Bidding and Investment in Wholesale Electricity Markets: Pay-as-Bid versus Uniform-Price Auctions

Bert Willems [*]   Yueting Yu [†‡]

May 16, 2022

Abstract

We compare two auction formats for wholesale electricity markets: uniform-price and pay-as-bid auctions, and study short-run bidding behavior and long-run investment incentives. A perfect competition model with a continuum of generation technologies and uncertain, elastic demand is developed. Pay-as-bid auctions are shown to be inefficient because consumers' willingness to pay exceeds the marginal cost and producers' long-run investment incentives are distorted. The generation mix relies too little on baseload technologies. Consumers' surplus is lower.

JEL Codes: D44, D47, L94

Keywords: Power Markets, Market design, Uniform-price auctions, Pay-as-bid auctions, Investment.

[*]Tilburg University and Toulouse School of Economics. Email: b.r.r.willems@tilburguniversity.edu

[†]Tilburg University. Email: y.yu@tilburguniversity.edu

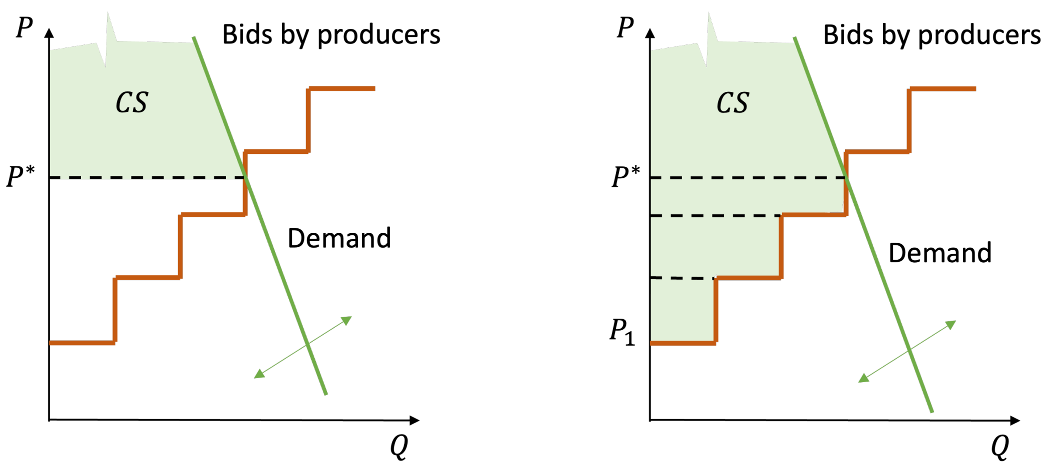
[‡]This is a preliminary version of our paper updated on May 16 2022.

1. Introduction

Electricity cannot be stored in a large scale. This means in wholesale electricity markets, production and consumption levels must match at any moment in time. The unique characteristic distinguishes electricity market from traditional financial markets and other commodity markets, and poses significant challenges for designing electricity market auctions.

As a result of unprecedented high electricity prices in 2021 and 2022, the EU Agency for the Cooperation of Energy Regulators (ACER) is considering an alternative price formation models to replace the current uniform-price auctions to provide reliable and affordable electricity (ACER, 2021, European Commission, 2022). The main goal is to decouple the electricity price from the marginal technology which are gas and coal. Pay-as-bid auctions are among the alternatives.

To organize real-time wholesale electricity markets, there are two alternative approaches: uniform-price auctions and pay-as-bid auctions. After submitting a quantity and the corresponding price, the system operator chooses the lowest bids to balance the aggregate demand and supply. The price of the last winning producer selected becomes the market clearing price. In uniform-price auctions, all winning producers receive the market clearing price, whereas in pay-as-bid auctions, winning producers are "paid as they bid" respectively.



(a) A uniform-price auction (UPA). **(b)** A pay-as-bid auction (PABA).

Figure 1: Comparison of uniform-price auctions and pay-as-bid auctions.

For an expositional purpose, Figure 1 depicts how these two auctions work. Bids by producers are represented by the increasing step function in orange. The green downward sloping curve shows the aggregate demand. The market clears at the intersection (p^* in Figure 1), where the aggregate demand balances the supply. In a uniform-price auction, also known as clearing-price auction, all accepted bids are paid the price of the marginal offer, so-called market-clearing price, while in a pay-as-bid auction (aka discriminatory price auctions), all winning producers are remunerated at their bidding prices, respectively (i.e., p_1, p_2 and p_3 in Figure 1b). The consumer surplus is illustrated by the green shaded region, which is the difference between consumer's willingness to pay and the price(s) received by winning producers. The average price paid by consumers depends less on the marginal technology under pay-as-bid auctions.

However, which auction design results into a higher consumer surplus is not obvious. Producers' decisions on how aggressively they bid depend on the auction format. The marginal bid as well as the market clearing price may not be the same.

In this paper, we compare uniform-price auctions and pay-as-bid auctions in wholesale electricity markets, and we examine the bidding behaviors and price-cost mark-ups in the short run, as well as producers' investment incentives and generation portfolio in the long run. We characterize the market equilibrium in the setting of perfect competition. Our model captures some important features of wholesale electricity markets. The demand is uncertain and elastic. Demand elasticity is relevant for long-run . Meanwhile, we model multiple generation technologies by introducing a continuum of generation technologies, from base-load to peak-load.

We show that in comparison to uniform-price auctions, pay-as-bid auctions are inefficient from both short-run and long-run perspectives. In the short run, consumers' willingness to pay exceeds the producers' marginal costs. In the long run, we find that the revenue of base-load producers is depressed during high demand periods, which distorts the generation mix. Additionally, consumer surplus is lower under pay-as-bid auctions. In a nutshell, we find that uniform-price auctions provide the producers with correct short-run price signals and induce sufficient investment incentives in the long run. Therefore, we believe that ACER's proposal of switching to pay-as-bid auctions requires scrutinizing. The switch may achieve its goal of lowering prices in the short run, but brings more inefficiency problems in the long run.

Which auction design is privileged remains controversial. The frequent observations of extreme price spikes and constantly increasing electricity prices with uniform-price auctions have raised a great many concerns that the inefficiency of the market system attributes to the auction design. Uniform-price auctions are more subject to it, considering that producers have no incentives to deviate from the collusive strategies. Evans and Green (2003) find empirical evidence of reduced prices after introducing pay-as-bid auctions.

Departing from the expectation that pay-as-bid auctions would intensify the market competition and decrease average prices, however, multiple drawbacks of pay-as-bid auctions have been identified in the literature. To name a few, first of all, pay-as-bid auctions cannot assure production efficiency and hence are unlikely to combat the soaring price situation (Wolfram, 1999, Kahn et al., 2001). Under pay-as-bid auctions, it is not always the case that producers with lower marginal costs are accepted in priority. Producers need to bid above their marginal costs to recover their investment cost, and also below the market-clearing price to get accepted. Small firms, even with efficient technologies may end up with an inefficient merit order, if they cannot afford the cost of making predictions. This means producers that benefit from the design are those who own more information about predicting the market-clearing price, rather than those with most efficient technologies.

Second, conducting pay-as-bid auctions drives the market further away from the goal of creating a competitive wholesale electricity market. On one hand, dominant producers gain informational advantages under pay-as-bid auctions (Bower and Bunn, 2001). These dominant firms can exert their market power not only by withholding capacities and raising the price, but more seriously by manipulating the price setting mechanism under the complex auction design (Harbord and McCoy, 2000). On the other hand, the introduction of pay-as-bid auctions would bring difficulties for regulators to identify the exercise of market power, since firms greatly change bidding behaviors that become less transparent (Kahn et al., 2001). Moreover, the switch would even exacerbate the fear of regulatory uncertainty that further discourages the investment in capacity (Tierney et al., 2008).

Our research contributes to the existing literature mainly in two themes. The first theme is auction theory in multiunit settings. Intense study has been undertaken to compare two alternative auction designs in the context of wholesale electricity markets; see Table 1. In this line of research, researchers have been mostly focusing on the comparison of short-run performances. Federico and Rahman (2003) construct a competitive model with multiple technologies, and they find that for a given demand realization, infra-marginal plants are remunerated less, while marginal plants receive more compared to uniform-price auctions. Hence, pay-as-bid auctions are less efficient but increase consumer surplus. Both Holmberg (2009) and Fabra et al. (2006) construct models with market power and a single technology. Holmberg (2009) applies a supply function equilibrium model, while Fabra et al. (2006) considers a duopoly model, where the two suppliers submit a single price offer for their entire capacity. These two papers end up with the same conclusion that average prices in pay-as-bid auctions are lower than those in uniform-price auctions and consumer surplus is higher. Among this whole set of literature, Federico and Rahman (2003) is closest to ours, with the setting of elastic demand and perfect competition.

However, most analysis focuses on comparing short-run performances of the two auctions, and there is a lack of adequate understanding of the interaction between producers' short-run and long-run decision making. Markets must provide the right investment signals to promote the long-run efficiency. Our research bridges this gap by building a tractable model to investigate the long-term efficiency of the two alternative auction designs on electricity market efficiency for one end of the spectrum of market structure – perfect competition. Similar to our paper, Fabra et al. (2011) also involve capacity investment in their research. They extend the work of Fabra et al. (2006) and find that pay-as-bid auctions generally lead to more competitive behaviors and lower prices than uniform-price auctions. The aggregate capacity stays the same under reasonable assumptions. Our result is in line with this, but on top of that, by introducing a continuum of generation technologies, we find the generation mix is inefficient with pay-as-bid auctions.

Moreover, we have demand responsiveness in our model and conclude a decrease in welfare under pay-as-bid auctions, which, in many of these literature, e.g. Fabra et al. (2006, 2011), Holmberg (2009), demand is inelastic. In general, the debate about which multi-unit auction mechanism generates superior outcomes in the context of electricity market is far from settled.

	Demand	**CS**	**W**	**Invest**	**Model**
Federico & Rahman '03	elastic	+	–	no	perf. comp monop.
Holmberg '09	inelastic	+	=	no	oligopoly SFE
Fabra et al. '06	inelastic	+	=	no	duopoly
Fabra et al. '11	inelastic	+	=	yes, 1 tech	duopoly
Our paper	**elastic**	**–**	**–**	**yes, ∞ tech**	**perf. comp.**

Table 1: Literature on auction theory in multiunit settings. CS: consumer surplus, W: welfare.

This paper also adds to the literature on the optimal portfolio choice in wholesale electricity markets. Since electricity is non-storable, a mix of generation technologies has to be selected and the respective prices for each technology have to be determined to efficiently serve the uncertain and unpredictable demand. The first papers in this literature can be dated back to Boiteux (1960) and Steiner (1957) that demonstrate the optimal capacity chosen by a monopolistic generator in the peak-load problem. Joskow and Tirole (2007) introduce competition that fits the electricity market reality after liberalization, and derive the optimal price setting and investment decision. Zöttl (2010) further investigate how market power affects firms' investment decision by comparing the cases of perfect competition and monopoly. Our paper illustrates how to choose the optimal generation mix in an elastic-demand setting under perfect competition. We compare the market outcomes where firms choose their capacities in the first stage, and in the second stage, the remunerated generators are selected by two alternative auctions.

The rest of our paper is organized as follows. Section 2 sets up our model. Section 3 characterizes the bidding equilibrium and the investment equilibrium for uniform-price auctions and pay-as-bid auctions. Then we illustrate and compare the equilibrium outcomes of the two alternative auction formats with a functional-form model in Section 4. Finally, we draw some conclusions in Section 5.

2. The Model

This section presents a competitive power market that consists of two stages: producers first make long-term investment decisions and then participate in a multi-unit spot market auction by submitting a supply bid. We consider both uniform-price and a pay-as-bid auctions. Consumers are price-takers and are represented by a stochastic demand function. Bids are assumed to be submitted before demand is realized, so bids are long-lived. In the remainder of this section, we first describe the technological assumptions regarding production technologies and the demand side, market clearing and then present the equilibrium concepts for spot market bidding and investment decisions.

2.1 Supply

We assume an atomistic market structure with a continuum of electricity producers that are price-takers in the spot market auction and are free to enter the market by investing in a generation technology of their choice. There are no entry barriers.

Each producer invests in an infinitesimal capacity unit dG, which can produce output dq.

A generation technology is represented by its marginal cost c and an annualized investment costs k. We assume that a continuum of technologies are available, corresponding to marginal cost c on the half-open interval $(0, \hat{c}]$, where $\hat{c}$ is the technology with the highest marginal cost. The investment cost of a technology with marginal cost c is the function $k(c)$, which represents the technology frontier of all existing production technologies (Figure 2). It has the following properties. First, it decreases with technology c, $k'(c) < 0$. This implies that power plants with lower operating costs have higher fixed costs and vice versa. For instance, to construct a nuclear plant, the investment cost $k(c)$ is relatively high, while the marginal cost is relatively low. On the other hand, some conventional energy sources such as oil, coal and natural gas, have lower investment costs, but due to relatively high fuel and carbon prices, operating costs are higher. The trade-off between the fixed investment cost and the operating cost implies that none of the technologies is strictly dominated by others. Second, $k(c)$ is convex, since it describes the technology frontier. If the function would not be convex, then there exists a technology c which is dominated by a linear combination of two technologies c_1 and c_2, and hence, the technology c should not be part of the frontier. Moreover, we make the following assumption on the shape of the investment cost function.

Assumption 1. *The technology frontier $k(c)$ is assumed to be twice continuously differentiable, convex and log-concave:*

$$\frac{d^2k(c)}{dc^2} > 0 \qquad \textit{and} \qquad \frac{d^2 \ln(k(c))}{dc^2} < 0.$$

Hence, the technology frontier is convex, but not too much. Note that log-concavity implies that the elasticity of the technology frontier

$$\epsilon_k(c) = d\ln(k(c))/d\ln(c)$$

is downward sloping, i.e. $\epsilon_k'(c) < 0$.

In addition, we make INADA-like assumptions, which guarantee interior solutions: the production technology with zero marginal costs $c = 0$ has prohibitively expensive capital costs $k(0) = \infty$ and $k'(0) = -\infty$. For the highest available technology, $\hat{c}$, we assume that the capital cost is zero: $k(\hat{c}) = 0$. It is common in electricity simulation models of generation portfolios to model consumer rationing (rolling blackout), as a fictitious supply-side technology with a zero investment cost and a marginal cost equal to the Value of Lost Load (VOLL). The VOLL corresponds to the maximum that consumers want to pay to prevent blackouts. We assume all consumers have a homogeneous VOLL. Hence, the parameter $\hat{c}$ can be set equal to the VOLL, $\hat{c} = \text{VOLL}$.[1]

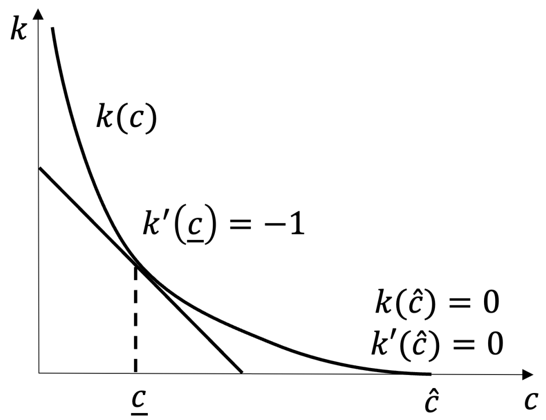


Figure 2: Technology frontier $k(c)$ with $k'(c) \leq 0,\ k''(c) > 0\ \forall c \in (0, \hat{c}]$. $\underline{c}$ stands for the "always-on" technology.

The aggregate investment by producers with marginal costs equal or less than c in the equilibrium is represented by the installed capacity function $G(c)$. By design, a producer with technology c can only invest strictly positive amounts, i.e. $dG(c) > 0$. We use c_0 to indicate the lowest marginal cost such that the investment level starts to be positive with $c > c_0$, i.e. $G(c) \geq 0,\ \forall c \in [c_0, \hat{c}]$.

In the spot market auction, a producer with marginal cost c submits a single bid $b(c)$ for its entire infinitesimal capacity $dG(c)$. We assume that the bidding function $b(c)$ is monotonic. Hence, the producers' investment and bidding strategies can be summarized as $\{G(c), b(c)\}$.

2.2 Demand

As for the demand side, we assume that consumers are price-takers represented by a stochastic inverse demand function with additive price shocks

$$p = P(q) + \varepsilon, \tag{1}$$

[1]It is common in electricity simulation models for generation portfolios to use consumer rationing (by creating black-outs), which is an admittedly crude demand-side technology, as an additional fictitious supply technology with zero investment costs and a marginal cost equal to the Value of Last Load (VOLL).

This corresponds to the state contingent gross utility function $V(q, \varepsilon)$:

$$V(q, \varepsilon) = \int_0^q P(t)dt + \varepsilon \cdot q. \tag{2}$$

Without loss of generality, we normalize demand by setting $P(0) = 0$, such that the demand shock ε is the intercept of the demand function. The demand shock ε follows a known cumulative distribution function $F(\varepsilon)$ over the interval $[\underline{\varepsilon}, \bar{\varepsilon}]$. Denote the quantile function of the demand shock by $\mathcal{Q}(\cdot) = F^{-1}(\cdot)$. The deterministic inverse demand $P(q)$ and the demand shock distribution $F(\varepsilon)$ are known before producers submit bids, but demand shock realization ε is not.

We impose an assumption on the distribution function.

Assumption 2. *The distribution function of demand shocks $\mathcal{Q}$ and the investment cost k satisfy*

$$\mathcal{Q}'(1 + k') > \frac{1}{k''}.$$

Note that this condition is in the primitives of the model the shock distribution ($\mathcal{Q} = F^{-1}$) and the technology frontier $k(c)$. This condition typically requires that the distribution function has a thin upper tail.

Example 1. *For an exponential distribution $F(\varepsilon) = 1 - \exp(-\lambda\varepsilon)$, Assumption 2 implies*

$$-\frac{k''}{k'} > \lambda,$$

where $-k''/k'$ measures the rate at which marginal investment cost decreases when the marginal cost increases by one euro [2]*. λ has to be small enough, such that the rate of decay is always higher than λ.*

Example 2. *If $F(\varepsilon)$ is uniformly distributed on $[\underline{\varepsilon}, \bar{\varepsilon}]$ and $\underset{\sim}{c} = \underline{c}, \tilde{c} = \bar{c}$, with the quadratic investment cost function:*

$$k(c) = \frac{1}{2}\frac{(\bar{c} - c)^2}{\bar{c} - \underline{c}},$$

Assumption 2 requires

$$\bar{\varepsilon} - \underline{\varepsilon} > \bar{c} - \underline{c},$$

indicating that the uncertainty interval of the demand shock distribution needs to be large enough. Additionally, it is noteworthy that for general cost function

[2]Generally, the growth rate for a function $f(x)$ is defined as $\frac{df(x)}{dx} \cdot \frac{1}{f(x)}$.

$k(c) = \frac{1}{\gamma+1}\frac{(\bar{c}-c)^{\gamma+1}}{\bar{c}-\underline{c}}$ *with* $\gamma < 1$, *the uniform demand shock distribution seems not to work.*

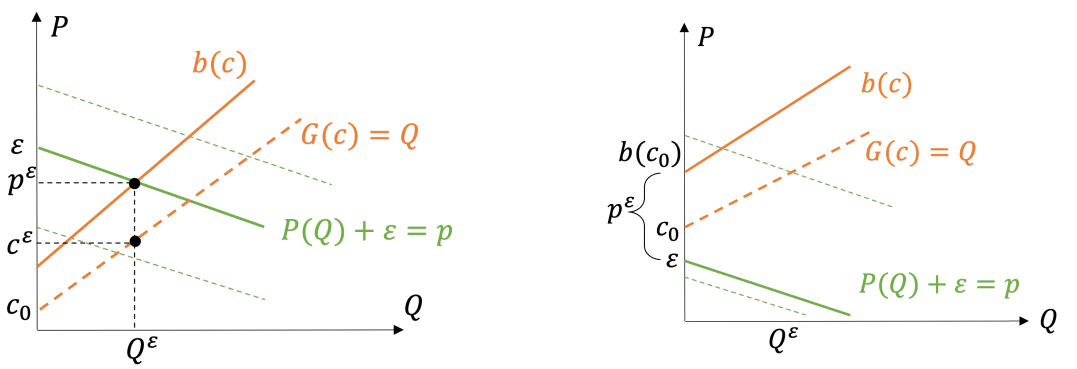
2.3 Market clearing

Based on the producers' investment and bidding strategies $\{G(c), b(c)\}$ and after observing the demand shock ε, the auctioneer determines the equilibrium price and aggregate production quantity by clearing the market. This is shown in Figure 3a.

The auctioneer clears the market and determines the market clearing price p^ε, the clearing quantity q^ε, and the marginal technology c^ε by solving:

$$P(q^\varepsilon) + \varepsilon = b(c^\varepsilon) = p^\varepsilon \qquad \text{and} \qquad q^\varepsilon = G(c^\varepsilon) \tag{3}$$

For small demand shocks ε, the market does not clear and the production level is zero, $q^\varepsilon = 0$. This happens when the demand intercept ε is below the intercept of the supply curve $b(c_0)$. That is, the willingness to pay for the first unit of production is less than the first supply bid. The equilibrium price p^ε is no longer uniquely defined, but lies in between the intercepts: $p^\varepsilon \in (\varepsilon, b(c_0))$. See Figure 3b.



(a) Market clearing with the realized demand shock ε.

(b) No market clearing exists with small demand shock ε.

For convenience, from now on, we will index different states of the world not by the demand shock ε, but by the marginal power plant c. Mathematically, the market clearing condition in the state where producers with technology c are marginal then becomes

$$p(c) = b(c) = P(G(c)) + \varepsilon(c), \tag{4}$$

where the function $\varepsilon(c)$ describes the size of the demand shock for technology c to be marginal.

2.4 Competitive Market Equilibrium

Before constructing the competitive market equilibrium, we first introduce some notations.

A producer with technology c will be selected by the auctioneer to produce when the realized demand shock ε is larger than the shock where its technology would be marginal: $\varepsilon > \varepsilon(c)$. The capacity factor h of a producer with technology c is the likelihood it will be producing and is given by

$$h(c) = \Pr[\varepsilon > \varepsilon(c)] = 1 - F(\varepsilon(c)). \tag{5}$$

The technology $\underline{c}$ that corresponds to $h(\underline{c}) = 1$ (see Figure 2) stands for the "always-on" technology. No technology $c < \underline{c}$ can ever be profitable, as it would run over 100% of the time. A producer with marginal cost $\underline{c}$ has the lowest total cost

$$\underline{c} \cdot 1 + k(\underline{c}).$$

The first-order condition pins down this technology at $h(\underline{c}) = 1$. Recall that Equation 5 links the capacity factor and the distribution of demand shocks. A capacity factor of 1 indicates that the technology $\underline{c}$ is marginal with the smallest possible demand shock $\underline{\varepsilon}$, i.e. $\varepsilon(\underline{c}) = \underline{\varepsilon}$. On the other hand, we define the shutdown boundary marginal cost as $\bar{c}$, which corresponds to $\bar{\varepsilon}$, i.e. $\varepsilon(\bar{c}) = \bar{\varepsilon}$. Technologies with very high marginal costs $c \geq \bar{c}$ will have a zero capacity factor $h(c) = 0$, as they would only become marginal with very high demand shocks which occur with zero probability.

In a uniform-price auction, all producers receive the market clearing price when producing and a producer with technology c collects an expected revenue T of

$$T^{PAB}(c) = \int_{c}^{\infty} b(t)dh(t), \tag{6}$$

while in a pay-as-bid auction, each producer simply sells at its own bid and the producer with technology c receives an expected revenue T of

$$T^{UP}(c) = b(c) \cdot h(c). \tag{7}$$

The total per unit expected profit of a producer with technology c is then:

$$\pi(c) = T(c) - c \cdot h(c) - k(c). \tag{8}$$

It is equal to expected revenue minus expected operating costs and investment cost.

The equilibrium price p is stochastic. Let $Z(p)$ be the cumulative distribution of the price. This price distribution depends on aggregate investments, bidding by producers and the market clearing condition. It is determined by the following condition:

$$Z(p(c)) = F(\varepsilon(c)). \tag{9}$$

The expected consumer surplus is equal to the expected gross utility minus the expected payments to producers:

$$U = \int_0^\infty V(G(c), \varepsilon(c))dh(c) - \int_0^\infty T(c)dG(c) \tag{10}$$

Given the producers' investment and bidding strategies $\{G(c), b(c)\}$, Equations 4 to 9 describe market outcomes, expected profits, and the price distribution. We now formally define the competitive market equilibrium as an extension of the Walrasian equilibrium.

Definition 1 (Investment equilibrium)**.** The set of functions $G(c)$, $b(c)$, $\varepsilon(c)$, $p(c)$ and $Z(p)$ constitutes the investment equilibrium, if

- **short-run:** taking the price distribution $Z(p)$ as given, a producer with marginal cost c finds it optimal to bid $b = b(c)$;
- **long-run:** a producer with marginal cost c makes zero expected profit $\pi(c) = 0$;
- **market clearing:** prices $p(c)$, demand shocks $\varepsilon(c)$ and the price distribution $Z(p)$ are consistent with market clearing (Equations 4 and 9 are satisfied).

The short-run condition is different for pay-as-bid and uniform-price auctions. In pay-as-bid auctions, when the producer bids bid b, it makes a mark-up $b - c$, and is selected with a probability $Z(b)$. The competitive bid should satisfy:

$$b^{PAB}(c) = \arg\max_b (b - c)(1 - Z(b)), \tag{11}$$

In uniform-price auctions, when the producer bids b it imposes a mark-up $p-c$ as long as its bid is below the price $p \geq b$. The competitive bid should satisfy:

$$b^{UP}(c) = \arg\max_{b} \int_{b}^{\infty} (p-c)dZ(p). \tag{12}$$

Instead of looking at the long-run equilibrium, we can also study the short-run equilibrium. Here we take $G(c)$ as exogenous, and drop the long-term condition from the definition of the competitive equilibrium.

Definition 2 (Bidding equilibrium)**.** The set of functions $b(c)$, $\varepsilon(c)$, $p(c)$ and $Z(p)$ constitutes a short-run market equilibrium for an exogenous level of investments $G(c)$ if

- **short-run:** a producer with marginal cost c which takes the price distribution $Z(p)$ as given, finds it optimal to bid $b = b(c)$;
- **market clearing:** prices $p(c)$, demand shocks $\varepsilon(c)$ and price distribution $Z(p)$ are consistent with market clearing (Equations 4 and 9 are satisfied).

3. Analysis

Now we set out to examine the equilibrium under the two alternative auction designs.

Before that, we introduce some notations and terminology that will be used throughout the paper. The baseload investment is represented by $\underset{\sim}{c} = \max\{\underline{c}, c_0\}$. It corresponds either to the marginal cost $\underline{c}$ that is associated with the lowest possible demand shock, or to the technology c_0 where the installed capacity $G(c)$ starts to be positive. The peak-load investment $\tilde{c}$ is defined by $\min\{\bar{c}, \hat{c}\}$, where $\bar{c}$ corresponds to the highest possible demand shock $\bar{\varepsilon}$, $h(\bar{\varepsilon}) = 0$, and $\hat{c}$ corresponds to VOLL. Now, we start with the short-run competitive bidding equilibrium.

3.1 Bidding equilibrium

Lemma 1 (PABA short-run optimal bid; Federico and Rahman 2003)**.** *In pay-as-bid auctions, a producer with technology c charges a mark-up which is equal to the inverse hazard rate of the equilibrium price distribution:*

$$b(c) = c + \frac{1 - Z(b(c))}{Z'(b(c))}.$$

Proof. Taking the first-order condition of Equation 11 w.r.t. $b(c)$, we obtain the optimal bidding strategy in Lemma 1. □

Relating to single-unit auctions, the pay-as-bid auction is similar to the first-price auction. By choosing $b(c)$, producers are faced with a trade-off between the mark-up on cost and the probability of getting accepted. Bidding higher increases the earning but lowers the likelihood of the offer being accepted. The optimal bidding strategy implies that all producers with marginal cost $c \in [\underline{c}, \bar{c})$ set positive mark-ups in pay-as-bid auctions, expect marginal producers that correspond to the highest demand realization, who will set the prices exactly at the marginal costs $(1 - Z(b(\bar{c})) = 0)$.

However, the price distribution $Z(p)$ is endogenous and depends on the bid function $b(c)$. Combining the inverse hazard rate formula and the market clearing conditions, we find Proposition 1

Proposition 1 (PABA short-run optimal bidding strategy)**.** *In a pay-as-bid auction, the optimal bid $b(c)$ and the capacity factor $h(c)$ are determined by the following set of equations:*

$$h(c) = -\frac{d}{dc}\left[(b(c) - c)h(c)\right], \tag{13}$$

$$b(c) = \mathcal{Q}(1 - h(c)) + P(G(c)), \tag{14}$$

together with the boundary condition that for technology $\bar{c}$ that corresponds to the highest demand shock with capacity factor $h(\bar{c}) = 0$, the mark-up is zero: $b(\bar{c}) = \bar{c}$. The set of equation is determined on the interval $[\max\{c_0, \underline{c}\}, \hat{c}]$, which corresponds to the technologies with positive investment levels $c > c_0$ and which lie on the margin with positive probability $\underline{c} \leq c \leq \bar{c}$. Marginal cost $\underline{c}$ is determined as the technology which has an equilibrium capacity factor of 1, $h(\underline{c}) = 1$

Proof. Together with Equation 4, 5 and 9, we have

$$1 - Z(b(c)) = h(c). \tag{15}$$

Taking derivatives w.r.t. c on both sides gives

$$Z'(b(c)) = -\frac{h'(c)}{b'(c)}. \tag{16}$$

Plugging Equation 15 and 16 into Lemma 1, we have

$$b(c) = c - \frac{h(c)b'(c)}{h'(c)} \Rightarrow (b(c) - c)h'(c) - (b'(c) - 1)h(c) = h(c),$$

which gives the differential equation 13. Equation 14 can be obtained by combining Equation 4 and 5. □

The next proposition directly follows from the first-order condition of the profit function for uniform-price auctions (Equation 12).

Proposition 2 (UPA optimal bidding strategy)**.** *In a uniform-price auction, the optimal bid $b(c)$ and the capacity factor $h(c)$ are determined by the following set of equations:*

$$\begin{aligned} b(c) &= c, \\ b(c) &= \mathcal{Q}(1 - h(c)) + P(G(c)). \end{aligned}$$

Uniform-price auctions are similar to second-price auctions, if we relate to single-unit auctions. Producers bid aggressively, since they are not paying the exact amount that they bid instead of the first rejected bid. Bidding higher does not directly reduce the earning, but lowers the likelihood of winning the auction. The optimal bidding strategy reveals that all producers will bid exactly their marginal costs. In this way, producers that have a higher valuation will be first assigned to operate, so uniform-price auctions are efficient.

3.2 Investment equilibrium

Before constructing the investment equilibrium for the two auction formats, we need the following lemma.

Lemma 2 (Local incentive constraint)**.** *Independent of auction format, the capacity factor satisfies the free entry condition. This implies*

$$h(c) = -\frac{dk(c)}{dc}. \tag{17}$$

Proof. Recall that the profit of a producer with technology c is given by

$$\pi(c) = T(c) - c \cdot h(c) - k(c).$$

Taking the total differential, we have

$$\frac{d\pi}{dc} = \frac{\partial \pi}{\partial T} \cdot T'(c) + \frac{\partial \pi}{\partial h} \cdot h'(c) + \frac{\partial \pi}{\partial k} \cdot k'(c) + \frac{\partial \pi}{\partial c}. \tag{18}$$

By the envelope theorem, the sum of the first two terms on the right hand side of Equation 18 equals zero. In the long run, all generators earn zero profit due to the free entry:

$$\pi(c) = 0 \Rightarrow \frac{d\pi}{dc} = 0$$

Hence,

$$\frac{\partial \pi}{\partial k} \cdot k'(c) + \frac{\partial \pi}{\partial c} = 0 \;\Rightarrow\; h(c) = -\frac{dk(c)}{dc}.$$

□

Pay-as-bid auctions can be regarded as a screening model. Each producer has private information about its marginal cost. Typically, the producers with the highest marginal costs will trade the same amount as in the first best (Fudenberg and Tirole, 1991). However, in uniform-price auctions, producers truthfully bid their marginal cost, and in this way, the formerly hidden information about costs is revealed. The auctioneer has complete and perfect information.

In following propositions, we characterize the investment equilibrium for the two alternative auction designs, respectively. We mainly focus on producers' bidding behaviors and investment decisions in the equilibrium.

Proposition 3 (PABA investment equilibrium)**.** *In the investment equilibrium of a pay-as-bid auction, producers with marginal cost $c \in [\underset{\sim}{c}, \tilde{c})$ invest. The cumulative installed capacity $G(c)$ satisfies*

$$P(G(c)) = c + \frac{k(c)}{h(c)} - \mathcal{Q}(1 + k'(c)), \tag{19}$$

and the bid follows

$$b(c) = c + \frac{k(c)}{h(c)}, \tag{20}$$

where the mark-up depends on the elasticity of the investment $k(c)$.

Proof. We combine Equation 13 with the free entry condition 17 $h(c) = -k'(c)$ and find:

$$\frac{d}{dc}\left[(b(c) - c)k'(c)\right] = -k'(c). \tag{21}$$

Integrating this differential equation over the interval $[c, \tilde{c}]$, and using the fact that that there is no distortion at the top $b(\tilde{c}) = \tilde{c}$ and that the capital cost $k(c)$ at $\tilde{c}$

is zero, we find that the bidding mark-up is equal to $-k(c)/k'(c)$:

$$b(c) = c - \frac{k(c)}{k'(c)} = c + \frac{k(c)}{h(c)}.$$

Plugging the optimal bid into the market clearing condition 4, we obtain the equation that pins down the cumulative installed capacity $G(c)$ in Equation 19. □

Taking the derivative of Equation 20 and using the properties of the technology frontier shows that slope of the bidding function

$$b' = \frac{k''k}{k'^2}$$

is positive. The mark-ups are decreasing, following the property of the investment cost function in Assumption 1. The slope of the bidding function $b'(c)$ is in the range of $(0, 1)$. All technologies will be used in equilibrium, if G is strictly increasing: $G'(c) > 0 \; \forall c \in [\underset{\sim}{c}, \tilde{c}]$. To see this, firstly, recall that the log-concavity property of the technology frontier in Assumption 1 implies

$$\frac{1}{k''} > \frac{k}{k'^2}.$$

Taking the derivative of 14, and using the fact that $P' < 0$, we see that $G' > 0$ requires

$$\mathcal{Q}'(1 + k') > \frac{k}{k'^2}.$$

Thanks to Assumption 2, this is satisfied.

The following lemma follows directly from Proposition 3.

Lemma 3. *The Lerner index in pay-as-bid auctions is the reciprocal of the elasticity $\epsilon_k(c)$ of investment costs:*

$$L = \frac{b(c) - c}{c} = \frac{k(c)}{|k'(c)| \cdot c} := \frac{1}{\varepsilon_k(c)}.$$

The Lerner index indicates that firms are able to charge over its marginal cost in the competitive equilibrium. However, the mark-up is not due to market power. It is necessary to recoup the investment cost with the mark-up.

Proposition 4 (UPA investment equilibrium)**.** *In the investment equilibrium of a uniform-price auction, producers with marginal costs $c \in [\underset{\sim}{c}, \tilde{c})$ make investment.*

The cumulative installed capacity $G(c)$ satisfies

$$P(G(c)) = c - \mathcal{Q}(1 + k'(c)).$$

A producer with technology c bids

$$b(c) = c.$$

Proof. First, we check $G(c)$ is increasing:

$$G' > 0 \Leftrightarrow \mathcal{Q}'(1 + k') > \frac{1}{k''}.$$

This is guaranteed by Assumption 2. Next, we prove that a producer with technology c_o would truthfully report his marginal cost and bid according to the optimal bidding function $b(c)$. Suppose not and the producer reports c or bid $b(c)$ instead and its profit is given by

$$\pi(c, c_o) = \pi(c_o) + h(c_o)(c_o - c) - k(c) + k(c_o),$$

where $\pi(c_o) = 0$ due to the free entry condition. The producer would optimally choose his report:

$$\frac{\partial \pi(c, c_o)}{\partial c} = -h(c_o) - k'(c) = 0,$$

which is achieve at $c = c_o$, implying that the producer would truthfully report his marginal cost. Then check the second-order condition for all c

$$\frac{\partial^2 \pi(c, c_o)}{\partial c^2} = -k''(c) \leq 0.$$

Hence, producers would not deviate and truthfully report their marginal costs. □

4. A Functional-Form Model

We now look in more details with a functional-form model. We make some reasonable assumptions to provide some insights of our model framework. In particular, we compare the two auction formats in terms of firms' bidding behaviors and investment incentives.

Assume that the demand function is linear, i.e. $P(q) = -\rho q$ with $\rho > 0$. The investment cost function is convex

$$k(c) = \frac{\alpha}{\gamma + 1} \frac{(\bar{c} - c)^{\gamma+1}}{\bar{c} - \underline{c}},$$

with $\alpha > 0$, which satisfies Assumption 1. The demand shocks follow the exponential distribution

$$F(\varepsilon) = 1 - exp(-\lambda\varepsilon), \; \lambda > 0$$

on the interval $[0, \infty)$. Following Example 1, for an increasing installed capacity function $G(c)$, the exponential distribution of demand shocks indicates that the rate of decay of marginal investment cost is always higher than the parameter λ.

The producers' optimal bidding strategy and installed capacity in pay-as-bid auctions versus uniform-price auctions can be illustrated as in Figure 4.

In the short-run bidding equilibrium, producers always set prices exactly at their marginal cost level in uniform-price auctions. On the other hand, producers make positive mark-ups in pay-as-bid auctions. There is no distortion only at the top for producers with highest-marginal cost technology and hence, more consumer surplus is extracted in pay-as-bid auctions.

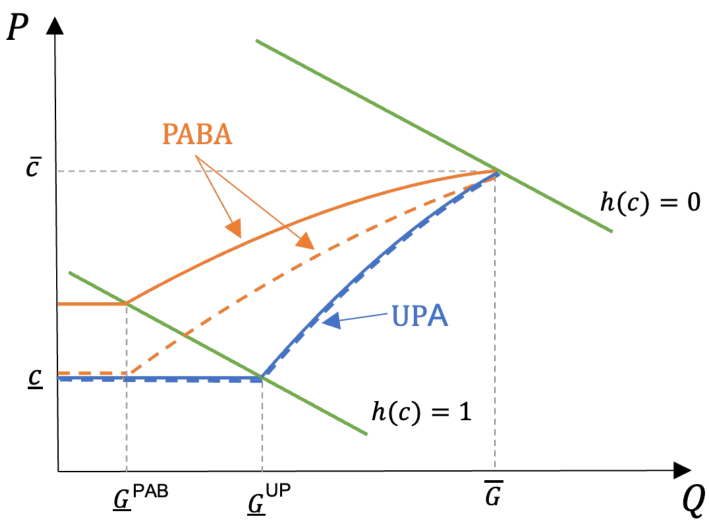


Figure 4: Comparison of bidding (solid) and portfolios (dashed) between pay-as-bid auctions and uniform-price auctions.

Figure 4 illustrates the optimal bidding strategies and optimal installed capacity in pay-as-bid auctions versus uniform-price auctions. Our results show that uniform-price auction is efficient, and it generates similar outcomes as peak-load pricing in Boiteux (1960). However, this is not the case for pay-as-bid auctions. The equilibrium price in the pay-as-bid auction is higher for each infinitesimal unit of production and the initial installed capacity is lower for pay-as-bid auctions. The two auction formats end up with the same equilibrium price and the same installed capacity at the top. Aggregate investments in uniform-price auctions first-order stochastically dominate investments in pay-as-bid auctions, except for the technology at the top, that is,

$$G^{PAB}(c) \leq G^{UP}(c) \text{ with equality for } c = \bar{c}.$$

The Lerner index in pay-as-bid auctions is the reciprocal of the elasticity $\epsilon_k(c)$ of investment costs:

$$L = \frac{b(c) - c}{c} = \frac{k(c)}{|k'(c)|c} := \frac{1}{\epsilon_k(c)}.$$

The inefficiency here does not necessarily originate from market power. Producers mark up on their marginal costs so as to recoup their investment costs. This implies that inefficiency could come from the market design itself.

To compare producers' investment incentives between the two auction formats, aggregate investments are identical as in Fabra et al. (2011) in both auctions. However, the generation mix is distorted. In pay-as-bid auctions, there are fewer investments in the baseload capacity, while all intermediate technologies are over-invested.

All producers make zero profits, and hence, consumer surplus equals welfare in our model. For high demand realizations, consumer surplus is higher and the volume stays the same in pay-as-bid auctions. Consumers pay less. On the other hand, for low demand realizations, consumer surplus is lower, since the volume is smaller and prices are higher in pay-as-bid auctions. Uniform-price auctions are efficient with higher consumer surplus compared to pay-as-bid auctions. We can gain some intuition from the result: if consumers are heterogenous, pay-as-bid auctions benefit consumers that consume a lot when demand is high.

It is also noteworthy that the revenue equivalence theorem does not hold in our model. This is because the demand here is elastic. By using different auction formats, the total size of the cake does not stay the same. Hence, the two auction formats do not give the same expected revenue.

5. Conclusions

Our paper is motivated by the recent proposal of ACER that suggests replacing the current wholesale electricity market design from uniform-price auctions to other possible alternatives. The purpose of this work is to address the question how auction formats affect electricity producers' long-run and short-run decision-making and to gain an all-round understanding of the impact of auction formats on market outcomes and social welfare. To the best of our knowledge, not much literature dedicates to modeling the investment incentives of a variety of technologies in the auctions. To make this happen, we construct a perfect-competitive model with a continuum of generation technologies and consider the demand responsiveness and stochasticity. As an example, we turn to a functional-form with some reasonable assumptions to illustrate our main results.

Our research demonstrates that inefficiency does not necessarily originate from market power. It could come from market design. We compare the two auction formats and examine producers' bidding behaviors and price-cost mark-ups in the short run, as well as the investment and generation portfolio in the long run. Our results show that pay-as-bid auctions are inefficient. In the short run, producers' willingness to pay exceeds their marginal costs, indicating that allocative inefficiency arises. In the long run, the revenue of base-load producers is depressed at high demand realizations, resulting in the decrease in investment incentives and the distortion of generation mix. Also, consumer surplus is lower in pay-as-bid auctions.

For future research, we consider to allow for some bunching in the investment so that our model could include the more general case where some intermediate technologies are not used. Additionally, the market structure we focus on in this work is perfect competition. However, in the discussion of market performance, market design and market structure shall not be isolated. So on top of perfect competition, we would like to introduce market power to the generation side and gain more understanding in designing the electricity market.

References

ACER (2021). ACER's Preliminary Assessment of Europe's high energy prices and the current wholesale electricity market design - Main energy price drivers, outlook and key market characteristics.

Boiteux, M. (1960). Peak-load pricing. *The Journal of Business*, 33(2):157–179.

Bower, J. and Bunn, D. (2001). Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the england and wales electricity market. *Journal of economic dynamics and control*, 25(3-4):561–592.

European Commission (2022). REPowerEU: Joint European action for more affordable, secure and sustainable energy.

Evans, J. and Green, R. C. (2003). Why did british electricity prices fall after 1998?

Fabra, N., von der Fehr, N.-H., and Harbord, D. (2006). Designing electricity auctions. *The RAND Journal of Economics*, 37(1):23–46.

Fabra, N., von der Fehr, N.-H. M., and de Frutos, M.-A. (2011). Market design and investment incentives. *The Economic Journal*, 121(557):1340–1360.

Federico, G. and Rahman, D. (2003). Bidding in an electricity pay-as-bid auction. *Journal of Regulatory Economics*, 24(2):175–211.

Fudenberg, D. and Tirole, J. (1991). *Game theory*. MIT press.

Harbord, D. and McCoy, C. (2000). Mis-designing the uk electricity market? *European Competition Law Review*, 21(5):258–260.

Holmberg, P. (2009). Supply function equilibria of pay-as-bid auctions. *Journal of Regulatory Economics*, 36(2):154–177.

Joskow, P. and Tirole, J. (2007). Reliability and competitive electricity markets. *The RAND Journal of Economics*, 38(1):60–84.

Kahn, A. E., Cramton, P. C., Porter, R. H., and Tabors, R. D. (2001). Uniform pricing or pay-as-bid pricing: a dilemma for California and beyond. *The Electricity Journal*, 14(6):70–79.

Steiner, P. O. (1957). Peak loads and efficient pricing. *The Quarterly Journal of Economics*, 71(4):585–610.

Tierney, S. F., Schatzki, T., and Mukerji, R. (2008). Uniform-pricing versus pay-as-bid in wholesale electricity markets: does it make a difference? *New York ISO*.

Wolfram, C. D. (1999). Electricity markets: Should the rest of the world adopt the united kingdom's reforms. *Regulation*, 22:48.

Zöttl, G. (2010). A framework of peak load pricing with strategic firms. *Operations Research*, 58(6):1637–1649.