 1994. The allocation of resources in the presence of indivisibilities. The Journal of Economic Perspectives 111–128.

Vazquez, C., 2003. Modelos de casación de ofertas en mercados eléctricos. Tesis Doctoral. Universidad Pontificia Comillas, Madrid.

Vickrey, W., 1961. Counterspeculation, Auctions, and Competitive Sealed Tenders. The Journal of Finance 16, 8–37. doi:10.1111/j.1540-6261.1961.tb02789.x

Wilson, R., 2002. Architecture of Power Markets. Econometrica 70, 1299–1340.

Wilson, R.B., 1997. Activity Rules for the California PX Electricity Auction. 3rd Annual Research Conference on Electricity Industry Restructuring, University of California, Berkeley, CA (March 1998).